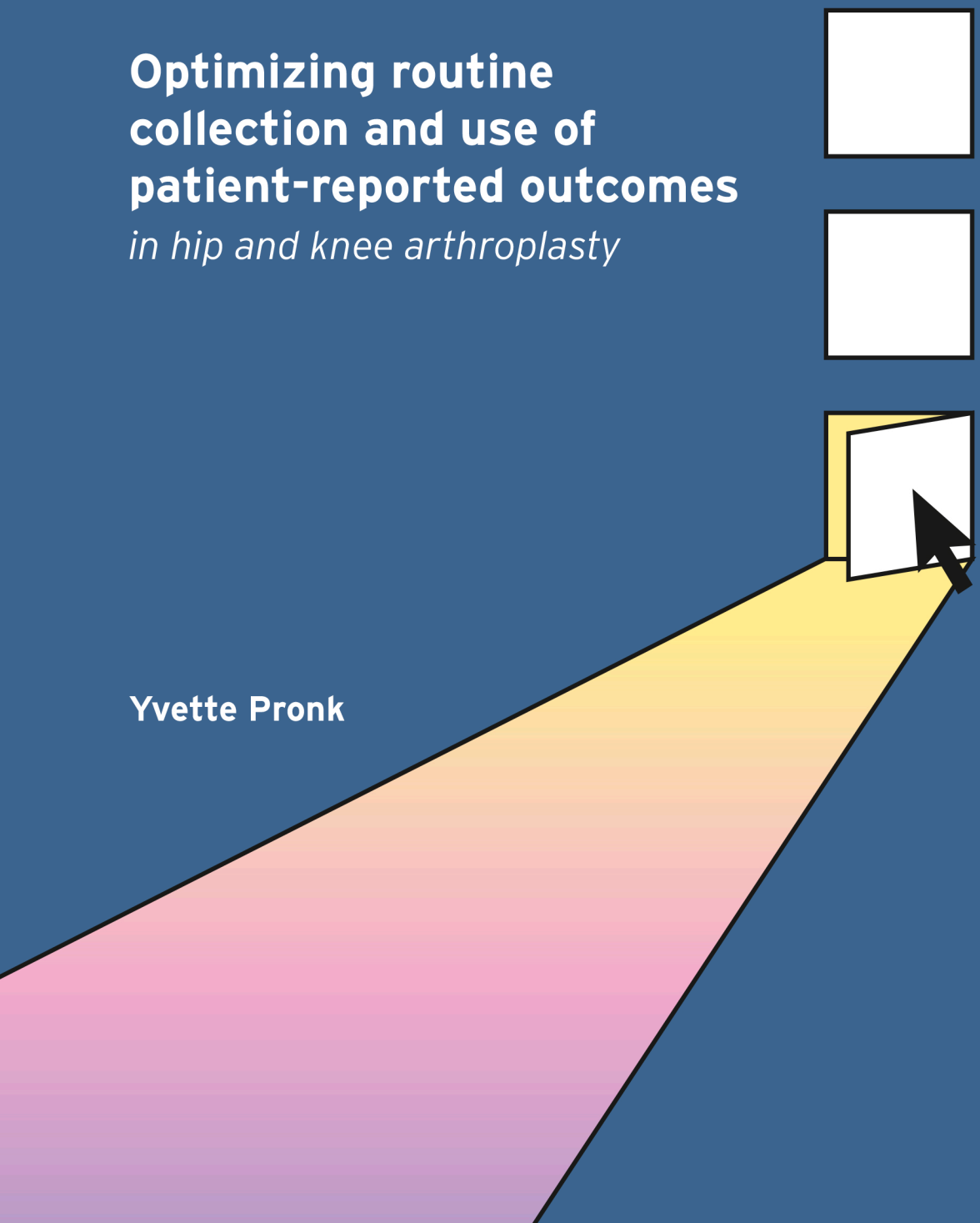
Optimizing routine collection and use of patient-reported outcomes

in hip and knee arthroplasty



Yvette Pronk

Optimizing routine collection and use of patient-reported outcomes in hip and knee arthroplasty

Yvette Pronk

Colofon

Publication of this thesis was kindly supported by:
Kliniek ViaSana
Radboud Universiteit
Nederlandse Orthopaedische Vereniging
Interactive Studios

Printing Ridderprint, ridderprint.nl
Layout and design Indah Hijmans, persoonlijkproefschrift.nl & Hilde Peters, vanhilde.nl
Cover design and illustration Hilde Peters, vanhilde.nl

ISBN 978-94-6483-412-3

Optimizing routine collection and use of patient-reported outcomes in hip and knee arthroplasty

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 15 december 2023
om 12.30 uur precies

door

Yvette Pronk
geboren op 1 juni 1991
te 's-Hertogenbosch

Promotor

Prof. dr. B.W. Schreurs

Copromotoren

Dr. W. van der Weegen (Anna Ziekenhuis)

Dr. J.M. Brinkman (Kliniek ViaSana)

Dr. P. Pilot (IMUKA)

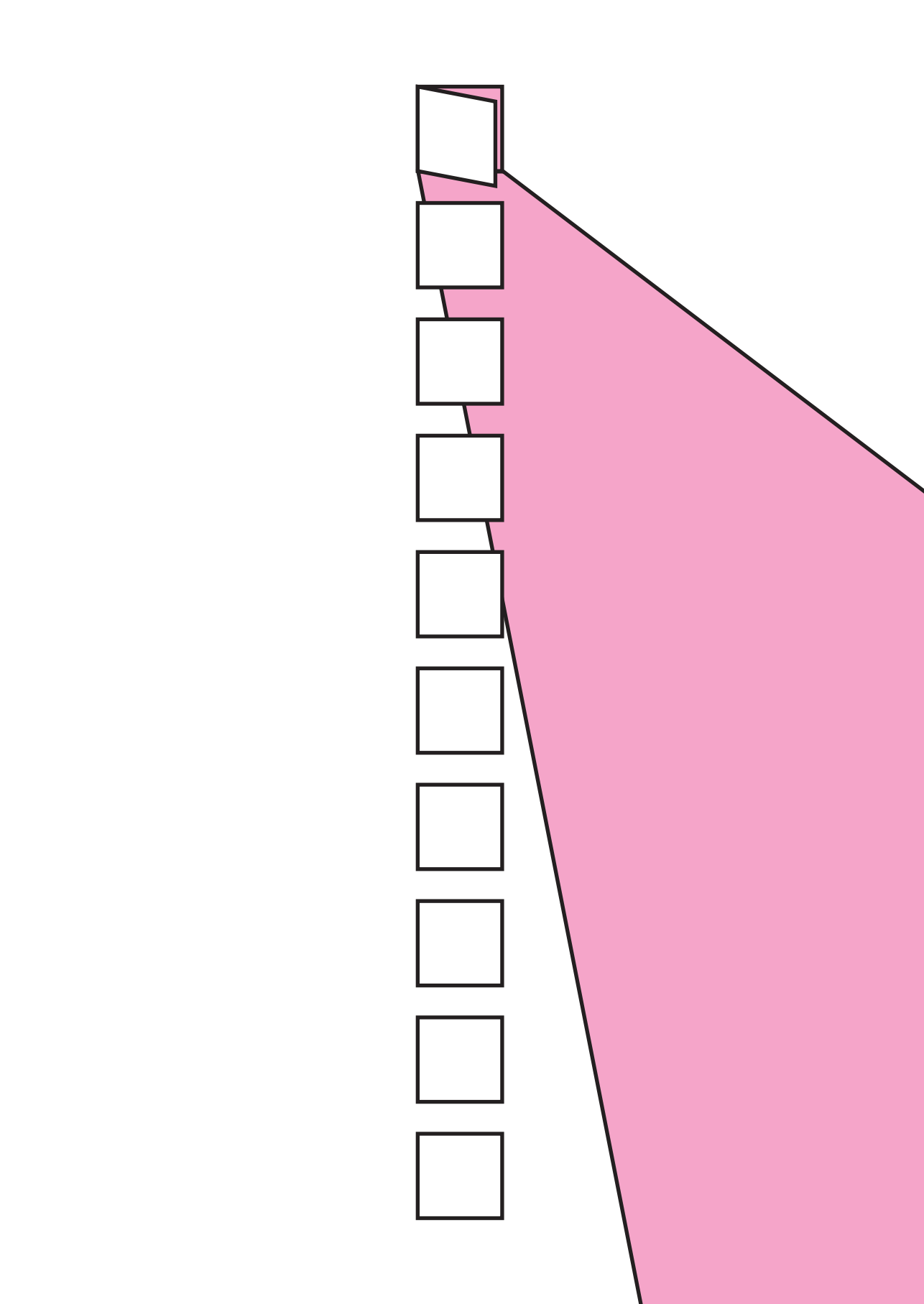
Manuscriptcommissie

Prof. dr. P.J. van der Wees

Prof. dr. D.L. Gerritsen

Prof. dr. R.W. Poolman (Universiteit Leiden)

Contents



CHAPTER 1

General introduction

Traditionally, orthopaedic surgeons were satisfied with the outcome of an arthroplasty if the prosthesis alignment was correct, and the implant was well fixed and balanced. They considered the long term outcome to be optimal if an excellent implant survival was obtained. However, patients are satisfied with the outcome if their pain is relieved, their function is restored and their quality of life has improved [1–4]. With the shift to a more patient centred orthopaedic health care, measuring patient-reported outcomes (PROs) using selected patient-reported outcome measures (PROMs) has increased [5].

PROs are subjective outcomes, such as pain relief or functional improvement, scored directly by the patient without interpretation by others. PROMs are questionnaires to measure these PROs. Currently, PROMs are seen as the gold standard for measuring outcomes from a patients' perspective.

In multiple countries collecting PROs of hip and knee arthroplasty patients to evaluate and to improve health care is recommended or even mandatory. Unfortunately, daily practice shows a large diversity between health care institutions in their success of PRO collection and in how they use PROs to improve health care. Wide ranges in response rates (RRs) are observed in both the Dutch and international arthroplasty registers [6, 7]. A recent study among surgeons concluded that the most important constraint on implementing PRO collection was costs [8]. Although collecting PROs of all patients seems the most desired situation, a more realistic approach is to study which RR at what costs is feasible (Chapter 3). To draw valid conclusions on PROs and to justify the costs for PRO collection, a certain RR on PROMs is needed. However, this called minimum RR (MRR) remains unknown. Despite many unknown factors, it is important to create a scientific starting point for this discussion (Chapter 4).

When PRO collection is optimized, it is logical that PROs should be used for the aim they are collected for. In the Netherlands, these PROs are publicly available to create transparency of the delivered care [7]. However, it is questionable if the aim 'improving arthroplasty health care' is achieved at the Dutch national level (Chapter 5). Ideas and examples on how PROs could be used to optimize health care are available [9–11]. However, scientific investigated examples from daily health care of routinely PRO use to optimize arthroplasty health care are lacking. PROs could be used to gain knowledge useful for shared decision making and for making recommendations to stakeholders (Chapters 6 and 7), in health care evaluation (Chapter 6), to guide patients after surgery (Chapter 8) and to triage patients to their suitable type of consultation (Chapter 9).

As mentioned, the routine PRO collection and use have been increased, however, multiple PRO-related questions remain unanswered. From that perspective, it is questionable if the PRO collection and use executed nowadays are justifiable from an ethical and value-based health care perspective. Although examples and recommendations how to collect [12–20] and use PROs exist [9–11], scientific evidence on how to optimize routine PRO collection and how to optimize health care with routine use of PROs in hip and knee arthroplasty is lacking.

AIM OF THIS THESIS

The aim of this thesis is to investigate how routine PRO collection can be optimized (part I) and subsequently how health care can be optimized with routine use of PROs (part II) in hip and knee arthroplasty.

OUTLINE OF THIS THESIS

Background information on the substantial health care burden called osteoarthritis, arthroplasty registries, PROs and PROMs, and current practice on PRO collection and PRO use is presented in **chapter 2**. Based on the two aims, this thesis is divided in two parts.

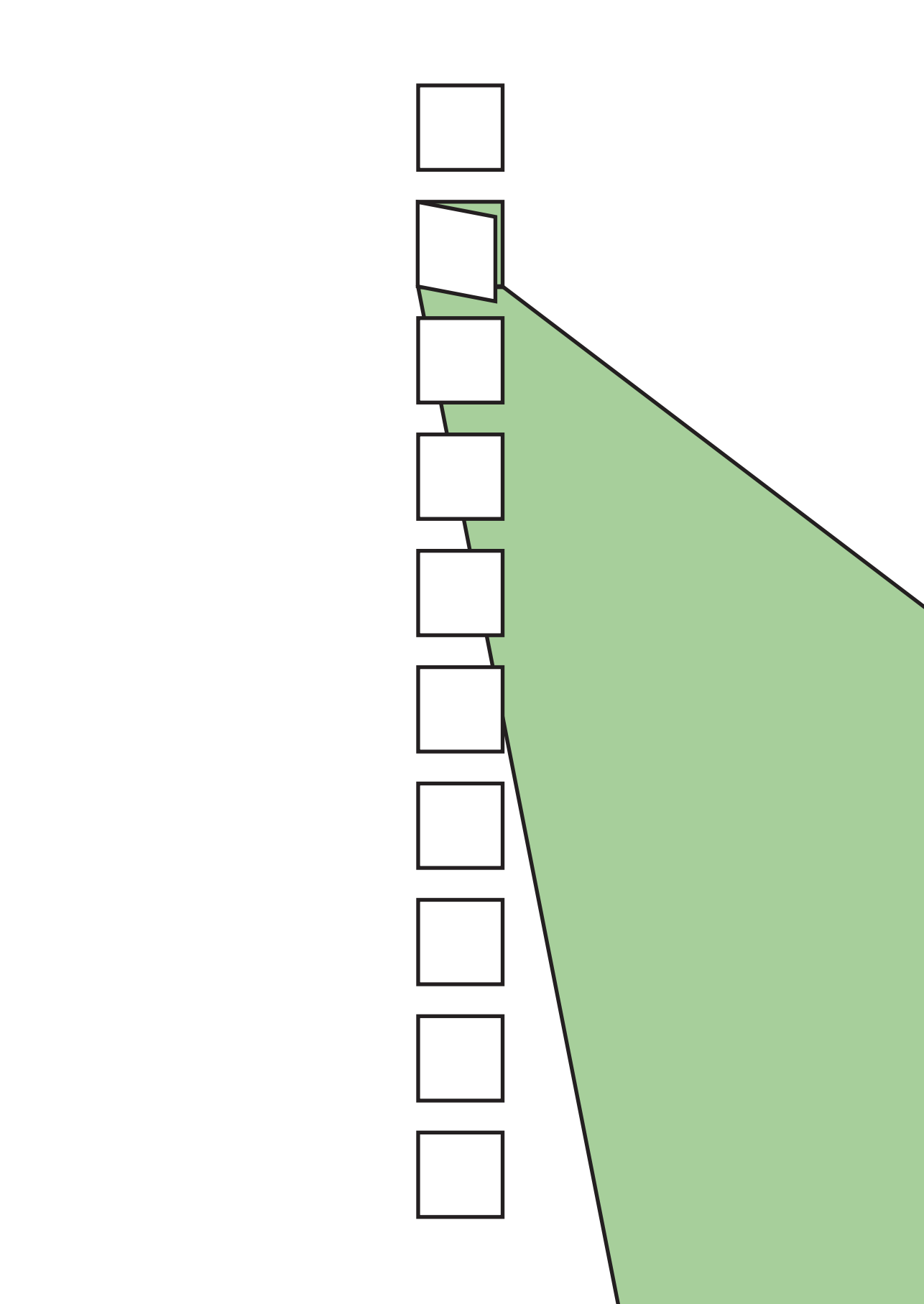
Part I of this thesis (Chapters 3 and 4) is focused on how routine PRO collection could be optimized. What RR is achievable against which costs? This is described for all orthopaedic surgical procedures as well as specified for total hip arthroplasty (THA), total knee arthroplasty (TKA) and unicompartmental knee arthroplasty (UKA), and anterior cruciate ligament reconstruction in **chapter 3**. This retrospective cohort study with prospectively collected data describes the RR and costs for automated PRO collection alone compared to combined automated and manual collection. **Chapter 4** provides an insight into the MRR on PROMs needed to adequately evaluate THAs in a retrospective study with prospective collected data.

Part II of this thesis (Chapters 5 to 9) focuses on how routine use of PROs could be helpful in optimizing health care in hip and knee arthroplasty. In **chapter 5** a longitudinal study with Dutch national THA indicator datasets, publicly available since 2016, examines if the goal of improving THA health care by evaluating outcomes from a patients' perspective based on PROs is achieved. Thereafter, scientifically investigated examples are given how PROs could be used in clinical practice to optimize health care. In **chapter 6**, PROs are used to compare two different UKA implant designs in a single high-volume surgeon, retrospective cohort study with prospectively collected data study. This study investigates which of the two frequently used implants, a mobile or fixed bearing design, should be used in daily health care. **Chapter 7** focuses on predicting patient satisfaction after a TKA using patient characteristics and preoperative PROs in a retrospective cohort study with prospectively collected data. In **chapter 8** a randomized controlled trial is described on the effects of an eHealth app (PainCoach app) on pain control and opiate use during the first 2 weeks at home after a TKA. In response to a patient's input of the pain experienced, as a PRO, the PainCoach app gives advice. Performed during COVID-19 pandemic, in **chapter 9** a pilot study with expert panels and a retrospective cohort with prospectively collected data describes the use of PROs to create a tool to triage THA patients to hospital or video consultation for their 6 weeks postoperative control visit.

The summary and general discussion are presented in **chapter 10**. The main findings are highlighted and reflected on, and recommendations for future steps are suggested.

REFERENCES

1. Söderman P, Malchau H, Herberts P, Zügner R, Regnér H, Garellick G (2001) Outcome after total hip arthroplasty: Part II. Disease-specific follow-up and the Swedish National Total Hip Arthroplasty Register. Acta Orthop Scand 72:113–119
2. Söderman P, Malchau H, Herberts P (2000) Outcome after total hip arthroplasty: Part I. General health evaluation in relation to definition of failure in the Swedish National Total Hip Arthroplasty register. Acta Orthop Scand 71:354–359
3. Garellick G, Malchau H, Herberts P (2000) Survival of Hip Replacements. Clin Orthop Relat Res 375:157–167
4. Britton AR, Murray DW, Bulstrode CJ, McPherson K, Denham RA (1997) Pain Levels After Total Hip Replacement: Their Use As Endpoints For Survival Analysis. J Bone Jt Surg 79:93–98
5. Wylde V, Blom AW (2011) The failure of survivorship. J Bone Joint Surg Br 93-B:569–570
6. Bohm ER, Kirby S, Trepman E, et al (2021) Collection and Reporting of Patient-reported Outcome Measures in Arthroplasty Registries: Multinational Survey and Recommendations. Clin Orthop Relat Res 479:2151–2166
7. Zorginstituut Nederland. Zorginzicht - Open data Medisch-specialistische zorg. https://www.zorginzicht.nl/openbare-data/open-data-ziekenhuizen-en-zelfstandige-behandelcentra---medisch-specialistische-zorg. Accessed 24 Nov 2021
8. Joeris A, Knoll C, Kalampoki V, Blumenthal A, Gaskell G (2018) Patient-reported outcome measurements in clinical routine of trauma, spine and craniomaxillofacial surgeons: between expectations and reality: a survey among 1212 surgeons. BMJ Open 8:e020629
9. Van Schie P, Van Bodegom-Vos L, Zijdeman TM, Nelissen RGHH, Marang-Van De Mheen PJ (2021) Awareness of performance on outcomes after total hip and knee arthroplasty among Dutch orthopedic surgeons: how to improve feedback from arthroplasty registries. Acta Orthop 92:54–61
10. Boyce MB, Browne JP, Greenhalgh J (2014) The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. BMJ Qual Saf 23:508–518
11. Baumhauer JF, Bozic KJ (2016) Value-based Healthcare: Patient-reported Outcomes in Clinical Decision Making. Clin Orthop Relat Res 474:1375–1378
12. Ho A, Purdie C, Tirosh O, Tran P (2019) Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. Patient Relat Outcome Meas 10:217–226
13. Patel J, Lee JH, Li Z, SooHoo NF, Bozic K, Huddleston JI (2015) Predictors of Low Patient-Reported Outcomes Response Rates in the California Joint Replacement Registry. J Arthroplasty 30:2071–2075
14. Tariq MB, Vega JF, Westermann R, Jones M, Spindler KP (2019) Arthroplasty studies with greater than 1000 participants: analysis of follow-up methods. Arthroplast Today 5:243–250
15. Slover JD, Karia RJ, Hauer C, Gelber Z, Band PA, Graham J (2015) Feasibility of integrating standardized patient-reported outcomes in orthopedic care. Am J Manag Care 21:e494-500
16. Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I, Cooper R, Felix LM, Pratap S (2009) Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2010:2009–2011
17. Hunter J, Corcoran K, Leeder S, Phelps K (2013) Is it time to abandon paper? The use of emails and the Internet for health services research - a cost-effectiveness and qualitative study. J Eval Clin Pract 19:855-61
18. Brook EM, Glerum KM, Higgins LD, Matzkin EG (2017) Implementing Patient-Reported Outcome Measures in Your Practice: Pearls and Pitfalls. Am J Orthop (Belle Mead NJ) 46:273–278
19. Franklin PD, Chenok KE, Lavalee D, Love R, Paxton L, Segal C, Holve E (2017) Framework To Guide The Collection And Use Of Patient-Reported Outcome Measures In The Learning Healthcare System. eGEMs (Generating Evid Methods to Improv patient outcomes) 5:17
20. Rolfson O, Bohm E, Franklin P, et al (2016) Patient-reported outcome measures in arthroplasty registries: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty RegistriesPart II. Recommendations for selection, administration, and analysis. Acta Orthop 87:9–23



CHAPTER 2

Background

Osteoarthritis (OA), one of the major chronic diseases in the world, results in pain and functional restrictions. Hip and knee arthroplasties are successful treatments for patients with end-stage OA. Arthroplasty registries gain insights into the quality of the delivered health care. In the last decade, multiple registries add patient-reported outcomes (PROs) collected with selected patient-reported outcome measures (PROMs) to extend this insight with a patients' perspective.

SUBSTANTIAL HEALTH CARE BURDEN: OSTEOARTHRITIS

In the Netherlands alone, more than 1.5 million people suffer from OA. Half of them suffer from knee OA and one-third suffer from hip OA [1]. These numbers will increase with 40% between 2015 and 2040 based on the increase of the population, the ageing of the population, and the expectation of more people suffering from overweight [2]. When conservative treatments fail, total hip arthroplasty (THA) and total knee arthroplasty (TKA) are effective treatments for patients with end-stage hip or knee OA to relieve pain, restore function and improve quality of life. If OA is only located medial or lateral of the knee joint, an unicompartmental knee arthroplasty (UKA) can be a successful treatment option. The outcome of an arthroplasty depends on multiple factors such as surgeon, patient, type of implant, surgical approach, pain management and rehabilitation. Implant survival has been the most commonly reported outcome variable. Valuable insights in implant survival are achieved with arthroplasty registries. However, more factors should be taken into account when assessing the quality of the delivered arthroplasty health care.

ARTHROPLASTY REGISTRIES

Multiple joint arthroplasty registries are united in the International Society of Arthroplasty Registries (ISAR). The ISAR aims to improve outcomes for arthroplasty patients worldwide [3]. In the Netherlands, all arthroplasties performed are registered in the Dutch arthroplasty register (LROI). The two main goals of this registry are tracing implants in case of calamity and gaining insight into quality of the delivered health care [4]. This registry was developed by the Dutch orthopaedic association in 2007. In 2019, 33,253 THAs were performed of which 85% for the diagnosis primary OA [5, 6]. Furthermore, 25,881 TKAs and 4,892 UKAs were performed of which 97% for the diagnosis primary OA [7, 8]. Since 2014 health care institutions are required to register PROs in the Dutch arthroplasty register. The Dutch orthopaedic association has developed his first PROMs advice in 2012 which has been updated in 2020 [9, 10]. The ISAR PROMs Working Group evaluates and advises on best practices in the selection, administration and interpretation of these PROs and PROMs. It also supports the adoption and use of PROMs and PROs arthroplasty registries worldwide [11].

The relevance of registry data depends largely on the quality of the collected data such as coverage and completeness of registration, validity and reliability of metrics, as well as on patients' response rate (RR) on questionnaires. Completeness of the registered arthroplasties

in the Dutch arthroplasty register, based on the hospital information system, is 99% [13]. In total, 93% of the THAs and 96% of the TKAs and UKAs are valid registered procedures in the Dutch arthroplasty register [12]. However, these percentages are without PROs.

PROS AND PROMS

Long term implant survival rate is of less importance to arthroplasty patients. They are more focused on pain relief and functional improvement to increase their participation in daily activities [14–17]. Pain relief and functional improvement are examples of PROs. PROs are subjective outcomes scored directly by the patient without interpretation by others. The term outcome may be confusing as it implies a measurement that occurs only after an intervention. Valuable outcomes involve the measurement of change and require repeated measures, so before and after an intervention or during a certain time [11].

PROs are measured using PROMs. PROMs are questionnaires that can be suitable for the general population (generic PROMs) or specific for a certain diagnosis or disease or patient group (disease-specific PROMs). Furthermore, these questionnaires can exist of multiple questions or a single question. PROMs were initially restricted to clinical research. Nowadays, PROMs have become an internationally accepted method and a gold standard to gain insight into outcomes from a patients' perspective.

PROs and PROMs in the Netherlands

Measuring and creating transparency of quality of health care gained more attention in the Netherlands in 2004 [18]. The government advised health insurance companies to focus not only on price and product, but also on quality of health care. Since 2006, there is a shift towards a more patient centred perspective in health care. As a result, the use of PROMs to measure PROs has increased [19].

In 2012, the Dutch orthopaedic association advised Dutch health care institutions to begin PRO collection of patients diagnosed with primary OA and planned for an arthroplasty with selected PROMs on selected measurement time points including ranges for valid response periods (table 1) [20]. The Dutch orthopaedic association aimed to improve health care by evaluating outcomes from a patients' perspective [9].

Table 1 Dutch hip- and knee arthroplasty PRO(M)s set

		Preoperative (≤ 182 days before surgery)	**3 months postoperative** (63-110 days)	**6 months postoperative** (154-210 days)	**12 months postoperative** (323-407 days)
PRO	**PROM**	*THA, TKA and UKA*	*THA*	*TKA and UKA*	*THA, TKA and UKA*
Pain at rest	NRS pain	x	x	x	x
	Anchor question		x	x	x
Pain during activity	NRS pain	x	x	x	x
Quality of life	EQ-5D-3L and later EQ-5D-5L	x	x	x	x
Physical functioning	HOOS-PS or KOOS-PS	x	x	x	x
	Anchor question		x	x	x
Pain and physical functioning	OHS or OKS	x	x	x	x
Satisfaction	NRS satisfaction		x	x	x

EQ-5D-3L = EuroQol 5 dimensions 3 level questionnaire, EQ-5D-5L = EuroQol 5 dimensions 5 level questionnaire, HOOS-PS = Hip disability and Osteoarthritis Outcome Score - Physical Function Shortform, KOOS-PS = Knee disability and Osteoarthritis Outcome Score - Physical Function Shortform, NRS = Numeric Rating Scale, OHS = Oxford Hip Score questionnaire, OKS = Oxford Knee Score questionnaire

From 2014 on, health care institutions are required to register PROs in the Dutch arthroplasty register [4]. Two years later, on advice of the Dutch orthopaedic association, PROs became a mandatory part of the national defined THA and TKA indicator sets hosted by a Dutch governmental organisation. These results are publicly available to create transparency of the delivered care [21].

Recently, in 2020, the Dutch orthopaedic association PROMs advice has been updated based on existing research and on lessons learned on PRO(M)s in the Netherlands [10]. This updated advice is mainly a proposal for further research.

PRO collection: current practice

Currently, 50% of the arthroplasty registries united in ISAR capture preoperative and postoperative PROs of the patients [22]. The annual reports of the Dutch arthroplasty register reported a THA preoperative RR of 51% in 2016 with a small improvement to 65% in 2019. Percentage of patients responding both preoperatively and 3 month postoperatively is increased from 34% to 44% respectively [23]. Same numbers are reported for a TKA [24]. Unfortunately, there is a wide range in RRs which indicates a large diversity in PRO collection in the Netherlands.

To quantify the success of PRO collection, RR could be calculated. RR refers to the proportion of responders in relation to the number of patients who receive the questionnaire. More precisely, RR is calculated by dividing the number of returned questionnaires on surgical procedures completed

partly or totally by the number of surgical procedures minus the number of surgical procedures of patients who are deceased (returned questionnaires / (surgical procedures – surgical procedures of patients who are deceased)) [25]. An 100% RR is unrealistic as patients reply the PROMs voluntary. Furthermore, because of logistic reasons (for example no or no valid (e)mail address) health care institutions might not be able to reach all patients.

Achieving high RRs depends on the method in PRO collection chosen. Making PRO collection a part of the routine care, using a PROMs digital administration station in the health care institutions and collecting via multiple sources (for example traditional mail and email) are the keys to high response rates [26–28]. In arthroplasty patients, a critical factor is to guarantee that PROs are collected preoperatively as it results in a 3 times higher chance of collecting the PROs 3 months after surgery and even a 15 times higher chance at 12 months [29]. Maintaining high postoperative RRs is crucial as non-responding patients can introduce a bias which results in incomparable PROs if the non-responders are different than the responders [27, 30] and missing data are not at random [31]. It seems logical that a higher RR requires more effort and money. A recent study among surgeons concluded that the most important constraint against implementing PRO collection was costs [32]. Therefore, a clear comprehending is needed of which RR is achievable and at what costs in daily orthopaedic health care.

To draw valid conclusions on PROs, a certain RR on PROMs is needed to both obtain an accurate outcome and ensure generalizability [33]. The ISAR PROMs Working Group proposed a RR of at least 60%. They mention that this advice is only based on the external difficulties to collect PROs that may be unrelated to survey logistics and the requirement of ≥60% for a survey study [25, 34]. This proposed RR of at least 60% is not based on scientific evidence yet.

PRO use to optimize health care: current practice

Multiple arthroplasty registries have incorporated PROs with the aim of improving health care by evaluating outcomes from a patients' perspective [9, 22, 35]. However, previous studies emphasize that there is no definitive evidence yet that this goal is achieved [36–40]. Ideas on how PROs could be used to optimize health care are available: to compare surgical procedures or implants used for the same treatment, to evaluate adaptations in a treatment or rehabilitation, to evaluate the effect of care over the years, to compare surgeons, to evaluate a new surgeon, to compare treatments appropriate for the same patient group, to guide patients in their rehabilitation, to inform patients about what to expect of their pain relief and functional improvement, to enhance shared decision making, to predict outcomes and to select patients appropriate for a treatment [41]. However, there is a lack of scientific based examples of routine PRO use to optimize health care from arthroplasty patients' perspective in daily health care. These examples are needed to inspire stakeholders what they could achieve if they routinely use PROs to optimize health care.

Potential examples of routine PRO use in daily orthopaedic health care

In daily orthopaedic health care, PROs are not always incorporated in health care evaluations to optimize health care. Based on previous studies, two different implant designs used for an UKA

(mobile bearing and fixed bearing) show excellent functional outcomes, implant survival rates and complication rates [42–46]. However, from a patients' perspective it is unknown if one of these implants is preferred. Comparing PROs for both implants will result in knowledge which is of potential use in shared decision making between orthopaedic surgeon and patient, and in making recommendations to surgeons and other stakeholders on implant choice.

Unfortunately, up to 20% of patients are dissatisfied with their TKA end result [47–54]. Patient characteristics and preoperative PROs are potential predictors of satisfaction one year after TKA. Being able to predict the outcome preoperatively might reduce the number of less satisfied patients. Furthermore, it could gain knowledge to enhance shared decision making and for making recommendations to stakeholders. It could, therefore, be an example of using PROs to optimize health care.

Another daily orthopaedic health care situation is created due to the successful fast-track surgery procedures, namely a shorter hospital stay after a TKA of 1 or 2 days. Patients need to take responsibility for their aftercare shortly after surgery. They feel uncertain and left alone after early discharge, which could increase anxiety and affect their pain coping and subsequent management [55, 56]. Patients might need more individualized guidance in their pain management in their first period at home. As pain is a PRO and pain management is based on the level of pain mentioned, using PROs to guide patients individually in pain control during the first period at home after a TKA could be an example of using PROs to optimize health care.

The digital transformation in health care has been accelerated by the COVID-19 pandemic. Video consultation has become the alternative for traditional hospital consultation. The number of these hospital consultations has dropped by 30%, and the number of teleconsultations has increased 5-fold [57]. Currently, video consultation provides health care institutions and health care professionals the opportunity to increase office efficacy and cost-effectiveness in an era of decreasing reimbursements and increasing time constraints [58–60]. From a patients' perspective, it can also improve care efficacy and patient satisfaction as well as eliminating travel time and expenses [57, 58]. However, not all patients might benefit from video consultation, and it is unknown how to select patients suitable for video consultation. Using PROs to select patients suitable for hospital or video consultation could be an example of using PROs to optimize health care.

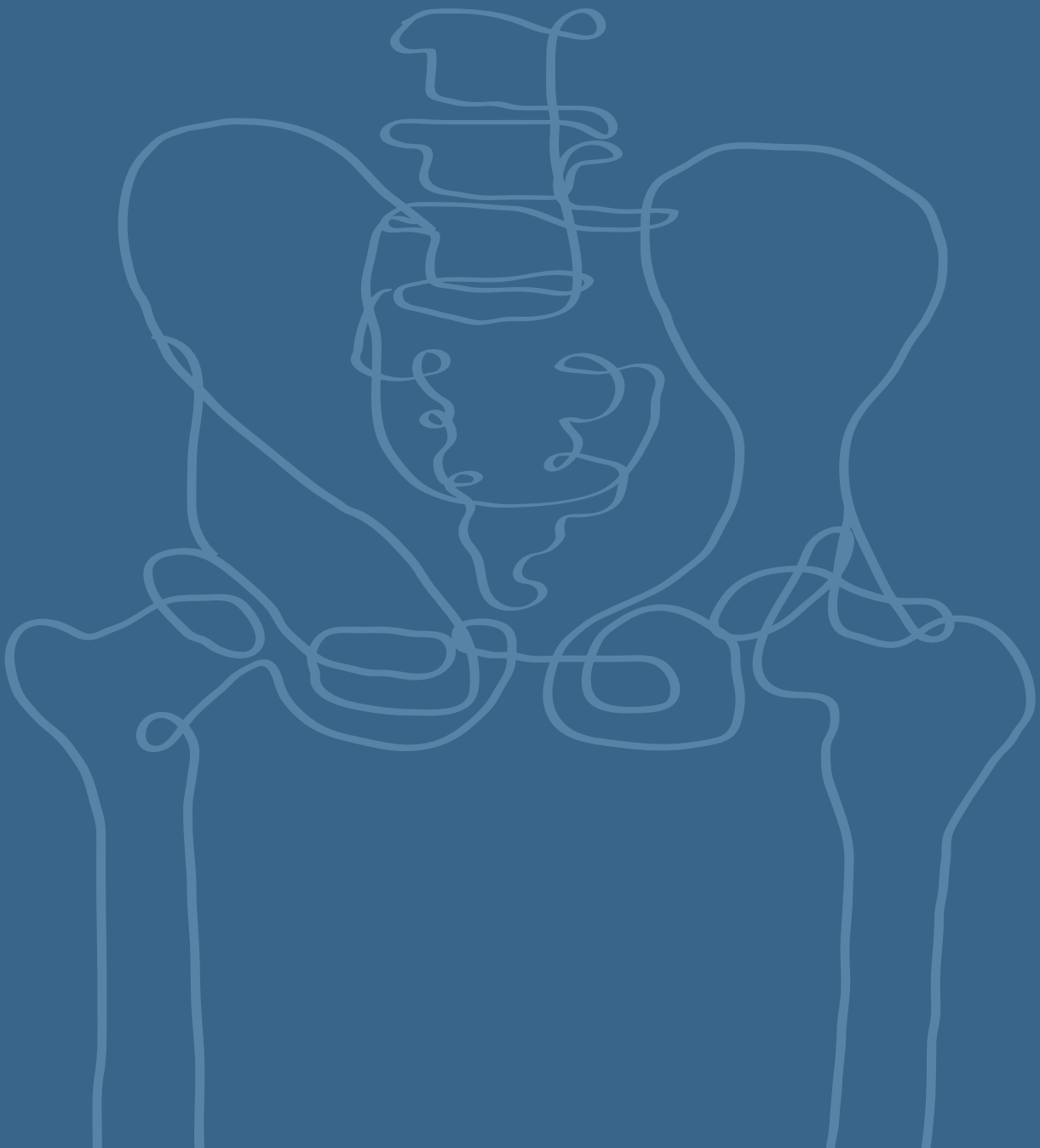
In conclusion, routine PRO collection and use have been increased, however, multiple PRO-related questions remain unanswered.

REFERENCES

1. Ministerie van Volksgezondheid Welzijn en Sport. Artrose. https://www.volksgezondheidenzorg.info/onderwerp/artrose/cijfers-context/huidige-situatie#node-prevalentie-en-aantal-nieuwe-gevallen-van-artrose-huisartsenpraktijk. Accessed 7 Jul 2022
2. Nederlandse Orthopaedische Vereniging (2019) Dossier artrose. Zorg voor Beweg 22–27
3. International Society of Arthroplasty Registries. Improving outcomes for individuals receiving joint replacement surgery worldwide. https://www.isarhome.org/. Accessed 7 Jul 2022
4. Nederlandse Orthopaedische Vereniging. Wat is de LROI? https://www.lroi.nl/over-de-lroi/wat-is-de-lroi. Accessed 1 Aug 2018
5. LROI. Number of registered hip arthroplasties. https://www.lroi-report.nl/hip/numbers/. Accessed 7 Jul 2022
6. LROI. Total Hip Arthroplasty Demographics. https://www.lroi-report.nl/hip/total-hip-arthroplasty/demographics/. Accessed 4 Apr 2022
7. LROI. Number of registered knee arthroplasties. https://www.lroi-report.nl/knee/numbers/. Accessed 7 Jul 2022
8. LROI. Knee Arthroplasty Demographics. https://www.lroi-report.nl/knee/patient-characteristics/. Accessed 7 Jul 2022
9. Nederlandse Orthopaedische Vereniging (2012) Patient Reported Outcome Measures: Advies Nederlandse Orthopaedische Vereniging (NOV). https://www.orthopeden.org/downloads/32/advies-proms-orthopedie.pdf. Accessed 1 Aug 2018
10. Nederlandse Orthopaedische Vereniging (2020) NOV PROMs-advies orthopedie 2020. https://www.orthopeden.org/downloads/775/nov-proms-advies.pdf. Accessed 7 Jul 2022
11. Rolfson O, Eresian Chenok K, Bohm E, et al (2016) Patient-reported outcome measures in arthroplasty registries: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries: Part I. Overview and rationale for patient-reported outcome measures. Acta Orthop 87:3–8
12. LROI. Data Quality - Validity. https://www.lroi-report.nl/data-quality/validity/. Accessed 7 Jul 2022
13. LROI. Data Quality - Completeness. https://www.lroi-report.nl/data-quality/completeness/. Accessed 20 Dec 2021
14. Söderman P, Malchau H, Herberts P, Zügner R, Regnér H, Garellick G (2001) Outcome after total hip arthroplasty: Part II. Disease-specific follow-up and the Swedish National Total Hip Arthroplasty Register. Acta Orthop Scand 72:113–119
15. Söderman P, Malchau H, Herberts P (2000) Outcome after total hip arthroplasty: Part I. General health evaluation in relation to definition of failure in the Swedish National Total Hip Arthroplasty register. Acta Orthop Scand 71:354–359
16. Garellick G, Malchau H, Herberts P (2000) Survival of Hip Replacements. Clin Orthop Relat Res 375:157–167
17. Britton AR, Murray DW, Bulstrode CJ, McPherson K, Denham RA (1997) Pain Levels After Total Hip Replacement: Their Use As Endpoints For Survival Analysis. J Bone Jt Surg 79:93–98
18. Willems R (2004) Hier werk je veilig, of hier werk je niet. Sneller Beter - De veiligheid in de zorg. Eindrapportage Shell Nederland. Den Haag
19. Wylde V, Blom AW (2011) The failure of survivorship. J Bone Joint Surg Br 93-B:569–570
20. Nederlandse Orthopaedische Vereniging. PROMs. https://www.lroi.nl/invoerders/registreren/proms. Accessed 1 Aug 2018
21. Zorginstituut Nederland. Zorginzicht - Open data Medisch-specialistische zorg. https://www.zorginzicht.nl/openbare-data/open-data-ziekenhuizen-en-zelfstandige-behandelcentra---medisch-specialistische-zorg. Accessed 24 Nov 2021
22. Bohm ER, Kirby S, Trepman E, et al (2021) Collection and Reporting of Patient-reported Outcome Measures in Arthroplasty Registries: Multinational Survey and Recommendations. Clin Orthop Relat Res 479:2151–2166
23. LROI. Hip PROMs Response. https://www.lroi-report.nl/hip/proms/response/. Accessed 2 Dec 2021
24. LROI. Knee PROMs Response. https://www.lroi-report.nl/knee/proms/response/. Accessed 2 Dec 2021

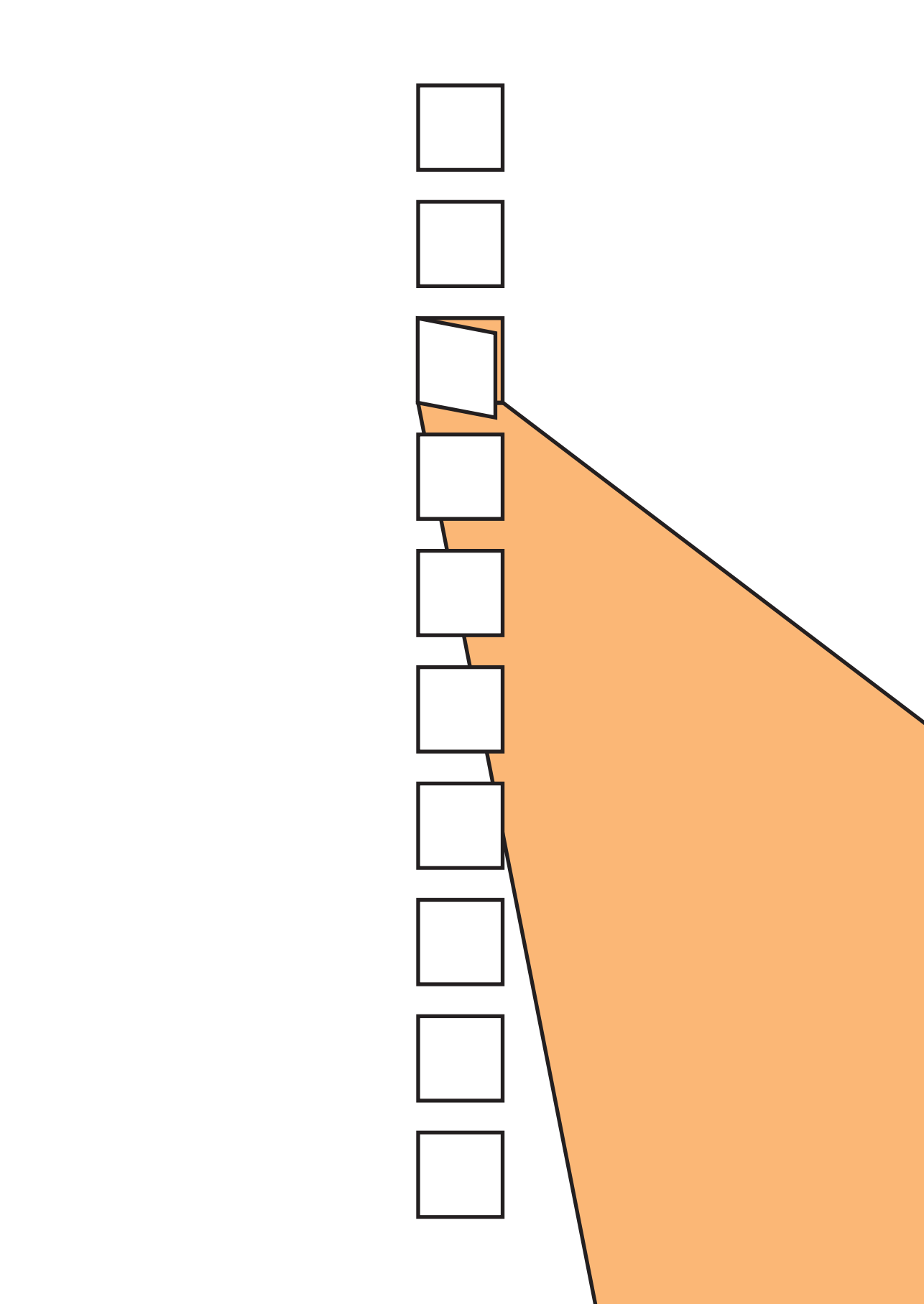
25. Rolfson O, Bohm E, Franklin P, et al (2016) Patient-reported outcome measures in arthroplasty registries: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty RegistriesPart II. Recommendations for selection, administration, and analysis. Acta Orthop 87:9–23
26. Ho A, Purdie C, Tirosh O, Tran P (2019) Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. Patient Relat Outcome Meas 10:217–226
27. Tariq MB, Vega JF, Westermann R, Jones M, Spindler KP (2019) Arthroplasty studies with greater than 1000 participants: analysis of follow-up methods. Arthroplast Today 5:243–250
28. Slover JD, Karia RJ, Hauer C, Gelber Z, Band PA, Graham J (2015) Feasibility of integrating standardized patient-reported outcomes in orthopedic care. Am J Manag Care 21:e494-500
29. Patel J, Lee JH, Li Z, SooHoo NF, Bozic K, Huddleston JI (2015) Predictors of Low Patient-Reported Outcomes Response Rates in the California Joint Replacement Registry. J Arthroplasty 30:2071–2075
30. Norquist BM, Goldberg BA, Matsen FA (2000) Challenges in Evaluating Patients Lost to Follow-up in Clinical Studies of Rotator Cuff Tears. J Bone Jt Surgery-American Vol 82:838–842
31. Kristman V, Manno M, Côté P (2003) Loss to Follow-Up in Cohort Studies: How Much is Too Much? Eur J Epidemiol 19:751–760
32. Joeris A, Knoll C, Kalampoki V, Blumenthal A, Gaskell G (2018) Patient-reported outcome measurements in clinical routine of trauma, spine and craniomaxillofacial surgeons: between expectations and reality: a survey among 1212 surgeons. BMJ Open 8:e020629
33. Paulsen A (2014) Patient reported outcomes in hip arthroplasty registries. Dan Med J 61:B4845
34. JAMA. Internal Medicine Instructions for Authors. https://jamanetwork.com/journals/jamainternalmedicine/pages/instructions-for-authors. Accessed 1 Aug 2018
35. Rolfson O, Eresian Chenok K, Bohm E, et al (2016) Patient-reported outcome measures in arthroplasty registries. Acta Orthop 87:3–8
36. Damman OC, Jani A, Jong BA, et al (2020) The use of PROMs and shared decision-making in medical encounters with patients: An opportunity to deliver value-based health care to patients. J Eval Clin Pract 26:524–540
37. Kotronoulas G, Kearney N, Maguire R, Harrow A, Di Domenico D, Croy S, MacGillivray S (2014) What Is the Value of the Routine Use of Patient-Reported Outcome Measures Toward Improvement of Patient Outcomes, Processes of Care, and Health Service Outcomes in Cancer Care? A Systematic Review of Controlled Trials. J Clin Oncol 32:1480–1501
38. Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, Black N (2018) How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. J Patient-Reported Outcomes 2:42
39. Greenhalgh J, Dalkin S, Gibbons E, Wright J, Valderas JM, Meads D, Black N (2018) How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis. J Health Serv Res Policy 23:57–65
40. Gutacker N, Bojke C, Daidone S, Devlin NJ, Parkin D, Street A (2013) Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. Health Econ 22:931–947
41. Wilson I, Bohm E, Lübbeke A, Lyman S, Overgaard S, Rolfson O, W-Dahl A, Wilkinson M, Dunbar M (2019) Orthopaedic registries with patient-reported outcome measures. EFORT Open Rev 4:357–367
42. Cao Z, Niu C, Gong C, Sun Y, Xie J, Song Y (2019) Comparison of Fixed-Bearing and Mobile-Bearing Unicompartmental Knee Arthroplasty: A Systematic Review and Meta-Analysis. J Arthroplasty 34:3114-3123.e3
43. Huang F, Wu D, Chang J, Zhang C, Qin K, Liao F, Yin Z (2021) A Comparison of Mobile- and Fixed-Bearing Unicompartmental Knee Arthroplasties in the Treatment of Medial Knee Osteoarthritis: A Systematic Review and Meta-analysis of 1,861 Patients. J Knee Surg 34:434–443
44. Peersman G, Stuyts B, van den Langenbergh T, Cartier P, Fennema P (2015) Fixed- versus mobile-bearing UKA: a systematic review and meta-analysis. Knee Surgery, Sport Traumatol Arthrosc 23:3296–3305
45. The New Zealand Joint Registry (2019) The New Zealand Joint Registry Twenty Year Report January 1999 To December 2018. https://nzoa.org.nz/system/files/DH8328_NZJR_2019_Repor t_v4_7Nov19.pdf. Accessed 11 Mar 2020
46. Von Keudell A, Sodha S, Collins J, Minas T, Fitz W, Gomoll A (2014) Patient satisfaction after primary total and unicompartmental knee arthroplasty: An age-dependent analysis. Knee 21:180–184

47. Ali A, Lindstrand A, Sundberg M, Flivik G (2017) Preoperative Anxiety and Depression Correlate With Dissatisfaction After Total Knee Arthroplasty: A Prospective Longitudinal Cohort Study of 186 Patients, With 4-Year Follow-Up. J Arthroplasty 32:767–770
48. Baker PN, Rushton S, Jameson SS, Reed M, Gregg P, Deehan DJ (2013) Patient satisfaction with total knee replacement cannot be predicted from pre-operative variables alone. Bone Joint J 95-B:1359–1365
49. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ (2010) Patient Satisfaction after Total Knee Arthroplasty: Who is Satisfied and Who is Not? Clin Orthop Relat Res 468:57–63
50. Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ (2018) The rate and predictors of patient satisfaction after total knee arthroplasty are influenced by the focus of the question. Bone Joint J 100-B:740–748
51. Noble PC, Conditt MA, Cook KF, Mathis KB (2006) The John Insall Award: Patient Expectations Affect Satisfaction with Total Knee Arthroplasty. Clin Orthop Relat Res 452:35–43
52. Van Onsem S, Van Der Straeten C, Arnout N, Deprez P, Van Damme G, Victor J (2016) A New Prediction Model for Patient Satisfaction After Total Knee Arthroplasty. J Arthroplasty 31:2660-2667.e1
53. Schnurr C, Jarrous M, Güdden I, Eysel P, König DP (2013) Pre-operative arthritis severity as a predictor for total knee arthroplasty patients' satisfaction. Int Orthop 37:1257–1261
54. Scott CEH, Howie CR, MacDonald D, Biant LC (2010) Predicting dissatisfaction following total knee replacement. J Bone Joint Surg Br 92-B:1253–1258
55. Specht K, Agerskov H, Kjaersgaard-Andersen P, Jester R, Pedersen BD (2018) Patients' experiences during the first 12 weeks after discharge in fast-track hip and knee arthroplasty – a qualitative study. Int J Orthop Trauma Nurs 31:13–19
56. Sjøveian AKH, Leegaard M (2017) Hip and knee arthroplasty - patient's experiences of pain and rehabilitation after discharge from hospital. Int J Orthop Trauma Nurs 27:28–35
57. Merkes K, Hagenaars N (2020) Corona: katalysator of struikelblok voor groenere ziekenhuiszorg? https://gupta-strategists.nl/studies/corona-katalysator-of-struikelblok-voor-groenere-ziekenhuiszorg. Accessed 13 Nov 2021
58. Hollander JE, Carr BG (2020) Virtually Perfect? Telemedicine for Covid-19. N Engl J Med 382:1679–1681
59. O'Connor CM, Anoushiravani AA, DiCaprio MR, Healy WL, Iorio R (2020) Economic Recovery After the COVID-19 Pandemic: Resuming Elective Orthopedic Surgery and Total Joint Arthroplasty. J Arthroplasty 35:S32–S36
60. Kane LT, Thakar O, Jamgochian G, Lazarus MD, Abboud JA, Namdari S, Horneff JG (2020) The role of telehealth as a platform for postoperative visits following rotator cuff repair: a prospective, randomized controlled trial. J Shoulder Elb Surg 29:775–783



PART I

Optimizing the routine patient-reported outcome collection



CHAPTER 3

Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection

Y. Pronk, P. Pilot, J.M. Brinkman, R.J. van Heerwaarden and W. van der Weegen

J Patient Rep Outcomes. 2019 Jun 3;3(1):31

ABSTRACT

Background

The response rate on patient-reported outcome measurements (PROMs) necessary to adequately evaluate a treatment and improve patient care is unknown. Hospitals generally aim for the highest possible response rate without insight into the increase in costs involved. This study aimed to investigate which PROMs response rate is achievable in relation to the costs in an orthopaedic practice.

Methods

In an observational study, patients planned for orthopaedic surgery were asked to participate per surgical procedure (5769 surgical procedures at 5300 patients). Patient-reported outcomes (PROs) collection with a digital online automated PROMs collection system (minimal effort) was compared to a combined automated system and manual collection (maximal effort). Response rate was calculated preoperative and at two postoperative time points separately, and on all three time points together. Costs were calculated for the study period, per year and per surgical procedure. Calculations were executed for all surgical procedures and for three subgroups: knee arthroplasty, hip arthroplasty and anterior cruciate ligament reconstruction (ACLR).

Results

Using maximal effort the response rate increased for all surgical procedures compared to minimal effort; the preoperative response rate from 86% to 100% and the postoperative response rates from 55% to 83% (3 or 6 months) and 53% to 83% (12 months). Concerning the response at all three time points for all surgical procedures, minimal effort resulted in 44% response rate and increased to 76% with maximal effort. Lowest postoperative response rates were found in the ACLR group for both maximal and minimal effort. A costs difference of €5.55–€5.98 per surgical procedure between maximal and minimal effort was found.

Conclusions

A two times higher PROMs response rate for patients responding at all three time points (44% versus 76%) is achievable with maximal effort compared to the use of an automated PROMs collection system only. Manual collection adds a cost of €5.5–€6 per surgical procedure to automated PROMs collection alone. It is debatable if these additional costs are justifiable from a value-based health care perspective as the response rate for adequate evaluation of a treatment is still unknown.

INTRODUCTION

From a patient's perspective, implant survival may not be the best measure of surgery success. Instead, pain reduction, functional improvement and quality of life are important [1–4]. With this shift towards a more patient-centred perspective in health care, there is an increase in the use of Patient-Reported Outcome Measurements (PROMs) [5]. PROMs are questionnaires that assess health status from patient's perspective and focus on pain, function, quality of life and/or satisfaction. This has resulted in the addition of patient-reported outcomes (PROs) to (national) arthroplasty registries for evaluating treatments and improving patient care. Since 2007 all Dutch hospitals have registered their implanted prostheses in a national registry and in 2012 the Dutch Orthopaedic Association (NOV) advised hospitals to add PROs collected by selected PROMs [6, 7]. This resulted in the first PROMs indicator which obliges hospitals to collect PROs of all hip arthroplasty patients. The first part of this indicator is a process indicator as it focusses on the achieved response rate.

To achieve the goal of evaluating treatments and improving patient care a certain level of response rate is necessary to ensure generalizability and to minimize selection bias of the collected PROs [8]. Unfortunately, there is no clear consensus of what rate is acceptable. The International Society of Arthroplasty Registries (ISAR) PROMs Working Group proposed a response rate of at least 60% [9, 10]. That percentage is based on what is considered a sufficient response rate in survey research [11]. In 2017, the Dutch arthroplasty registry reported an average preoperative response rate of 54%, ranging from 5% to 99% [12].

Although PROs are an important component of health outcome and several authors have reported tips and tricks regarding PROs collection [13–15], even specific for orthopaedic practice [9, 16], this wide range in response rate reported by the Dutch arthroplasty registry shows that the implementation and integration of PROs collection into orthopaedic practice has its challenges. Generally, hospitals strive for an as high as possible response rate without having an insight into the increase in costs involved and not knowing if their response rate justifies the expenses made.

Therefore, a clear understanding is needed of which response rate is achievable and at what costs. The aim of this study was to investigate which PROMs response rate is achievable in relation to the costs for PROs collection in an orthopaedic practice.

METHODS

Setting and inclusion

PROs collection was performed in a medium-sized-orthopaedic hospital (Kliniek ViaSana, Mill, the Netherlands). Between January 2014 and June 2015, 5300 orthopaedic patients that underwent in total 5769 surgical procedures, characterised by aged 12 years and older, American Society of Anaesthesiologists (ASA) classification of I or II, and body mass index (BMI) ≤35 kg/m^2, were followed.

Patients were informed and asked by their surgeon's receptionist to participate in PROs collection and to allow further scientific analysis using their anonymised data. All included patients signed the informed consent form. PROMs sets were based on the type of surgery performed and included those that were mandatory as set out by the NOV [10]. All sets had comparable length and linguistic difficulty. Retrospective analysis was executed on the prospectively collected data. This study was approved by the district medical ethics committee (N18.156).

Data collection

Patients registered and completed their preoperative PROMs on a computer using a web-based survey of a digital, online, automated system for collecting PROs (OnlinePROMs, Interactive Studios, Rosmalen, the Netherlands) directly after consultation in the hospital. In case they needed assistance or could not handle a computer, an (admission administrator) employee was available to provide instructions or hand out paper forms. Before surgery, completeness of the PROMs was checked by the PROMs administrator and in case of incomplete PROMs, a paper form was handed out to the patient at the day of surgery to collect the missing PROs (response check). After surgery the PROMs administrator manually entered the date of surgery in the automated system; by doing so, postoperative PROMs were automatically sent by email 3 or 6, and 12 months after surgery. In case of non-response an automatic reminder was sent after 7 days. In case no email address was registered, the PROMs were sent by postal service and included an invitation letter and a stamped self-addressed envelope. This was all done by the PROMs administrator who received a notification by the automated system to execute this. If the patient did not respond after two invitations by email, the PROMs administrator automatically received a notification by the system to send a third invitation per postal service. All returned forms were manually entered in the automated system by the PROMs administrator. All questions in the automated system were mandatory. In total, per surgical procedure the patient was invited to complete the PROMs at three time points: preoperatively, at 3 or 6 months postoperatively, and at 12 months postoperatively.

Data analysis

After data collection, per surgical procedure and per time point patients were allocated to two groups: the minimal effort or the maximal effort group. Patients for which PROs were collected only using the automated system were included in the minimal effort group. For this group, additional manual labour was only needed for entering the date of surgery. The maximal effort group included all patients where extra manual labour was needed: response check, PROMs sent by postal service, third invitations sent by postal service and remaining tasks. These remaining tasks consisted of answering patients phone calls or emails, or correcting administrative errors such as wrong email addresses.

Response rate and costs

Response rate was calculated by dividing the number of returned questionnaires completed partly or totally by the number of surgical procedures minus the number of surgical procedures of patients who were deceased (returned questionnaires / (surgical procedures – surgical procedures of patients who were deceased)) [9]. Reasons for loss to follow-up were reported. First, response

rate was calculated per time point. Second, it was calculated for all three time points together. Response at all three time points was defined as when per surgical procedure a patient returned the PROMs at all three time points: preoperatively, 3 or 6 months postoperatively, ánd 12 months postoperatively. When there was no returned questionnaire on one or more time points, this was defined as no response at all three time points. Completion rate per time point was calculated by dividing the number of returned questionnaires completed totally by the number of surgical procedures minus the number of surgical procedures of patients who were deceased (totally completed returned questionnaires / (surgical procedures – surgical procedures of patients who were deceased)). Costs were calculated for the entire study period, per surgical procedure and per year. Costs consisted of the license fee for the automated PROMs system (€7500,- per year), pay for two computers on which the registration and completion of the preoperative PROMs was done (€1600,- over 5 years), costs for paper forms including sending per postal service (€0.08 per sheet of paper, €0.07 per envelope and €10.000 per year for sending), and staff employment costs: PROMs administrator (€22.1 per hour), surgeon's receptionist (€21.1 per hour) and admission administrator (€22.1 per hour). The amount of time needed for all specific manual tasks in the collection process was estimated. Response rate and costs were calculated for all surgical procedures and for three patient groups as subgroups: total hip arthroplasty (THA), total or unicompartmental knee arthroplasty (TKA&UKA) and anterior cruciate ligament reconstruction (ACLR). Baseline demographic data were collected from the electronic patient records.

RESULTS

Between January 2014 and June 2015, all 5300 patients planned for 5769 surgeries were included of which only 2 times a patient declined participation, therefore 5767 surgical procedures (100%) of 5298 patients were available for participating PROMs (Fig. 1).

Characteristics

The characteristics of the 5769 surgical procedures as well as the subgroups qualifications are listed in Table 1.

Response rate

With maximal effort for PROs collection the response rate increased for all surgical procedures compared to minimal effort, the preoperative response rate from 86% to 100% and the postoperative response rates from 55% to 83% (3 or 6 months) and 53% to 83% (12 months) (Fig. 2a). The lowest postoperative response rates were found in the ACLR group for both maximal and minimal effort compared to the other groups (Fig. 2). For all surgical procedures minimal effort resulted in 44% response rate at all three time points. An increased in response rate to 76% was reached with maximal effort (Fig. 2a). Various differences in response rates between the subgroups were found (Fig. 2b-d).

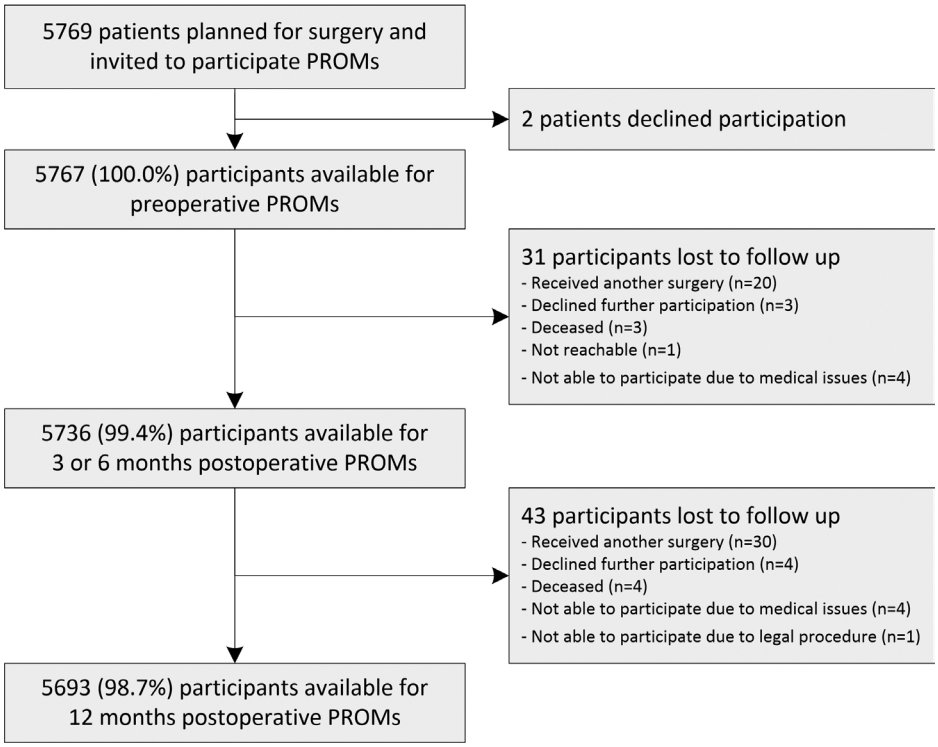


Fig. 1 Flowchart

Note: PROMs indicates patient-reported outcome measurements. n indicates number

Of all the additional tasks performed in the maximal effort group sending a third invitation by postal service after no response on two automated email invitations resulted in the highest extra response rate in all surgical procedures and in all the three subgroups ranging from 13% to 27% (Fig. 2).

Regarding the completion rate, maximal effort for PROs collection resulted in 100% preoperative completion rate compared to 86% with minimal effort, 81% compared to 54% 3 or 6 months postoperatively and 79% in comparison with 52% 12 months postoperatively respectively.

Table 1 Characteristics for all surgical procedures and the THA, TKA&UKA and ACLR subgroups

	All surgical procedures (n=5769)	THA (n=535)	TKA&UKA (n=742)	ACLR (n=430)
Age (y, mean ± SD)	50.3 ± 15.8	64.7 ± 8.3	64.3 ± 7.8	27.4 ± 9.5
BMI (kg/m^2, mean ± SD)	26.0 ± 3.6	25.9 ± 3.5	28.0 ± 3.5	23.8 ± 2.9
Gender – female (n, (%))	2715 (47.1%)	339 (63.4%)	377 (50.8%)	138 (32.1%)
ASA – II (n, (%))	1986 (34.4%)	264 (49.4%)	438 (59.0%)	25 (5.81%)

Note: THA indicates total hip arthroplasty. TKA indicates total knee arthroplasty. UKA indicates unicompartmental knee arthroplasty. ACLR indicates anterior cruciate ligament reconstruction. y indicates year. SD indicates standard deviation. BMI indicates body mass index. kg/m^2 indicates kilogram per square meter. n indicates number. ASA indicates American Society of Anaesthesiologists classification

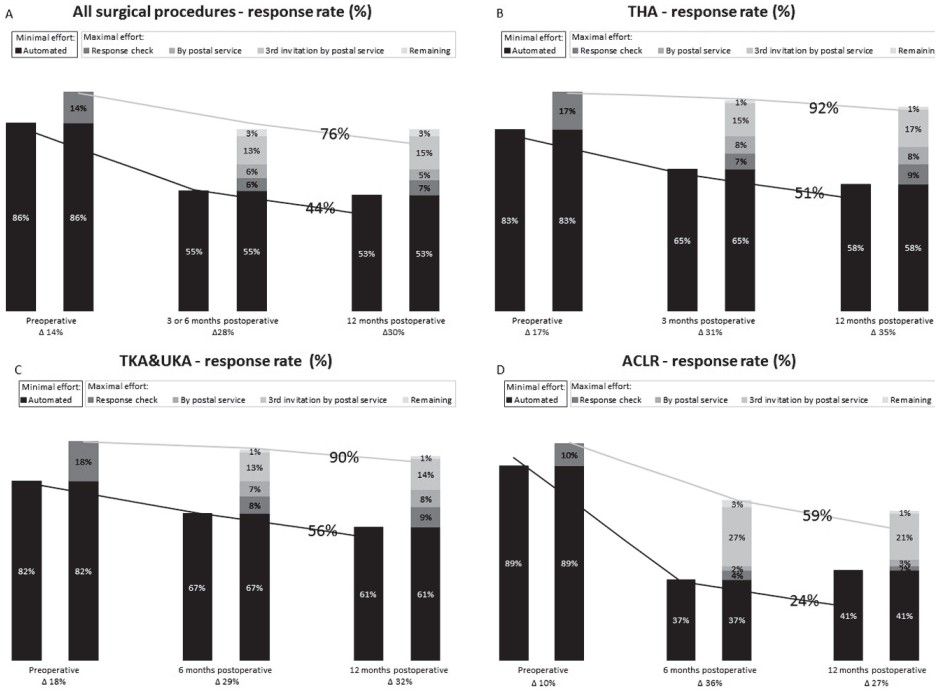


Fig. 2 Response rates (%) of PROs collection: minimal effort versus maximal effort

Represented for all surgical procedures (a) and the THA (b), TKA&UKA (c) and ACLR (d) subgroups. The line represents the response rate at all three time points. Note: PROs indicates patient-reported outcomes. THA indicates total hip arthroplasty. TKA indicates total knee arthroplasty. UKA indicates unicompartmental knee arthroplasty. ACLR indicates anterior cruciate ligament reconstruction

Table 2 Costs of PROs collection: minimal effort versus maximal effort

	All surgical procedures (n=5769)		THA (n=535)		TKA&UKA (n=742)		ACLR (n=430)	
	Minimal effort	**Maximal effort**	**Minimal effort**	**Maximal effort**	**Minimal effort**	**Maximal effort**	**Minimal effort**	**Maximal effort**
License fee for automated system (€)	10,875	10,875	10,875	10,875	10,875	10,875	10,875	10,875
Two computers for registration and completion of the preoperative PROMs (€)	480	480	480	480	480	480	480	480
Staff (€)								
PROMs administrator	1725	17,254	160	1599	217	2165	129	1285
Surgeon's receptionist	2027	2027	188	188	261	261	151	151
Admission administrator	7510	7510	696	696	966	966	560	560
Paper forms and sending (€)	462	17,936	43	1573	59	2264	34	1447
In total (€)	**23,079**	**56,081**	**12,442**	**15,411**	**12,857**	**17,011**	**12,229**	**14,799**
Per surgical procedure (€)	**4.00**	**9.72**	**23.26**	**28.81**	**17.33**	**22.93**	**28.44**	**34.42**
Per year (€)	**15,479**	**37,481**	**8283**	**10,263**	**8581**	**11,350**	**8176**	**9889**

Represented for all surgical procedures and the THA, TKA&UKA and ACLR subgroups. Note: PROs indicates patient-reported outcomes, THA indicates total hip arthroplasty, TKA indicates total knee arthroplasty, UKA indicates unicompartmental knee arthroplasty, ACLR indicates anterior cruciate ligament reconstruction, € indicates euro, n indicates number

Costs

Costs associated with collecting PROs with maximal effort for all surgical procedures increased to €56,081 compared to €23,079 with minimal effort; €9.72 versus €4.00 per surgical procedure and €37,481 versus €15,479 per year. In all surgical procedures and in the three subgroups, the calculated difference per surgical procedure between minimal and maximal effort ranged between €5.55 and €5.98. Costs per surgical procedure in the three subgroups were the highest in the ACLR group for both minimal (€28.44) and maximal effort (€34.42) compared to the other subgroups (Table 2).

DISCUSSION

This study aimed to investigate which PROMs response rate is achievable in relation to the costs for PROs collection in an orthopaedic practice. Collecting PROs with maximal effort for all surgical procedures resulted in a preoperative response rate increasing from 86% reachable with minimal effort to the optimal of 100%, and at the two postoperative time points from 53% or 55% to 85%. Furthermore, with maximal effort a two times higher response rate for patients responding at all three time points was achievable compared to only using a digital online automated PROMs collection system as minimal effort. Both achieved with two times higher costs (€4 to €10 per surgical procedure). These additional costs of €6 per surgical procedure were found for all surgical procedures as well as in the subgroups. Regarding these subgroups, lowest response rates and highest costs were found in the ACLR group with both maximal and minimal effort.

The only two previous orthopaedic studies that use a digital online automated PROMs collection system reported 43% response 6 months after knee surgery for patellar instability, ligament, cartilage, or meniscus injury [17] and 92% after elbow arthroplasty [18]. Howard et al. found similar rates related to the ACLR patients (37%) as the most comparable group of the current study. However, only 9% of their patients responded at all time points [17], which is less compared to the 24% in the present study. Viveen et al. used the same automated system and reported a similar response rate to this study, but calculated it by dividing the number of returned PROMs by the number of sent PROMs [18]. In studies outside of orthopaedics, response rates of web-based surveys vary greatly between 14% and 83% [19–23]. Web-based surveys are said to be cost-effective [14], have a decreased risk of errors and missing values [24] and are favoured [25] compared to paper forms. In the current study, only using an automated system, the ISAR PROMs Working Group proposed response rate of at least 60% was reached for the preoperative collected PROs [9], but not postoperatively for all surgical procedures, ACLR and THA at 12 months. Regarding at least 60% on all preoperative ánd postoperative time points, none of the four groups reached this threshold while using an automated system only. Using maximal effort in collecting PROs this ISAR threshold is almost achieved as it resulted in at least 68% for one single time point and at least 59% response at all three time points. This shows that alternatives beside an automated system as minimal effort to complete PROMs are needed to improve response rate [14, 26, 27] and to reach the proposed threshold of 60%. Similarly, Rolfson et al. concluded that

only using web-based surveys in THA patients results in an insufficient response rate of 49%, and it is unable to replace PROs collection with paper forms in PROs collection with an automated system only as the PROs and patient demographics for being a respondent differ between both ways of collection [24]. The sending of a 3rd invitation by postal service after no response was received on two email invitations, as a part of maximal effort, had the highest impact (≥13% extra response rate) on improving postoperative response rate and should be added to any automated collection system in order to achieve the ISAR threshold on every single time point. To achieve the proposed threshold for response at all three time points, maximal effort is needed. The downside of this is that maximal effort increased costs.

A recent study among trauma and orthopaedic surgeons concluded that one of the two most important constraints against implementing PROMs was costs [28]. Previous studies reported $2.00-$6.39 (€1.70–€5.50) per respondent using an automated system [19, 29] reaching a lower response rate (between 14% and 21%) compared to the current study. In the present study, collecting PROs was €6 per surgical procedure more expensive with maximal effort. The smaller the number of surgical procedures, the fixed costs such as the license fee for an automated system and hardware weigh heavier, as shown by the smaller ACLR group that was more expensive per surgical procedure compared to all surgical procedures included. Therefore, to consider the value of adding costs of €6 per surgical procedure to achieve higher response rate, the size of the hospital or patient group involved should be taking into account. Regarding the different patient groups, the THA and TKA&UKA patients had the highest pre- and postoperative response rates and had the lowest costs to collect PROs. This might be explained by their more compliant attitude to their surgeon [30]. The younger ACLR patients showed to be more inclined to handle computers due to their high preoperative response rate by using only an automated system [19]. However, their postoperative response rates with an automated system only were lower compared to the older patient groups. It might be that the age group of ACLR patients already get too many emails, so they were more aware of responding due to an invitation by postal service, as seen in the higher response rates on a 3rd invitation by postal service. Furthermore, the ACLR patients were mainly male patients who are reported to be more likely to respond by postal service [19, 26]. Younger [18, 19, 31] and male [19, 30, 31] patients in general are the most challenging group; they are less likely to respond at all. This also explains the higher costs for the ACLR patient in the current study. To ensure wider acceptance and to improve the response rate, postal service as additional effort is advised in younger and male patients [14, 26, 27]; again with the downside of higher costs.

Little is known about the costs made to collect PROs in relation to the benefit of collecting PROs. The present study shows the considerable costs to achieve high response rates; knowing that these costs are even without costs for data analysis and improvement strategies, which is expected to result in reducing costs. From a value based health care perspective, it is questionable if the costs made to collect PROs, and the additional costs for improving the response rate, are justifiable. The most important question might not be how many response is needed, but how representative the respondents are for the hospital or patient group in question [32]. It could

very well be that a more homogeneous patient population in a specific setting requires a lower response rate compared to a more heterogeneous patient population in another setting. It is questionable that a quality indicator is set on achieved response rate without actually knowing the threshold.

To the authors knowledge, this is the first study clarifying the achievable response rate on PROMs versus the associated health care costs in a medium sized orthopaedic practice. It provides other hospitals insights into what costs they might expect for collecting PROs in their hospital setting or patient groups using minimal and maximal effort. A limitation of this study was that the amount of time needed for all specific manual tasks in the collection process was not exactly measured but was estimated.

CONCLUSIONS

A two times higher PROMs response rate for patients responding at all time points is achievable with maximal effort compared to the use of a digital online automated PROMs collection system only for PROs collection in an orthopaedic practice. Manual collection adds a cost of €6 per surgical procedure to automated PROMs collection alone. As the response rate for adequate evaluation of a treatment is still unknown it is questionable if these additional costs are justifiable from a value-based health care perspective.

REFERENCES

1. Söderman, P., Malchau, H., Herberts, P., Zügner, R., Regnér, H., & Garellick, G. (2001). Outcome after total hip arthroplasty: Part II. Disease-specific followup and the Swedish National Total hip Arthroplasty Register. Acta orthopaedica Scandinavica, 72, 113–119.
2. Söderman, P., Malchau, H., & Herberts, P. (2000). Outcome after total hip arthroplasty: Part I. general health evaluation in relation to definition of failure in the Swedish National Total hip Arthroplasty register. Acta Orthopaedica Scandinavica, 71, 354–359.
3. Garellick, G., Malchau, H., & Herberts, P. (2000). Survival of hip replacements. A comparison of a randomized trial and a registry. Clinical Orthopaedics and Related Research, 375, 157–167.
4. Britton, A. R., Murray, D. W., Bulstrode, C. J., McPherson, K., & Denham, R. A. (1997). Pain levels after total hip replacement: Their use as endpoints for survival analysis. The Journal of bone and joint surgery. British, 79, 93–98.
5. Wylde, V., & Blom, A. W. (2011). The failure of survivorship. Bone Joint J, 93B, 569–570.
6. NOV (2012) Patient Reported Outcome Measures. https://www.orthopeden.org/downloads/32/advies-proms-orthopedie.pdf. Accessed 1 Aug 2018.
7. Nederlandse Orthopaedische Vereniging Wat is de LROI? https://www.lroi.nl/over-de-lroi/wat-is-de-lroi. Accessed 1 Aug 2018.
8. Paulsen, A. (2014). Patient reported outcomes in hip arthroplasty registries. Dan Med J, 61, B4845.
9. Rolfson, O., Bohm, E., Franklin, P., et al. (2016). Patient-reported outcome measures in arthroplasty registries: Report of the patient-reported outcome measures working Group of the International Society of arthroplasty RegistriesPart II. Recommendations for selection, administration, and analysis. Acta orthopaedica, 87, 9–23.
10. Nederlandse Orthopaedische Vereniging PROMs. https://www.lroi.nl/invoerders/registreren/proms. Accessed 1 Aug 2018.
11. JAMA Network Instruction for Authors. https://jamanetwork.com/journals/jama/pages/instructions-for-authors#SecReportsofSurveyResearch. Accessed 13 Mar 2019.
12. Nederlandse Orthopaedische Vereniging (2017) LROI-rapportage Response. http://www.lroi-rapportage.nl/hip-proms-response. Accessed 1 Aug 2018.
13. Edwards, P. J., Roberts, I., Clarke, M. J., DiGuiseppi, C., Wentz, R., Kwan, I., Cooper, R., Felix, L. M., & Pratap, S. (2009). Methods to increase response to postal and electronic questionnaires (review). Cochrane database of systematic reviews, 3A, 2009–2011.
14. Hunter, J., Corcoran, K., Leeder, S., & Phelps, K. (2013). Is it time to abandon paper? The use of emails and the internet for health services research – a cost-effectiveness and qualitative study. Journal of Evaluation in Clinical Practice, 19, 855–861.
15. Franklin, P., Chenok, K., & Lavalee, D. Love R eGEMs Of Patient-Reported Outcome Measures In The Learning Healthcare System (Vol. 5, pp. 1–17).
16. Brook, E. M., Glerum, K. M., Higgins, L. D., & Matzkin, E. G. (2017). Implementing patient-reported outcome measures in your practice: Pearls and pitfalls. American journal of orthopedics, 46, 273–278.
17. Howard, J. S., Toonstra, J. L., Meade, A. R., Whale Conley, C. E., & Mattacola, C. G. (2016). Feasibility of conducting a web-based survey of patientreported outcomes and rehabilitation progress. Digital health, 2, 205520761664484.
18. Viveen, J., Prkic, A., The B, Koenraadt, K. L. M., & Eygendaal, D. (2016). Effect of introducing an online system on the follow-up of elbow arthroplasty. World journal of orthopedics, 7, 826–831.
19. Bergeson, S. C., Gray, J., Ehrmantraut, L. A., Laibson, T., & Hays, R. D. (2013). Comparing web-based with mail survey Administration of the Consumer Assessment of healthcare providers and systems (CAHPS®) clinician and group survey. Primary health care: open access. https://doi.org/10.4172/2167-1079.1000132.
20. Hoonakker, P., & Carayon, P. (2009). Questionnaire survey nonresponse: A comparison of postal mail and internet surveys. Intl. Journal of Human–Computer Interaction, 25, 348–373.
21. Chen, L., & Goodson, P. (1924). Web-based survey of US health educators: Challenges and lessons. American journal of health behavior, 34, 3–11.

22. Basch, E., Artz, D., Dulko, D., Scher, K., Sabbatini, P., Hensley, M., Mitra, N., Speakman, J., McCabe, M., & Schrag, D. (2005). Patient online self-reporting of toxicity symptoms during chemotherapy. Journal of clinIcal oncology, 23, 3552–3561.

23. Vitale, M. G., Vitale, M. A., Lehmann, C. L., Hyman, J. E., Roye, D. P., Skaggs, D. L., Schmitz, M. L., Sponseller, P. D., & Flynn, J. M. (2006). Towards a National Pediatric Musculoskeletal. Journal of Pediatric Orthopaedics, 26, 151–156.

24. Rolfson, O., Salomonsson, R., Dahlberg, L. E., & Garellick, G. (2011). Internetbased follow-up questionnaire for measuring patient-reported outcome after total hip replacement surgery-reliability and response rate. Value in Health, 14, 316–321.

25. McCluskey, S., & Topping, A. E. (2011). Increasing response rates to lifestyle surveys: A pragmatic evidence review. Perspect Public Health, 131, 89–94.

26. Derby, D. C., Haan, A., & Wood, K. (2011). Data quality assurance: An analysis of patient non-response. International journal of health care quality assurance, 24, 198–210.

27. Bhinder, S., Chowdhury, N., Granton, J., Krahn, M., Tullis, D. E., Waddell, T. K., & Singer, L. G. (2010). Feasibility of internet-based health-related quality of life data collection in a large patient cohort. Journal of medical Internet research, 12, e35.

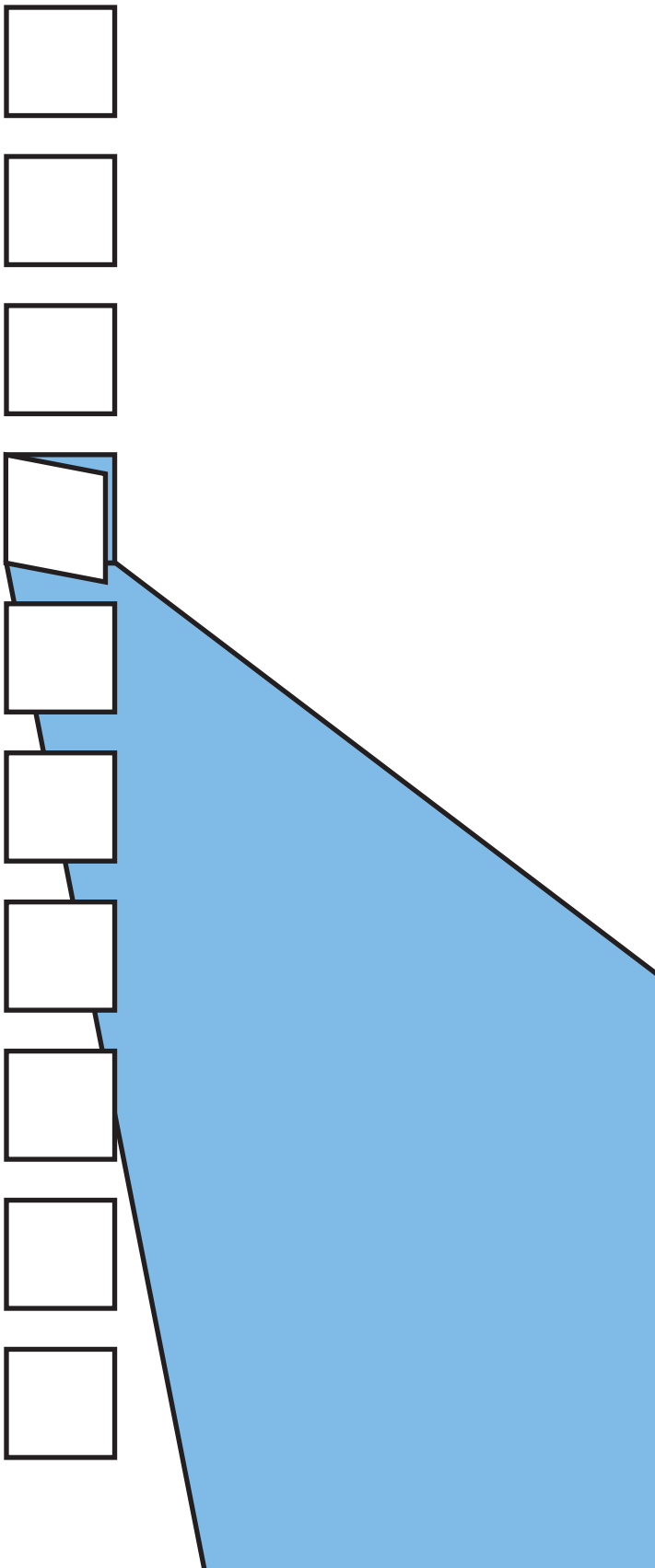
28. Joeris, A., Knoll, C., Kalampoki, V., Blumenthal, A., Gaskell, G., & Kalampoki, V. (2018). Patient-reported outcome measurements in clinical routine of trauma , spine and craniomaxillofacial surgeons: Between expectations and reality: A survey among 1212 surgeons. BMJ Open, 8, 1–11.

29. Barr, P. J., Forcino, R. C., Thompson, R., Ozanne, E. M., Arend, R., Castaldo, M. G., O'Malley, A. J., & Elwyn, G. (2017). Evaluating CollaboRATE in a clinical setting: Analysis of mode effects on scores, response rates and costs of data collection. BMJ Open, 7, 1–9.

30. Choi, J. K., Geller, J. A., Patrick, D. A., Wang, W., & Macaulay, W. (2015). How are those "lost to follow-up" patients really doing? A compliance comparison in arthroplasty patients. World journal of orthopedics, 6, 150.

31. Hutchings, A., Neuburger, J., Grosse Frie, K., Black, N., & van der Meulen, J. (2012). Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. Health and quality of life outcomes, 10, 34.

32. How many is enough? The quest for an acceptable survey response rate | Bright Ideas. https://kkbiersdorff.wordpress.com/2009/09/16/how-many-isenough/. Accessed 9 Jul 2018.



CHAPTER 4

What is the minimum response rate on patient-reported outcome measures needed to adequately evaluate total hip arthroplasties?

Y. Pronk, W. van der Weegen, R. Vos, J.M. Brinkman, R.J. van Heerwaarden and P. Pilot

Health Qual Life Outcomes. 2020 Dec 2;18(1):379

ABSTRACT

Background

Unknown is which response rate on patient-reported outcome measures (PROMs) is needed to both obtain an accurate outcome and ensure generalizability in evaluating total hip arthroplasty (THA) procedures. Without an evidence based minimum response rate (MRR) on THA PROMs, it is possible that hospitals report invalid patient-reported outcomes (PROs) due to a too low response rate. Alternatively, hospitals may invest too much in achieving an unnecessary high response rate. The aim of this study is to gain an insight into the MRR on PROMs needed to adequately evaluate THA procedures from a clinical perspective.

Methods

Retrospective study on prospective collected data of primary, elective THA procedures was performed. MRR was investigated for each PROM (NRS pain at rest, NRS pain during activity, EQ-5D-3L, HOOS-PS, anchor function, OHS, anchor pain and NRS satisfaction) separately to calculate the primary outcome: MRR for the THA PROMs set. MRR on a PROM needed to have (condition 1.) similar PRO change score (3 month score minus preoperative score) including confidence interval, (condition 2.) maintaining the influence of each change score predictor and (condition 3.) equal distribution of each predictor, as those of a 100% PROM response rate group. Per PROM, a 100%-group was identified with all patients having the PRO change score. Randomly assessed groups of 90% till 10% response rate (in total 90 groups) were compared with the 100%-group. Linear mixed model analyses and linear regressions were executed.

Results

The MRR for the THA PROMs set was 100% (range: 70–100% per PROM). The first condition resulted in a MRR of 60%, the second condition in a MRR of 100% and the third condition in a MRR of 10%.

Conclusions

A 100% response rate on PROMs is needed in order to adequately evaluate THA procedures from a clinical perspective. All stakeholders using THA PROs should be aware that 100% of the THA patients should respond on both preoperative and 3 month postoperative PROMs. For now, taking the first step in improving evaluation of THA for quality control by achieving at least two of the three conditions of MRR, advised is to require a response rate on PROMs of 60% as the lower limit.

INTRODUCTION

Total hip arthroplasty (THA) is performed to relieve pain, restore function and improve quality of life in patients with end-stage osteoarthritis. Patient-reported outcome measures (PROMs) gain insight into these results from a patients' perspective. Nowadays, patient-reported outcomes (PROs) are collected on a large scale to evaluate THAs in hospitals and to compare THA health care between hospitals. PROs are seen as useful information to reflect on the clinical work executed as even on clinicians' own executed care to improve patient care.

To draw valid conclusions on these evaluations, a certain response rate on PROMs is needed to both obtain an accurate outcome and ensure generalizability [1]. This minimum response rate (MRR), however, is unknown. The PROMs working group of International Society of Arthroplasty Registries (ISAR) advises a MRR of 60%. They mention that this is only based on the external difficulties to collect PROs that may be unrelated to survey logistics and the requirement of ≥60% for a survey study [2–4], however, without any further scientific evidence.

Since 2014, when THA PROs collection became mandatory in the Netherlands, huge differences are observed in response rate while comparing outcomes between Dutch hospitals; ranging from 10 to 100% preoperatively and from 2 to 95% at 3 months postoperatively [5, 6]. One might assume that these differences conceal a high risk of bias affecting the THA evaluation with PROs.

Achieving high PROMs response rate on multiple time points has proven to be even more challenging [7]. Even though automated collection systems are available, using these systems alone results in a moderate THA PROMs response rate on multiple time points (51%). A high response rate (> 90%) can be achieved with extra manual effort as sending paper questionnaires, but at an extra cost of around €6.0 per patient [7]. From a value-based health care perspective, it is debatable if these additional costs are justified as the MRR on PROMs for adequate evaluation of THA is unknown.

Without an evidence based MRR on THA PROMs, it is possible that hospitals report invalid PROs due to a too low response rate. Alternatively, hospitals may invest too much in achieving an unnecessary high response rate. Therefore, the aim of this study is to gain an insight into the MRR on PROMs needed to adequately evaluate THA procedures from a clinical perspective.

METHODS

A single centre retrospective study on prospective collected data from primary elective THA procedures was performed. THA procedures had been performed between March 2015 and December 2016 by three experienced high-volume orthopaedic surgeons in medium sized orthopaedic hospital (Kliniek ViaSana, Mill, the Netherlands). Patients were characterised by having an American Society of Anaesthesiologists (ASA) score of I or II, and a body mass index

(BMI) of ≤35. Before each THA procedure, patients were informed, and asked to participate in PROs collection and to allow further scientific analysis using their anonymised data. All patients gave written informed consent. This study was approved by the district medical ethics committee (N18.156).

PROs collection

The THA PROMs set included the mandatory PROMs as set out by the Dutch Orthopaedic Association (NOV) (Table 1) [4]. PROMs were collected preoperatively and at 3 months postoperatively with maximal effort to achieve 100% response rate [7]. PROs collection was preferably electronic using a digital, online, automated system (OnlinePROMs, Interactive Studios, Rosmalen, the Netherlands) with all questions obliged. In case patients were not or less able to handle a computer, paper questionnaires were sent by postal service. A maximum of three invitations to complete the questionnaires were sent. Patients with incomplete paper questionnaires were followed up by phone to complete all questionnaires [7]. Reasons for missing data were reported.

Table 1 Required and additional THA preoperative and 3 month postoperative PROMs [4]

THA PROMs set	PROM	Preoperative	3 months postoperative
Required PROMs	Pain by Numeric Rating Scale (NRS) – at rest (0 = no pain and 10 = unbearable pain)	✓	✓
	Pain by NRS – during activity (0 = no pain and 10 = unbearable pain)	✓	✓
	Quality of life by 3-level version of EuroQol 5 dimensions (EQ-5D-3L) (EQ VAS: 0 = worst imaginable health state and 100 = best imaginable health state; EQ-5D descriptive system: 0 = dead and 1 = healthy)	✓	✓
	Physical functioning by Hip disability and Osteoarthritis Outcome Score-Physical function Short-form (HOOS-PS) (0 = no difficulty and 100 = extreme difficulty) [8,9]	✓	✓
	Anchor hip function (1 = very much deteriorated and 7 = very much improved)		✓
Additional PROMs	Hip specific function and pain by Oxford Hip Score (OHS) (0 = least difficulty and 48 = most difficulty) [10]	✓	✓
	Anchor hip pain (1 = very much deteriorated and 7 = very much improved)		✓
	Satisfaction by NRS (0 = very dissatisfied and 10 = very satisfied)		✓

PROMs: patient-reported outcome measures; THA: total hip arthroplasty

Minimum response rate

The primary outcome was the MRR on the THA PROMs set, both required and additional PROMs, to adequately evaluate the results of THA. From a clinical perspective, evaluating the results of THA means evaluating the improvement patients made from before THA to a certain moment after THA. Minimal clinical important difference (MCID) does not yet exist for most THA PROMs, therefore, the change score was used as the best alternative. Three month change score (3 month score minus preoperative score) was utilized as this is a part of the Dutch PROMs indicator. Anchor questions regarding hip function and pain, and satisfaction question already measure a change, so these 3 month scores were seen as a change score.

The change score could be influenced by variables reported as predictors in previous studies: gender [11– 13], age on the day of surgery [14–17], BMI [15, 18], Charnley score [11–13], comorbidity [12, 15] and anxiety [13, 19]. If a predictor influences the change score of the total THA patient group in this study (100% response rate group), this influence should be observed in smaller groups (lower response rate groups) as well to maintain the effect of the predictor on the change score. Furthermore, these predictors (for example gender) should exist of the same proportion (for example females and males) at a lower response rate to maintain a generalizable sample of the total THA patient group.

Therefore, the MRR was investigated for each PROM total- or subscore separately to calculate the MRR for the THA PROMs set. The MRR on a PROM needed to have (condition 1.) the similar change score including confidence interval (CI), (condition 2.) maintaining the influence of each change score predictor and (condition 3.) the equal distribution of each predictor as those of a 100% PROM response rate group. Regarding the THA PROMs set included, only quality of life measured using the 3-level version of EuroQol 5 dimensions (EQ-5D-3L) existed of two subscores instead of one totalscore (Table 1).

Besides PROs, patients characteristics including the known THA PROs predictors were assessed. Gender, age on the day of surgery (years), preoperative BMI (kg/m^2), Charnley score (A, B1, B2, C), comorbidity (yes/no), ASA (I/II), osteoarthritis as diagnosis (yes/no) and complication (yes/no) were collected from the electronic patient records. Preoperative anxiety was measured using question 5 of the EQ-5D-3L of which answers 2 (moderately anxious or depressed) and 3 (extremely anxious or depressed) were grouped as having anxiety.

Patient selection

A THA procedure was included when the patient signed informed consent form, was a valid responder and had a change score on one of the PROMs. A response was considered valid if the patient responded within the NOV selected time period (preoperative questionnaires: maximum 182 day before surgery; 3 month questionnaires: between 63 and 110 days after surgery) [4]. There were no exclusion criteria.

Data analysis

Missing items were recalculated to complete the questionnaire if this was allowed according to the instrument-specific guidelines of the used questionnaires. To investigate if there was any difference between included and excluded THA procedures in patients characteristics including the predictors and preoperative PROs, independent t-tests or Mann–Whitney U tests for continuous variables were executed depending on the normal distribution of the data investigated using Shapiro–Wilk tests of normality and histograms, or Pearson's chi-square or Fisher's exact tests for categorical variables. Furthermore, variance patterns with respect to heteroscedasticity were investigated.

As missing PROs data are rarely MCAR and it was not sure if it was MAR of MNAR, to adopt an appropriate analytical strategy, three type of strategies were executed and results of the linear mix model analysis were compared: complete case analysis (MCAR or MAR), multiple impute missing data analysis with 200 imputations (MCAR or MAR) and sensitivity analyses (MNAR) [2]. These analyses were executed on the HOOS-PS which showed to have the most missing data. As no big deviations were found, complete case analysis was adapted in further analyses.

For each PROM total- or subscore, a 100%-group was identified with all included patients having the change score. Of this 100%-group, 10 times a random group of 90%, 10 times a random group of 80%, and so on for 70%, 60%, 50%, 40%, 30%, 20% and 10% were created (in total 91 groups). These groups were coded by the response rate and a random group number (for example 90,02). Linear mixed model analysis was used to assess differences in each PRO preoperatively and at 3 months postoperatively to investigate the change score of the 100%-group corrected by the 6 predictors. An unstructured covariance structure for the two repeated measures was used. This analysis method accounts for baseline differences and dependencies between repeated measures, and allowing unequal variances across groups. For PROs with one measurement (anchor questions hip function and pain, and satisfaction), this change score was analysed executing linear regression. P-values of the 6 predictors were checked. To compare the change score and the p-values of the predictors with all groups, in each group the same linear mixed model analysis or linear regression was performed. All group change scores with 95% CI or range were visualised in a graph (MRR condition 1). Regarding the predictors, defined was that 8 or more of the 10 groups of a certain response rate needed to have the same statistically significant or non-significant level as the 100%-group to be adequate (MRR condition 2).

To compare equal distribution of each predictor in each group to the 100%-group, Pearson's chi-square or Fisher's exact tests were performed. Defined as adequate was that 8 or more of the 10 groups of a certain response rate had to have an equal distribution of a predictor (MRR condition 3). For this step, both age and BMI were transformed to categorical variables. Age was recorded to 5 groups: < 50 years, 50–59 years, 60–69 years, 70–79 years and ≥80 years. BMI was categorised to underweight (≤ 18.5), normal weight (> 18.5–25.0), overweight (> 25.0–30.0) and obesity (> 30.0–40.0) [20].

For all statistical analyses, an alpha of 0.05 was considered statistically significant and IBM SPSS Statistics 25.0 (IBM Corporation, U.S.) was used.

RESULTS

During the study period 622 THA procedures (592 patients) were performed of which 616 (99.8%) were valid responders preoperatively and 557 (92.2%) at 3 months. Finally, 552 (88.8%) THA procedures were included. Main reasons for exclusion were no response preoperative and/or at 3 months postoperatively (n=36 (5.8%)) and a response outside the valid preoperative and/or at 3 month postoperative response period (n=30 (4.8%)). Of the 552 included THA procedures, 474 had all change scores available, the remaining 78 at least one (Fig. 1). No statistical significant differences regarding patients characteristics and preoperative PROs were found between the included and excluded THA procedures (Table 2).

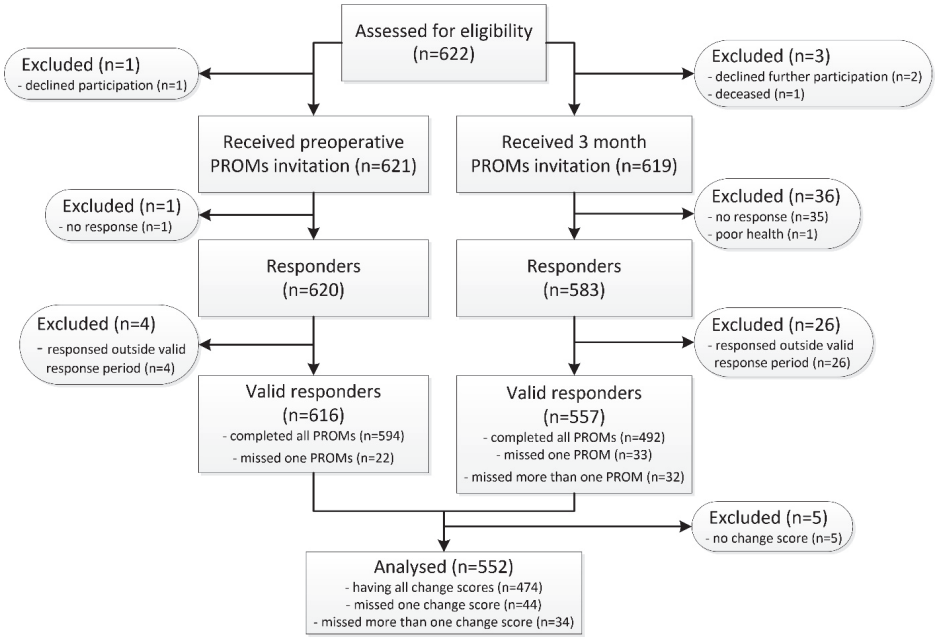


Fig. 1 Study flowchart

n: number; PROMs: patient-reported outcome measures

Table 2 Patients characteristics and preoperative PROs of included and excluded THA procedures

Patients characteristics and preoperative PROs	Included THA procedures n=552	Excluded THA procedures n=70	p-value
ASA classification (II; n (%))	284 (51%)	28 (40%)	0.071
Age on date of surgery (years; median (IQR))	66 (60-71)	65 (55-74)	0.805
BMI (kg/m²; median (IQR))	26.00 (23.90-28.41)	26.29 (24.48-28.13)	0.389
Gender (male; n (%))	209 (38%)	31 (44%)	0.298
Diagnosis (osteoarthritis; n (%))	486 (88%)	60 (86%)	0.575
Charnley score (n (%))			0.064
A – one hip joint affected	135 (24%)	15 (21%)	
B1 – both hip joints affected	245 (44%)	23 (33%)	
B2 – contralateral hip joint with a total hip prothesis	110 (20%)	17 (24%)	
C – multiple joints affected	62 (11%)	15 (21%)	
Comorbidity (yes, n (%))	178 (32%)	23 (33%)	0.918
Anxiety (n (%))	123 (22%)	20 (29%)	0.188
Preoperative NRS pain at rest (median (IQR))	6 (4-7)	5 (4-7)	0.543
Preoperative NRS pain during activity (median (IQR))	8 (7-9)	8 (7-9)	0.363
Preoperative EQ-5D descriptive system (median (IQR))	0.693 (0.310-0.775)	0.693 (0.335-0.775)	0.625
Preoperative EQ VAS (median (IQR))	80 (60-87)	77 (66-85)	0.960
Preoperative HOOS-PS (median (IQR))	46.1 (37.7-55.9)	50.8 (41.7-55.9)	0.341
Preoperative OHS (median (IQR))	24 (18-29)	24 (17-29)	0.454
Complication (yes, n (%))	33 (6%)	8 (11%)	0.118

ASA: American Society of Anaesthesiologists; BMI: body mass index; EQ-5D descriptive system: EuroQol 5 dimensions descriptive system; EQ VAS: EuroQol Visual Analogical scale; HOOS-PS: Hip disability and Osteoarthritis Outcome Score-Physical function Short-form; NRS: numeric rating scale; OHS: Oxford Hip Score; PROs: patient-reported outcomes; THA: total hip arthroplasty

Missing data

Most of the 78 patients, who had not all change scores, had no HOOS-PS change score due to missing data in the HOOS-PS 3 month questionnaire (n=59 (10.7%)) or had no EQ VAS change score due to missing data in the EQ VAS question at 3 months (n=31 (5.6%)). Main reason for missing data on this HOOS-PS 3 month questionnaire was about the item running. Patients were advised not to run after THA surgery and the question asked to indicate the degree of difficulty experienced in performing this activity.

Different strategies for missing data were executed. Mixed model analysis with complete cases reported a mean HOOS-PS change score of -32.4 (CI: −34.1–−30.8) (n=480), with multiple imputed

missing data a mean of −32.5 (CI: −32.6–−32.4), with imputed worst scores a mean of −33.2 (CI: −34.9–−31.5) (n=552) and with imputed best scores a mean of -29.1 (CI: −31.1–−27.1) (n=552). Maximum difference between these strategies was 4.1 points for the change score resulted in a 2.1% difference on the HOOS-PS change score scale of −100 to 100. The CI ranged from 0.2 to 4.0 in size. Only in the analysis with imputed worst scores, the predictor anxiety was not a significant predictor (p=0.053) and age was (p=0.001). The estimate changes of the predictors were, however, similar in all analyses. Based on these small differences found, complete case analysis was adapted in further analyses.

MRR for NRS pain at rest

In the 100% NRS-pain-at-rest-group the mean change score was −4.4 (CI: −4.6–−4.2) (n=551) which was no longer similar when the response rate dropped below 30%. Mean change score in the 20%-groups was −4.4 (CI: −4.8–−3.9). This score was similar and the CI was 2.3 times (230%) greater (0.9 versus 0.4) compared to the 100%-group (Fig. 2; condition 1). Gender (p=0.001), comorbidity (p=0.041), age (p=0.002) and BMI (p=0.018) were significant predictors in the 100%-group which remained down to and including the 60%, 100%, 60% and 100%-group respectively. Charnley score and anxiety remained no significant predictors down to and including the 10%-groups (condition 2). Equal distributions of all predictors were observed down to the 10%-groups inclusive compared to the 100%-group (Table 3; condition 3).

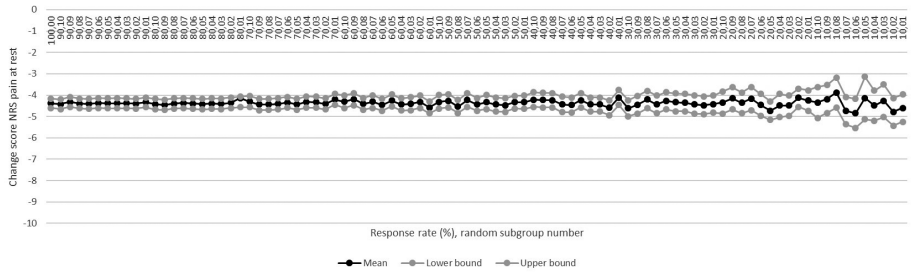


Fig. 2 Mean NRS pain at rest change score per group

NRS: numeric rating scale

Table 3 Number of NRS pain at rest groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.435	x	0.001	x	0.041	x	0.631	x	0.002	x	0.018	x
90% (n)	0	10	10	10	7	10	0	10	10	10	7	10
80% (n)	0	10	10	10	5	10	0	10	10	10	7	10
70% (n)	0	10	8	10	3	10	0	10	10	10	3	10
60% (n)	0	10	9	10	1	10	0	10	8	10	5	10
50% (n)	1	10	7	10	2	10	0	10	6	10	2	10
40% (n)	0	10	6	10	3	10	0	10	3	10	4	10
30% (n)	2	10	3	10	0	10	1	10	1	10	5	10
20% (n)	1	10	2	10	0	10	0	10	4	10	2	10
10% (n)	0	9	1	10	2	10	0	10	2	10	2	10

BMI: body mass index

MRR for NRS pain during activity

The mean change score of −5.4 (CI: −5.6–−5.2) (n=551) found in the 100% NRS-pain-during-activity-group was observed down to and including the 30%-groups. In the 20%-groups, the mean change score was −5.4 (CI: −5.9–−4.9). Compared to the 100%-group, this score was similar and the CI was 2.5 times (250%) greater (1.0 versus 0.4) (Additional file 1, Fig. 1; condition 1). Gender (p=0.000) and age (p=0.000) were significant predictors for this change score in the 100%-group which remained down to and including the 40% and 60%-groups respectively. BMI remained a non-significant predictor down to the 100%-group. The other predictors stayed non-significant predictors in all groups (condition 2). Down to the 10%-groups inclusive, equal distribution of all predictors was found compared to the 100%-group (Additional file 1, Table 1; condition 3).

MRR for EQ-5D-3L

EQ-5D descriptive system

The mean change score of 0.250 (CI: 0.225–0.274) in the 100% EQ-5D descriptive system group (n=544) was observed down to and including the 30%-groups. The 20%-groups reported a mean change score of 0.249 (CI: 0.195–0.303). This score differed 0.001 points (0.4%) and the CI was 2.2 times (220%) greater (0.108 versus 0.049) compared to the 100%-group (Additional file 1, Fig. 2; condition 1). Regarding the significant predictors, gender (p=0.001) was found to be a significant predictor down to the 50%-groups inclusive, anxiety (p=0.000) to 10%, age (p=0.004) to 80% and BMI (p=0.019) to 100%. Comorbidity remained a non-significant predictor down to and including the 60%-groups (condition 2). All predictors were equal distributed down to the 10%-groups inclusive compared to the 100%-group (Additional file 1, Table 2; condition 3).

EQ VAS

The 100% EQ VAS group had a mean EQ VAS change score of 7.1 (CI: 5.3–8.8) (n=521) and showed to remain similar down to and including the 40%-groups. Mean change score in the 30%-groups was 7.2 (CI: 4.0–10.5). Compared to the 100%-group, this score differed 0.1 point (1.4%) and the CI was 1.9 times (190%) greater (6.5 versus 3.5) (Additional file 1, Fig. 3; condition 1). Gender (p=0.001), comorbidity (p=0.003) and anxiety (p=0.000) were significant predictors in the 100%-group and down to the 70%, 60% and 50%-groups inclusive respectively. The other predictors remained non-significant predictors in all groups (condition 2). Equal distribution was found down to and including the 10%-groups for all predictors compared to the 100%-group (Additional file 1, Table 3; condition 3).

MRR for HOOS-PS

The mean change score of the 100% HOOS-PS group was −32.4 (CI: −34.1–−30.8) (n=480) and found to be similar down to and including the 40%-groups. The 30%-groups reported a mean change score of −32.2 (CI: −35.1–−29.2). This score differed 0.2 points (0.6%) and the CI was 1.8 times (180%) greater (5.9 versus 3.3) compared to the 100%-group (Additional file 1, Fig. 4; condition 1). Significant predictors were gender (p=0.000) and anxiety (p=0.003) which both remained down to the 60%-groups inclusive. Charnley score and BMI stayed non-significant predictors down to the 60% and 90%-groups inclusive respectively (condition 2). All predictors

were equally distributed down to and including the 10%-groups compared to the 100%-group (Additional file 1, Table 4; condition 3).

MRR anchor hip function

The mean anchor hip function was 5.8 (CI: 5.3–6.2) in the 100%-group (n=540) and showed to be similar down to and including the 60%-groups. Regarding the 50%-groups, the mean score was 5.8 (CI: 5.2–6.4). This score was similar and the CI was 1.3 times (133%) greater (1.2 vs. 0.9) compared to the 100%-group (Fig. 3; condition 1). In the 100%-group, there were no significant predictors which remained down to and including the 60%-groups for gender and for comorbidity, the 90%-groups for BMI and the 10%-groups for the other predictors (condition 2). Equal distribution was found in all predictors down to the 10%-groups inclusive compared to the 100%-group (Table 4; condition 3).

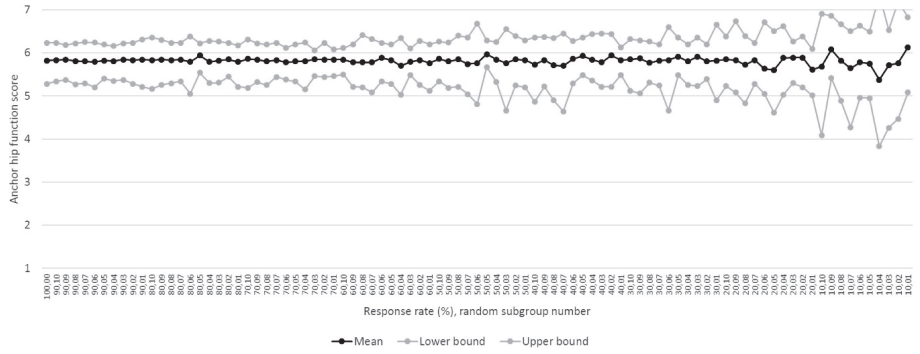


Fig. 3 Mean anchor hip function score per group

Table 4 Number of anchor function groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.339	x	0.113	x	0.248	x	0.341	x	0.496	x	0.051	x
90% (n)	0	10	0	10	0	10	0	10	0	10	2	10
80% (n)	0	10	1	10	0	10	0	10	0	10	4	10
70% (n)	0	10	2	10	0	10	0	10	0	10	2	10
60% (n)	0	10	0	10	1	10	0	10	0	10	0	10
50% (n)	1	10	3	10	3	10	1	10	0	10	2	10
40% (n)	0	10	2	10	0	10	1	10	0	10	3	10
30% (n)	1	10	0	10	0	10	0	10	0	10	2	10
20% (n)	1	10	0	8	0	10	0	10	0	10	1	10
10% (n)	1	10	0	8	1	9	0	10	1	10	3	10

BMI: body mass index

MRR for OHS

In the 100% OHS group a mean change score of 16.4 (CI: 15.7–17.1) was found (n=542) and observed to be similar down to and including the 30%-groups. The 20%-groups had a mean change score of 16.0 (CI: 14.4–17.6). Compared to the 100%-group, this score differed 0.4 points (2.4%) and the CI was 2.3 times (230%) greater (3.2 vs. 1.4) (Fig. 4; condition 1). Regarding the predictors, gender (p=0.000), anxiety (p=0.000), age (p=0.016) and BMI (p=0.001) were significant predictors in the 100%-group which remained down to the 50%, 30%, 100% and 50%-groups inclusive respectively. Both Charnley score and comorbidity stayed non-significant predictors (condition 2). Down to and including the 10%-groups, all predictors showed to have an equal distribution compared to the 100%-group (Table 5; condition 3).

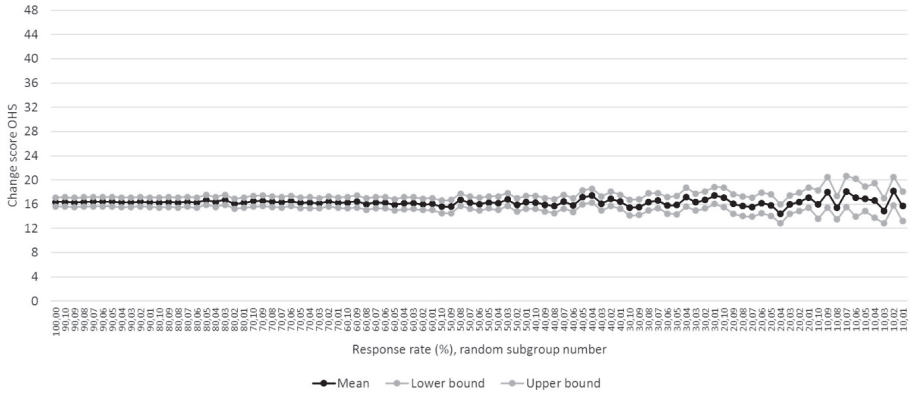


Fig. 4 Mean OHS change score per group

OHS: Oxford Hip Score

Table 5 Number of OHS groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.306	x	0.000	x	0.234	x	0.000	x	0.016	x	0.001	x
90% (n)	0	10	10	10	0	10	10	10	6	10	10	10
80% (n)	0	10	10	10	0	10	10	10	8	10	10	10
70% (n)	0	10	10	10	1	10	10	10	4	10	9	10
60% (n)	1	10	9	10	1	10	10	10	6	10	8	10
50% (n)	1	10	9	10	0	10	10	9	4	10	9	10
40% (n)	1	10	7	10	0	10	10	10	5	10	4	10
30% (n)	0	10	6	10	0	10	10	10	2	10	3	10
20% (n)	1	10	7	10	0	10	2	9	2	10	3	10
10% (n)	0	10	1	10	1	10	1	10	1	9	1	10

BMI: body mass index

MRR for anchor hip pain

The 100% anchor hip pain group had a mean score of 6.2 (CI: 5.7–6.5) (n=539) and showed to be similar down to and including the 50%-groups. The 40%-groups had a mean score of 6.2 (CI: 5.7–6.6). This score was similar and the CI was 1.1 times (110%) greater (0.9 versus 0.8) compared to the 100%-group (Additional file 1, Fig. 5; condition 1). Significant predictors of this score were gender (p=0.040) and comorbidity (p=0.022) in the 100%-group, both remaining significant down to the 100%-group inclusive. The other predictors stayed non-significant predictors in all groups (condition 2). Down to and including the 10%-groups, all predictors were equally distributed compared to the 100%-group (Additional file 1, Table 5; condition 3).

MRR for satisfaction

The mean NRS satisfaction score in the 100%-group was 8.5 (CI: 7.5–9.3) (n=537) and was observed to be similar down to and including the 60%-groups. The 50%-groups reported a mean score of 8.6 (CI: 7.5–9.4). This score differed 0.1 points (1.2%) and the CI was 1.1 (110%) greater (1.9 versus 1.8) compared to the 100%-group (Additional file 1, Fig. 6; condition 1). In the 100%-group, gender (p=0.013) and BMI (p=0.029) were significant predictors which stayed down to and including the 90% and 100%-group respectively. Age and the other predictors remained non-significant predictors down to the 30% and 100%-group inclusive respectively (condition 2). Compared to the 100%-group, equal distribution was found in all predictors down to the 10%-groups inclusive (Additional file 1, Table 6; condition 3).

Table 6 MRR for each THA PROM including per complied condition

THA PROMs set	PROM	1.Similar change score (%)	2.Maintaining influence of predictors (%)	3.Equal distribution of predictors (%)	MRR (%)
Required	NRS pain at rest	30	100	10	100
	NRS pain during activity	30	100	10	100
	EQ-5D-3L				
	EQ-5D descriptive system	30	100	10	100
	EQ VAS	40	70	10	70
	HOOS-PS	40	90	10	90
	Anchor hip function	60	90	10	90
Required set		**60**	**100**	**10**	**100**
Additional	OHS	30	100	10	100
	Anchor hip pain	50	100	10	100
	Satisfaction	60	100	1	100
Total set		**60**	**100**	**10**	**100**

EQ-5D descriptive system: EuroQol 5 dimensions descriptive system; EQ VAS: EuroQol Visual Analogue Scale; HOOS-PS: Hip disability and Osteoarthritis Outcome Score-Physical function Short-form; MRR: minimum response rate; NRS: numeric rating scale; OHS: Oxford Hip Score; PROMs: patient-reported outcome measures; THA: total hip arthroplasty

MRR for THA PROMs set

To investigate the MRR of the THA PROMs set, summarized: condition 1 resulted in a MRR of 60% (30–60%) for both the total THA PROMs set as only the required THA PROMs set, condition 2 in a MRR of 100% (70–100%) respectively and condition 3 in a MRR of 10% (10–10%) respectively. MRR per PROM ranged from 70 to 100% (Table 6).

DISCUSSION

Gaining an insight into the response rate on PROMs needed to adequately evaluate THA procedures from a clinical perspective was the aim of this study. Results show that for the Dutch THA PROMs set a 100% (range: 70% to 100% per PROM) response rate is needed. It was not possible to lower this MRR of 100% due to not maintaining the influence of each change score predictor at a lower response rate (condition 2). Still measuring the similar change score (condition 1) resulted in a MRR of 60% and still maintaining equal distribution of each predictor (condition 3) in a MRR of 10%.

In many countries, PROs are measured routinely and incorporated into arthroplasty registers. PROs are evaluated in hospitals, compared between hospitals and even financial incentives are based on these outcomes. For each hospital as even for each clinician, PROs are seen as useful information to reflect on the clinical work executed to improve patient care. From a clinical perspective, for adequate evaluation of THA with PROs a response rate of 100% is needed, shown by the current study (Table 6). This means that 100% of the THA patients should respond on the preoperative PROMs as well as on the 3 month postoperative PROMs. However, it is impossible to achieve this in clinical practice. None of the hospitals reached a 100% response rate on THA PROMs preoperatively as well as postoperatively; mean reported response rate on both time points is 37% in the Dutch register and 79% in the Swedish register [6, 21].

A first step in improving THA evaluation with PROs from a clinical perspective for quality control can be made by achieving at least two of the three MRR conditions (Table 6). This results in a MRR of 60% as the lower limit of evaluating THA outcome using PROs meaning 60% of the patients should be a responder on the preoperative as well as on the 3 month postoperative PROMs. Advised is to discard PROs collected below 60% to prevent for both invalid in-hospital evaluation as for invalid comparisons between hospitals. As a consequence, to achieve the lower limit of 60%, ISAR should tighten up their MRR advice and hospitals should increase their response rates beyond 60% if they are not there yet.

Interestingly, to a certain extent lower response rates are acceptable provided that MCIDs are evaluated [22]. Comparison between PROs of patients with lumbar discectomy incorporated into the Swedish spine register with PROs of the same patient population of a single hospital showed significant different change scores in PROs, but all within the MCIDs [22]. It could be that in the present study the observed differences in change scores in lower response rate groups

compared to the 100%-group are still within clinical relevant difference. However, yet no MCIDs or comparable values are available for most THA PROMs as even the best method to determine them [23, 24]. One study investigated and reported the 6 month OHS MCID at group level of around 11 points [25]. Comparing this with the results of the present study, MRR for OHS could be 10% instead of 30% (Fig. 4). The current study should be repeated when these MCIDs based on a golden standard method to determine them are known.

Although practice shows difficulties in achieving high response rates, response rates of > 80% are achievable in orthopaedic patients [7, 26–29]. It is even shown to be feasible to achieve > 90% response rate in busy orthopaedic hospitals, urban and rural, using a digital collection system without any major disruption to the clinical work flow [29]. As seen in the current study, ASA classification and Charnley score were almost significant predictors for being a responder or not. However, achieving high response rates depends more on the method in PROs collection chosen. Making PROs collection a part of routine care, using a PROMs digital administration station in the hospital and collecting via multiple sources (for example mail and email) are the keys to high response rates [7, 27–29]. In arthroplasty patients, a critical factor is making sure PROs are collected preoperatively as it results in a 3 times more chance of collecting the PROs 3 months after surgery and even a 15 times more chance at 12 months [30]. Maintaining high postoperative response rates is crucial as non-responding patients can introduce bias which results in incomparable PROs if the non-responders are different than the responders [28, 31] and missing data are not at random [32]. Therefore, it is advised that hospitals should take the winners in effort and costs in this method to at least reach the lower limit of 60% response rate.

For this first study tackling the methodological challenge in investigating the required response rate to ensure THA PROs could be used to adequately evaluate THA procedures from a clinical perspective, several assumptions had to be made to create a starting point in clarifying this issue. This study used the change scores at 3 months postoperatively (towards preoperative). Complexity exists as this study should be repeated for change scores at 12 and 24 months postoperatively towards preoperative and even at 12 and 24 months postoperatively towards 3 months postoperatively to have a more complete answer. Acquiring a complete PROMs dataset including also 12 and 24 months results is even more challenging than a dataset including only preoperative and 3 months results. The method chosen for this challenge was considered as the only option due to unequal variances and unknown MCIDs. Future research should investigate if the MCIDs instead of change scores remain similar in lower response rates when these MCIDs are available. Another assumption made was that all three conditions are of the same value. Future research should investigate if this is indeed the case. Case-mix is important in investigating MRR. Based on previous literature, six predictors were incorporated in all three conditions besides only correcting for them to adjust the change score in condition 1. As case-mix is another methodological challenge, future research should take the next step in the influence of the case-mix on the MRR (for example interaction between predictors). As another strength, different strategies for dealing with missing data were checked to see if there were substantial deviations. As a limitation, the results of the present study are not completely generalizable as the included

patients were characterised with ASA I-II and BMI below 35, which represent around 80% of the total THA population [20]. Patients with higher ASA classification and a higher BMI mostly score worse on the THA PROMs [33]. Adding this group to the study group of the current study will result in a more heterogeneous patient group. The mean change score will be lower and a larger CI is expected. It would be harder to comply the MRR conditions in lower response rate groups. Therefore, the MRR will be higher. Expected is that the more homogeneous the patient group is, the lower the MRR could be. Therefore, external validation of the results in a variety of hospitals settings is needed. This study was executed in a medium sized orthopaedic hospital. Another suggestion for further research is to investigate the minimum response number instead of MRR (percentage) as hospitals could be small or large in THA volume. Expected is that a combination of number and percentage is needed.

In general, PROs collection has already begun to yield results. However, there is still much work to do until significant benefits with respect to evaluating THA and improving patient care are found [34, 35]. Studies such as the present study are important, since PROs are increasingly transparent and publicly available while current validity is questionable without sufficient scientific evidence on the possible effects of (in)complete PROs collection. Health care providers, decision makers and payers are often unaware of these effects.

CONCLUSIONS

To adequately evaluate THA procedures from a clinical perspective in theory a response rate on PROMs of 100% is needed. All stakeholders using THA PROs should be aware that 100% of the THA patients should respond on both preoperative and 3 month postoperative PROMs to measure similar change scores, to keep the influence of each change score predictor and to maintain a representative random sample of THA patients. For now, taking the first step in improving evaluation of THA for quality control, advised is to require that 60% of the THA patients should be responders on both time points as the lower limit in evaluating THA PROs.

REFERENCES

1. Paulsen A. Patient reported outcomes in hip arthroplasty registries. Dan Med J. 2014;61:B4845.
2. Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, et al. Patient-reported outcome measures in arthroplasty registries: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. Acta Orthop. 2016;87:9–23.
3. JAMA. Instructions for Authors. https://jamanetwork.com/journals/jama/pages/instructions-for-authors. Accessed Aug 2018.
4. Nederlandse Orthopaedische Vereniging. PROMs. https://www.lroi.nl/invoerders/registreren/proms. Accessed Aug 2018.
5. Nederlandse Orthopaedische Vereniging. Pre-operative PROMs response percentage hip. http://www.lroi-rapportage.nl/hip-2018-proms-hip-2018-proms-response-pre-operative-proms. Accessed Aug 2019.
6. Nederlandse Orthopaedische Vereniging. Three months postoperative PROMs response percentage hip. http://www.lroi-rapportage.nl/hip-2018-proms-hip-2018-proms-response-three-months-postoperative-proms. Accessed Aug 2019.
7. Pronk Y, Pilot P, Brinkman JM, van Heerwaarden RJ, van der Weegen W. Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection. J Patient-Reported Outcomes. 2019;3:1–8.
8. Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert JF, 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;17:843–7.
9. Davis AM, Perruccio AV, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative. Osteoarthr Cartil. 2008;16:551–9.
10. Gosens T, Hoefnagels NHM, De Vet RCW, Dhert WJA, Van Langelaan EJ, Bulstra SK, 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;76:204–11.
11. Gordon M, Frumento P, Sköldenberg O, Greene M, Garellick G, Rolfson O. Women in Charnley class C fail to improve in mobility to a higher degree after total hip replacement. Acta Orthop. 2014;85:335–41.
12. Greene ME, Rolfson O, Nemes S, Gordon M, Malchau H, Garellick G. Education attainment is associated with patient-reported outcomes: Findings from the Swedish hip arthroplasty register. Clin Orthop Relat Res. 2014;472:1868–76.
13. Rolfson O, Dahlberg LE, Nilsson JÅ, Malchau H, Garellick G. Variables determining outcome in total hip replacement surgery. J Bone Jt Surg - Ser B. 2009;91:157–61.
14. Gordon M, Greene M, Frumento P, Rolfson O, Garellick G, Stark A. Age- and health-related quality of life after total hip replacement. Acta Orthop. 2014;85:244–9.
15. Judge A, Arden NK, Batra RN, Thomas G, Beard D, Javaid MK, et al. The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-replacement surgery: a prospective cohort study. BMJ Open. 2013;3:1–11.
16. Aalund PK, Glassou EN, Hansen TB. The impact of age and preoperative health-related quality of life on patient-reported improvements after total hip arthroplasty. Clin Interv Aging. 2017;12:1951–6.
17. Clement ND, MacDonald D, Howie CR, Biant LC. The outcome of primary total hip and knee arthroplasty in patients aged 80 years or more. J Bone Jt Surg - Ser B. 2011;93 B:1265–70.
18. Judge A, Batra RN, Thomas GE, Beard D, Javaid MK, Murray DW, et al. Body mass index is not a clinically meaningful predictor of patient reported outcomes of primary hip replacement surgery: Prospective cohort study. Osteoarthr Cartil. 2014;22:431–9.
19. Duivenvoorden T, Vissers MM, Verhaar JAN, Busschbach JJV, Gosens T, Bloem RM, et al. Anxiety and depressive symptoms before and after total hip and knee arthroplasty: a prospective multicentre study. Osteoarthr Cartil. 2013;21:1834–40.
20. Nederlandse Orthopaedische Vereniging. Patients characteristics by diagnosis - primairy THA. http://www.lroi-rapportage.nl/hip-total-hip-arthroplasty-demographics-patient-characteristics-by-diagnosis2018. Accessed Aug 2019.

21. Rolfson O, Karrholm J, Dahlberg LE, Garellick G. Patient-reported outcomes in the Swedish Hip Arthroplasty Register: Results of a nationwide prospective observational study. Bone Joint J. 2011;93-B:867–75.
22. Elkan P, Lagerbäck T, Möller H, Gerdhem P. Response rate does not affect patient-reported outcome after lumbar discectomy. Eur Spine J. 2018;27:1538–46.
23. Çelik D, Çoban Ö, Kılıçoğlu Ö. Minimal clinically important difference of commonly used hip-, knee-, foot-, and ankle-specific questionnaires: a systematic review. J Clin Epidemiol. 2019;113:44–57.
24. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol. 2010;63:524–34.
25. Beard DJ, Harris K, Dawson J, Doll H, Murray DW, Carr AJ, et al. Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. J Clin Epidemiol. 2015;68:73–9.
26. Viveen J, Prkic A, The B, Koenraadt KLM, Eygendaal D. Effect of introducing an online system on the follow-up of elbow arthroplasty. World J Orthop. 2016;7:826–31.
27. Ho A, Purdie C, Tirosh O, Tran P. Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. Patient Relat Outcome Meas. 2019;10:217–26.
28. Tariq MB, Vega JF, Westermann R, Jones M, Spindler KP. Arthroplasty studies with greater than 1000 participants: analysis of follow-up methods. Arthroplast Today. 2019;5:243–50.
29. Slover JD, Karia RJ, Hauer C, Gelber Z, Band PA, Graham J. Feasibility of integrating standardized patient-reported outcomes in orthopedic care. Am J Manag Care. 2015;21:e494-500.
30. Patel J, Lee JH, Li Z, SooHoo NF, Bozic K, Huddleston JI. Predictors of low patient-reported outcomes response rates in the california joint replacement registry. J Arthroplasty. 2015;30:2071–5.
31. Norquist BM, Goldberg BA, Matsen FA. Challenges in evaluating patients lost to follow-up in clinical studies of rotator cuff tears. J Bone Jt Surg - Ser A. 2000;82:838–42.
32. Kristman V, Manno M, Côté P. Loss to follow-up in cohort studies: how much is too much? Eur J Epidemiol. 2004;19:751–60.
33. Peters RM, van Steenbergen LN, Stewart RE, Stevens M, Rijk PC, Bulstra SK, et al. Which patients improve most after total hip arthroplasty? Influence of patient characteristics on patient-reported outcome measures of 22,357 total hip arthroplasties in the Dutch Arthroplasty Register. HIP Int. 2020. https://doi.org/10.1177/1120700020913208.
34. Greenhalgh J, Dalkin S, Gibbons E, Wright J, Valderas JM, Meads D, et al. How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis. J Heal Serv Res Policy. 2018;23:57–65.
35. Wilson I, Bohm E, Lübbeke A, Lyman S, Overgaard S, Rolfson O, et al. Orthopaedic registries with patient-reported outcome measures. EFFORT Open Rev. 2019;4:357–67.

ADDITIONAL FILE 1

Additional figures and tables

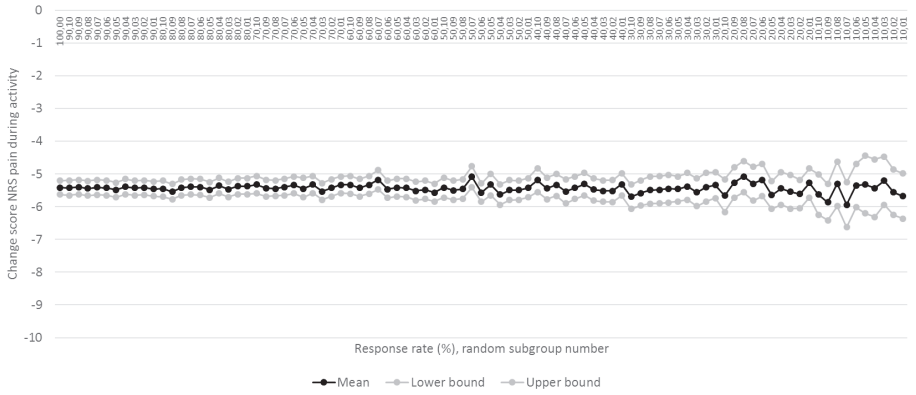


Fig. 1 Mean NRS pain during activity change score per group

NRS: numeric rating scale

Table 1 Number of NRS pain during activity groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.564	x	0.000	x	0.532	x	0.444	x	0.000	x	0.051	x
90% (n)	0	10	10	10	0	10	0	10	10	10	5	10
80% (n)	0	10	10	10	0	10	0	10	10	10	3	10
70% (n)	0	10	10	10	0	10	0	10	10	10	1	10
60% (n)	0	10	10	10	0	10	1	10	10	10	1	10
50% (n)	1	10	9	10	0	10	1	10	7	10	3	10
40% (n)	0	10	9	10	0	10	1	10	3	10	3	10
30% (n)	0	10	7	10	0	10	0	10	4	10	4	10
20% (n)	0	10	2	10	1	10	0	10	4	10	2	10
10% (n)	0	9	0	10	1	10	0	10	2	10	1	10

BMI: body mass index; NRS: numeric rating scale

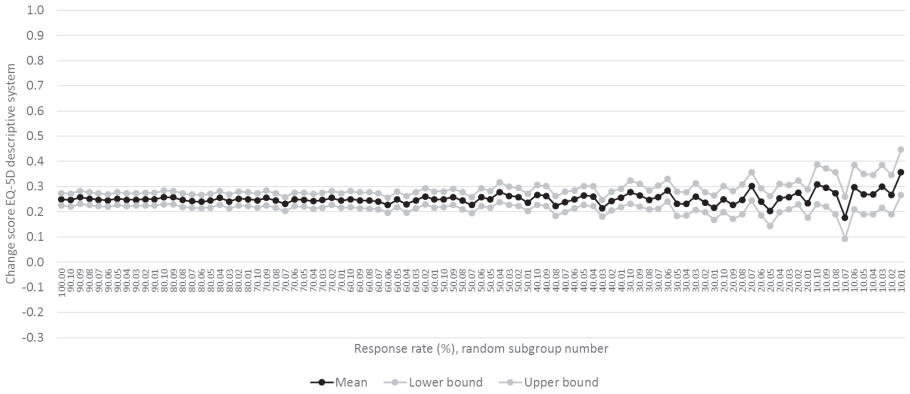


Fig. 2 Mean change score EQ-5D descriptive system per group

EQ-5D descriptive system: EuroQol 5 dimensions descriptive system

Table 2 Number of EQ-5D descriptive system groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.429	x	0.001	x	0.162	x	0.000	x	0.004	x	0.019	x
90% (n)	0	10	10	10	1	10	10	10	10	10	5	10
80% (n)	0	10	10	10	0	10	10	10	10	10	5	10
70% (n)	0	10	10	10	1	10	10	10	6	10	5	10
60% (n)	1	10	9	10	1	10	10	10	7	10	4	10
50% (n)	0	10	8	10	3	10	10	10	4	10	3	10
40% (n)	1	10	4	10	2	10	10	10	5	10	3	10
30% (n)	1	10	6	10	0	10	10	10	5	10	3	10
20% (n)	0	9	2	9	0	10	10	10	2	10	1	10
10% (n)	1	10	0	10	0	10	2	9	0	10	1	10

BMI: body mass index; EQ-5D descriptive system: EuroQol 5 dimensions descriptive system; NRS: numeric rating scale

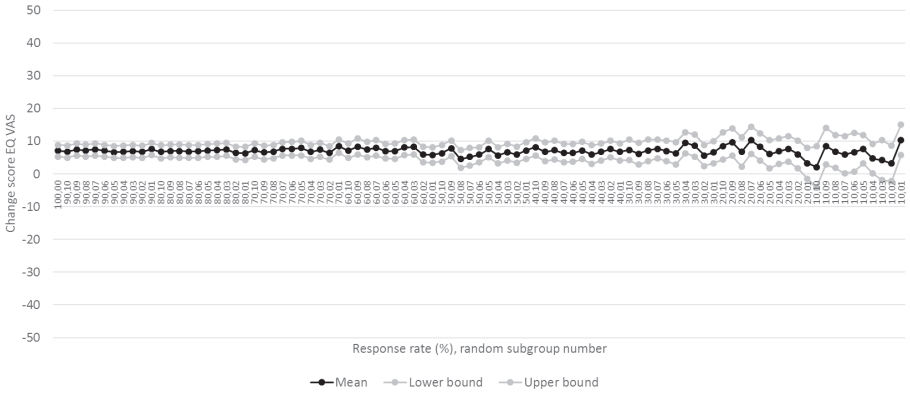


Fig. 3 Mean EQ VAS change score per group

EQ VAS: EuroQol Visual Analogue Scale

Table 3 Number of EQ VAS groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.722	x	0.001	x	0.003	x	0.000	x	0.245	x	0.821	x
90% (n)	0	10	10	10	10	10	10	10	0	10	0	10
80% (n)	0	10	10	10	10	10	10	10	1	10	0	10
70% (n)	0	10	10	10	9	10	10	10	0	10	0	10
60% (n)	0	10	7	10	9	9	10	10	1	10	0	10
50% (n)	1	10	7	10	6	10	10	10	1	10	0	10
40% (n)	0	10	6	10	5	10	7	10	0	10	0	10
30% (n)	0	10	4	10	2	10	8	10	0	10	0	10
20% (n)	0	10	2	10	5	10	4	9	0	10	0	10
10% (n)	1	10	2	9	1	9	1	10	1	10	1	10

BMI: body mass index; EQ VAS: EuroQol Visual Analogue Scale; NRS: numeric rating scale

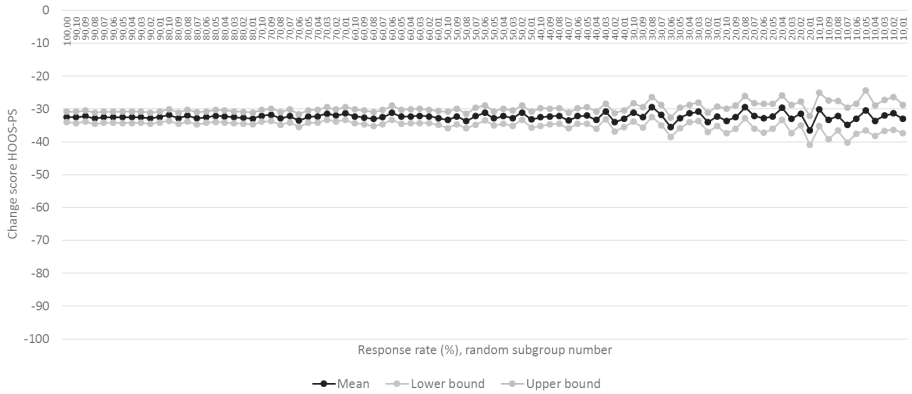


Fig. 4 Mean HOOS-PS change score per group

HOOS-PS: Hip disability and Osteoarthritis Outcome Score-Physical function Short-form

Table 4 Number of HOOS-PS groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.421	x	0.000	x	0.144	x	0.003	x	0.115	x	0.090	x
90% (n)	0	10	10	10	0	10	10	10	0	10	1	10
80% (n)	1	10	10	10	1	10	10	10	0	10	3	10
70% (n)	0	10	10	10	1	10	8	10	2	10	3	10
60% (n)	0	10	10	10	0	10	8	10	1	10	1	10
50% (n)	3	10	7	10	1	10	6	10	0	10	0	10
40% (n)	0	10	8	10	0	10	4	10	0	10	1	10
30% (n)	1	10	6	10	0	10	4	10	2	10	3	10
20% (n)	0	10	2	10	2	10	4	10	2	10	0	10
10% (n)	0	10	0	9	1	9	2	10	0	10	0	10

BMI: body mass index; HOOS-PS: Hip disability and Osteoarthritis Outcome Score-Physical function Short-form; NRS: numeric rating scale

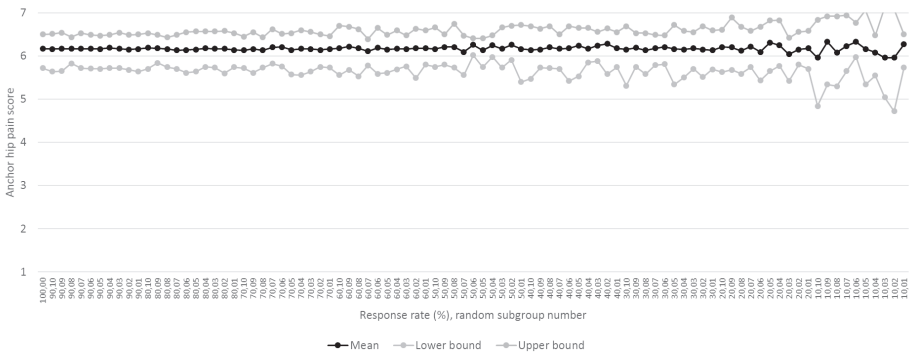


Fig. 5 Mean anchor hip pain score per group

Table 5 Number of anchor hip pain groups per response rate with predictors as significant predictor or equal distribution

	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.595	x	0.040	x	0.022	x	0.156	x	0.514	x	0.350	x
90% (n)	0	10	2	10	6	10	1	10	0	10	0	10
80% (n)	0	10	5	10	4	10	1	10	0	10	0	10
70% (n)	0	10	3	10	5	10	0	10	0	10	0	10
60% (n)	0	10	1	10	6	10	1	10	0	10	0	10
50% (n)	0	10	1	10	2	10	0	10	0	10	1	10
40% (n)	0	10	2	10	1	10	2	10	0	10	2	10
30% (n)	0	10	0	10	1	10	0	10	0	10	2	10
20% (n)	0	10	1	10	1	10	0	10	0	10	0	10
10% (n)	0	9	0	10	4	10	0	9	0	10	2	9

BMI: body mass index; NRS: numeric rating scale

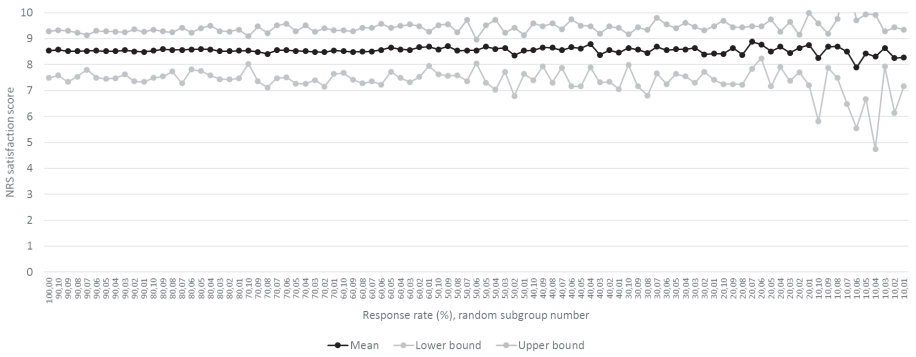


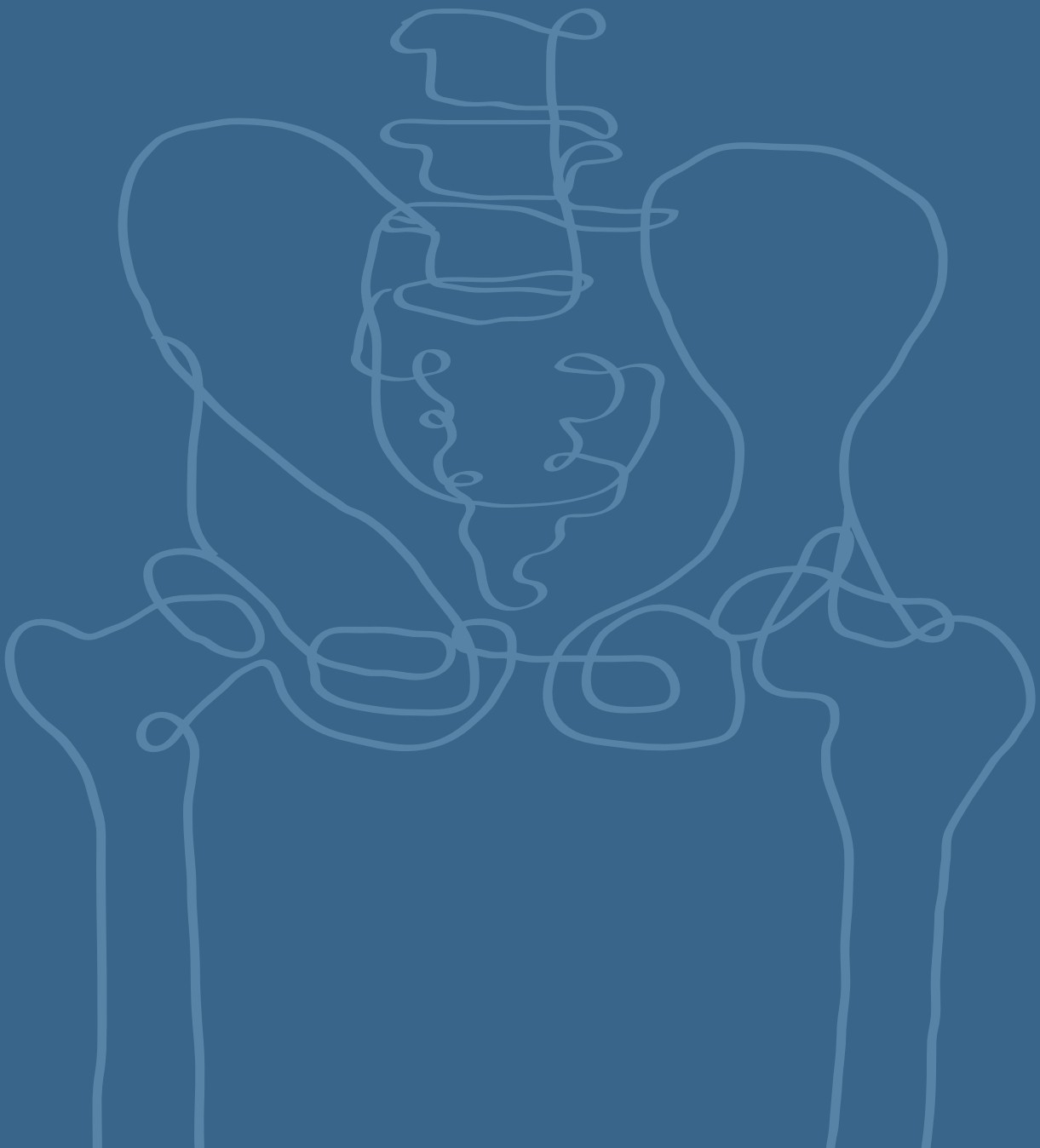
Fig. 6 Mean NRS satisfaction score per group

NRS: numeric rating scale

Table 6 Number of NRS satisfaction groups per response rate with predictors as significant predictor or equal distribution

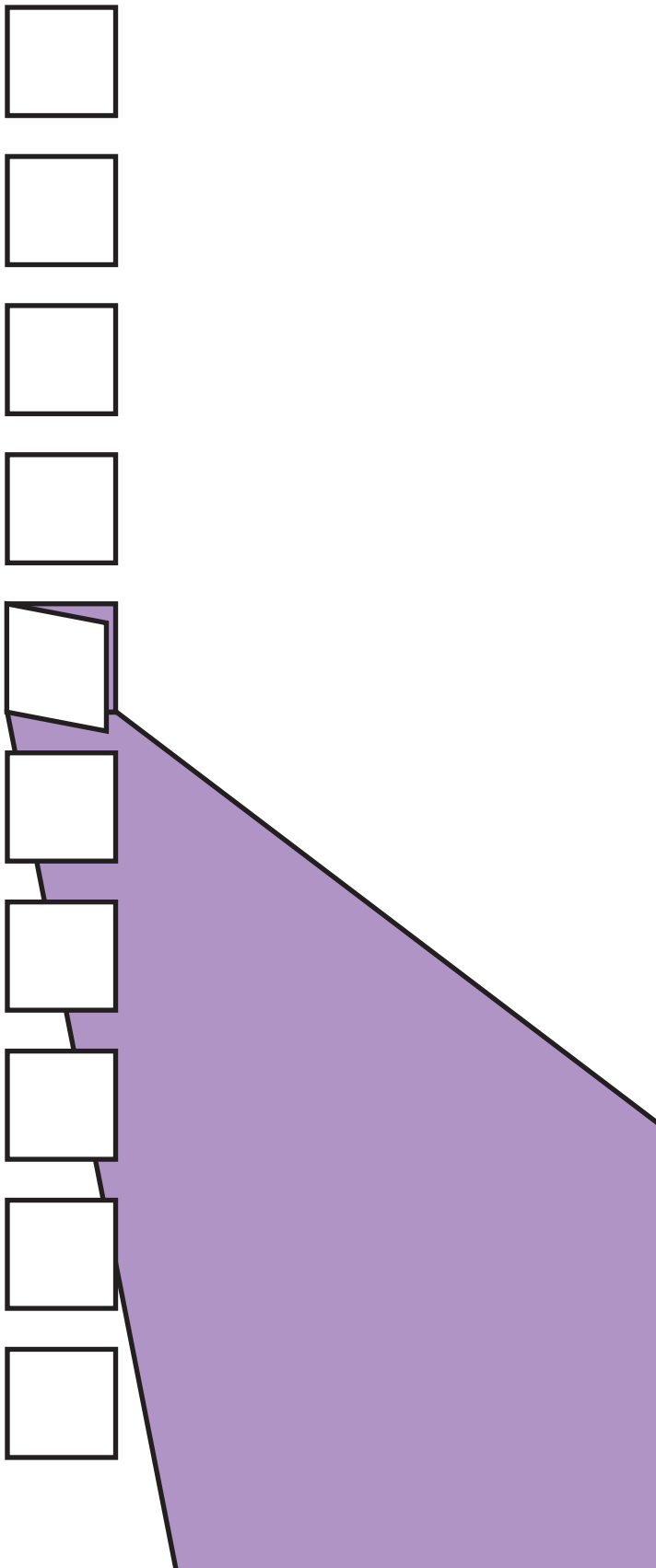
	Charnley score		Gender		Comorbidity		Anxiety		Age		BMI	
	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?	Significant predictor?	Equal distribution?
100% (p-value)	0.787	x	0.013	x	0.184	x	0.075	x	0.734	x	0.029	x
90% (n)	0	10	10	10	0	10	1	10	0	10	4	10
80% (n)	0	10	7	10	1	10	2	10	0	10	5	10
70% (n)	0	10	5	10	1	10	1	10	0	10	4	10
60% (n)	0	10	7	10	2	10	1	10	0	10	4	10
50% (n)	0	10	3	10	1	10	2	10	0	10	3	10
40% (n)	0	10	3	10	1	10	0	10	1	10	3	10
30% (n)	2	10	3	10	0	10	2	10	0	10	0	10
20% (n)	0	10	1	10	0	8	0	10	3	10	0	10
10% (n)	0	9	0	9	0	9	1	9	1	10	0	10

BMI: body mass index; NRS: numeric rating scale



PART II

Optimizing health care with routine use of patient-reported outcomes



CHAPTER 5

Quality of total hip arthroplasty health care based on four years of patient-reported outcomes in the Netherlands

Y. Pronk, W. van der Weegen, B.W. Schreurs and P. Pilot

Health Qual Life Outcomes. 2023 Mar 14;21(1):24

ABSTRACT

Background

Joint arthroplasty registries have incorporated patient-reported outcomes (PROs) to evaluate outcomes from a patients' perspective to improve total hip arthroplasty (THA). To draw valid conclusions on PROs, a minimum response rate (RR) of 60% is advised. This study investigated (1) if the quality of THA health care based on PROs improved over the years in the Netherlands, (2) if RRs improved over the years, and (3) difference in PROs over the years in hospitals with RR≥60% compared to RR<60%.

Methods

Longitudinal study with publicly available datasets from 2016 to 2019. Primary outcome was increase/decrease in PRO change scores including 95%CI ranges over the years between preoperatively and 3 months postoperatively (pre-3 m), and 12 months postoperatively (pre-12 m). Improved quality of health care was arbitrary defined as when ≥3 of 4 included scores or ranges were statistically significant improved. Secondary outcome was increase/decrease in RRs over the years. Subgroups RR≥60% and RR<60% were compared.

Results

Hospitals (%) collecting THA PROs increased from 78 to 92%. EQ VAS change score increased over the years, and 95%CI ranges of EQ VAS, EQ-5D descriptive system and NRS pain during activity decreased over the years at pre-3 m ($p<0.05$). All THA pre-12 m PRO change scores and 95%CI ranges remained equal ($p>0.05$). Pre-3 m RR remained equal (around 43%, $p=0.107$) and pre-12 m RR decreased 9% (49% to 40%, $p=0.008$). Pre-3 m subgroup RR≥60% was too small to analyse (5%). No difference was found between pre-12 m subgroups (RR≥60% = 16%), $p>0.05$).

Conclusions

Quality of THA health care based on PROs seems equal in the Netherlands between 2016 and 2019. Although more hospitals participated in PRO collection, low RRs with large IQRs are observed and only 16% of the hospitals achieved the advised RR≥60%. Multiple recommendations are provided to improve PRO collection and use.

INTRODUCTION

Total hip arthroplasty (THA) is an effective treatment for patients with end-stage hip osteoarthritis. Traditionally, THA is surgically successful if alignment is correct, and the implant well fixed and stable. The long term outcome is considered optimal if excellent implant survival is obtained. However, patients are mainly satisfied if their pain is relieved, their function is restored, their quality of life has improved and they can participate in daily activities. To measure these outcomes, collection of patient-reported outcomes (PROs) by selected patient-reported outcome measures (PROMs) has become an internationally accepted method.

Multiple national joint arthroplasty registries have incorporated PROs to evaluate the outcomes from a patients' perspective to improve THA health care [1–3]. The Dutch arthroplasty register (LROI) incorporated PROs of patients diagnosed with hip osteoarthritis since 2014. In the Netherlands, these PROs are also a mandatory part of a national defined indicator set since 2016. These results are publicly available to create transparency of the delivered care [4]. To improve health care, hospitals could use these publicly available PROs to benchmark themselves. Furthermore, surgeons could use these data to inform their patients what to expect of a treatment and to facilitate shared decision making. Moreover, health insurance companies could use PROs in their negotiations with hospitals. However, previous studies emphasize that there is no definitive evidence yet that the goal of improving health care by evaluating PROs is achieved [5–9].

Informing patients on what PRO results to expect, discussing with patients what PRO results are achieved and proactively following up on deviating PRO results are examples of how to incorporate PROs in daily practice, which might lead to improved quality of THA health care from the patients' perspective.

Collecting PROs to adequately evaluate THAs involves effort and budget [10]. Nowadays, 50% of the worldwide existing national joint arthroplasty registries capture preoperative and postoperative PROs of the patients [3]. Multiple national joint arthroplasty registries do not achieve the advised minimum RR of 60% yet [3, 11, 12]. So, investing effort and budget to collect these data in its current form could be questioned, especially if it is unclear if the quality of health care is improved by collecting and using PROs.

It was hypothesized that evaluating PROs will result in improved quality of THA health care from a patients' perspective, which should be reflected in better PROs and higher RRs over the years. Therefore, the primary aim of this study was to investigate if the quality of THA health care from a patients' perspective based on PROs improved over the years since the mandatory introduction of the PROM indicators in the Netherlands in 2016. Secondary aims were to investigate (1) if PROM RRs improved over the years, and (2) if there was a difference in PROs over the years between hospitals which achieved the advised minimum RR of 60% compared to hospitals that did not. Better PROs from hospitals with a RR≥60% were expected.

METHODS

For this longitudinal study, the publicly available Dutch national THA indicator datasets were downloaded (https://www.zorginzicht.nl/openbare-data/open-datamedisch-specialistische-revalidatie). Datasets were included from the start of the PROM indicators in 2016 up to and including 2019. Although the datasets of 2020 and 2021 were available, these datasets were not included due to unknown effect of the COVID-19 pandemic on the quality of health care.

In case of hospitals with multiple locations, these locations were considered as separate entities. Hospitals were included when they were present in all included datasets. Reasons for not being present in all included datasets could be merging of hospitals, bankruptcy or newly hospitals started up after 2016. Hospitals were excluded when in the data quality rapports, published by a governmental institution (Zorginstituut Nederland, Diemen, the Netherlands) each year [13], problems with the data quality was mentioned, for example: two locations of one hospital sent in the same scores.

Dutch national indicator datasets

The PROM indicators are part of the Dutch national THA indicator dataset. The PROM indicators are (1) the preoperative response rate, (2) the preoperative score per PROM and (3) the change scores between preoperative and multiple postoperative measurement time points per PROM [14]. The THA PROM set used is the mandatory PROM set of the Dutch Orthopaedic Association [15]. Hospitals had to collect or upload the PROs for all patients diagnosed with hip osteoarthritis in the Dutch arthroplasty register (LROI). The Dutch arthroplasty register data scientists calculated the numbers of the PROM indicators including correction for case mix (gender, age, Charnley score, smoking, ASA, preoperative PRO and BMI) when calculating change scores. This method was the same for all hospitals. Hospitals were asked to verify the data, which, after approval, were sent to Zorginstituut Nederland. This institution published the datasets online.

From these datasets the following data were collected per year, per hospital, per preoperative or change PROM measurement time point and per PRO: number of THAs with a score, mean score, 95% confidence interval (95%CI) lower bound and 95%CI upper bound. Furthermore, per year and per hospital the number of performed primary THAs, and the number of surgeons performing these surgeries were collected. The numbers of performed THA and surgeons were based on all THA patients, not only on patients diagnosed with hip osteoarthritis (85% of all THA patients) [16].

Outcomes

The primary outcome was the increase or decrease in PRO change scores including 95%CI ranges over the years. The four included PROs were pain at rest, pain during activity, quality of life and physical functioning. Pain at rest and pain during activity were both measured using a Numeric Rating Scale (NRS) question scored from 0 (no pain) to 10 (severe pain). NRS are well correlated and sensitive for pain assessment including osteoarthritic knee pain and are preferred over Visual Analogue Scales by the elderly population [17–19]. A decrease in the score

was defined as an improvement in these PROs over the years. Quality of life was assessed with 3-level version of EuroQol 5 dimensions questionnaire (EQ-5D-3L) which existed of two subscores: EQ-5D descriptive system with the highest score 1 defined as healthy, and EQ visual analogue scale (EQ VAS) scored from 0 (worst imaginable health state) to 100 (best imaginable health state) [20]. An increase in both subscores was defined as an improvement in this PRO over the years. Physical functioning was measured using Hip disability and Osteoarthritis Outcome Score-Physical Function Shortform (HOOS-PS) on a scale from 0 (no difficulty) to 100 (extreme difficulty) [21, 22]. Although HOOS-PS has to be used with care, it was a mandatory PRO from the 2012 guideline on PRO collection from the Dutch Orthopedic Association [1, 23]. A decrease in this score was defined as an improvement in this PRO over the years. The 95%CI range was calculated by 95%CI upper bound minus 95%CI lower bound. A decreased 95%CI range was defined as an improvement over the years. The included change scores and 95%CI ranges were between preoperatively and 3 months postoperatively (pre-3 m), and between preoperatively and 12 months postoperatively (pre-12 m). As a minimal clinically important difference (MCID) is not available for most PROs [24, 25] and to answer the primary aim based on the same method per PRO, improved quality of health care over the years was defined as when ≥3 of the 4 included PRO change scores or 95%CI ranges were statistically significant improved over the years. As the EQ-5D descriptive system and EQ VAS were two subscores of one PROM for one PRO, both counted for 0.5.

The first secondary outcome was the increase or decrease in PROM RRs over the years. RR was calculated by dividing the highest number of performed THAs with a PRO preoperative score or change score by the number of performed THAs multiplied by 0.85 and, thereafter, multiplied by 100. By multiplying with 0.85 a correction was made for the difference between the number of performed THAs (all patients) and the number of performed THAs with a PRO score (patients diagnosed with hip osteoarthritis, 85% [16]). RR was calculated for response on the preoperative measurement (pre RR), for response on both preoperatively and 3 months postoperatively measurements (pre-3 m RR), and for response on both preoperatively and 12 months postoperatively measurements (pre-12 m RR). The second secondary outcome was increase or decrease in PRO change scores including 95%CI ranges over the years between hospitals which achieved the advised minimum RR of 60% and hospitals that did not. Per calculated RR, hospitals were allocated to subgroup RR≥60% or subgroup RR<60%. Hospitals needed to have a RR≥60% in all four years for allocation to the subgroup RR≥60%.

Statistical analysis

Based on the data quality rapports published by Zorginstituut Nederland, unlikely outliers were recoded into missing values. Statistical analyses were performed using SPSS version 26.0 (IBM Corp, Armonk, New York). Results were reported in mean and standard deviation (SD), median and interquartile range (IQR) or number (n) and percentage (%) based on the test performed.

Differences in the number of performed THAs and the number of surgeons performing these surgeries between included and excluded hospitals were investigated. Distribution of the data

was investigated using Shapiro–Wilk tests of normality. Mann–Whitney U tests were used for these non-parametric distributed data.

Of the included hospitals, for each PRO at pre-3 m or pre-12 m, normal distribution of the change score and 95%CI range were investigated using Shapiro–Wilk tests of normality. For the primary aim change score and 95%CI range of each PRO at pre-3 m or pre-12 m were analysed on the overall rate of increase or decrease over the years using linear mixed model analyses. For the secondary aims linear mixed model analyses were executed to investigate the overall rate of increase or decrease of PROM RR over the years for each RR, and to investigate the overall rate of increase or decrease of each PRO change score and 95%CI range between both subgroups. When the percentage of included hospitals in the subgroups RR≥60% or RR<60% were below 10%, these analyses were not executed. The linear mixed model analyses included correction for differences between included and excluded hospitals. Continuous variables were centralized to create a more interpretable intercept.

RESULTS

Between 2016 and 2019 124,810 THAs were implanted. In these four years THA data of 109 unique hospitals were published. This number of 109 is partly based on merging hospitals and new hospital registrations. The number of hospitals per year was rather constant: mean 92 hospitals per year (2016: 92, 2017: 95, 2018: 91, 2019: 90). The number of hospitals collecting PROs increased from 72 (72/92, 78%) in 2016 to 83 (83/90, 92%) in 2019. Median pre RRs were between 55% (IQR 39%) and 70% (IQR 38%), median pre-3 m RRs were between 36% (IQR 32%) and 48% (IQR 33%) and median pre-12 m RRs were between 41% (IQR 43%) and 48% (IQR 55%) (Fig. 1).

Included hospitals

Out of mean 92 hospitals per year, 73 (79%) hospitals were included for further analyses. Main reason for exclusion was that no data was available in one or more years (21%). Most of these hospitals (12%) missed more than one year of data. Included hospitals performed statistically significant more THAs by statistically significant more surgeons compared to excluded hospitals (THAs: 352 (240–503) versus 147 (36–238), p<0.001; surgeons: 5 (4–7) versus 3 (2–5), p<0.001).

Main results

Of the 4 THA PRO change scores and 95%CI ranges at pre-3 m, EQ VAS change score increased over the years (0.5 of 4) (p=0.008) defined as EQ VAS change score improved over the years. The 95%CI ranges of EQ-5D-3L (both EQ VAS and EQ-5D descriptive system) and NRS pain during activity decreased over the years (2 of 4) (all p<0.001) defined as these 95%CI ranges improved over the years. All THA PRO change scores and 95%CI ranges remained equal over the years at pre-12 m (p>0.05) (Table 1).

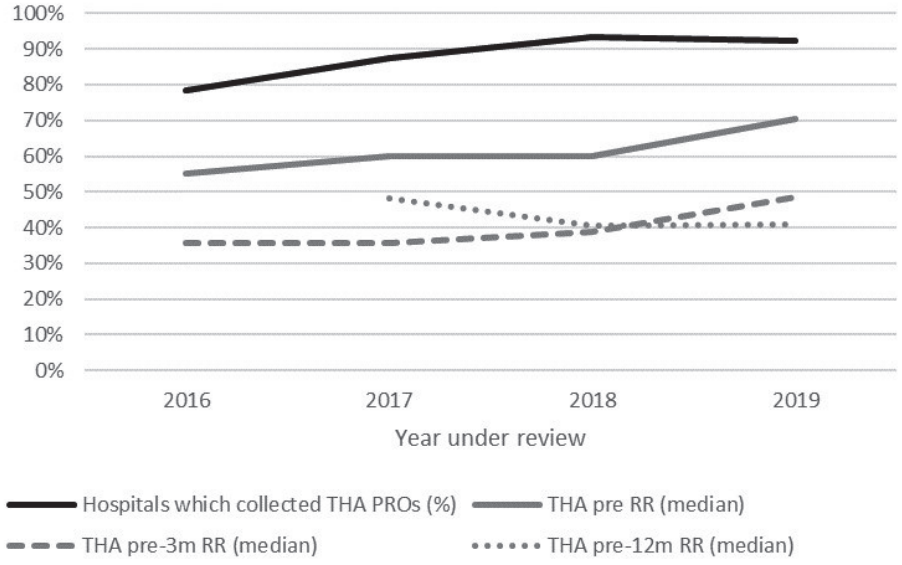


Fig. 1 Hospitals which collected THA PROs, and THA PROM RR per measurement time point per year

In 2016, pre-12m RR data was not available yet.

Pre = preoperative; Pre-12 m = between preoperatively and 12 months postoperatively; Pre-3 m = between preoperatively and 3 months postoperatively; PROs = patient-reported outcomes; RR = response rate; THA = total hip arthroplasty

The pre-3 m RR remained equal (p=0.107) and pre-12 m RR decreased over the years (p=0.008) (Fig. 2). At pre-3 m the subgroup RR≥60% was too small (n=4, 5%) to answer the second secondary study aim. At pre-12 m the subgroup RR≥60% (16%) reported equal PRO change scores and 95%CI ranges over the years compared to the subgroup RR<60% (p>0.05) (Table 2).

Detailed results: PRO change score and 95%CI range

At pre-3 m EQ VAS change score increased statistically significant over the years (intercept: 10.67 (9.47–11.87), 2016: −2.25 (−3.69 to −0.81), 2017: −1.52 (−2.91 to −0.14, 2018: 0.09 (−1.03–1.21), 2019: 0; p=0.008). Furthermore, EQ VAS 95%CI range significantly decreased over the years (intercept: 6.44 (5.48–7.41), 2016: 10.61 (6.50–14.72), 2017: 1.98 (0.61–3.36), 2018: 0.54 (−0.18–1.26), 2019: 0; p<0.001). EQ-5D descriptive system 95%CI range significantly decreased over the years (intercept: 0.107 (0.087–0.127), 2016: 0.080 (0.036–0.123), 2017: 0.015 (−0.016–0.045), 2018: 0.015 (−0.006–0.035), 2019: 0; p<0.001). For NRS pain during activity, the 95%CI range significantly decreased over the years (intercept: 0.74 (0.64–0.83), 2016: 0.82 (0.41–1.24), 2017: 0.30 (0.05–0.55), 2018: 0.11 (0.06–0.17), 2019: 0; p<0.001). All PRO change scores and 95%CI ranges remained equal over the years at pre-12 m (Table 1).

Table 1 Median change scores including median 95%CI ranges per PRO and per year

	Pre-3 m					Pre-12 m				
	2016	**2017**	**2018**	**2019**	**p-value***	**2016**	**2017**	**2018**	**2019**	**p-value***
Hospitals (n (%))	55 (75.34)	68 (93.15)	68 (93.15)	67 (91.78)		n.a.	64 (87.67)	68 (93.15)	69 (94.52)	
NRS pain at rest (change score (95%CI range))	4.06 (0.87)[a]	4.08 (0.93)	4.02 (0.84)	4.05 (0.70)[e]	0.545 (0.101)	n.a.	4.28 (0.85)[b]	4.28 (0.84)[f]	4.33 (0.74)[g]	0.446 (0.053)
NRS pain during activity (change score (95%CI range))	5.08 (1.05)[a]	5.09 (0.72)	5.10 (0.67)	5.07 (0.53)[e]	0.377 (<0.001)	n.a.	5.76 (0.62)[b]	5.81 (0.62)[f]	5.76 (0.59)[g]	0.623 (0.081)
HOOS-PS (change score (95%CI range))	29.38 (7.00)[a]	30.76 (6.30)	30.60 (6.45)	30.19 (5.43)	0.066 (0.078)	n.a.	34.72 (5.79)[d]	34.57 (5.52)	35.26 (5.44)	0.694 (0.300)
EQ-5D descriptive system (change score (95%CI range))	0.250 (0.130)[a]	0.250 (0.095)	0.260 (0.090)[f]	0.250 (0.070)	0.090 (<0.001)	n.a.	0.290 (0.080)[d]	0.290 (0.080)[f]	0.280 (0.080)[f]	0.281 (0.220)
EQ VAS (change score (95%CI range))	8.97 (11.90)[a]	10.41 (5.70)[f]	10.97 (5.91)[f]	11.28 (4.59)	0.008 (<0.001)	n.a.	10.57 (5.92)[c]	10.64 (5.40)[f]	11.53 (5.55)[f]	0.603 (0.183)

*p-values of the overall rate of increase or decrease over the years are presented.

EQ-5D descriptive system = EuroQol 5 dimensions descriptive system; EQ-5D VAS = EuroQol visual analogue scale; HOOS-PS = Hip disability and Osteoarthritis Outcome Score - Physical Function Shortform; n.a. = not available as PROs collection was started the year before, so no 12 month scores were available yet; NRS = numeric rating scale; Pre-12 m = between preoperatively and 12 months postoperatively; Pre-3 m = between preoperatively and 3 months postoperatively

[a]n=54, [b]n=61, [c]n=62, [d]n=63, [e]n=66, [f]n=67, [g]n=68

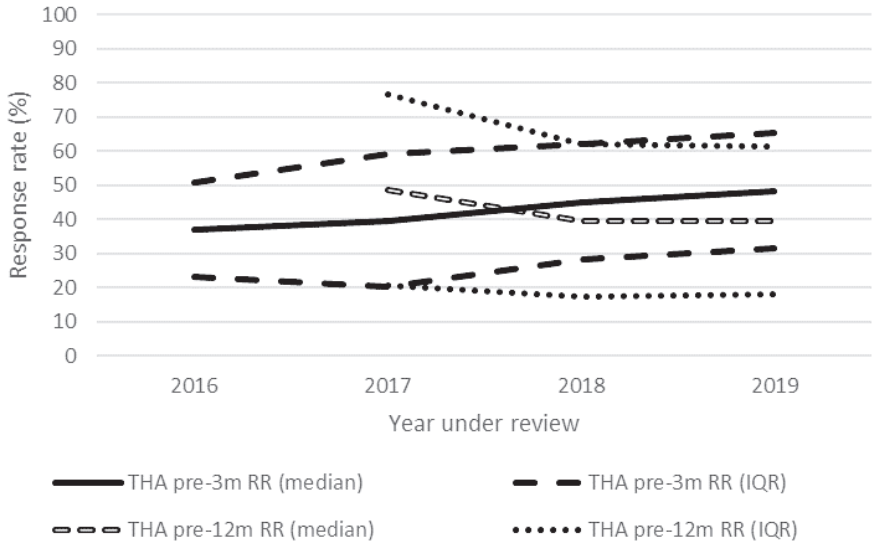


Fig. 2 THA PROM RR per measurement time point per year of included hospitals

In 2016, pre-12m RR data was not available yet.

Pre-12 m = between preoperatively and 12 months postoperatively; Pre-3 m = between preoperatively and 3 months postoperatively; RR = response rate; THA = total hip arthroplasty

Detailed results: PROM response rate

The number of hospitals collecting THA PROs increased from 55 (75%) in 2016 to 67 (92%) in 2019. The pre-3 m RR remained equal over the years (around 43%, p=0.107). The pre-12 m RR statistically significant decreased over the years from 49% (IQR 56%) in 2017 to 40% (IQR 43%) in 2019 (intercept: 43.96 (37.65–50.27), 2017: 8.00 (0.87–15.13), 2018: 0.08 (−4.82–4.98), 2019: 0; p=0.008) (Fig. 2).

Detailed results: subgroup response rate ≥60% compared to subgroup response rate <60%

The subgroup RR≥60% comprised of a minimum of 8 (11%) to a maximum of 22 (30%) hospitals per year at pre-3 m, and a minimum of 22 (30%) to a maximum of 27 (37%) hospitals per year at pre-12 m. In total 4 (5%) hospitals reached RR≥60% all years at pre-3 m and 12 (16%) hospitals all years at pre-12 m. At pre-3 m the subgroup RR≥60% was too small to answer the second secondary study aim. At pre-12 m all PRO change scores and 95%CI ranges remained equal over the years between both subgroups (p>0.05). In each year median PRO 95%CI ranges were smaller in the subgroup RR≥60% compared to the subgroup RR<60% (Table 2).

Table 2 Median change scores including median 95%CI ranges in subgroups RR≥60% and RR<60% at pre-12m

	RR≥60% (n=12 (16%))			RR<60% (n=57 (78%))			p-value*
	2017	2018	2019	2017 (n=52)	2018 (n=56)	2019 (n=57)	
NRS pain at rest (change score (95%CI range))	4.31 (0.55)	4.35 (0.52)	4.27 (0.53)	4.26 (0.94)	4.27 (1.00)	4.34 (0.82)	0.673 (0.610)
NRS pain during activity (change score (95%CI range))	5.80 (0.43)	5.74 (0.43)	5.73 (0.43)	5.75 (0.73)	5.83 (0.76)	5.76 (0.67)	0.885 (0.678)
HOOS-PS (change score (95%CI range))	34.08 (3.70)	35.67 (3.54)	35.79 (3.60)	34.72 (7.35)	34.48 (6.46)	35.26 (5.87)	0.378 (0.339)
EQ-5D descriptive system (change score (95%CI range))	0.290 (0.050)	0.290 (0.050)	0.285 (0.055)	0.290 (0.100)	0.290 (0.100)	0.280 (0.090)	0.758 (0.218)
EQ VAS (change score (95%CI range))	9.12 (3.91)	11.82 (3.86)	11.68 (3.79)	10.69 (6.46)	10.51 (6.25)	11.53 (5.81)	0.522 (0.079)

* p-values of the difference in the subgroups on overall rate of increase or decrease over the years are presented.

EQ-5D descriptive system = EuroQol 5 dimensions descriptive system; EQ-5D VAS = EuroQol visual analogue scale; HOOS-PS = Hip disability and Osteoarthritis Outcome Score - Physical Function Shortform; NRS = numeric rating scale; RR = response rate

DISCUSSION

The primary aim of this study was to investigate if the quality of THA health care from a patients' perspective based on PROs improved over the years since the mandatory introduction of the PROM indicators in the Netherlands in 2016. Secondary aims were to investigate (1) if PROM RRs improved over the years, and (2) if there was a difference in PROs over the years between hospitals which achieved the advised minimum RR of 60% compared to hospitals that did not. Main results show that of the 4 THA PRO change scores, only EQ VAS change score improved over the years (0.5 of 4) at pre-3 m. Regarding their 95%CI ranges, EQ VAS, EQ-5D descriptive system and NRS pain during activity improved over the years (2 of 4). At pre-12 m all THA PRO change scores and 95%CI ranges remained equal over the years. These results mean that since the mandatory introduction of the PROMs the quality of THA health care from a patients' perspective based on PROs remained equal at both pre-3 m and pre-12 m (<3 of 4). Although the percentage of hospitals collecting PROs increased, low RRs with large IQRs were observed. The pre-3 m RR remained equal and, disappointingly, the pre-12 m RR decreased over the years. At pre-3 m the subgroup with sufficient PROs at all years (RR≥60%) was very small (5%) hampering the second secondary aim. Interestingly, at pre-12 m this subgroup (16%) reported equal PRO change scores and 95%CI ranges over the years compared to the subgroup without sufficient PROs (RR<60%).

The quality of THA health care from a patients' perspective based on PROs remained equal over the years in the Netherlands between 2016 and 2019, while improvement of quality of health care is the desirable direction. Maybe more years are needed to achieve a detectable improvement. However, a previous single center cohort study on twenty year data of Dutch THA patients executed trends over time analyses and also reported, in general, no improvement over time [26]. Interestingly, in the present study, two PRO 95%CI ranges (EQ-5D-3L (both EQ VAS and EQ-5D descriptive system) and NRS pain during activity) decreased over the years at pre-3 m. Decreased 95%CI ranges mean smaller 95%CI ranges, so less positive and negative outliers, which could be interpreted as an improvement. However, decreased 95%CI ranges could also be the result of more hospitals collecting PROs as more data generally results in smaller 95%CI ranges [27].

The statistical power of large datasets, as is common in data retrieved from national joint registries, has inherent pitfalls. This includes the possibility of reaching statistical significance for a score difference, with this score difference being (much) smaller than the minimal clinical relevant difference, which is the only relevant outcome from the perspective of the patient.

It was hypothesized that PRO collection and transparency of PROs lead in PRO evaluation, which will result in improved future PROs and subsequently improved health care. However, it remains unknown if hospitals use the collected PROs to evaluate (and improve) health care. Collection is mandatory, however, using aggregated or individual PROs in daily practice to evaluate THA health care is not. For evaluation an intrinsic motivation of surgeons, hospitals and other stakeholders is needed [28]. The Dutch Orthopaedic Association uses implant information from the Dutch

arthroplasty register (LROI) for an outlier analysis including conversations with hospitals if needed [29, 30]. It is recommended to include an outlier procedure on PROs and RRs. If hospitals only collect to comply with mandatory PRO collection, no better understanding of the patients' perspective nor improvement of quality of health care will be likely, while the costs and burden involved with PRO collection remain.

With and without excluded hospitals, low median pre-3 m RRs and pre-12 m RRs (<49%) were observed which indicates low quality of PRO data. Improvement is seen in the percentage of hospitals collecting PROs (around 15%). However, of the included hospitals, pre-12 m RR decreased 9% over the years which is worrisome. Besides the low RRs, large IQRs (56%) were observed. This reveals a large diversity in PRO collection in the Netherlands. To comply with mandatory PRO collection for registries and the Dutch PROM indicators, hospitals need a minimum RR of only 1%. However, there is evidence that for a sufficient evaluation of THAs a minimum RR of 60% is advised [11, 12]. A first exploration by the present study shows that hospitals achieving this 60% at pre-12 m have equal PRO change scores and 95%CI ranges over the years compared to hospitals that do not. Interestingly, PRO 95%CI ranges seem twice as small for hospitals with a RR≥60%. This indicates that less outliers are expected in hospitals achieving RR≥60%. However, these results are based on aggregated scores per hospital per year. Further analyses on individual scores per patient per hospital per year are needed before conclusions on differences between hospitals achieving RR≥60% and RR<60% could be made.

The low quality of PRO data based on RR is a point of concern. Only 5% of the hospitals achieved the advised RR≥60% at pre-3 m and only 16% at pre-12 m. Therefore, it is questionable if a conclusion on quality of THA health care from a patients' perspective based on PROs over the years could be made. Continuing PRO collection in its current form, including the involved effort and costs, might not be justifiable from an ethical and value based health care perspective.

So, in what direction should PRO collection and use develop to improve quality of THA health care from a patients' perspective? Firstly, investigate if stakeholders use collected PROs to evaluate THA health care. It is assumed that if PROs are made available, they will be used. However, studies examining this assumption have found limited use of PROs. Main reasons according to surgeons are a lack of knowledge on how to use PROs in daily health care, the perception that PROs do not provide actionable information, and because gathering and handling of PROs add work to an already busy schedule [31, 32]. In addition, orthopaedic surgeons state that using PROs on an individual patient level is difficult based on logistical barriers (access and display issues, time required) and perceptual barriers (concerns about patients understanding, and validity and reliability of measures). They prefer to talk with patients about personal outcomes. However, they mention that using PROs on an aggregated level is valuable for hospitals and individual surgeons [33]. Secondly, support stakeholders to evaluate THA outcomes from a patients' perspective using the already existing multiple examples and recommendations how to use the PROs [34, 35]. Thirdly, investigate how all stakeholders rate the quality of THA health care provided today. Of course, improvement is always desirable, however, there might be a consensus that the delivered

quality is of such a high level that improvement is unlikely or that the desired improvement is not value-based. Fourthly, increase the RRs to at least 60% to improve the data quality. Multiple recommendations to improve RRs already exist [10, 36–42]. Fifthly, evaluate the set aim(s) of PROs. Maybe the goal of improving health care from a patients' perspective is not achievable or not formulated well. Each aim sets different requirements for the PRO(M)s, time points of collecting PROs and statistical analysis. The primary aim is the basis. Although PROMs are the gold standard to measure outcomes from a patients' perspective at this moment, maybe other instruments are needed to achieve the goal set. These five points need to be part of a coordinated effort of all stakeholders to improve PRO collection and use.

As a strength of the present study, a first exploration is presented on the goal of improving THA health care by evaluating outcomes from a patients' perspective in the Netherlands. Moreover, as the Dutch arthroplasty registry reported comparable results to multiple other national joint arthroplasty registries [3], similar results are expected for PRO collection in other countries around the world. In a previous review of registry based studies reporting PRO response rates there was also concern on the large variation and downward trend of PROM response rates [43]. Furthermore, each year the same method for the calculated data in the used public available datasets was used including correction for case mix. As a limitation, due to these used public available datasets, data on if hospitals use the collected PROs to evaluate and, if necessary, to improve their health care were missing. Moreover, only aggregated data of hospitals were available. Furthermore, as a MCID is not available for most PROs [24, 25], the authors needed to define improved quality of health care over the years from a statistical perspective. Future studies should focus on if stakeholders use collected PROs to evaluate THA health care, how all stakeholders rate the quality of health care provided today and if other instruments instead of PROMs are needed to achieve the goal of improving health care from a patients' perspective.

CONCLUSIONS

The quality of THA health care from a patients' perspective based on PROs seems equal in the Netherlands between 2016 and 2019. Although the percentage of hospitals collecting THA PROs increased, low RRs with large IQRs reveal a large diversity in PRO collection. Only 16% of the Dutch hospitals have sufficient PROs to evaluate THAs from a patients' perspective at 12 months (RR≥60%). Based on these observations, it is questionable if a conclusion on quality of THA health care based on PROs could be made. Similar results are expected for PRO collection in other countries around the world. Multiple recommendations are provided to improve PRO collection and use. A coordinated effort of all stakeholders should be initiated to improve PRO collection and use.

REFERENCES

1. Nederlandse Orthopaedische Vereniging. Patient Reported Outcome Measures: Advies Nederlandse Orthopaedische Vereniging (NOV) [Internet]. 2012 [cited 2018 Aug 1]. Available from: https://www.orthopeden.org/downloads/32/advies-proms-orthopedie.pdf
2. Rolfson O, Eresian Chenok K, Bohm E, Lübbeke A, Denissen G, Dunn J, et al. Patient-reported outcome measures in arthroplasty registries. Acta Orthop. 2016;87(1):3–8
3. Bohm ER, Kirby S, Trepman E, Hallstrom BR, Rolfson O, Wilkinson JM, et al. Collection and reporting of patient-reported outcome measures in arthroplasty registries: multinational survey and recommendations. Clin Orthop Relat Res. 2021;479(10):2151–66. https://doi.org/10.1097/CORR.0000000000001852
4. Zorginstituut Nederland. Zorginzicht - Transparantiekalender [Internet]. Available from: https://www.zorginzicht.nl/transparantiekalender
5. Damman OC, Jani A, Jong BA, Becker A, Metz MJ, Bruijne MC, et al. The use of PROMs and shared decision-making in medical encounters with patients: an opportunity to deliver value-based health care to patients. J Eval Clin Pract. 2020;26(2):524–40. https://doi.org/10.1111/jep.13321
6. Kotronoulas G, Kearney N, Maguire R, Harrow A, Di Domenico D, Croy S, et al. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? a systematic review of controlled trials. J Clin Oncol. 2014;32(14):1480–501. https://doi.org/10.1200/JCO.2013.53.5948
7. Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. J Patient-Reported Outcomes. 2018;2(1):42. https://doi.org/10.1186/s41687-018-0061-6
8. Greenhalgh J, Dalkin S, Gibbons E, Wright J, Valderas JM, Meads D, et al. How do aggregated patient-reported outcome measures data stimulate health care improvement? A realist synthesis. J Health Serv Res Policy. 2018;23(1):57–65. https://doi.org/10.1177/1355819617740925
9. Gutacker N, Bojke C, Daidone S, Devlin NJ, Parkin D, Street A. Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. Health Econ. 2013;22(8):931–47. https://doi.org/10.1002/hec.2871
10. Pronk Y, Pilot P, Brinkman JM, van Heerwaarden RJ, van der Weegen W. Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection. J Patient-Reported Outcomes. 2019;3(1):31. https://doi.org/10.1186/s41687-019-0121-6
11. Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, et al. Patient-reported outcome measures in arthroplasty registries: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. Acta Orthop. 2016;87(June):9–23
12. Pronk Y, van der Weegen W, Vos R, Brinkman J-M, van Heerwaarden RJ, Pilot P. What is the minimum response rate on patient-reported outcome measures needed to adequately evaluate total hip arthroplasties? Health Qual Life Outcomes. 2020;18(1):379. https://doi.org/10.1186/s12955-020-01628-1
13. Zorginstituut Nederland. Zorginzicht - Open data Medisch-specialistische zorg [Internet]. [cited 2021 Nov 24]. Available from: https://www.zorginzicht.nl/openbare-data/open-data-ziekenhuizen-en-zelfstandige-behandelcentra---medisch-specialistische-zorg
14. Zorginstituut Nederland. Zorginzicht - Heupprothese indicatoren [Internet]. [cited 2020 Nov 16]. Available from: https://www.zorginzicht.nl/kwaliteitsinstrumenten/heupprothese-indicatoren
15. Nederlandse Orthopaedische Vereniging. PROMs [Internet]. [cited 2018 Aug 1]. Available from: https://www.lroi.nl/invoerders/registreren/proms
16. LROI. Total Hip Arthroplasty Demographics [Internet]. [cited 2022 Apr 4]. Available from: https://www.lroi-report.nl/hip/total-hip-arthroplasty/demographics/
17. Taylor LJ, Harris J, Epps CD, Herr K. Psychometric evaluation of selected pain intensity scales for use with cognitively impaired and cognitively intact older adults. Rehabil Nurs. 2005;30(2):55–61
18. Alghadir AH, Anwer S, Iqbal A, Iqbal ZA. Test-retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. J Pain Res. 2018;26(11):851–6
19. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, editors. Handbook of Pain Assessment. Guilford Press; 2011. p. 19–44

20. EuroQol Group. EuroQol - a new facility for the measurement of health related quality of life. Health Policy (New York). 1990;16(3):199–208
21. Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert J-F, 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):843–7. https://doi.org/10.1016/j.joca.2009.01.005
22. Davis AM, Perruccio AV, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative. Osteoarthr Cartil. 2008;16(5):551–9
23. Braaksma C, Wolterbeek N, Veen MR, Prinsen CAC, Ostelo RWJG. Systematic review and meta-analysis of measurement properties of the Hip disability and Osteoarthritis Outcome Score-Physical Function Shortform (HOOS-PS) and the Knee Injury and Osteoarthritis Outcome Score-Physical Function Shortform (KOOS-PS). Osteoarthr Cartil. 2020;28(12):1525–38
24. Çelik D, Çoban Ö, Kılıçoğlu Ö. Minimal clinically important difference of commonly used hip-, knee-, foot-, and ankle-specific questionnaires: a systematic review. J Clin Epidemiol. 2019;113:44–57. https://doi.org/10.1016/j.jclinepi.2019.04.017
25. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol. 2010;63(5):524–34. https://doi.org/10.1016/j.jclinepi.2009.08.010
26. van der Wees PJ, Wammes JJG, Akkermans RP, Koetsenruijter J, Westert GP, van Kampen A, et al. Patient-reported health outcomes after total hip and knee surgery in a Dutch University Hospital Setting: results of twenty years clinical registry. BMC Musculoskelet Disord. 2017;18(1):97. https://doi.org/10.1186/s12891-017-1455-y
27. Schünemann HJ, Vist GE, Higgins JPT, Santesso N, Deeks JJ, Glasziou P, et al. Chapter 15: Interpreting results and drawing conclusions. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 6 [Internet]. Cochrane; 2021. Available from: https://training.cochrane.org/handbook/current/chapter-15#section-15-3-1
28. Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. Med Care. 2003;41(Supplement):I-30–I—38
29. LROI. Completeness of registering hospitals and completeness of registered arthroplasties in the LROI based on the hospital information system in 2020 [Internet]. 2021 [cited 2021 Dec 20]. Available from: https://www.lroi-report.nl/data-quality/completeness/
30. Nederlandse Orthopaedische Vereniging. Procedure Uitkomstanalyse [Internet]. [cited 2021 Dec 20]. Available from: https://www.orthopeden.org/kwaliteit/kwaliteitsbeleid/procedure-uitkomstanalyse
31. Alshehri F, Alarabi A, Alharthi M, Alanazi T, Alohali A, Alsaleem M. Use of patient-reported outcome measures (PROMs) by orthopedic surgeons in Saudi Arabia. J Orthop Surg Res. 2020;15(1):598
32. Mou D, Sisodia RC, Castillo-Angeles M, Ladin K, Bergmark RW, Pusic AL, et al. The surgeon's perceived value of patient-reported outcome measures (PROMs). Ann Surg. 022;275(3):500–5. https://doi.org/10.1097/SLA.0000000000004253
33. Whitebird RR, Solberg LI, Ziegenfuss JY, Norton CK, Chrenka EA, Swiontkowski M, et al. What do orthopaedists believe is needed for incorporating patient-reported outcome measures into clinical care? a qualitative study. Clin Orthop Relat Res. 2022;480(4):680–7. https://doi.org/10.1097/CORR.0000000000002059
34. Van Schie P, Van Bodegom-Vos L, Zijdeman TM, Nelissen RGHH, Marang Van De Mheen PJ. Awareness of performance on outcomes after total hip and knee arthroplasty among Dutch orthopedic surgeons: how to improve feedback from arthroplasty registries. Acta Orthop. 2021;92(1):54–61
35. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. BMJ Qual Saf. 2014;23(6):508–18. https://doi.org/10.1136/bmjqs-2013-002524
36. Ho A, Purdie C, Tirosh O, Tran P. Improving the response rate of patient reported outcome measures in an Australian tertiary metropolitan hospital. Patient Relat Outcome Meas. 2019;10:217–26
37. Patel J, Lee JH, Li Z, SooHoo NF, Bozic K, Huddleston JI. Predictors of low patient-reported outcomes response rates in the California joint replacement registry. J Arthroplasty. 2015;30(12):2071–5
38. Tariq MB, Vega JF, Westermann R, Jones M, Spindler KP. Arthroplasty studies with greater than 1000 participants: analysis of follow-up methods. Arthroplasty Today. 2019;5(2):243–50

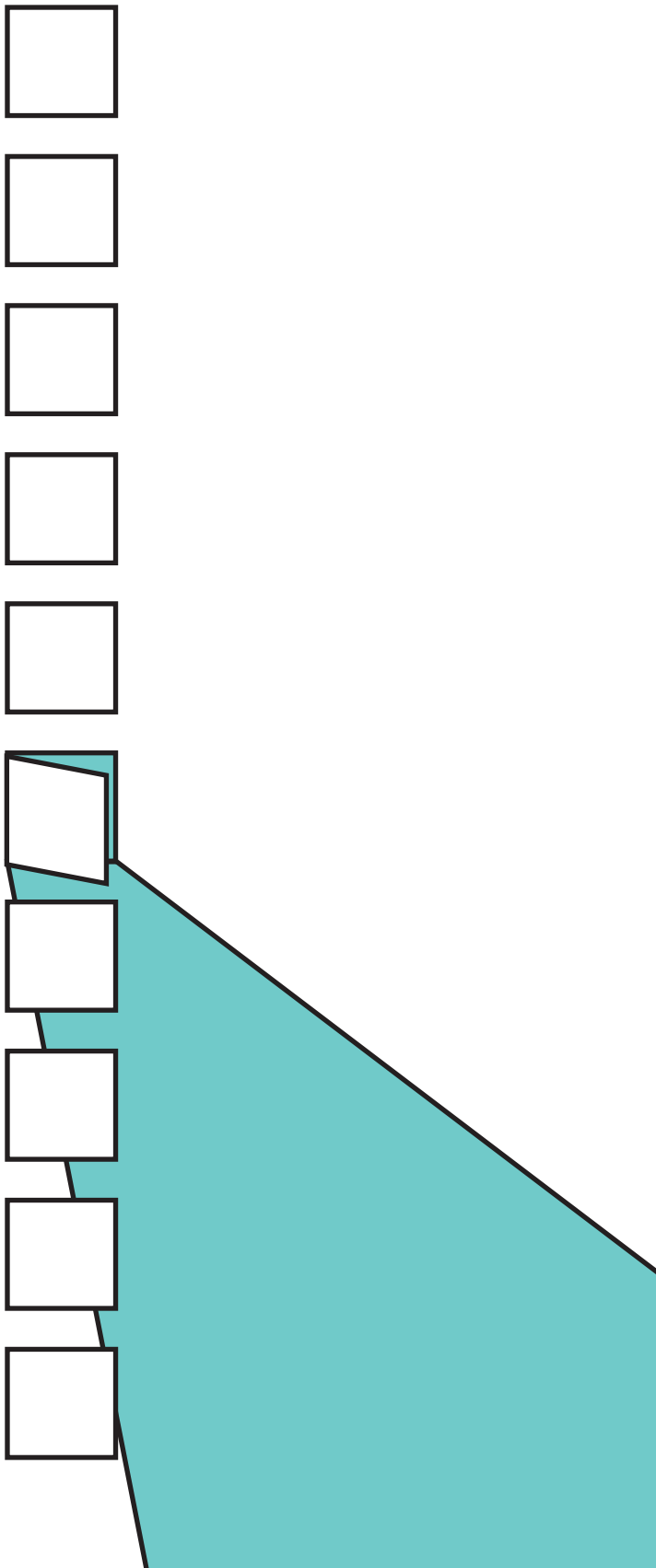
39. Slover JD, Karia RJ, Hauer C, Gelber Z, Band PA, Graham J. Feasibility of integrating standardized patient-reported outcomes in orthopedic care. Am J Manag Care. 2015;21(8):e494-500

40. Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev. 2009;2010(1):2009–11. https://doi.org/10.1002/14651858.MR000008.pub4

41. Hunter J, Corcoran K, Leeder S, Phelps K. Is it time to abandon paper? The use of emails and the Internet for health services research—a cost-effectiveness and qualitative study. J Eval Clin Pract. 2012. https://doi.org/10.1111/j.1365-2753.2012.01864.x

42. Brook EM, Glerum KM, Higgins LD, Matzkin EG. Implementing patient reported outcome measures in your practice: pearls and pitfalls. Am J Orthop (Belle Mead NJ). 2017;46(6):273–8

43. Wang K, Eftang CN, Jakobsen RB, Årøen A. Review of response rates over time in registry-based studies using patient-reported outcome measures. BMJ Open. 2020;10(8):e030808



CHAPTER 6

No difference in patient satisfaction after mobile bearing or fixed bearing medial unicompartmental knee arthroplasty

Y. Pronk, A.A.M. Paters and J.M. Brinkman

Knee Surg Sports Traumatol Arthrosc. 2021 Mar;29(3):947-954

ABSTRACT

Background

Medial unicompartmental knee arthroplasty (UKA) has excellent survival rates using one of the two implant designs: mobile bearing (MB) or fixed bearing (FB). There is a lack of studies comparing patient-reported outcomes (PROs) of both implants. This study aimed to document and compare PROs of MB UKA to FB UKA at 6, 12 and 24 months after surgery.

Methods

A single high-volume surgeon, retrospective cohort study with prospectively collected data of two groups of UKA patients, with a MB (n = 66) or FB (n = 97) implant. Primary outcome was patient satisfaction (0–10; NRS). Secondary outcomes were pain at rest (NRS), pain during activity (NRS), function (OKS, KOOS-PS), quality of life (EQ-5D-3L), anchor pain, anchor function and anchor recovery. PROs were collected 6, 12 and 24 months postoperatively. The complication rate and revision rate within one year after surgery were recorded.

Results

For the MB group, the median NRS satisfaction score was 9.0 (8.0–10.0) compared to 9.0 (8.0–9.5) for the FB group at 6 months ($p = 0.620$). Similar scores were found at 12 and 24 months; both MB 9.0 (8.0–10.0) and FB 9.0 (8.0–10.0) ($p = 0.556$ and $p = 0.522$, respectively). There were no statistically significant differences between MB and FB groups in all secondary outcomes postoperatively.

Conclusion

Medial UKA performed by a high-volume surgeon, using a MB or a FB implant, results in excellent patient satisfaction, pain relief, functional improvement and quality of life improvement at 6, 12 and 24 months after surgery. The recommendation and use of one over the other is not justified based on the outcomes in the current study.

INTRODUCTION

Medial unicompartmental knee arthroplasty (UKA) can be a successful treatment option for patients with end-stage medial knee osteoarthritis (OA). With strict patient selection and accurate implant positioning, excellent results can be achieved [5].

Two different implant designs for UKA are used, those with a mobile bearing (MB) and a fixed bearing (FB). Each implant has its reported advantages and disadvantages; excellent functional outcomes, implant survival rates and complication rates for both (MB and FB) have been reported [6, 16, 32, 36, 38]. Most studies, however, do not specifically mention surgeon volume [6, 16, 32]. Revision rate appears to be affected by surgeon volume [24, 25].

From patient's perspective the revision rate might not define the procedure's success or failure, but rather being satisfied with the outcome, and experiencing pain relief, functional improvement and improvement in quality of life [7, 14, 21, 39]. Previous studies that compared these patient-reported outcomes (PROs) of both MB and FB implants mainly focused on functional outcome and used many different patient-reported outcome measures (PROMs) to investigate the outcome [6, 16, 32]. International and national institutions, such as the Dutch Orthopaedic Association (NOV), advise which PROMs should be used to evaluate outcomes in knee arthroplasty health care [31]. There is, however, a lack of studies that compare MB and FB implants using PROs on patient satisfaction, pain and quality of life; measured with PROMs advised by these institutions. These outcomes are of potential use in shared decision making between surgeon and patient, and in making recommendations to orthopaedic surgeons and other stakeholders on implant choice.

Therefore, the aim of this study was to document and compare PROs of MB UKA to FB UKA at 6, 12 and 24 months after surgery. The primary outcome was patient satisfaction and the secondary outcomes were pain, function, quality of life, complication rate and early revision rate. It is hypothesized that both implant designs would offer similar patient satisfaction score.

METHODS

A single-centre, single-surgeon, retrospective study with prospectively collected data was performed at a medium-sized orthopaedic hospital. Patients characterized with an American Society of Anaesthesiologists (ASA) score of I–II, a body mass index (BMI) of ≤ 35 and a plan to undergo primary medial UKA between January 2016 and March 2018 were enrolled. Surgeries were performed by a high-volume (mean 90 UKAs per year; 3 UKAs out of 10 knee arthroplasty surgeries per year) [24, 25], non-designer, orthopaedic surgeon (JMB). All surgeries were performed under spinal anaesthesia and a tourniquet was used.

From February 2016 to December 2016, an MB implant (MB group) and from January 2017 to March 2018, a FB implant was used (FB group) as a constructive series. All patients in both groups

met the Oxford criteria for UKA. These criteria include debilitating knee pain combined with end-stage isolated anteromedial OA (Kellgren Lawrence grade 4 [19]), retained full thickness of the cartilage in the lateral compartment, fixed flexion contracture < 10°, correctable varus deformity of < 10°, and intact cruciate and medial collateral ligaments [13]. Patients were included in this study if they signed the informed consent form preoperatively to allow further scientific analysis using their anonymised data. There were no exclusion criteria.

Implants

Standard surgical technique as described by the manufacturer was used. The MB implant has a mobile polyethylene insert whereas the polyethylene insert of the FB implant is fixed to the tibial baseplate. The MB implant used was the Phase III cementless Oxford UKA (ZimmerBiomet Ltd., Bridgend, UK). The FB implant used was the Physica ZUK UKA (LIMA Corporate UD, Italy).

Outcomes and measurements

Primary outcome was patient satisfaction at 6, 12 and 24 months, measured using a Numeric Rating Scale (NRS) for satisfaction (0–10, 0 dissatisfied to 10 very satisfied). The question asked was 'How satisfied are you (in general) with the results of your knee surgery?'. Secondary outcomes were pain at rest, pain during activity, function, quality of life, anchor pain, anchor function and anchor recovery at 6, 12 and 24 months, and complication rate and revision rate within 12 months. Pain at rest and pain during activity were assessed using NRS pain (0–10, 0 no pain to 10 worst possible pain) [15]. Function was measured using the Oxford Knee Score (OKS; 0–48, 0 most severe symptoms to 48 least severe symptoms) [8], and the Knee injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS; 0–100, 0 no difficulty to 100 extreme difficulty) [33]. Quality of life was assessed using the 3-level version of EuroQol 5 Dimensions (EQ-5D-3L) consisting of two parts: EQ visual analogue scale (EQ VAS; 0–100, 0 worst imaginable health state to 100 best imaginable health state) and EQ-5D descriptive system [9]. Anchor questions on pain (1–7, 1 very much deteriorated to 7 very much improved), function (1–7, 1 very much deteriorated to 7 very much improved) and recovery measured using the general perceived recovery questionnaire (GPR; 0–6, 0 worse than ever to 6 fully recovered) [18] were asked. Complications and revisions within 12 months were collected from the electronic patient records.

These PROMs and time points (preoperatively, and 6, 12 and 24 months postoperatively) are set out as mandatory by the NOV [28]. All PROs were collected using a digital, online system (OnlinePROMs, Interactive Studios, Rosmalen, The Netherlands). After registration in this system, patients received invitations by email to complete the questionnaires. In case a patient was not able to handle a computer, invitations were sent by mail. Two reminders were sent in case of non-response.

Patient characteristics

Age [20, 23, 38], gender, BMI, ASA scores, Charnley scores and the side of the surgery were collected from the electronic patient records. Preoperative pain, function [22], quality of life [12, 22] and anxiety were recorded, using the preoperative questionnaire sent by the PROs collection system. Anxiety was assessed using question 5 of the EQ-5D-3L, answers were combined into anxious ('I am moderately anxious' and 'I am very anxious') and not anxious.

No statistically significant differences were found in the patient characteristics between the MB and FB groups (Table 1). Response rate on the preoperative PROMs was 100% in both groups.

Table 1 Patient characteristics

	MB group (n = 66)	FB group (n = 97)	p value
Demographic			
Age (years) [mean (SD)]	61.4 (7.8)	61.2 (7.7)	0.831
Gender—male [n (%)]	35 (53.0)	54 (55.7)	0.740
BMI (kg/m^2) [median (IQR)]	26.0 (24.3–30.9)	27.8 (24.9–31.3)	0.134
ASA score—I [n (%)]	38 (57.6)	49 (50.5)	0.375
Charnley score [n (%)]			0.157
One knee affected with OA	32 (48.5)	90 (55.2)	
Both knees affected with OA	15 (22.7)	37 (22.7)	
Contra lateral KA	17 (25.8)	29 (17.8)	
Multiple joints affected with OA	2 (3.0)	7 (4.3)	
Localisation—L [n (%)]	33 (50.0)	48 (49.5)	0.948
PROs preoperatively			
NRS pain score [median (IQR)]			
At rest	5.0 (3.0–7.3)	6.0 (3.0–7.0)	0.571
During activity	8.0 (7.0–9.0)	8.0 (6.0–9.0)	0.270
OKS [mean (SD)]	25.0 (8.6)	25.7 (7.1)	0.586
KOOS-PS score [median (IQR)]	47.3 (40.3–57.9)	44.0 (40.3–54.4)	0.338
EQ-5D-3L [median (IQR)]			
EQ-5D descriptive system	0.775 (0.298–0.783)	0.775 (0.651–0.775)	0.907
EQ VAS	80.0 (75.0–90.0)	80.0 (70.0–90.0)	0.563
Anxiety [n (%)]	11 (16.7)	18 (18.6)	0.757

MB mobile bearing implant, ***FB*** fixed bearing implant, ***SD*** standard deviation, ***BMI*** body mass index, ***kg/m^2*** kilogramme per square metre, ***IQR*** interquartile range, ***ASA*** American Society of Anaesthesiologists score, ***OA*** osteoarthritis, ***KA*** knee arthroplasty, ***L*** left, ***PROs*** patient-reported outcomes, ***NRS*** Numeric Rating Scale, ***OKS*** Oxford Knee Score, ***KOOS-PS*** Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form, ***EQ-5D-3L*** EuroQol 5 dimensions 3-level version, ***EQ VAS*** EuroQol Visual Analogue Scale

Flowchart

A total of 163 patients planned to undergo primary medial UKA met the inclusion criteria. Of these 163 patients, 66 were allocated to the MB group and 97 to the FB group. Two (3.0%) patients in the MB group were lost to follow-up, declining further participation after 6 months and deceased after 12 months. One (1.0%) patient in the FB group was lost to follow-up after 12 months, declining further participation. Analysis was performed on 66 patients in the MB group and 97 in the FB group (Fig. 1).

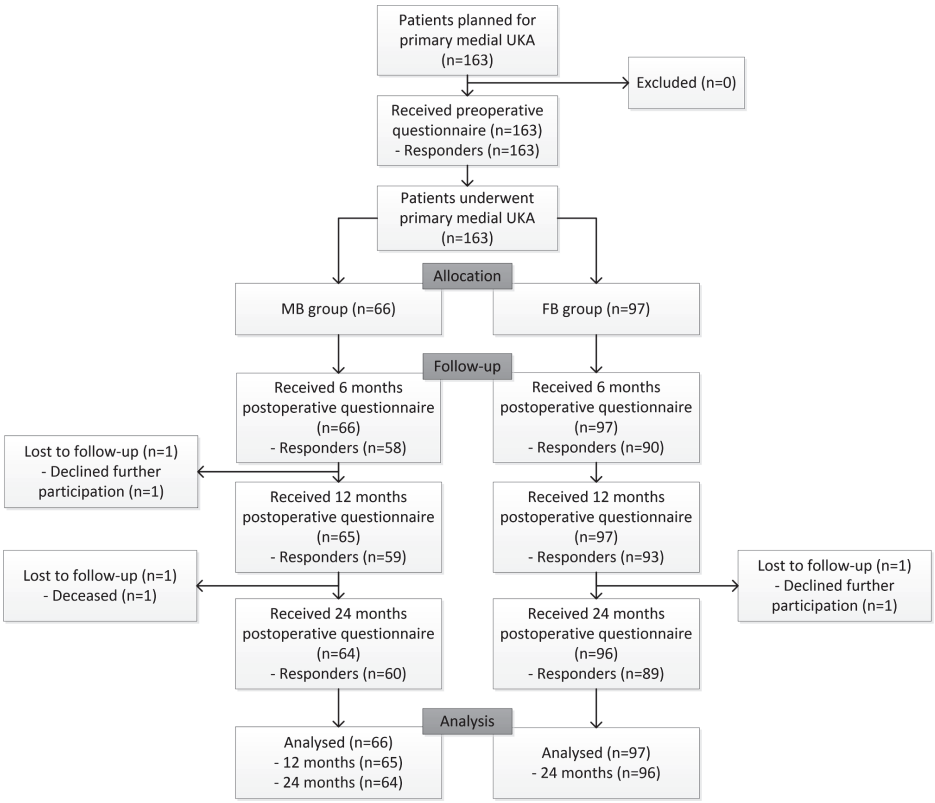


Fig. 1 Flowchart

UKA unicompartmental knee arthroplasty, ***MB*** mobile bearing implant, ***FB*** fixed bearing implant

Ethics

All patients signed informed consent to allow further scientific analysis using their anonymised data and thus the institutional review board of Kliniek ViaSana deemed that formal approval was not required for this study (2018-11).

Statistical analysis

Results were reported in mean and standard deviation (SD), median and interquartile range (IQR) or number (n) and percentage (%) based on the test performed. To investigate if there was any difference in patient characteristics between both groups preoperatively and to compare both groups on the postoperative outcomes, first continuous variables were checked for normal distribution. Second, independent t tests or Mann–Whitney U tests for continuous variables were executed depending on the normal distribution of the data or Pearson's Chi square or Fisher's exact tests for categorical variables. An alpha of 0.05 was considered statistically significant. Statistical analysis was performed using IBM SPSS Statistics version 25.0 (IBM Corp, Armonk, New York). A post hoc power analysis (sample size calculation) was not performed as it would have no meaning in a retrospective study design. Especially if non-significant p values were found between both groups as these p values always correspond to a low post hoc power. At best, this power will be slightly larger than 50% for p values ≥ 0.05 [31].

RESULTS

The postoperative PROMs response rate was 87.8% (n = 58) and 92.8% (n = 90) at 6 months, 90.8% (n = 59) and 95.9% (n = 93) at 12 months, and 93.8% (n = 60) and 92.7% (n = 89) at 24 months for the MB and FB groups, respectively (Fig. 1).

For the MB group the median NRS satisfaction score was 9.0 (8.0–10.0) compared to 9.0 (8.0–9.5) for the FB group at 6 months (p = 0.620). Similar scores were found at 12 and 24 months; MB 9.0 (8.0–10.0) and FB 9.0 (8.0–10.0) (p = 0.556), and MB 9.0 (8.0–10.0) and FB 9.0 (8.0–10.0) (p = 0.522), respectively. There were no statistically significant differences found between groups in all secondary outcomes at 6, 12 and 24 months (Table 2). In the MB group, there was 1 (1.5%) complication: 1 deep infection. In the FB group, 2 (2.1%) complications occurred: 1 deep infection and in one patient, bone cement debris was removed. Both knees with a deep infection were treated with debridement and implant retention, in both further follow-up was uneventful and outcome was not different.

Table 2 Outcomes in the MB and FB groups 6, 12 and 24 months postoperatively

	MB group	FB group	p value
6 months postoperatively	**(n = 66)**	**(n = 97)**	
NRS satisfaction score [median (IQR)]	9.0 (8.0-10.0)	9.0 (8.0-9.5)	0.620
NRS pain score [median (IQR)]			
At rest	0.0 (0.0-2.0)	1.0 (0.0-2.0)	0.442
During activity	2.0 (0.0-4.0)	2.0 (1.0-3.0)	0.879
OKS [median (IQR)]	42.0 (38.0-45.0)	41.0 (36.0-45.0)	0.487
KOOS-PS score [mean (SD)]	25.0 (11.9)	27.9 (11.8)	0.153
EQ-5D-3L [median (IQR)]			
EQ-5D descriptive system	0.895 (0.807-1.000)	0.897 (0.807-1.000)	0.852
EQ VAS	86.5 (75.0-90.8)	82.0 (75.0-91.5)	1.000
Anchor question [median (IQR)]			
Pain	6.0 (6.0-7.0)	6.0 (6.0-7.0)	0.645
Function	6.0 (6.0-7.0)	6.0 (6.0-7.0)	0.438
GPR	5.0 (5.0-6.0)	5.0 (5.0-5.5)	0.441
12 months postoperatively	**(n = 65)**	**(n = 97)**	
NRS satisfaction score [median (IQR)]	9.0 (8.0-10.0)	9.0 (8.0-10.0)	0.566
NRS pain score [median (IQR)]			
At rest	0.0 (0.0-1.0)	0.0 (0.0-2.0)	0.193
During activity	1.0 (0.0-3.0)	1.0 (0.0-3.0)	0.832
OKS [median (IQR)]	43.0 (38.0-46.0)	44.0 (37.0-46.0)	0.774
KOOS-PS score [median (IQR)]	24.9 (14.8-29.7)	27.5 (14.8-35.3)	0.306
EQ-5D-3L [median (IQR)]			
EQ-5D descriptive system	1.000 (0.775-1.000)	1.000 (0.811-1.000)	0.366
EQ VAS	82.5 (78.0-90.0)	80.0 (74.8-90.0)	0.714
Anchor question [median (IQR)]			
Pain	6.0 (6.0-7.0)	6.0 (6.0-7.0)	0.807
Function	6.5 (6.0-7.0)	6.0 (6.0-7.0)	0.602
GPR	5.0 (5.0-6.0)	5.5 (5.0-6.0)	0.343
Complication rate [n (%)]	1 (1.5)	2 (2.1)	1.000
Revision rate [n (%)]	0 (0.0)	0 (0.0)	1.000
24 months postoperatively	**(n = 64)**	**(n = 96)**	
NRS satisfaction score [median (IQR)]	9.0 (8.0-10.0)	9.0 (8.0-10.0)	0.522
NRS pain score [median (IQR)]			
At rest	0.0 (0.0-1.8)	0.0 (0.0-1.0)	0.826
During activity	1.0 (0.0-3.0)	1.0 (0.0-3.0)	0.710
OKS [median (IQR)]	44.0 (38.3-46.0)	44.0 (39.0-47.0)	0.882

Table 2 Outcomes in the MB and FB groups 6, 12 and 24 months postoperatively (continued)

	MB group	FB group	p value
KOOS-PS score [median (IQR)]	24.9 (10.5–33.2)	22.00 (14.8–31.8)	0.580
EQ-5D-3L [median (IQR)]			
EQ-5D descriptive system	0.949 (0.807–1.000)	1.000 (0.819–1.000)	0.351
EQ VAS	84.0 (71.8–90.0)	81.0 (76.0–91.0)	0.517
Anchor question [median (IQR)]			
Pain	6.0 (6.0–7.0)	6.0 (6.0–7.0)	0.246
Function	6.0 (6.0–7.0)	6.0 (6.0–7.0)	0.387
GPR	5.0 (5.0–6.0)	6.0 (5.0–6.0)	0.459

MB mobile bearing implant, ***FB*** fixed bearing implant, ***NRS*** Numeric Rating Scale, ***IQR*** interquartile range, ***OKS*** Oxford Knee Score, ***KOOS-PS*** Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form, ***SD*** standard deviation, ***EQ-5D-3L*** EuroQol 5 dimensions 3-level version, ***EQ VAS*** EuroQol Visual Analogue Scale, ***GPR*** General Perceived Recovery

DISCUSSION

The most important findings of the current study were very high patient satisfaction scores in both MB and FB groups of a 9.0 out of 10.0 at 6 months as well as at 12 and 24 months, without statistically significant differences between groups. Low pain, high functional and high quality of life scores were found at 6, 12 and 24 months, again without statistically significant differences between groups.

Previous study outcomes showed that 92% to 94% of UKA patients are satisfied after UKA [4, 10, 22, 37]. Comparison to the current study is marred by differences in follow-up time, multiple surgeons and no reported data on surgeon volume. The Swedish Knee Arthroplasty Register reported that 89% was satisfied at one year after surgery [35], similar to results in the current study. Implant design had no effect on outcome in the current study, similar to what has been reported by others; albeit at a minimum follow-up of 2 years, with surgeries performed by more than one surgeon [2, 3, 29]. Because of poorer reliability and validity, single-item questions used for measuring satisfaction are more vulnerable to random measurement errors and unknown biases [1, 34]. Therefore, anchor questions were used in the current study to consolidate these satisfaction scores. These anchor questions reported large improvements in function, pain relief and good recovery using both MB and FB implants, with no difference between groups.

Previously, the lack of uniform use of PROMs to measure clinical outcomes has made it difficult to compare PROs between implants [6, 16, 32]. Even more important, making it more difficult to reliably advice patients on which implant to choose. The International Consortium for Health Outcomes Measurement (ICHOM) [17] and registries, such as the English National Joint Registry (NJR) [30] and the Dutch Registry (LROI) [28], have recommended the use of the OKS and/or KOOS-PS to measure the patient-reported functional outcome in knee arthroplasty patients. Only

two previous studies compared MB to FB implants using OKS scores and showed no differences in outcome [11, 29], similar to the outcome in the current study. For quality of life, ICHOM [17], NJR [30], LROI [28] and the International Society of Arthroplasty Registries (ISAR) PROMs working group [34] have advised the use of the EQ-5D or the 12-Item Short Form Survey (SF-12). SF-12 scores were reported to be equal between both implants [3, 29]. Scores for both EQ-5D and the NRS pain questionnaires [17, 28], comparing MB to FB implants, have not been previously reported. The current study showed no differences in pain, function and quality of life between MB and FB at 6, 12 and 24 months after UKA.

These results can be used in shared decision making in the care of patients presenting with medial unicompartmental end-stage knee OA. Provided the surgeon is a high-volume surgeon, both MB and FB resulted in excellent PROs in the current study. As the used MB implant (Oxford) and FB implant (Physica ZUK) in the current study are the most commonly used implants of each design type among others in the Netherlands [26] and Sweden [35], this study provides accurate input for daily practice. It provides the patients and surgeons with useful information on PROs of the most commonly used UKA implants.

As a strength of this study, high response rates on PROMs were achieved resulting in representative outcomes. Moreover, anchor questions were used to consolidate the satisfaction scores. Furthermore, all UKA surgeries were performed by a high-volume and non-designer single surgeon, eliminating any bias and/or differences caused by a difference in surgical skill. As a limitation, the current study has a retrospective, nonrandomized design. The actual preferred sample size is unknown; the sample size of this study is larger than previous studies [2, 3, 29]. It is questionable if the results are generalizable, as only patients characterized by a BMI ≤ 35 and ASA score of I–II were included in the current study. It should be mentioned, however, that these characteristics apply to 90% of the total UKA population [27]. Future studies should focus on a larger sample size of the total UKA population. A prospective level 1 study should be performed to confirm our results.

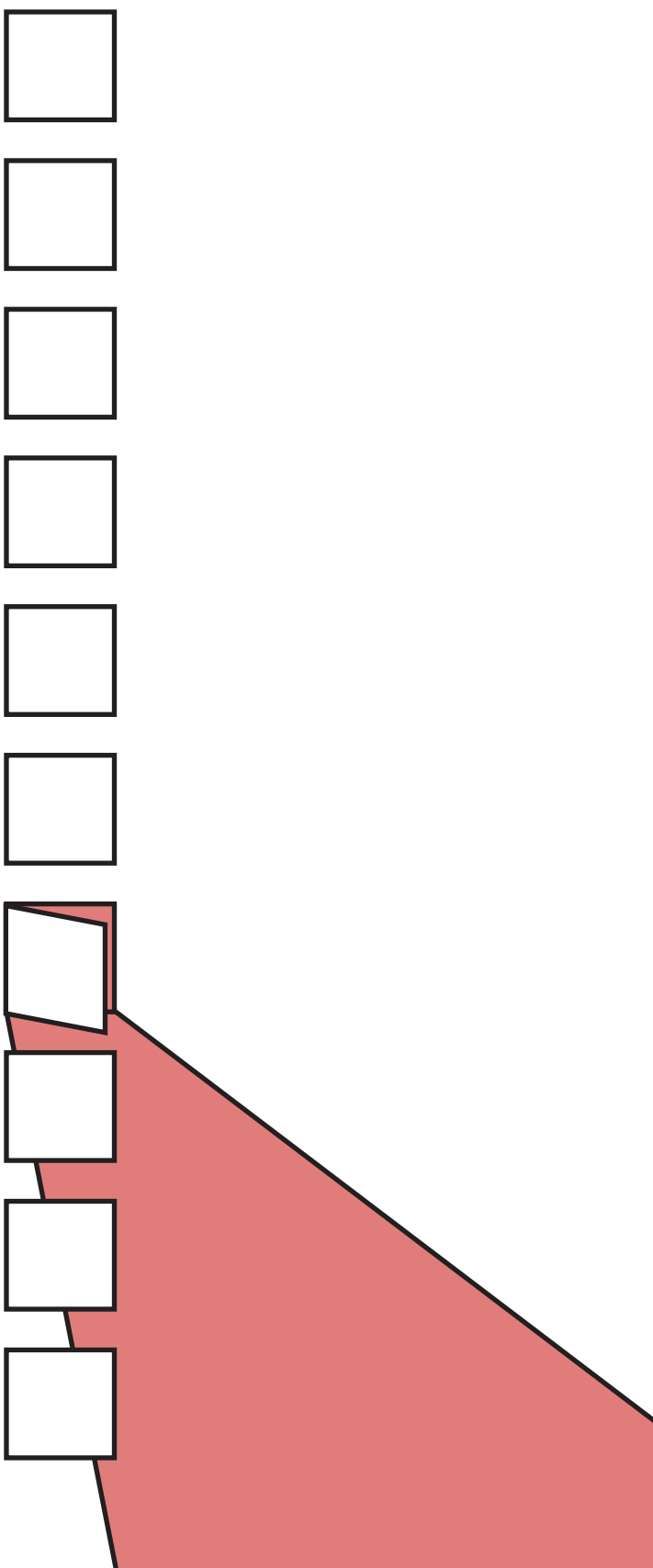
CONCLUSION

Medial UKA performed by a high-volume surgeon, using a MB implant or a FB implant, results in excellent patient satisfaction, pain relief, functional improvement and quality of life improvement at 6, 12 and 24 months after surgery. The recommendation and use of one over the other is not justified based on the outcomes in the current study.

REFERENCES

1. Andrawis JP, Chenok KE, Bozic KJ (2013) Health policy implications of outcomes measurement in orthopaedics. Clin Orthop Relat Res 471(11):3475–3481
2. Bhattacharya R, Scott CEH, Morris HE, Wade F, Nutton RW (2012) Survivorship and patient satisfaction of a fixed bearing unicompartmental knee arthroplasty incorporating an all polyethylene tibial component. Knee 19(4):348–351
3. Biau DJ, Greidanus NV, Garbuz DS, Masri BA (2013) No difference in quality-of-life outcomes after mobile and fixed-bearing medial unicompartmental knee replacement. J Arthroplasty 28(2):220–226
4. Biswal S, Brighton RW (2010) Results of unicompartmental knee arthroplasty with cemented, fixed-bearing prosthesis using minimally invasive surgery. J Arthroplasty 25(5):721–727
5. Campi S, Tibrewal S, Cuthbert R, Tibrewal SB (2018) Unicompartmental knee replacement—current perspectives. J Clin Orthop Trauma 9(1):17–23
6. Cao ZW, Niu CL, Gong CZ, Sun Y, Xie JH, Song YL (2019) Comparison of fixed-bearing and mobile-bearing unicompartmental knee arthroplasty: a systematic review and meta-analysis. J Arthroplasty 34(12):3114–3123
7. Choi YJ, Ra HJ (2016) Patient satisfaction after total knee arthroplasty. Knee Surg Relat Res 28(1):1–15
8. Dawson J, Fitzpatrick R, Murray D, Carr A (1998) Questionnaire on the perceptions of patients about total knee replacement. J Bone Jt Surg Br 80(1):63–69
9. EuroQol Group (1990) EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 16(3):199–208
10. Felts E, Parratte S, Pauly V, Aubaniac JM, Argenson JN (2010) Function and quality of life following medial unicompartmental knee arthroplasty in patients 60 years of age or younger. Orthop Traumatol Surg Res 96(8):861–867
11. Gleeson RE, Evans R, Ackroyd CE, Webb J, Newman JH (2004) Fixed or mobile bearing unicompartmental knee replacement? A comparative cohort study. Knee 11(5):379–384
12. Goh GSH, Liow MHL, Pang HN, Tay DKJ, Lo NN, Yeo SJ (2018) Patients with poor baseline mental health undergoing unicompartmental knee arthroplasty have poorer outcomes. J Arthroplasty 33(8):2428–2434
13. Goodfellow J, O'Connor J, Dodd C, Murray D (2006) Unicompartmental arthroplasty with the Oxford Knee. Oxford University Press, New York
14. Harris IA, Harris AM, Naylor JM, Adie S, Mittal R, Dao AT (2013) Discordance between patient and surgeon satisfaction after total joint arthroplasty. J Arthroplasty 28(5):722–727
15. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, Fainsinger R, Aass N, Kaasa S (2011) Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: A systematic literature review. J Pain Symptom Manage 41(6):1073–1093
16. Huang F, Wu D, Chang J, Zhang C, Qin K, Liao F, Yin Z (2019) A comparison of mobile- and fixed-bearing unicompartmental knee arthroplasties in the treatment of medial knee osteoarthritis: a systematic review and meta-analysis of 1,861 patients. J Knee Surg. https://doi.org/10.1055/s-0039-1697901
17. ICHOM (nd) Hip and Knee Osteoarthritis Standard Set. https:// www.ichom.org/portfolio/hip-knee-osteoarthritis/. Accessed 12 Mar 2020
18. Kamper SJ, Ostelo RWJG, Knol DL, Maher CG, de Vet HCW, Hancock MJ (2010) Global perceived effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol 63(7):760–766
19. Kellgren JH, Lawrence JS (1957) Radiological assessment of osteoarthrosis. Ann Rheum Dis 16(4):494–502
20. Koh IJ, Kim MS, Sohn S, Song KY, Choi NY, Jung H, In Y (2019) Predictive factors for satisfaction after contemporary unicompartmental knee arthroplasty and high tibial osteotomy in isolated medial femorotibial osteoarthritis. Orthop Traumatol Surg Res 105(1):77–83
21. Lau RL, Gandhi R, Mahomed S, Mahomed N (2012) Patient satisfaction after total knee and hip arthroplasty. Clin Geriatr Med 28(3):349–365
22. Lee M, Huang Y, Chong HC, Ning Y, Lo NN, Yeo SJ (2016) Predicting satisfaction for unicompartmental knee arthroplasty patients in an asian population. J Arthroplasty 31(8):1706–1710

23. Liddle AD, Judge A, Pandit H, Murray DW (2014) Determinants of revision and functional outcome following unicompartmental knee replacement. Osteoarthr Cartil 22(9):1241–1250
24. Liddle AD, Pandit H, Judge A, Murray DW (2015) Optimal usage of unicompartmental knee arthroplasty: a study of 41 986 cases from the national joint registry for England and Wales. Bone Jt J 97-B(11):1506–1511
25. Liddle AD, Pandit H, Judge A, Murray DW (2016) Effect of surgical caseload on revision rate following total and unicompartmental knee replacement. J Bone Jt Surg Am 98(1):1–8
26. LROI (nd) The most frequently registered primary unicondylar knee arthroplasties in The Netherlands in 2018. (2018a) https:// www.lroi-rapportage.nl/knee-primary-knee-arthroplasty-unicondylar-knee-arthroplasty-most-frequently-registered-unicondylar-knee-prostheses. Accessed 12 Mar 2020
27. LROI (nd) Patient characteristics of all patients with a registered primary knee arthroplasty by type of knee arthroplasty in The Netherlands in 2018. (2018b) https://www.lroi-rapportage.nl/ knee-primary-knee-arthroplasty-demographics-patient-characteristics-by-type-of-knee-prosthesis19. Accessed 12 Mar 2020
28. Nederlandse Orthopaedische Vereniging (nd) Registratieformulieren. https://www.lroi.nl/invoerders/registreren/registratieformulieren. Accessed 1 Aug 2018
29. Neufeld ME, Albers A, Greidanus NV, Garbuz DS, Masri BA (2018) A comparison of mobile and fixed-bearing unicompartmental knee arthroplasty at a minimum 10-year follow-up. J Arthroplasty 33(6):1713–1718
30. NHS (nd) Patient Reported Outcome Measures (PROMs). https:// digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms. Accessed 12 Mar 2020
31. NOV (nd) PROMs. https://www.orthopeden.org/kwaliteit/kwaliteitsbeleid/proms. Accessed 20 Apr 2020
32. Peersman G, Stuyts B, Vandenlangenbergh T, Cartier P, Fennema P (2015) Fixed- versus mobile-bearing UKA: a systematic review and meta-analysis. Knee Surg Sport Traumatol Arthrosc 23(11):3296–3305
33. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, Roos EM, Jordan JM, Maillefert JF, Dougados M, Davis AM (2008) The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS)—an OARSI/OMERACT initiative. Osteo arthr Cartil 16(5):542–550
34. Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, Dunn J, Eresian Chenok K, Dunbar M, Overgaard S, Garellick G, Lübbeke A (2016) Patient-reported outcome measures in arthroplasty registries: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. Acta Orthop 87(1):9–23
35. Swedish Knee Arthroplasty Register (2019) Annual Report 2019. https://www.myknee.se/pdf/SVK_2019_1.0_Eng.pdf. Accessed 13 Mar 2020
36. The New Zealand Joint Registry (2019) The New Zealand Joint Registry twenty year report January 1999 to December 2018. https://nzoa.org.nz/system/files/DH8328_NZJR_2019_Repor t_v4_7Nov19.pdf. Accessed 11 Mar 2020
37. van der List JP, Zuiderbaan HA, Pearle AD (2016) Why do medial unicompartmental knee arthroplasties fail today? J Arthroplasty 31(5):1016–1021
38. Von Keudell A, Sodha S, Collins J, Minas T, Fitz W, Gomoll AH (2014) Patient satisfaction after primary total and unicompartmental knee arthroplasty: an age-dependent analysis. Knee 21(1):180–184
39. Wylde V, Blom AW, Whitehouse SL, Taylor AH, Pattison GT, Bannister GC (2009) Patient-reported outcomes after total hip and knee arthroplasty Comparison of midterm results. J Arthroplasty 24(2):210–216



CHAPTER 7

Is patient satisfaction after total knee arthroplasty predictable using patient characteristics and preoperative patient-reported outcomes?

Y. Pronk, M.C.W.M. Peters and J.M. Brinkman

J Arthroplasty. 2021 Jul;36(7):2458-2465

ABSTRACT

Background

Dissatisfaction after total knee arthroplasty (TKA) remains a difficult problem. Patient characteristics and preoperative patient-reported outcomes (PROs) are potential predictors of satisfaction one year after TKA. Being able to predict the outcome preoperatively might reduce the number of less satisfied patients.

Methods

A retrospective cohort study on prospectively collected data of 1239 primary TKA patients (ASA I-II, BMI <35) was performed. Primary outcome was degree of patient satisfaction one year after TKA (Numeric Rating Scale (NRS) 0-10). Secondary outcomes were degree of patient satisfaction six months and two years after TKA and being dissatisfied (NRS 0-6) or satisfied (NRS 7-10) at all three time points. Multivariate linear and binary logistic regression analyses were executed with patient characteristics and preoperative PROs as potential predictors.

Results

One year after TKA, median NRS satisfaction score was 9.0 (8.0-10.0) and 1117 (90.2%) patients were satisfied. BMI, degree of medial cartilage damage, previous knee surgery, Knee injury and Osteoarthritis Outcome Score-Physical Function Short Form score, EQ VAS score, and anxiety were identified as predictors of the degree of patient satisfaction ($P = .000$, $R^2 = 0.027$). Models on secondary outcomes reported R^2 of 1.7%-7.1% ($P < .05$). All models showed bad agreement between observed and predicted values for lower NRS satisfaction scores and being dissatisfied.

Conclusion

The degree of patient satisfaction and the chance of being dissatisfied or satisfied six months, one, and two years after TKA are predictable by patient characteristics and preoperative PROs but not at a reliability level that is clinically useful.

INTRODUCTION

Total knee arthroplasty (TKA) is a successful treatment option for patients with end-stage osteoarthritis (OA) [1]. Time to reach the end result is approximately six to twelve months [2]. TKA surgically is a success if alignment is correct, the implant well fixed and balanced, and long-term outcome is considered good as long as no revision has been performed. However, these objective outcomes are quite possibly not in line with patients' definition of a successful outcome after TKA. Subjective patient-reported outcomes (PROs) can be measured using patient-reported outcome measures (PROMs). PROs, therefore, also need to be evaluated to be able to consider the surgery a success. Unfortunately, up to 20% of patients are dissatisfied with their end result [3-10].

Several preoperative and postoperative predictors for patient (dis)satisfaction after TKA have been found in previous studies (Table 1). However, these studies included potential predictors that can only be measured after surgery [3-5,7,9-18], did not combine potential predictors into one prediction model [19], did not included preoperative PROs [20], or were based on a recovery period <1 year [8]. Knowledge of preoperative predictors of patient satisfaction could enhance shared decision-making between patient and surgeon, optimize personalized health care, and lower the overall percentage of dissatisfied patients.

Therefore, the primary aim of the present study was to investigate which patient characteristics and preoperative PROs are predictors of the degree of patient satisfaction one year after TKA. The secondary aim was to investigate which patient characteristics and preoperative PROs are predictors of the degree of patient satisfaction six months and two years after TKA and the chance of being dissatisfied or satisfied at all three time points separately. The hypothesis was that patient satisfaction can be predicted preoperatively using patient characteristics and preoperatively collected PROs and that PROs have an important role in this prediction.

Table 1 Predictors of patient (dis)satisfaction after TKA, focused on dissatisfaction

Predictor	Associated with dissatisfaction
Preoperative	
Sex	Female [8,14,20,21]
Age	Contradicted associations found [5,7,8,13–15,21]
BMI	Higher BMI [15,21]
ASA score	ASA score > 2 [17]
SES	Lower SES [20]
Diagnosis	Diagnosis of osteoarthritis [14]
Cartilage damage	Lower KL score [9], lower radiological joint narrowing score [15]
Previous knee surgery	History of previous knee surgery [21]
Comorbidity	Presence of comorbidity [6,21]

Table 1 Predictors of patient (dis)satisfaction after TKA, focused on dissatisfaction (continued)

Predictor	Associated with dissatisfaction
Waiting time	Prolonged waiting time [16]
Preoperative pain	Contradicted associations found [5,8,18]
Health status	Contradicted associations found [4,6,10]
Anxiety or depression	Presence of anxiety or depression [3,4,6,8,16,21]
Postoperative	
Length of stay	Longer length of stay [3]
Complication	Having a deep prosthetic infection [3], having a complication [5,15,17]
Stiffness	Presence of stiffness [5,7]
Improvement	Poor pain relief and/or function improvement [4,5,10,11,18]
Expectations	No fulfilment of expectations [5,11,12]

ASA score; American Society of Anesthesiologist score, BMI; body mass index, KL score; Kellgren-Lawrence score, SES; social economic status, TKA; total knee arthroplasty

METHODS

A retrospective cohort study with prospectively collected data of primary TKA patients was performed. Patients underwent surgery between July 2015 and December 2017 in a medium sized orthopedic clinic. Guidelines of the Dutch Health Regulatory Agency (IGJ) state that patients with an American Society of Anesthesiologists (ASA) score >2 and a BMI >35 are not allowed to be operated on in the present study's clinic. Therefore all patients in the present study had an ASA score of I-II and a body mass index (BMI) 35 kg/m^2.

Five high-volume orthopedic surgeons performed all surgeries. Standard surgical technique as described by the implant manufacturer was used. In all patients, the same implant, either the posterior-stabilized or cruciate-retaining version based on surgeon's preference, was used (Nexgen, Zimmer Biomet Ltd., Bridgend, UK). The length of stay was 1 or 2 days.

Patients were included if they signed the informed consent form preoperatively to allow further scientific analysis using their anonymized data. Therefore, the institutional review board ruled that formal approval was not required for this study. Patients were excluded if they had no Numeric Rating Scale (NRS) satisfaction score one year after TKA.

Outcomes

Primary outcome was the degree of patient satisfaction one year after TKA measured using a nonvalidated NRS satisfaction question, with a score of 0 being very dissatisfied and 10 being very satisfied. The question asked was "How satisfied are you (in general) with the results of your knee surgery?". To validate these scores, three anchor questions were inquired one year after TKA:

functional improvement on a 7-point Likert scale ranging from 1 (very much deteriorated) to 7 (very much improved), pain relief on a 7-point Likert scale ranging from 1 (very much deteriorated) to 7 (very much improved), and degree of recovery measured using general perceived recovery question with a 7-point Likert scale ranging from 0 (worse than ever) to 6 (fully recovered) [22]. If patients scored high on NRS satisfaction question and low on the three anchor questions, or the other way around, their NRS satisfaction scores were recoded into missing values.

Secondary outcomes were the degree of patient satisfaction six months and two years after TKA separately and the chance of being dissatisfied or satisfied at all three time points separately. Being dissatisfied or satisfied was based on dichotomizing the NRS satisfaction scores; scores from 0 to 6 were defined as dissatisfied and scores from 7 to 10 as satisfied.

Investigated potential predictors

Investigated potential predictors were based on predictors identified in previous studies (Table 1). Investigated potential predictive preoperative PROs were pain, function, quality of life, and anxiety. Pain at rest and pain during activity were both measured using a NRS question scored from 0 (no pain) to 10 (severe pain). Knee function and pain were assessed using the Oxford Knee Score questionnaire with scores ranging from 0 (most severe symptoms) to 48 (least symptoms) [23]. Furthermore, knee function on a scale from 0 (no difficulty) to 100 (extreme difficulty) was measured using the Knee injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS) questionnaire [24]. Quality of life was inquired using the 3-level version of EuroQol 5 Dimensions (EQ-5D-3L) questionnaire consisting of two parts: EQ visual analog scale (EQ VAS; 0-100, 0 worst imaginable health state to 100 best imaginable health state) and EQ-5D descriptive system [25]. Anxiety was assessed using question 5 of the EQ-5D-3L; answers were combined into having anxiety ("I am moderately anxious" and "I am very anxious") and no anxiety.

Investigated potential predictive patient characteristics were sex, age at surgery (years), BMI, ASA score (I or II), social economic status (low, average, or high), diagnosis (OA or other), degree of medial cartilage damage assessed with the Kellgren-Lawrence score (KL score; 0-4), degree of lateral cartilage damage (KL score; 0-4), degree of retropatellar cartilage damage (KL score; 0-4), Charnley score, history of previous knee surgery (yes or no), comorbidity (yes or no), and wait time between screening and surgery (days).

Measurements

PROs were primarily collected digitally (OnlinePROMs, Rosmalen, the Netherlands). PROs collection was performed as per the advice of the Dutch Orthopedic Association preoperatively (typically 4-8 weeks before surgery) and six months, one year, and two years postoperatively [26]. In case patients were not able to handle a computer, paper questionnaires were sent. A maximum of two reminders were sent to complete the PROMs [27].

Patient characteristics were collected from the electronic patient records. Social economic status was determined by the Netherlands Institute for Social Research using postal code [28].

Radiographs were assessed by the radiologist, the treating surgeon, and the presence of grade IV OA (or not) was confirmed and recorded at the time of surgery.

Statistical analysis

Analyses were performed using SPSS version 25.0 (IBM Corp; Armonk, New York). The results were reported in mean and standard deviation, median and interquartile range, or number (n) and percentage (%) based on the test performed. To investigate if there was any difference between the included and excluded patients, patient characteristics and preoperative PROs were compared between both groups. First, continuous variables were checked for normal distribution using skewness and histograms. Second, depending on the normal distribution of the data, independent t-tests or Mann-Whitney U tests for continuous variables were executed and Pearson's chi-square or Fisher's exact tests for categorical variables. Missing values were investigated, found to be below 15% per potential predictors (≤1.6%) and assessed as missing at random.

To obtain an impression which patient characteristics and preoperative PROs dissatisfied and satisfied patients differed, included patients were allocated to the subgroup very satisfied or the subgroup dissatisfied one year after TKA and compared. NRS satisfaction scores at one year were dichotomized in scores 0-6 for being dissatisfied and scores 9-10 for being very satisfied.

Prediction of NRS satisfaction score six months, one year, and two years after TKA as dependent variables separately with patient characteristics and preoperative PROs as independent variables (n = 20) was investigated using multivariate linear regression analyses with backward selection. Parameter estimates (b) and P-values were calculated. Variables with a P-value <.200 were considered for the final models. Before, assumptions were checked and met. Skewness was found for wait time, NRS pain score during activity, KOOS-PS score, EQ-5D descriptive system score, and EQ VAS score. These variables were transformed into a normal distribution by ^(1/2) or ^(1/3) in case of positive skewness and by ^3 in case of negative skewness (square root or cube root).

Prediction of the chance of being dissatisfied (coded 1) or satisfied (coded 0) six months, one year, and two years after TKA as dependent variables separately with patient characteristics and preoperative PROs as independent variables was investigated using multivariate binary logistic regression analyses with Wald backward selection. Variables with a P-value <.05 were considered for the final models. Before, assumptions were checked and met. Around 10% of the patients were dissatisfied at all three time points separately, meaning a maximum of 11 variables could be included. These variables were chosen based on previous analyses of multivariate linear regression; statistically significant and clinically relevant predictors were chosen for the binary model. An odds ratio (OR) > 1 indicated an increasing risk of being dissatisfied relative to being satisfied.

In depth analyses with cartilage damage as independent variable were executed. First, the degree of medial cartilage damage was dichotomized in KL score 4 (coded 1) and KL scores 0-3 (coded 0). All six multivariate regression analyses were executed using this independent variable instead

of the degree of medial cartilage damage. Second, the degree of medial and lateral cartilage damages were combined and dichotomized in medial KL score 4 (coded 1) and medial KL scores 0-3 with lateral KL score 4 (coded 0). Again, all six regression analyses were executed using this variable.

Although the focus was not on postoperative predictors, the influence of complications on satisfaction was also assessed. Variables were categorized into having a complication (coded 1) or not (coded 0) and having a deep prosthetic infection (coded 1) or not (coded 0); these variables were analyzed separately.

For the overall measurement of the predictive ability, models with explained variances (R^2) ≥25% were considered as strong performing models and with <10% as weak performing models. Scatter plots were created to investigate the agreement between the observed and predicted values.

RESULTS

Flowchart, patient characteristics, and preoperative PROs

A total of 1367 patients underwent primary TKA surgery of which 1239 patients (90.6%) were included. Main exclusion reason was not being a responder on the NRS satisfaction question at the one-year postoperative questionnaire (n = 95 (6.9%)). Because of conflicting NRS satisfaction question and the three anchor question scores, NRS satisfaction scores of 15 patients (1.2%) were coded into missing values. Of the included patients, 1235 (99.7%) responded on the preoperative PROMs, 1196 (96.5%) at six months, 1239 (100%) at one year, and 1175 (94.8%) on the two-year postoperative PROMs (Fig. 1). No statistically significant differences were found in patient characteristics and preoperative PROs between included and excluded patients (Table 2).

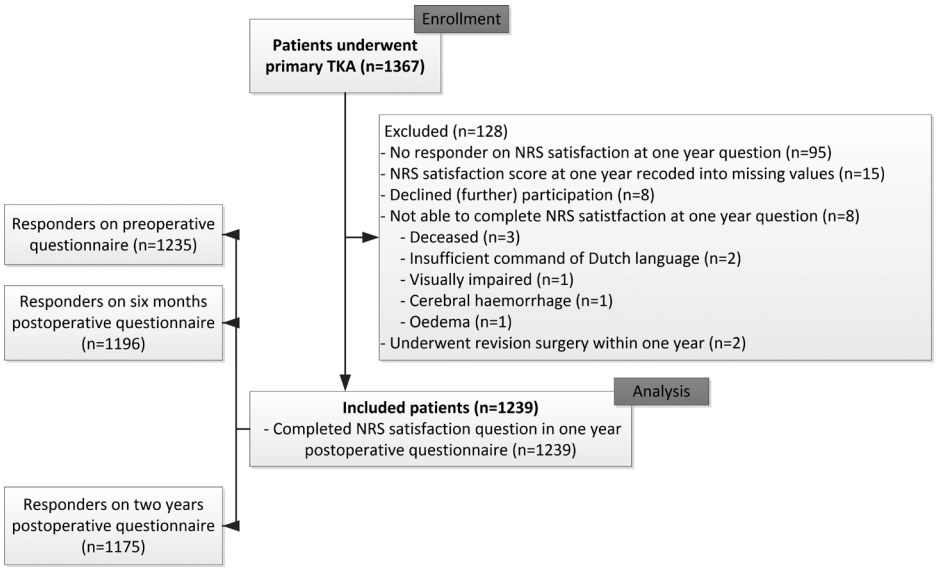


Fig. 1. Flowchart

n; number, NRS; numeric rating scale, TKA; total knee arthroplasty

Table 2 Patient characteristics and preoperative PROs of included and excluded TKA patients

	Included (n=1239)	Excluded (n=128)	p-value
Patient characteristics			
Sex (male), n (%)	609 (49.2)	71 (55.5)	0.174
Age (y), mean (SD)	65.3 (7.9)	64.1 (9.4)	0.112
BMI (kg/m^2), mean (SD)	28.1 (3.6)	28.1 (3.6)	0.835
ASA (I), n (%)	574 (46.3)	56 (43.8)	0.578
SES, n (%)			0.340
Low	134 (10.9)	11 (8.7)	
Average	981 (79.8)	107 (84.9)	
High	115 (9.3)	8 (6.3)	
Diagnosis (OA), n (%)	1188 (95.9)	122 (95.3)	0.954
Medial cartilage damage (KL score), n (%)			0.365
No presence of OA (0)	140 (11.3)	13 (10.2)	
Doubtful presence of OA (1)	7 (0.6)	1 (0.8)	
Minimal presence of OA (2)	77 (6.2)	4 (3.1)	
Moderate presence of OA (3)	136 (11.0)	10 (7.9)	
Severe presence of OA (4)	877 (70.9)	99 (78.0)	

Table 2 Patient characteristics and preoperative PROs of included and excluded TKA patients (continued)

	Included (n=1239)	Excluded (n=128)	p-value
Lateral cartilage damage (KL score), n (%)			0.836
No presence of OA (0)	357 (28.9)	34 (26.8)	
Doubtful presence of OA (1)	25 (2.0)	4 (3.1)	
Minimal presence of OA (2)	135 (10.9)	17 (13.4)	
Moderate presence of OA (3)	324 (26.2)	33 (26.0)	
Severe presence of OA (4)	396 (32.0)	39 (30.7)	
Retro-patellar cartilage damage (KL score), n (%)			0.472
No presence of OA (0)	258 (20.9)	28 (22.0)	
Doubtful presence of OA (1)	9 (0.7)	1 (0.8)	
Minimal presence of OA (2)	213 (17.2)	20 (15.7)	
Moderate presence of OA (3)	450 (36.4)	38 (29.9)	
Severe presence of OA (4)	307 (24.8)	40 (31.5)	
Charnley score, n (%)			0.326
One knee affected with OA	577 (46.6)	61 (47.7)	
Both knees affected with OA	291 (23.5)	22 (17.2)	
Contralateral TKA	250 (20.2)	32 (25.0)	
Multiple joints affected with OA	121 (9.8)	13 (10.2)	
Previous knee surgery, n (%)	706 (57.0)	78 (60.9)	0.394
Comorbidity, n (%)	591 (47.7)	70 (54.7)	0.132
Wait time (d), median (IQR)	65.0 (42.0-84.0)	67.5 (48.0-91.0)	0.263
Preoperative PROs			
NRS pain score in rest, mean (SD)	53.7 (24.4)	54.0 (23.5)	0.913
NRS pain score during activity, median (IQR)	80.0 (70.0-90.0)	80.0 (70.0-90.0)	0.538
OKS score, mean (SD)	24.2 (7.3)	23.6 (8.1)	0.389
KOOS-PS score, median (IQR)	48.5 (42.0-57.9)	51.2 (42.0-57.9)	0.528
EQ-5D-3L, median (IQR)			
EQ-5D descriptive system score	0.775 (0.516-0.775)	0.775 (0.298-0.778)	0.466
EQ VAS score	80.0 (70.0-90.0)	80.0 (65.0-86.0)	0.056
Anxiety, n (%)	230 (18.6)	26 (21.1)	0.741

ASA score; American Society of Anesthesiologists score, BMI; body mass index, EQ VAS; EuroQol visual analogue scale, EQ-5D descriptive system; EuroQol 5 Dimensions descriptive system, EQ-5D-3L; 3-level version of EuroQol 5 Dimensions, IQR; interquartile range, KL score; Kellgren-Lawrence score, KOOS-PS; Knee injury and Osteoarthritis Outcome Score – Physical function Short from, n; number, NRS; numeric rating scale, OA; osteoarthritis, OKS; Oxford Knee Score, PROs; patient-reported outcomes, SD; standard deviation, SES; social economic status, TKA; total knee arthroplasty

Dissatisfied versus very satisfied patients one year after TKA

The dissatisfied subgroup (n = 122 (9.8%)) was found to be statistically significantly different compared with the very satisfied subgroup (n = 805 (65.0%)) on the following patient characteristics and preoperative PROs: different distribution in Charnley score (P = .021), higher KOOS-PS score of 52.8 (44.0-57.9) than 49.4 (40.3-57.9) (P = .024), lower EQ-5D descriptive system score of 0.577 (0.298-0.775) than 0.654 (0.651-0.775) (P = .005), lower EQ VAS score of 74.2 (61.5-90.0) than 78.9 (72.0-90.0) (P = .005), and more patients having anxiety (n = 25 (20.7%) versus n = 126 (15.7%), P = .001).

Degree of patient satisfaction one year after TKA

Median NRS satisfaction score one year after TKA was 9.0 (8.0-10.0); most patients scored an 8 (n = 220 (17.8%)), 9 (n = 319 (25.7%)), or 10 (n = 486 (39.2%)).

The final multivariate linear regression model with NRS satisfaction score at one year as dependent outcome identified BMI (b = 0.032, P = .024), degree of medial cartilage damage (b = 0.057, P = .125), history of previous knee surgery (b = -0.137, P = .169), KOOS-PS score (b = -0.071, P = .187), EQ VAS score (b < 0.001, P = .000), and anxiety (b = -0.239, P = .067) as predictors (R^2 = 0.027, P = .000). Preoperatively BMI and EQ VAS score contributed statistically significant to predicting NRS satisfaction score (Table 3). Satisfaction scores were not influenced by surgeon (P = .456). The scatter plot showed a bad agreement between observed and predicted values for lower NRS satisfaction scores (Fig. 2).

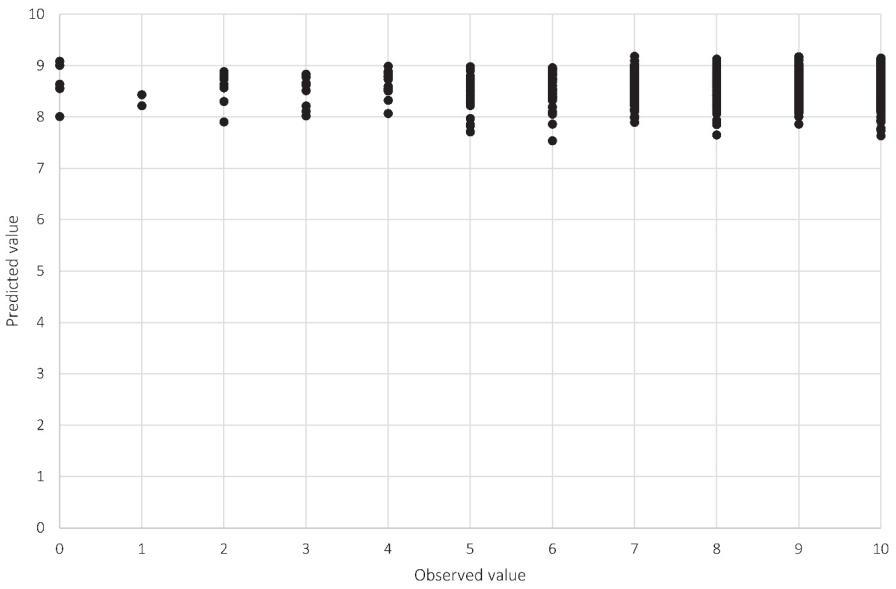


Fig. 2. Predicted versus observed value of NRS satisfaction score one year after TKA

Table 3 Final models of the multivariate linear regression analyses with NRS satisfaction score six months, one year and two years after TKA as dependent variables separately, and patient characteristics and preoperative PROs as independent variables (n=20)

		Final model				
Dependent variable	**Patients (n)**	**ANOVA (p-value)**	**R^2**	**Predictors**	**b (95%CI)**	**p-value**
NRS satisfaction score at six months	1144	0.000	0.046	Constant	6.299 (4.973 to 7.625)	**0.000**
				Age	0.009 (-0.005 to 0.023)	0.192
				BMI	0.032 (0.004 to 0.061)	**0.026**
				Medial cartilage damage	0.078 (0.003 to 0.152)	**0.042**
				Previous knee surgery	-0.180 (-0.388 to 0.028)	0.089
				Comorbidity	-0.138 (-0.339 to 0.064)	0.180
				T_EQ VAS	<0.001 (0.000 to 0.000)	**0.000**
				Anxiety	-0.324 (-0.589 to -0.060)	**0.016**
NRS satisfaction score at one year	1207	0.000	0.027	Constant	7.798 (6.718 to 8.878)	**0.000**
				BMI	0.032 (0.004 to 0.059)	**0.024**
				Medial cartilage damage	0.057 (-0.016 to 0.131)	0.125
				Previous knee surgery	-0.137 (-0.333 to 0.059)	0.169
				T_KOOS-PS	-0.071 (-0.176 to 0.034)	0.187
				T_EQ VAS	<0.001 (0.000 to 0.000)	**0.000**
				Anxiety	-0.239 (-0.459 to 0.017)	0.067
NRS satisfaction score at two years	1137	0.002	0.017	Constant	7.460 (6.433 to 8.487)	**0.000**
				ASA score	-0.179 (-0.414 to 0.056)	0.136
				BMI	0.033 (0.000 to 0.066)	**0.049**
				Medial cartilage damage	0.068 (-0.018 to 0.154)	0.121
				T_EQ VAS	<0.001 (0.000 to 0.000)	**0.024**
				Anxiety	-0.281 (-0.590 to 0.028)	0.075

Bold values denote p-values with statistical significance (P < .05).

95%CI; 95% confidence interval, ASA score; American Society of Anesthesiologists score, b; unstandardized B, BMI; body mass index, n; number, NRS; numeric rating scale, PROs; patient-reported outcomes, R^2; explained variance, T_EQ VAS; cube root transformed EuroQol visual analogue scale, T_KOOS-PS; square root transformed Knee injury and Osteoarthritis Outcome Score – Physical Function Short form, TKA; total knee arthroplasty

Degree of patient satisfaction six months and two years after TKA

Median NRS satisfaction score was 9.0 (8.0-10.0) six months after TKA and 9.0 (8.0-10.0) two years after TKA.

The final multivariate linear regression model with NRS satisfaction score at six months as dependent outcome identified age, BMI, degree of medial cartilage damage, history of previous knee surgery, comorbidity, EQ VAS score, and anxiety as predictors (R^2 = 0.046, P = .000). Preoperatively BMI, degree of medial cartilage damage, EQ VAS score, and anxiety contributed statistically significant to predicting NRS satisfaction score. Regarding NRS satisfaction score at two years as dependent outcome, the final multivariate linear regression model identified ASA score, BMI, degree of medial cartilage damage, EQ VAS score, and anxiety as predictors (R^2 = 0.017, P = .002). Preoperatively BMI and EQ VAS score contributed statistically significant to predicting NRS satisfaction score (Table 3). The satisfaction scores at six months (P = .774) and two years (P = .069) were not influenced by surgeon.

The chance of being dissatisfied or satisfied six months, one year, and two years after TKA

In total, 1033 (88.0%) patients were satisfied after six months, 1117 (90.2%) after one year, and 1041 (89.1%) after two years.

The final multivariate binary logistic regression model with being dissatisfied or satisfied at six months as dependent outcome identified BMI, degree of medial cartilage damage, Charnley score, EQ-5D descriptive system score, and EQ VAS score as predictors (R^2 = 0.071, P = .000). Goodness of fit test indicated a good fit (P = .273). The model correctly predicted 87.7% of the observed values: 99.7% for being satisfied and 1.4% for being dissatisfied. Regarding being dissatisfied or satisfied at one year as dependent outcome, the final multivariate binary logistic regression model identified degree of medial cartilage damage, Charnley score, and EQ-5D descriptive system score as predictors (R^2 = 0.057, P = .000). Goodness of fit test indicated a good fit (P = .115) and 90.2% of the observed values was correctly predicted (100.0% for being satisfied and 0.0% for being dissatisfied). The final multivariate binary logistic regression model with being dissatisfied or satisfied at two years as dependent outcome identified degree of medial cartilage damage as predictor (R^2 = 0.029, P = .038). Goodness of fit test indicated a good fit (P = .320) and 89.2% of the observed values was correctly predicted (100.0% for being satisfied and 0.0% for being dissatisfied) (Table 4). Outcomes at six months (P = .431), one year (P = .974), and two years after TKA (P = .428) were not influenced by surgeon.

In depth analysis of cartilage damage and complications

The degree of medial cartilage damage (KL score 4 vs KL scores 0-3) in the final multivariate linear regression model with NRS satisfaction score at six months and at one year and at two years as dependent outcomes was identified as a predictor (6 months: b = 0.221, P = .049; R^2 = 0.046, P = .000; one year: b = 0.221, P = .043; R^2 = 0.027, P = .000; and two years: b = 0.296, P = .022; R^2 = 0.019, P = .000). Regarding the final multivariate binary logistic regression model with being

dissatisfied or satisfied at one year as dependent outcome, this model identified it as a predictor (OR = 0.506, P = .001; R^2 = 0.050, P = .000). At 6 months and at 2 years, it did not (OR = 0.736, P = .123; R^2 = 0.062, P = .000 and OR = 0.687, P = .064; R^2 = 0.020, P = .040).

The degree of medial and lateral cartilage damages (medial KL score 4 vs medial KL scores 0-3 and lateral KL score 4) in the final multivariate linear regression model with NRS satisfaction score at six months and at one year and at two years as dependent outcomes was identified as a predictor (6 months: b = 0.279, P = .031; R^2 = 0.045, P = .000; one year: b = 0.306, P = .014; R^2 = 0.028, P = .000; and two years: b = 0.446, P = .003; R^2 = 0.025, P = .000). Regarding the final multivariate binary logistic regression model with being dissatisfied or satisfied at one year and two years as dependent outcomes, these models identified it as a predictor (one year: OR = 0.419, P = .000; R^2 = 0.059, P = .000 and two years: OR = 0.606, P = .026; R^2 = 0.030, P = .016). At 6 months, it did not (OR = 0.711, P = .134; R^2 = 0.059, P = .000).

There were not more complications in the patients who were excluded or had conflicting NRS and anchor question scores. Satisfaction scores in patients with a complication during the study period were below satisfaction scores for patients without a complication at six months (8.0 (6.0-9.0) vs 9.0 (8.0-10.0) (P < .001)), one year (9.0 (7.0-10.0) vs 9.0 (8.0-10.0) (P = .001)), and two years after TKA (9.0 (7.0-10.0) vs 9.0 (8.0-10.0) (P = .009)). Furthermore, a higher percentage of patients with a complication was dissatisfied at six months (19.7% vs 11.0% (P = .005)), one year (16.3% vs 9.0% (P = .008)), and two years after TKA (16.8% vs 10.0% (P = .025)) than patients without a complication. Having a complication or not did not affect the models' prediction ability. Having a deep prosthetic infection resulted in no significantly difference in outcome.

DISCUSSION

In the present study, patient satisfaction after TKA was measured on two different scales (continuous and dichotomous) at three different postoperative time points (six months, one year, and two years) and possible predictors of the outcome were investigated. The main result was that BMI, degree of medial cartilage damage, history of previous knee surgery, KOOS-PS score, EQ VAS score, and anxiety are predictors of the degree of patient satisfaction one year after TKA. Although this prediction model was statistically significant, the current model is not useful in clinical practice: explained variance was 2.7% (R^2 = 0.027) resulting in a weak performing model, and bad agreement was found between the observed and predicted values for the lower NRS satisfaction scores resulting in a model unable to predict these lower scores correctly. Regarding the secondary outcomes, the same results concerning no clinically useful models were found.

Table 4 Final models of the multivariate binary logistic regression analyses with being dissatisfied or satisfied six months, one year and two years after TKA as dependent variables separately, and patient characteristics and preoperative PROs as independent variables (n=11)

		Final model					
Dependent variable	**Patients (n)**	**Chi-square (p-value)**	**R^2**	**H-L (p-value)**	**Predictors**	**Odds ratio (95%CI)***	**p-value**
Dissatisfied or satisfied at six months	1151	0.000	0.071	0.273	BMI	0.937 (0.889 to 0.986)	**0.013**
					Medial cartilage damage (1)	5.869 (1.046 to 32.927)	**0.044**
					Charnley score (1)	1.702 (1.107 to 2.619)	**0.015**
					EQ-5D descriptive	0.399 (0.189 to 0.844)	**0.016**
					EQ VAS	0.984 (0.974 to 0.994)	**0.001**
Dissatisfied or satisfied at one year	1216	0.000	0.057	0.115	Medial cartilage damage (4)	0.543 (0.310 to 0.951)	**0.033**
					Charnley score (1)	2.020 (1.286 to 3.174)	**0.002**
					EQ-5D descriptive	0.345 (0.167 to 0.713)	**0.004**
Dissatisfied or satisfied at two years	1146	0.038	0.029	0.320	Medial cartilage damage (2)	2.450 (1.089 to 5.515)	**0.030**

Bold values denote p-values with statistical significance.

95%CI; 95% confidence interval, BMI; body mass index, EQ VAS; EuroQol visual analogue scale, EQ-5D descriptive system; EuroQol 5 Dimensions descriptive system, H-L; Hosmer and Lemeshow goodness of fit test, n; number, PROs; patient-reported outcomes, R^2; explained variance, TKA; total knee arthroplasty

Independent variables (n=11): age, BMI, ASA score, degree of medial cartilage damage, Charnley score, history of previous knee surgery, comorbidity, KOOS-PS score, EQ VAS score, EQ-5D descriptive system score and anxiety.

*An odds ratio (OR) > 1 indicated an increasing risk of being dissatisfied relative to being satisfied.

In line with previous studies, in the present study found median NRS satisfaction score was 9.0 (8.0-10.0) and less than 10% patients were dissatisfied one year after TKA [4,6,8]. Prediction of these satisfaction scores was investigated using previously reported preoperative predictors (Table 1). Lower BMI, lower degree of medial cartilage damage, lower quality of life scores, and having anxiety were found as predictors for lower patient satisfaction scores and/or being dissatisfied. Apart from BMI, these associations are in line with previous studies (Table 1). The somewhat contradictory finding in the present study with respect to BMI can possibly be explained by the fact that only patients with a BMI ≤35 were included. TKA patients with BMI <35 are more satisfied compared with patients with BMI >35 [29]. Complication risk increases with BMI >35 (especially >40), and having a complication is associated with lower patient satisfaction scores [5,15,17]. These patients at the higher BMI end are not in the present study and, therefore, the assessment of BMI as a predictor in the present study cannot be applied to the general TKA population. In the present study, the degree of medial cartilage damage was found to be a predictor of patient satisfaction, and the degree of lateral and/or retropatellar cartilage damage was not (Tables 3 and 4). This means that, with the models correcting for degree of cartilage damage in other parts of the knee joint, patient satisfaction mainly depended on the presence of KL score 4 medial cartilage damage. In depth analyses showed that patients with KL score 4 medial OA were more likely to be satisfied with their TKA, irrespective of the amount of damage elsewhere. Patients with KL score 4 lateral OA but without KL score 4 medial OA were less likely to be satisfied. However, not at a level that can be used to predict outcome preoperatively.

Although several patient characteristics and preoperative PROs were found to be predictors of the degree of patient satisfaction or the chance of being dissatisfied or satisfied, no reliable prediction for patient satisfaction after TKA could be made. All models were weak prediction models (R^2: 0.017 to 0.071) with a bad agreement between the observed and predicted values (Tables 3 and 4, and Fig. 2). Models showed a good prediction of higher degree of satisfaction and the chance of being satisfied (99.7%-100% is correctly predicted) (Fig. 2). However, weak or even no prediction was possible for lower degree of satisfaction and the chance of being dissatisfied (0.0%-1.4% is correctly predicted) (Fig. 2). Although other preoperative predictors were entered into the models, the same findings regarding weak prediction models (R^2: 0.1 to 0.2) for dichotomized patient satisfaction one year after TKA with only preoperative predictors were reported by others [6]. This means that the existing preoperative prediction models are not useful for clinical practice. It might just not be possible at all to preoperatively predict patient satisfaction after TKA.

In the present study, patients who had a complication were less likely to be satisfied with their TKA. Interestingly, the percentage of dissatisfied patients who had a complication 19.7% at six months decreased to 16.8% at 2 years. As stated by others, it might be that preoperatively measured predictors have minimal influence on patient satisfaction compared with postoperatively measured predictors, such as having a complication [4]. A key predictor seems to be fulfillment of expectations, which is a postoperatively measured predictor [30]. No fulfillment of preoperatively set expectations is associated with a 10.7 times higher risk of patient dissatisfaction one year after

TKA [5]. This might explain the higher percentage of dissatisfied patients who had a complication, as having a complication will clearly for most be outside of what is expected. Although this is a postoperative predictor, knowing patient expectations preoperatively could be critical to prevent patient dissatisfaction. While it might be not possible to preoperatively predict patient satisfaction in a manner useful for clinical practice, it instead might be possible to influence preoperatively set possibly unrealistic expectations to improve patient satisfaction after surgery. Further research in this area is needed.

In the present study, a low number of dissatisfied patients (less than 10% of the patients scored 6) and a high number of maximally satisfied patients (almost 40% scored the maximum score of 10) were found one year after TKA. This might make it difficult to use the current models to predict patient satisfaction as there are too few dissatisfied patients to make any distinction. Furthermore, possibly there is a ceiling effect of the NRS satisfaction question, although this has not been investigated yet. More importantly, there is no golden standard for measuring patient satisfaction; heterogeneous methods for measuring patient satisfaction are used [30-32] resulting in a multitude of different predictors [6,20]. So, the question is how reliably exactly patient satisfaction can be measured using current tools and thus predicted. Previous advice to improve the reliability of patient satisfaction measures has been to not use a single question as the focus of the question influences which predictors will be found [6,20,33,34].

Strength of the present study are as follows: the primary outcome was validated by using the ordinal anchor questions, patient satisfaction was investigated on a continuous scale and a dichotomous scale, and there were a high number of patients included (>1200) with few missing values (≤1.6%), and high PROM response rates (>90%). Furthermore, surgery was performed by high-volume surgeons, reducing the risk of revision [35]. No difference in outcomes was found between these surgeons and posterior-stabilized or cruciate-retaining implants [36]. As a limitation, because the low amount of dissatisfied patients, only 11 independent variables were included in the logistics regression analyses instead of the planned 20 variables. Furthermore, only those with ASA scores I or II and BMI ≤35 were included. These characteristics, however, apply to 80% of the total TKA population [37]. Future research should focus on a golden standard for measuring patient satisfaction and the development of tools to properly align patients' expectations.

CONCLUSION

The degree of patient satisfaction and the chance of being dissatisfied or satisfied six months and one and two years after TKA are predictable by patient characteristics and preoperative PROs but not at a reliability level that is clinically useful.

REFERENCES

1. Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. Best Pract Res Clin Rheumatol 2006;20:721e40. https://doi.org/10.1016/j.berh.2006.05.002.
2. Browne JP, Bastaki H, Dawson J. What is the optimal time point to assess patient-reported recovery after hip and knee replacement? A systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. Health Qual Life Outcomes 2013;11:1e7. https://doi.org/10.1186/1477-7525-11-128.
3. Ali A, Lindstrand A, Sundberg M, Flivik G. Preoperative anxiety and depression correlate with dissatisfaction after total knee arthroplasty: a prospective longitudinal cohort study of 186 patients, with 4-year follow-up. J Arthroplasty 2017;32:767e70. https://doi.org/10.1016/j.arth.2016.08.033.
4. Baker PN, Rushton S, Jameson SS, Reed M, Gregg P, Deehan DJ. Patient satisfaction with total knee replacement cannot be predicted from pre-operative variables alone: a cohort study from the national joint registry for England and wales. Bone Joint J 2013;95 B:1359e65. https://doi.org/10.1302/0301-620X.95B10.32281.
5. 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;468:57e63. https://doi.org/10.1007/s11999-009-1119-9.
6. Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ. The rate and predictors of patient satisfaction after total knee arthroplasty are influenced by the focus of the question: a standard satisfaction question is required. Bone Joint J 2018;100B:740e8. https://doi.org/10.1302/0301-620X.100B6.BJJ-2017-1292.R1.
7. Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: patient expectations affect satisfaction with total knee arthroplasty. Clin Orthop Relat Res 2006;452:35e43. https://doi.org/10.1097/01.blo.0000238825.63648.1e.
8. Van Onsem S, Van Der Straeten C, Arnout N, Deprez P, Van Damme G, Victor J. A new prediction model for patient satisfaction after total knee arthroplasty. J Arthroplasty 2016;31:2660e2667.e1. https://doi.org/10.1016/j.arth.2016.06.004.
9. Schnurr C, Jarrous M, Güdden I, Eysel P, Konig DP. Pre-operative arthritis severity as a predictor for total knee arthroplasty patients' satisfaction. Int Orthop 2013;37:1257e61. https://doi.org/10.1007/s00264-013-1862-0.
10. Scott CEH, Howie CR, MacDonald D, Biant LC. Predicting dissatisfaction following total knee replacement: a prospective study of 1217 patients. J Bone Joint Surg Br 2010;92:1253e8. https://doi.org/10.1302/0301-620X.92B9.24394.
11. Hamilton DF, Lane JV, Gaston P, Patton JT, MacDonald D, Simpson AHRW, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. BMJ Open 2013;3:e002525. https://doi.org/10.1136/bmjopen-2012-002525.
12. Lützner C, Postler A, Beyer F, Kirschner S, Lützner J. Fulfillment of expectations influence patient satisfaction 5 years after total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2019;27:2061e70. https://doi.org/10.1007/s00167-018-5320-9.
13. Matsuda S, Kawahara S, Okazaki K, Tashiro Y, Iwamoto Y. Postoperative alignment and ROM affect patient satisfaction after TKA knee. Clin Orthop Relat Res 2013;471:127e33. https://doi.org/10.1007/s11999-012-2533-y.
14. Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the national joint registry for England and Wales. J Bone Joint Surg Br 2007;89:893e900. https://doi.org/10.1302/0301-620X.89B7.19091.
15. Merle-Vincent F, Couris CM, Schott AM, Conrozier T, Piperno M, Mathieu P, et al. Factors predicting patient satisfaction 2 years after total knee arthroplasty for osteoarthritis. Jt Bone Spine 2011;78:383e6. https://doi.org/10.1016/j.jbspin.2010.11.013.
16. Lizaur-Utrilla A, Martinez-Mendez D, Miralles-Munoz FA, Marco-Gomez L, Lopez-Prats FA. Negative impact of waiting time for primary total knee arthroplasty on satisfaction and patient-reported outcome. Int Orthop 2016;40:2303e7. https://doi.org/10.1007/s00264-016-3209-0.

17. Khatib Y, Badge H, Xuan W, Naylor JM, Harris IA. Patient satisfaction and perception of success after total knee arthroplasty are more strongly associated with patient factors and complications than surgical or anaesthetic factors. Knee Surgery. Sport Traumatol Arthrosc 2020;28:3156e63. https:// doi.org/10.1007/s00167-019-05804-9.

18. Kim TK, Chang CB, Kang YG, Kim SJ, Seong SC. Causes and predictors of patient's dissatisfaction after uncomplicated total knee arthroplasty. J Arthroplasty 2009;24:263e71. https://doi.org/10.1016/j.arth.2007.11.005.

19. Kwon SK, Kang YG, Kim SJ, Chang CB, Seong SC, Kim TK. Correlations between commonly used clinical outcome scales and patient satisfaction after total knee arthroplasty. J Arthroplasty 2010;25:1125e30. https://doi.org/10.1016/j.arth.2009.10.015.

20. Barrack RL, Ruh EL, Chen J, Lombardi AV, Berend KR, Parvizi J, et al. Impact of socioeconomic factors on outcome of total knee arthroplasty knee. Clin Orthop Relat Res 2014;472:86e97. https://doi.org/10.1007/s11999-013-3002-y.

21. Fisher DA, Dierckman B, Watts MR, Davis K. Looks good but feels bad: factors that contribute to poor results after total knee arthroplasty. J Arthroplasty 2007;22:39e42. https://doi.org/10.1016/j.arth.2007.04.011.

22. Kamper SJ, Ostelo RWJG, Knol DL, Maher CG, de Vet HCW, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol 2010;63:760e766.e1. https://doi.org/10.1016/j.jclinepi.2009.09.009.

23. Haverkamp D, Breugem SJ, Sierevelt IN, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. Acta Orthop 2005;76:347e52.

24. Perruccio AV, Stefan Lohmander L, Canizares M, Tennant A, Hawker GA, Conaghan PG, 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 Cartil 2008;16:542e50. https://doi.org/10.1016/j.joca.2007.12.014.

25. The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy (New York) 1990;16:199e208. https://doi.org/10.1016/0168-8510(90)90421-9.

26. Nederlandse Orthopaedische Vereniging. PROMs. https://www.lroi.nl/invoerders/registreren/proms [accessed 01.08.18].

27. Pronk Y, Pilot P, Brinkman JM, van Heerwaarden RJ, van der Weegen W. Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection. J Patient Rep Outcomes 2019;3:1e8. https://doi.org/10.1186/s41687-019-0121-6.

28. Sociaal en Cultureel Planbureau. Sociaal-Economische Status per postcodegebied. https://bronnen.zorggegevens.nl/Bron?naam¼SociaalEconomische-Status-per-postcodegebied; 2017 [accessed 01.02.19].

29. Baker P, Muthumayandi K, Gerrand C, Kleim B, Bettinson K, Deehan D. Influence of body mass index (BMI) on functional improvements at 3 Years following total knee replacement: a retrospective cohort study. PLoS One 2013;8:e59079. https://doi.org/10.1371/journal.pone.0059079.

30. Kahlenberg CA, Nwachukwu BU, McLawhorn AS, Cross MB, Cornell CN, Padgett DE. Patient satisfaction after total knee replacement: a systematic review. HSS J 2018;14:192e201. https://doi.org/10.1007/s11420-018-9614-8.

31. 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;32:3854e60. https://doi.org/10.1016/j.arth.2017.07.021.

32. Gibon E, Goodman MJ, Goodman SB. Patient satisfaction after total knee arthroplasty: a realistic or imaginary goal? Orthop Clin North Am 2017;48:421e31. https://doi.org/10.1016/j.ocl.2017.06.001.

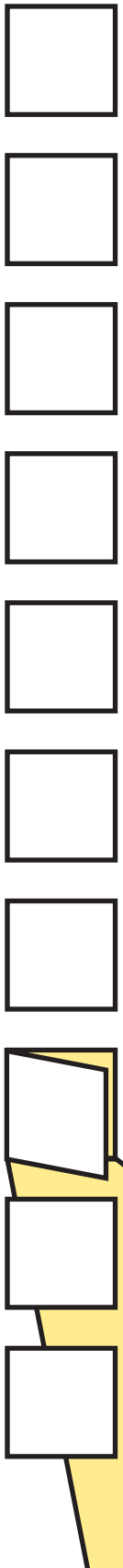
33. Woolhead GM, Donovan JL, Dieppe PA. Outcomes of total knee replacement: a qualitative study. Rheumatology 2005;44:1032e7. https://doi.org/10.1093/rheumatology/keh674.

34. Loth FL, Giesinger JM, Giesinger K, Howie CR, Hamilton DF. Single-item satisfaction scores mask large variations in pain, function and joint awareness in patients following total joint arthroplasty. Eur J Orthop Surg Traumatol 2020;30:267e74. https://doi.org/10.1007/s00590-019-02560-x.

35. Liddle AD, Pandit H, Judge A, Murray DW. Effect of surgical caseload on revision rate following total and unicompartmental knee replacement. J Bone Joint Surg Am 2016;98:1e8. https://doi.org/10.2106/JBJS.N.00487.

36. Mayne A, Harshavardhan HP, Johnston LR, Wang W, Jariwala A. Cruciate Retaining compared with Posterior Stabilised Nexgen total knee arthroplasty: results at 10 years in a matched cohort. Ann R Coll Surg Engl 2017;99:602e6. https://doi.org/10.1308/rcsann.2017.0086.

37. LROI. Patient characteristics of all patients with a registered primary knee arthroplasty by type of knee arthroplasty in The Netherlands in. https://www.lroi-rapportage.nl/knee-primary-knee-arthroplasty-demographics-patientcharacteristics-by-type-of-knee-prosthesis19; 2018 [accessed 12.03.20].



CHAPTER 8

Effectiveness of a mobile eHealth app in guiding patients in pain control and opiate use after total knee replacement: randomized controlled trial

Y. Pronk, M.C.W.M. Peters, A. Sheombar and J.M. Brinkman

JMIR Mhealth Uhealth. 2020 Mar 13;8(3):e16415

ABSTRACT

Background

Little is known about pain and opiate use at home directly after total knee replacement (TKR). Due to adverse effects, low opiate use is desired. An electronic health app (PainCoach) was developed to guide patients in pain control and opiate use.

Objective

The aim of this paper was to investigate the effects of the PainCoach app on pain control and opiate use in patients who underwent TKR during the first 2 weeks at home after surgery.

Methods

In an unblinded randomized controlled trial, patients scheduled for TKR were offline recruited and randomized to a PainCoach group or control group. In the PainCoach group, the PainCoach app was downloaded on each patient's smartphone or tablet. In response to the patient's input of the pain experienced, the PainCoach app gave advice on pain medication use, exercises/rest, and when to call the clinic. This advice was the same as that received during usual care. The control group received usual care. The primary outcomes were opiate use and visual analog scale (VAS) pain scores at rest, during activity, and at night during the first 2 weeks at home after surgery, which were collected daily from day 1 until 14 postoperatively by online questionnaires. The actual amount of app use was recorded, and active use was defined as ≥12 total app uses.

Results

The pain scores did not differ between the groups. The PainCoach group (n=38) used 23.2% less opiates (95% CI −38.3 to −4.4; P=.02) and 14.6% more acetaminophen (95% CI 8.2-21.3; P<.001) when compared with the findings in the control group (n=33). The PainCoach app was used 12 (IQR 4.5-22.0) times per patient. In the active PainCoach subgroup (n=19), the following were noted when compared with the findings in the control group: 4.1 times faster reduction of the VAS pain score during activity (95% CI −7.5 to −0.8; P=.02), 6.3 times faster reduction of the VAS pain score at night (95% CI −10.1 to −2.6; P=.001), 44.3% less opiate use (95% CI −59.4 to −23.5; P<.001), 76.3% less gabapentin use (95% CI −86.0 to −59.8; P<.001), and 21.0% more acetaminophen use (95% CI 12.6-30.0; P<.001).

Conclusions

The use of the PainCoach app contributes to reduced opiate use in the initial period at home after TKR. Active use of this app leads to a further reduction in opiate use and improved pain control.

Trial Registration

ClinicalTrials.gov NCT03961152; https://clinicaltrials.gov/ct2/show/NCT03961152

INTRODUCTION

Total knee replacement (TKR) is a successful treatment option for patients with end-stage knee osteoarthritis (OA) [1]. Moderate-to-severe pain after TKR can be expected [2,3]. Local infiltration anesthesia (LIA) techniques and so-called fast-track recovery programs have resulted in reduced pain and early mobilization, subsequently reducing the length of stay in hospital and increasing patient satisfaction [4-7]. Previous research established several factors associated with increased pain after TKR [8-19] (Table 1). Postoperative pain inhibits recovery, increases morbidity, and may result in chronic pain, ultimately limiting the effectiveness of TKR [6,20]. Therefore, pain should be controlled optimally both in the hospital and at home.

Although pain is usually under control during hospital stay, less is known about pain control in the initial period at home after TKR. Current pain management strategies include a combination of nonsteroidal anti-inflammatory drugs (NSAIDs), nonnarcotic medication, opiates, and exercise [4]. Although opiates are very effective for reducing pain, serious adverse effects, such as nausea, itching, reduced gut mobility, and urinary retention, often occur [21]. Addiction to opiates is an ever increasing problem and may ultimately lead to an increased risk of death [22]. The amount of opiate use should therefore be kept to a minimum. Orthopedic surgery, however, accounts for an estimated 8.8% of prolonged prescription opiate use [23]. Therefore, alternative pain management strategies are needed. Electronic health (eHealth) apps can be used to guide patients in improving their pain management strategies at home. An important benefit of these apps is that patients can access the information provided directly and anywhere whenever necessary [24-29]. The number of older adults with internet access and acceptance of internet-based interventions is increasing, and patients tend to remember up to 80% of the information acquired from interactive education [30,31].

With this in mind, to manage pain better and potentially decrease opiate use, an eHealth app named PainCoach was developed. This app aims to help patients control their pain better in the initial period at home after TKR, including optimal use of the available pain medication. This study aimed to determine the effects of PainCoach on pain control and opiate use in TKR patients in the first 2 weeks at home after surgery. The hypothesis was that the use of this app would decrease pain and opiate use.

Table 1 Factors associated with increased pain after total knee replacement

Factor	Association with increased pain after TKR[a]
Gender	Being female [8-12]
Age	Older age [8,10,13]
BMI[b]	Higher BMI[b] [8,10][a]
ASA[c] score	Higher ASA[c] score [10]
Pain catastrophization	Higher pain catastrophization score [12,14-17]
Comorbidity	Presence of comorbidities [8,10,13,18]
Previous knee surgery	Having a history of previous knee surgery [10]
Preoperative pain	Higher preoperative pain severity [8,12,18,19]
Social support	Poor social support [13]
Preoperative mental health	Poor preoperative mental health [8,10,13,18]

[a]TKR: total knee replacement. [b]BMI: body mass index. [c]ASA: American Society of Anaesthesiologists score

METHODS

Study design

An unblinded, randomized, controlled, single-center trial was performed at Kliniek ViaSana (Mill, The Netherlands). Patients with an American Society of Anesthesiologists (ASA) score of I-II, a body mass index (BMI) of ≤35, and a plan to undergo primary TKR between February and June 2016 were enrolled. Four experienced high-volume knee surgeons performed all surgeries, and three experienced anesthesiologists administered spinal anesthesia. The same type of TKR implant was used in all patients (NexGen LPS, ZimmerBiomet, Warsaw, Indiana). All surgeries were performed using a tourniquet. The pain management protocol consisted of preoperatively administered medication, LIA injections during surgery directly before cementing the implant, and a step-wise postoperative pain management protocol (Multimedia Appendix 1). Patients were excluded if they did not possess a smartphone or tablet, had a contraindication to any of the medications used in the study, did not have an email address, did not have internet at home, did not have a thorough command of the Dutch language, had memory disorders, or had surgery under general anesthesia. Patients were recruited over the phone by the research staff after being scheduled for primary TKR under spinal anesthesia, and contraindication to any of the medications used in the study and presence of memory disorders were checked by the anesthesiologists. Patients were asked over the phone if they possessed a smartphone or tablet, had an email address, had internet at home, and had a thorough command of the Dutch language. Patient information and informed consent were sent by postal service if a patient met the criteria and was interested to participate. Patients were considered lost to follow-up if they completed less than two postoperative questionnaires during the first 2 weeks at home. Power analysis (significance level: .05, power: 90%) showed that 35 patients would be needed in each

group to detect a difference of 10 points on a visual analog scale (VAS) for pain (VAS pain, 0-100). Written informed consent was obtained from all participants. The study was approved by the medical ethics committee of St. Anna Hospital (Geldrop, The Netherlands, Study ID: 5.12) and was registered at Clinicaltrials.gov retrospectively (ID: NCT03961152).

Randomization

Included unblinded patients were randomly assigned to the PainCoach or control group using lots presented in sealed opaque envelopes during admission. All lots were created and sealed by a researcher in the ratio of 1:1. A blinded nurse presented the envelops to a patient, and the patient selected one to complete randomization. All patients received the usual pain management care including pre-, peri-, and postoperative pain medication (Multimedia Appendix 1), participated in group information meetings, received an information booklet, and could contact the clinic at any time (24 hours a day/7 days a week) in case of any remaining questions. In the PainCoach group, in addition to receiving the aforementioned usual care, the PainCoach app (Interactive Studios, Rosmalen, The Netherlands) was downloaded on each patient's smartphone or tablet, using a unique download code. In this way, the PainCoach app was not available to the control group. An unblinded nurse provided the code and assisted the patient by completing the download process of the app during admission. The app gave the same advice as that during usual care. After only entering the date of surgery as patient data, the app allowed patients to input their pain level (no pain, bearable pain, unbearable pain, or untenable pain) whenever they wanted until day 14 after surgery. Based on the patient's input and taking into account the number of days after surgery, the app provided advice on pain medication use, physiotherapy exercises including videos, use of ice or heat packs, rest, immobilization of the operated leg, and when to call the clinic (Multimedia Appendix 2). Patients in the PainCoach group were not subjected to any treatment that was different from that in the control group (ie, advice on pain management was delivered in an extra and different way, but the pain medication itself was exactly the same for both groups). During the study, no major changes or revisions were made to the PainCoach app.

Outcomes and measurements

Beside the actual amount of app use, all the outcome measurements were assessed using a digital, online, automated collection system (OnlinePROMs, Interactive Studios, Rosmalen, The Netherlands), which automatically sent an invitation by email to complete an online questionnaire preoperatively, daily from day 1 to 14, and at 1 month postoperatively. In case of nonresponse to the preoperative or 1-month questionnaire, an automatic reminder was sent after 3 days. The invitation to complete the daily questionnaire was sent at 5 pm, and patients had access to the questionnaire until midnight.

The primary outcomes were opiate use and pain score of the operated knee at rest, during activity, and at night in the first 2 weeks at home after TKR. The pain score was measured on a VAS for pain, which ranged from 0 (no pain) to 100 (worst imaginable pain), preoperatively, daily from day 1 to 14, and at 1 month postoperatively [32-35]. Severe pain was defined as a VAS pain score

from 70 to 100. Opiate (oxycodon; 5 mg per tablet; different manufacturers) use was recorded in quantities per 24 hours from day 1 to 14.

The secondary outcomes in the first 2 weeks at home and 1 month after TKR included other pain medication use (ie, NSAIDs [diclofenac], acetaminophen, or gabapentin; different manufacturers), which was also recorded in quantities per 24 hours from day 1 to 14. Additionally, pain acceptance at rest, during activity, and at night was assessed with a happy smiley (acceptable pain) and a sad smiley (unacceptable pain) preoperatively, daily from day 1 to 14, and at 1 month postoperatively. Experiences with the executed recommended physiotherapy exercises were recorded daily from day 1 to 14 on a 3-item scale (did too much, exactly enough, or could have done more exercises). Moreover, function and quality of life were measured preoperatively and 1 month postoperatively. Knee function was assessed using the Knee Injury and Osteoarthritis Outcome Score–Physical Function Short-form (KOOS-PS) on a scale from 0 (no difficulty) to 100 (extreme difficulty) [36]. The Oxford Knee Score was used to measure combined function and pain on a scale from 0 (most severe symptoms) to 48 (least severe symptoms) [37]. Quality of life was measured using the EuroQol-5 Dimensions (EQ-5D) 3-level version (EQ-5D-3L) questionnaire consisting of the following two scores: EQ VAS score, which is assessed on a scale from 0 (worst imaginable health state) to 100 (best imaginable health state), and EQ-5D descriptive system [38]. The PainCoach app's perceived effectiveness (usability, added value, and likelihood of being recommended to others) was recorded on a 5-item scale ranging from totally agree to totally disagree at day 14 after surgery. Each downloaded app had its own app code that was used to record the actual amount of app use. As the admission period was generally 1 or 2 days, outcomes were measured until day 14 after surgery, and outcomes at home were investigated, the outcome active PainCoach app use was defined as using the app at least 12 times in total.

Preoperative opiate and other pain medication use, age, gender, ASA score, BMI, preoperative comorbidities, history of knee surgery on the same side, Charnley score, date of surgery, date of discharge, and complication data were collected from the electronic patient records. Pain coping, anxiety, education level, and marital status were determined preoperatively using an online questionnaire. Pain coping was measured using the pain coping and cognition list scored from 1 (totally disagree) to 6 (totally agree), and it had the following four categories: catastrophizing, pain coping, internal pain management, and external pain management [39].

Statistical analysis

Analysis was performed using SPSS version 25.0 (IBM Corp, Armonk, New York). All measured outcomes from day 1 until day 14 after surgery were recoded into measured outcomes for days at home by subtraction of the admission period. Patient characteristics were analyzed using descriptive statistics, and data were checked for normal distribution. Differences in mean, median, or percentage were tested using the independent two-sample t-test, Mann-Whitney U test, likelihood analysis, Fisher's test, or Pearson's chi-squared test, depending on the type of data. Mixed linear models were used to analyze the overall rate of decrease or increase for continuous data, and generalized linear models were used to analyze the percentage decrease or increase

for count and nominal data. Additional analysis was performed to compare the active PainCoach subgroup with the control group, with correction for differences in preoperative data. Statistical significance was set at P<.05, and trends were defined as .05<P<.10.

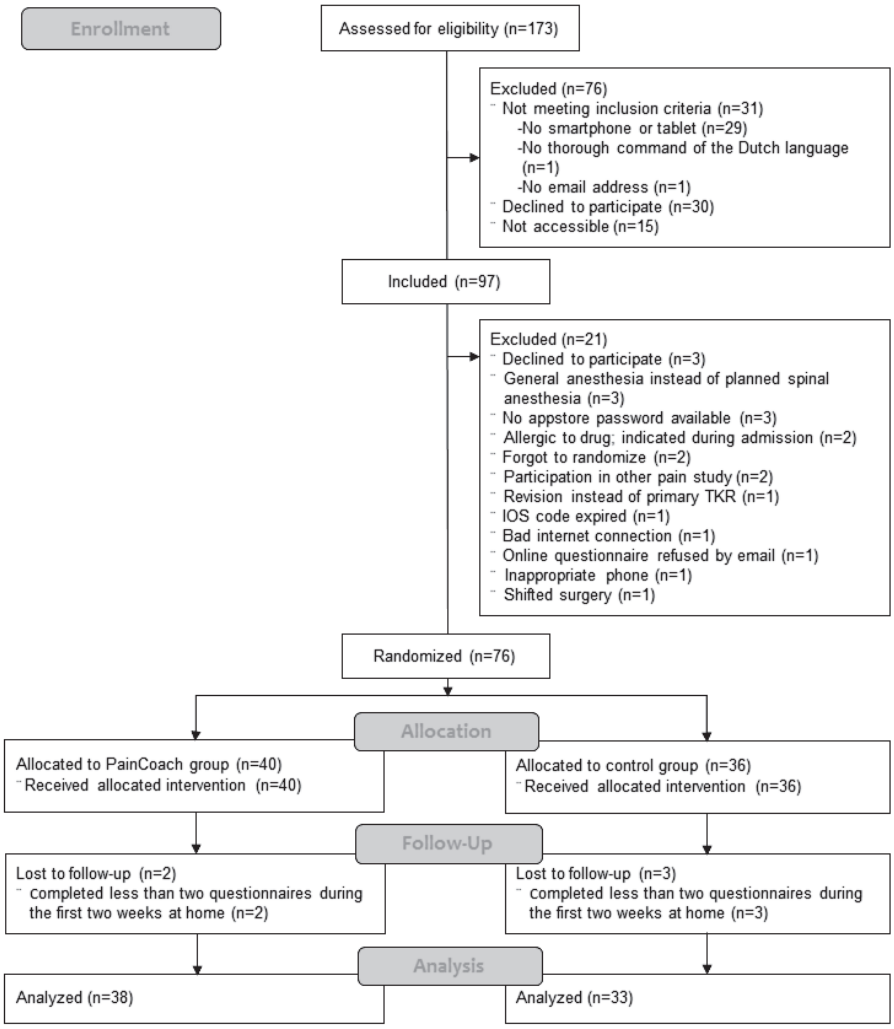


Fig. 1 Study flowchart

IOS: iPhone operating system; TKR: total knee replacement

RESULTS

Patient characteristics

A total of 97 patients were included, and of these, 76 patients were randomized. Because of loss to follow-up, the final analysis was performed with 71 patients (PainCoach group, n=38; control group, n=33) (Figure 1). The response rates for the daily questionnaires at home were 91% in the PainCoach group and 89% in the control group.

No statistically significant differences in patient characteristics were found between the PainCoach group and control group. The preoperative VAS pain score at night was significantly lower in the active PainCoach subgroup (n=19) than in the control group (P=.02) (Table 2).

Visual analog scale pain scores and opiate use

During the first 2 weeks at home, the PainCoach group had VAS pain scores of 17.0 (IQR 5.0-30.0) at rest, 20.0 (IQR 7.0-35.0) during activity, and 17.0 (IQR 4.0-37.0) at night. The control group had VAS pain scores of 20.0 (IQR 7.0-33.0) at rest, 21.0 (IQR 10.0-38.0) during activity, and 20.5 (IQR 8.0-40.0) at night. Pain was classified as severe on one or more days in 21% (8/38) of patients from the PainCoach group and 30% (10/33) of patients from the control group. No statistically significant differences were found between the two groups in terms of the VAS pain scores at rest, during activity, and at night (Figure 2A-C, Table 3). Regarding opiate use, the PainCoach group used a mean of 0.4 (SD 0.7) tablets a day and the control group used a mean of 0.5 (SD 0.8) tablets a day. Opiate use was significantly reduced by 23.2% in the PainCoach group when compared with the finding in the control group (95% CI −38.3 to −4.4; P=.02) (Figure 2A-C, Table 3). One month after surgery, no statistically significant differences in the VAS pain scores were found between the PainCoach group and control group (Table 4).

Table 2 Characteristics of patients the PainCoach group, active PainCoach subgroup, and control group

Characteristic	**1. PainCoach** (n=38)	**2. Active PainCoach** (n=19)	**3. Control** (n=33)	***P* value** (1 vs 3)	***P* value** (2 vs 3)
Gender (male), n (%)	23 (61)	13 (68)	19 (58)	.80	.44
Age (years), mean (SD)	62.6 (7.0)	62.8 (6.1)	64.6 (7.5)	.24	.38
BMI[a], mean (SD)	27.6 (3.5)	26.7 (3.4)	27.8 (3.0)	.83	.24
ASA[b] (I), n (%)	18 (47)	11 (58)	12 (36)	.35	.13
Preoperative comorbidities, n (%)	14 (37)	8 (42)	17 (52)	.21	.51
Preoperative prescription, n (%)					
NSAIDs[c]	5 (13)	3 (16)	6 (18)	.20	.48
Acetaminophen	1 (3)	0 (0)	1 (3)	.92	>.99
Opiate	3 (8)	0 (0)	0 (0)	.24	>.99
Gabapentin	0 (0)	0 (0)	0 (0)	>.99	>.99
Preoperative anxiety, n (%)				.59	>.99
No anxiety	33 (87)	18 (95)	30 (91)		
Some anxiety	5 (13)	1 (5)	3 (9)		
Much anxiety	0 (0)	0 (0)	0 (0)		
History of previous knee surgery same side, n (%)	27 (71)	15 (79)	21 (64)	.51	.25
Charnley score, n (%)				.64	.98
One knee affected with OA[d]	22 (58)	12 (63)	19 (58)		
Both knees affected with OA	7 (18)	3 (16)	6 (18)		
Contralateral TKR[e]	5 (13)	1 (5)	2 (6)		
Multiple joints affected with OA	4 (11)	3 (16)	6 (18)		
Education level, n (%)				.33	.30
Primary school	3 (8)	1 (5)	1 (3)		
Secondary school	14 (37)	7 (37)	10 (31)		
Tertiary school	21 (55)	11 (58)	21 (66)		
Marital status, n (%)				.19	.09
Married	29 (76)	18 (95)	22 (67)		
Other[f]	9 (24)	1 (5)	11 (33)		
Pain coping, mean (SD)					
Catastrophization	2.5 (0.7)	2.5 (0.8)	2.3 (0.6)	.15	.32
Pain coping	3.6 (1.0)	3.8 (1.0)	3.7 (0.8)	.68	.66
Internal pain management	4.1 (0.8)	4.2 (0.9)	3.9 (0.8)	.24	.29
External pain management	2.7 (0.8)	2.7 (0.7)	2.5 (0.8)	.31	.36

Table 2 Characteristics of patients the PainCoach group, active PainCoach subgroup, and control group (continued)

Characteristic	1. PainCoach (n=38)	2. Active PainCoach (n=19)	3. Control (n=33)	*P* value (1 vs 3)	*P* value (2 vs 3)
Preoperative VAS[g] pain, median (IQR[h])					
Knee at rest	33.0 (20.8-52.8)	33.0 (13.0-43.0)	32.0 (17.8-49.0)	.65	.82
Knee during active	60.5 (36.5-77.3)	57.0 (30.0-75.0)	60.0 (43.3-73.8)	.82	.69
Knee at night	20.5 (4.8-42.5)	15.0 (1.0-30.0)	35.5 (15.0-58.5)	.11	.02[i]
Preoperative acceptable pain, n (%)					
Knee at rest	29 (76)	16 (84)	27 (82)	.40	>.99
Knee during active	16 (42)	10 (53)	13 (39)	.90	.41
Knee at night	28 (74)	16 (84)	27 (82)	.28	>.99
Preoperative KOOS-PS[j], median (IQR)	47.3 (41.6-55.3)	46.1 (40.3-54.4)	48.5 (40.3-57.9)	.76	.96
Preoperative OKS[k], mean (SD)	25.3 (7.2)	27.0 (7.2)	24.8 (5.6)	.75	.23
Preoperative EQ-5D[l] descriptive system, median (IQR)	0.775 (0.471-0.783)	0.775 (0.516-0.807)	0.775 (0.651-0.807)	.27	.81
Preoperative EQ VAS[m], median (IQR)	86.0 (73.6-94.3)	87.0 (79.0-93.0)	86.0 (74.0-95.5)	.89	.72
Complications, n (%)	3 (8)	2 (11)	1 (3)	.62	.55

[a]BMI: body mass index. [b]ASA: American Society of Anesthesiologists. [c]NSAIDs: nonsteroidal anti-inflammatory drugs. [d]OA: osteoarthritis. [e]TKR: total knee replacement. [f]Other marital status: single, living together, divorced, widow(er), living apart together relationship, different. [g]VAS: visual analog scale. [h]IQR: interquartile range. [i]Significant difference (P<.05). [j]KOOS-PS: Knee Injury and Osteoarthritis Outcome Score–Physical Function Short-form. [k]OKS: Oxford Knee Score. [l]EQ-5D: EuroQol-5 Dimensions. [m]EQ VAS: EuroQol visual analog scale

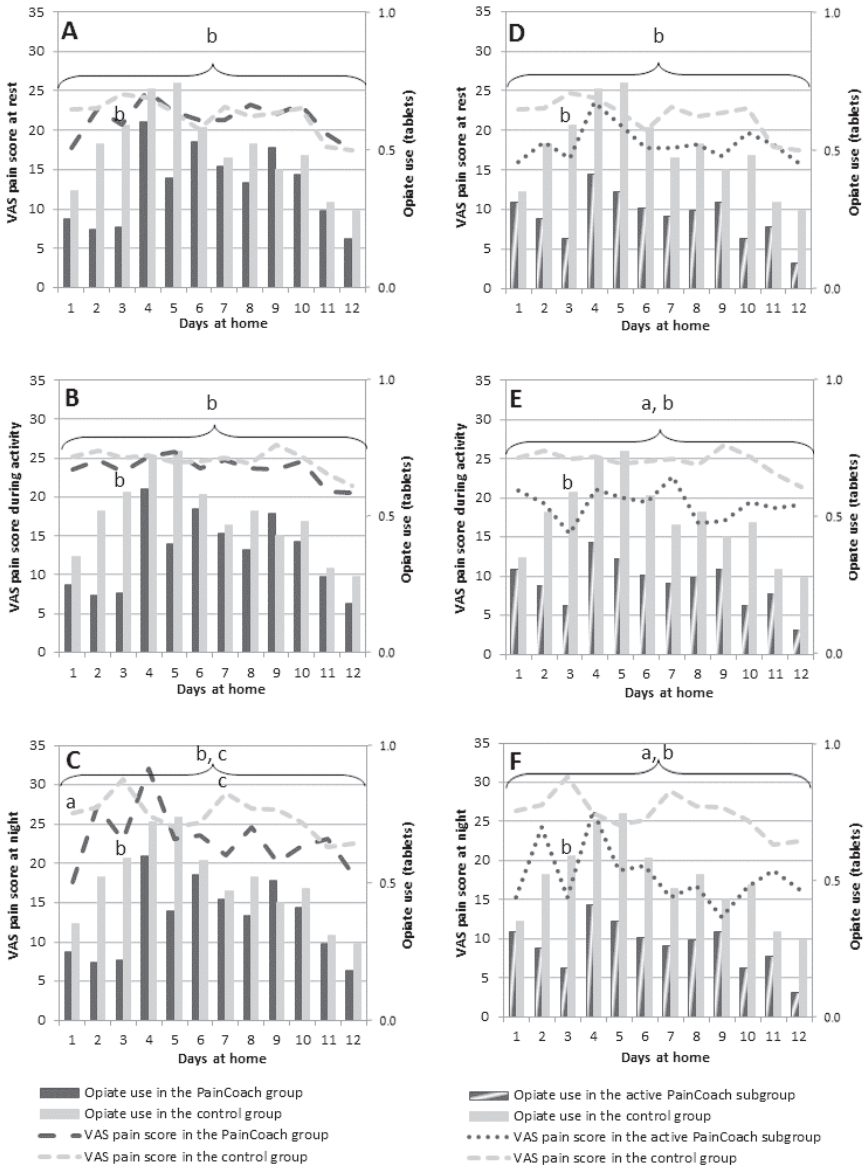


Fig. 2 VAS pain scores and opiate use in the PainCoach group and control group at rest (A), during activity (B), and at night (C) and in the active PainCoach subgroup and control group at rest (D), during activity (E), and at night (F) on separate days and in the overall first period at home.

a: significant difference in VAS pain (P<.05); b: significant difference in opiate use (P<.05); c: trend in VAS pain (.05<P<.10); VAS: visual analog scale

Table 3 Findings in the PainCoach group, active PainCoach subgroup, and control group during the first 2 weeks at home

Variable	1. PainCoach versus control	2. Active PainCoach versus control	*P* value (1)	*P* value (2)
VAS[a] pain, decrease or increase (rate)				
Knee at rest	↓ 0.3	↓ 1.9	.86	.27
Knee during active	↓ 1.0	↓ 4.1	.48	.02[b]
Knee at night	↓ 3.0	↓ 6.3	.06[c]	<.001[b]
Medication use, decrease or increase (%)				
Opiate	↓ 23.2	↓ 44.3	.02[b]	<.001[b]
NSAIDs[d]	↓ 9.2	↓ 12.8	.08[c]	.06[c]
Acetaminophen	↑ 14.6	↑ 21.0	<.001[b]	<.001[b]
Gabapentin	↑ 4.6	↓ 76.3	.71	<.001[b]
Acceptable pain, decrease or increase (%)				
Knee at rest	↓ 31.3	↓ 20.3	.11	.25
Knee during active	↓ 17.2	↑ 31.1	.40	.38
Knee at night	↓ 21.1	↑ 36.4	.21	.25
Experience with the executed recommended exercises - exactly enough, decrease or increase (%)	↓ 33.1	↓ 8.7	.02[b]	.67

[a]VAS: visual analog scale. [b]Significant difference (P<.05). [c]Trend (.05<P<.10). [d]NSAIDs: nonsteroidal anti-inflammatory drugs

Adjusted analyses showed that the active PainCoach subgroup had VAS pain scores of 10.0 (IQR 4.0-26.3) at rest, 12.0 (IQR 5.0-25.0) during activity, and 10.0 (IQR 2.8-28.0) at night during the first 2 weeks at home. Pain was reported as severe on one or more days in 16% (3/19) of patients from the active PainCoach subgroup. The VAS pain score during activity significantly decreased 4.1 times faster in the active PainCoach subgroup when compared with the finding in the control group (95% CI −7.5 to −0.8; P=.02) (Figure 2E, Table 3). The VAS pain score at night significantly decreased 6.3 times faster in the active PainCoach subgroup when compared with the finding in the control group (95% CI −10.1 to −2.6; P=.001) (Figure 2F, Table 3). The mean opiate use was 0.3 (SD 0.5) tablets a day in the active PainCoach subgroup. Opiate use was significantly reduced by 44.3% in the active PainCoach subgroup when compared with the finding in the control group (95% CI −59.4 to −23.5; P<.001) (Figure 2D-F, Table 3). One month after surgery, no statistically significant differences in VAS pain scores were found between the active PainCoach subgroup and control group (Table 4).

Table 4 Findings in the PainCoach group, active PainCoach subgroup, and control group 1 month after surgery

	1. PainCoach (n=38)	2. Active PainCoach (n=19)	3. Control (n=33)	*P* value (1vs3)	*P* value (2vs3)
VAS[a] pain, median (IQR[b])					
Knee at rest	11.5 (5.0-20.8)	11.5 (4.3-18.8)	10.0 (5.0-25.0)	.77	.53
Knee during active	14.0 (7.0-28.8)	12.5 (9.3-26.3)	15.0 (8.0-35.0)	.49	.59
Knee at night	15.0 (7.0-33.0)	15.0 (5.0-33.0)	15.0 (7.0-27.8)	.79	.89
Acceptable pain, n (%)					
Knee at rest	31 (96.9)	16 (100.0)	28 (96.6)	>.99	>.99
Knee during active	30 (93.8)	15 (93.8)	25 (86.2)	.41	.64
Knee at night	26 (81.3)	14 (87.5)	26 (89.7)	.48	>.99
KOOS-PS[c], mean (SD)	36.5 (10.5)	33.5 (8.4)	39.6 (9.8)	.24	.04[d]
OKS[e], mean (SD)	28.4 (8.4)	29.9 (9.1)	26.8 (6.2)	.42	.18
EQ-5D[f] descriptive system, median (IQR)	0.775 (0.693-0.843)	0.811 (0.775-0.857)	0.775 (0.651-0.811)	.34	.11
EQ VAS[g], median (IQR)	80.0 (70.0-90.0)	83.5 (70.0-90.0)	80.0 (65.5-89.5)	.56	.32

[a]VAS: visual analog scale. [b]IQR: interquartile range. [c]KOOS-PS: Knee Injury and Osteoarthritis Outcome Score–Physical Function Short-form. [d]Significant difference (P<.05). [e]OKS: Oxford Knee Score. [f]EQ-5D: EuroQol-5 Dimensions. [g]EQ VAS: EuroQol visual analog scale

Other pain medication use, pain acceptance, and experience with executed recommended exercises

In the PainCoach group, there was a statistically significant 14.6% increase in acetaminophen use (95% CI 8.2-21.3; P<.001) and no statistically significant differences in NSAID use and gabapentin use when compared with the findings in the control group during the first 2 weeks at home (Table 3). Overall pain medication use was below the advised maximum in both groups. Pain acceptance was 86.5% at rest, 86.5% during activity, and 79.4% at night in the PainCoach group and was 90.4% at rest, 88.6% during activity, and 83.0% at night in the control group, without statistically significant differences between the two groups. Regarding experience with executing recommended exercises, the PainCoach group had statistically significant 33.1% reduced experience with executing exactly enough exercises when compared with the findings in the control group (69.7% vs. 77.5%; 95% CI −52.0 to −6.7; P=.02) (Table 3). At 1 month after surgery, no statistically significant differences were found when comparing both groups (Table 4).

Adjusted analyses comparing the active PainCoach subgroup with the control group showed statistically significant 21.0% increased acetaminophen use in the active PainCoach subgroup (95% CI 12.6-30.0; P<.001) during the first 2 weeks at home. Additionally, the active PainCoach subgroup had statistically significant 76.3% decreased gabapentin use when compared with the findings in the control group (mean 0.1 [SD 0.3] tablets a day vs. 0.4 [SD 1.0] tablets a day;

95% CI −86.0 to −59.8; P<.001) (Table 3). In the active PainCoach subgroup, pain acceptance was 88.4% at rest, 90.9% during activity, and 87.4% at night. Regarding pain acceptance and experience with executing recommended exercises, no statistically significant differences were found between the active PainCoach subgroup and control group (Table 3). One month after surgery, the mean KOOS-PS was significantly lower in the active PainCoach subgroup (33.5 [SD 8.4]) than in the control group (39.6 [SD 9.8]) (P=.048) (Table 4).

PainCoach app use

Among 28 patients who provided appropriate responses, 25 (89%) reported ease of app use, 22 (79%) found that the app added value, and 22 (79%) would recommend the app to friends and family. The PainCoach app was used 12 (IQR 4.5-22.0) times per patient on 7 (IQR 4.0-9.0) days at home. The number of patients with at least one entry in the PainCoach app ranged from 11 (30%) to 26 (70%) per day at home (Multimedia Appendix 2). The app was most frequently used between 9 and 10 am and mostly for advice on bearable pain.

DISCUSSION

Principal findings

This study aimed to determine the effects of an eHealth app, the PainCoach app, on pain control and opiate use in patients who underwent TKR during the first 2 weeks at home after surgery. The hypothesis was that the app would decrease pain and opiate use. As indicated by the main findings, there was no statistically significant difference in pain scores between the two groups and opiate use was significantly reduced by 23.2% in the PainCoach group when compared with the finding in the control group. In the active PainCoach subgroup, however, pain during activity and at night significantly decreased 4.1 and 6.3 times faster, respectively, and opiate use significantly reduced by 44.3% when compared with the findings in the control group.

Overall, low pain scores and high levels of pain acceptance were found in this study. Only 21% (8/38) of patients in the PainCoach group and 30% (10/33) in the control group classified their pain as severe during one or more days at home. Other studies have stated that the most painful period after TKR surgery was the initial period at home, with 23%-30% of patients rating their average pain as severe [40,41]. Aside from the use of modern LIA techniques and a step-wise pain management protocol postoperatively, a possible explanation for the reported low pain and high acceptance scores in this study could be the guidance program that was provided to all patients who underwent TKR in Kliniek ViaSana. As less anxiety is associated with lower pain scores [14,19], the guidance provided might have resulted in less anxiety and therefore lower pain scores. The reported overall low pain scores also probably explain why no difference in pain scores was found between the PainCoach group and control group. Although overall pain scores were low, active use of the PainCoach app resulted in even lower pain scores during activity and at night when compared with the findings in the control group. These findings are in line with the results of a previous study showing that pain decreased by 0.7 points on a scale from 0 to 10 in

patients with OA after online "pain coping skills" training [29]. Others have stated that 80% of interactive information is remembered compared with 20% of auditory information and 40% of read information [30,42,43]. As the PainCoach app is an interactive tool, it is logical that active use will result in better use of the pain management strategies provided and subsequently lower pain scores.

Opiate addiction caused 74 deaths in the Netherlands in 2016, and this number is increasing each year [44]. Using the PainCoach app, opiate use reduced by 23.2%, and active PainCoach app use resulted in a further reduction (44.3%). Because of a lack of standardized opiate prescribing protocols in orthopedic surgery, it is difficult to compare the reported amount of opiate use in this study with that in other studies. In one available study, a daily average morphine dose at discharge of 155 (SD 63) mg was prescribed to patients who underwent TKR, which would be the equivalent of 11 tablets per day of the opiate used in this study (oxycodon, 5 mg per tablet) and is far above the average use of 0.4 opiate tablets per day in this study [45]. The low preoperative opiate use of patients in this study might have contributed to the low opiate use after surgery, as preoperative opiate use is a strong predictor for prolonged opiate use after TKR [42,46,47]. With lower opiate use, acetaminophen use was higher, with a 14.6% increase in the PainCoach group and 21.0% increase in the active PainCoach subgroup. It can be concluded that because of the advice provided by the PainCoach app, opiate use was substituted by acetaminophen use. Opiate use was only advised in the presence of severe enough reported pain in the app. Therefore, it is concluded that the app helps to reduce the risk of the adverse effects of opiate use [48,49].

A shorter hospital stay is associated with a higher burden among patients, who need to take responsibility for aftercare shortly after surgery. Recent studies have shown that patients feel uncertain and left alone after discharge, which could increase anxiety and affect their pain coping and subsequent management [50,51]. Patients might need more individualized guidance, and the PainCoach app was developed to satisfy this need. The app scored high on usability, likelihood of being recommended to others, and added value. The results of this study show that the PainCoach app is a successful pain management tool, and its active use is recommended for the best effects on pain and opiate use.

To our knowledge, this is the first randomized controlled trial to examine the effects of eHealth with regard to controlling pain and reducing opiate use after TKR. The strengths of this study are that the actual amount of app use was measured and because of the unique download codes adopted, it was not possible for the control group to use the PainCoach app. The shortcomings are that the additional analysis was underpowered and the cost-effectiveness of the PainCoach app was not investigated. Furthermore, as there is no short validated questionnaire in Dutch for measuring pain acceptance, an expert group decided to assess pain acceptance using happy and sad smileys as the best alternative. In the population of this study, opiate use was already low. The app might have a much stronger effect in patient populations where preoperative opiate use is much higher. It is questionable if the PainCoach app is effective in the overall TKR population, as this study investigated the effects in patients having ASA I-II and BMI ≤35, which represent

around 80% of the total TKR population [52,53]. Future research should focus on a larger sample size of the total TKR population, determination of the cost-effectiveness of the app, and use of the app in populations that have much higher preoperative opiate use.

CONCLUSIONS

The use of the PainCoach app contributes to reduced opiate use in the initial period at home after TKR. Active use of this app leads to further reduction in opiate use and improved pain control.

REFERENCES

1. Jüni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. Best Pract Res Clin Rheumatol 2006 Aug;20(4):721-740. [doi: 10.1016/j.berh.2006.05.002] [Medline: 16979535]
2. de Beer JD, Winemaker MJ, Donnelly GA, Miceli PC, Reiz JL, Harsanyi Z, et al. Efficacy and safety of controlled-release oxycodone and standard therapies for postoperative pain after knee or hip replacement. Can J Surg 2005 Aug;48(4):277-283 [FREE Full text] [Medline: 16149361]
3. Lin P. An evaluation of the effectiveness of relaxation therapy for patients receiving joint replacement surgery. J Clin Nurs 2012 Mar;21(5-6):601-608. [doi: 10.1111/j.1365-2702.2010.03406.x] [Medline: 21306457]
4. Dalury DF. A state-of-the-art pain protocol for total knee replacement. Arthroplast Today 2016 Mar;2(1):23-25 [FREE Full text] [doi: 10.1016/j.artd.2016.01.004] [Medline: 28326393]
5. Kehlet H, Wilmore DW. Evidence-based surgical care and the evolution of fast-track surgery. Ann Surg 2008 Aug;248(2):189-198. [doi: 10.1097/SLA.0b013e31817f2c1a] [Medline: 18650627]
6. Wu CL, Rowlingson AJ, Partin AW, Kalish MA, Courpas GE, Walsh PC, et al. Correlation of postoperative pain to quality of recovery in the immediate postoperative period. Reg Anesth Pain Med 2005;30(6):516-522 [FREE Full text] [doi: 10.1016/j.rapm.2005.07.190] [Medline: 16326335]
7. Trasolini NA, McKnight BM, Dorr LD. The Opioid Crisis and the Orthopedic Surgeon. J Arthroplasty 2018 Nov;33(11):3379-3382.e1. [doi: 10.1016/j.arth.2018.07.002] [Medline: 30075877]
8. Dowsey MM, Nikpour M, Choong PF. Outcomes following large joint arthroplasty: does socio-economic status matter? BMC Musculoskelet Disord 2014 May 06;15:148 [FREE Full text] [doi: 10.1186/1471-2474-15-148] [Medline: 24885773]
9. Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley JL. Sex, gender, and pain: a review of recent clinical and experimental findings. J Pain 2009 May;10(5):447-485 [FREE Full text] [doi: 10.1016/j.jpain.2008.12.001] [Medline: 19411059]
10. Jiang Y, Sanchez-Santos MT, Judge AD, Murray DW, Arden NK. Predictors of Patient-Reported Pain and Functional Outcomes Over 10 Years After Primary Total Knee Arthroplasty: A Prospective Cohort Study. J Arthroplasty 2017 Jan;32(1):92-100.e2 [FREE Full text] [doi: 10.1016/j.arth.2016.06.009] [Medline: 27444848]
11. Racine M, Tousignant-Laflamme Y, Kloda LA, Dion D, Dupuis G, Choinière M. A systematic literature review of 10 years of research on sex/gender and experimental pain perception - part 1: are there really differences between women and men? Pain 2012 Mar;153(3):602-618. [doi: 10.1016/j.pain.2011.11.025] [Medline: 22192712]
12. Sullivan M, Tanzer M, Stanish W, Fallaha M, Keefe FJ, Simmonds M, et al. Psychological determinants of problematic outcomes following Total Knee Arthroplasty. Pain 2009 May;143(1-2):123-129. [doi: 10.1016/j.pain.2009.02.011] [Medline: 19304392]
13. Escobar A, Quintana JM, Bilbao A, Azkárate J, Güenaga JI, Arenaza JC, et al. Effect of patient characteristics on reported outcomes after total knee replacement. Rheumatology (Oxford) 2007 Jan;46(1):112-119. [doi: 10.1093/rheumatology/kel184] [Medline: 16735451]
14. Blackburn J, Qureshi A, Amirfeyz R, Bannister G. Does preoperative anxiety and depression predict satisfaction after total knee replacement? Knee 2012 Oct;19(5):522-524. [doi: 10.1016/j.knee.2011.07.008] [Medline: 21846588]
15. 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 [FREE Full text] [doi: 10.1155/2009/273783] [Medline: 19714271]
16. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. Clin Orthop Relat Res 2010 Mar;468(3):798-806 [FREE Full text] [doi: 10.1007/s11999-009-0963-y] [Medline: 19585177]
17. Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra SM, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. Semin Arthritis Rheum 2012 Feb;41(4):576-588. [doi: 10.1016/j.semarthrit.2011.07.003] [Medline: 22035624]
18. Lingard EA, Riddle DL. Impact of psychological distress on pain and function following knee arthroplasty. J Bone Joint Surg Am 2007 Jun;89(6):1161-1169. [doi: 10.2106/JBJS.F.00914] [Medline: 17545417]

19. Brander VA, Stulberg SD, Adams AD, Harden RN, Bruehl S, Stanos SP, et al. Predicting total knee replacement pain: a prospective, observational study. Clin Orthop Relat Res 2003 Nov(416):27-36. [doi: 10.1097/01.blo.0000092983.12414.e9] [Medline: 14646737]

20. De Luca ML, Ciccarello M, Martorana M, Infantino D, Letizia Mauro G, Bonarelli S, et al. Pain monitoring and management in a rehabilitation setting after total joint replacement. Medicine (Baltimore) 2018 Oct;97(40):e12484 [FREE Full text] [doi: 10.1097/MD.0000000000012484] [Medline: 30290604]

21. Niemeläinen M, Kalliovalkama J, Aho AJ, Moilanen T, Eskelinen A. Single periarticular local infiltration analgesia reduces opiate consumption until 48 hours after total knee arthroplasty. A randomized placebo-controlled trial involving 56 patients. Acta Orthop 2014 Dec;85(6):614-619 [FREE Full text] [doi: 10.3109/17453674.2014.961399] [Medline: 25238439]

22. ReeceAS, Hulse GK. Impact of lifetime opioid exposure on arterial stiffness and vascular age: cross-sectional and longitudinal studies in men and women. BMJ Open 2014 Jun 02;4(6):e004521 [FREE Full text] [doi: 10.1136/bmjopen-2013-004521] [Medline: 24889849]

23. Schoenfeld AJ, Jiang W, Chaudhary MA, Scully RE, Koehlmoos T, Haider AH. Sustained Prescription Opioid Use Among Previously Opioid-Naive Patients Insured Through TRICARE (2006-2014). JAMA Surg 2017 Dec 01;152(12):1175-1176 [FREE Full text] [doi: 10.1001/jamasurg.2017.2628] [Medline: 28813584]

24. Currie M, Philip LJ, Roberts A. Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. BMC Health Serv Res 2015 Apr 16;15:162 [FREE Full text] [doi: 10.1186/s12913-015-0825-0] [Medline: 25888988]

25. Hochstenbach LM, Zwakhalen SM, Courtens AM, van Kleef M, de Witte LP. Feasibility of a mobile and web-based intervention to support self-management in outpatients with cancer pain. Eur J Oncol Nurs 2016 Aug;23:97-105 [FREE Full text] [doi: 10.1016/j.ejon.2016.03.009] [Medline: 27456381]

26. Jibb LA, Stevens BJ, Nathan PC, Seto E, Cafazzo JA, Johnston DL, et al. Implementation and preliminary effectiveness of a real-time pain management smartphone app for adolescents with cancer: A multicenter pilot clinical study. Pediatr Blood Cancer 2017 Oct;64(10). [doi: 10.1002/pbc.26554] [Medline: 28423223]

27. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: A systematic review. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1098-1109. [doi: 10.1016/j.berh.2017.04.002] [Medline: 29103552]

28. Naylor MR, Naud S, Keefe FJ, Helzer JE. Therapeutic Interactive Voice Response (TIVR) to reduce analgesic medication use for chronic pain management. J Pain 2010 Dec;11(12):1410-1419 [FREE Full text] [doi: 10.1016/j.jpain.2010.03.019] [Medline: 20620119]

29. Rini C, Porter LS, Somers TJ, McKee DC, DeVellis RF, Smith M, et al. Automated Internet-based pain coping skills training to manage osteoarthritis pain: a randomized controlled trial. Pain 2015 May;156(5):837-848 [FREE Full text] [doi: 10.1097/j.pain.0000000000000121] [Medline: 25734997]

30. Sandberg EH, Sharma R, Sandberg WS. Deficits in retention for verbally presented medical information. Anesthesiology 2012 Oct;117(4):772-779. [doi: 10.1097/ALN.0b013e31826a4b02] [Medline: 22902965]

31. Zickuhr K, Madden M. Pew Research Center's Internet & American Life Project. 2012. Older adults and internet use URL: https://www.sainetz.at/dokumente/studien/Older_adults_and_internet_use_2012.pdf [accessed 2018-02-08]

32. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983 Sep;17(1):45-56. [doi: 10.1016/0304-3959(83)90126-4] [Medline: 6226917]

33. Delgado DA, Lambert BS, Boutris N, McCulloch PC, Robbins AB, Moreno MR, et al. Validation of Digital Visual Analog Scale Pain Scoring With a Traditional Paper-based Visual Analog Scale in Adults. J Am Acad Orthop Surg Glob Res Rev 2018 Mar;2(3):e088 [FREE Full text] [doi: 10.5435/JAAOSGlobal-D-17-00088] [Medline: 30211382]

34. Chiu LY, Sun T, Ree R, Dunsmuir D, Dotto A, Ansermino JM, et al. The evaluation of smartphone versions of the visual analogue scale and numeric rating scale as postoperative pain assessment tools: a prospective randomized trial. Can J Anaesth 2019 Jun;66(6):706-715. [doi: 10.1007/s12630-019-01324-9] [Medline: 30796700]

35. Jamison RN, Gracely RH, Raymond SA, Levine JG, Marino B, Herrmann TJ, et al. Comparative study of electronic vs. paper VAS ratings: a randomized, crossover trial using healthy volunteers. Pain 2002;99(1):341-347. [doi: 10.1016/s0304-3959(02)00178-1]

36. Perruccio AV, Stefan LL, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for knee OA KOOS-Physical Function Shortform (KOOS-PS) - an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008 May;16(5):542-550 [FREE Full text] [doi: 10.1016/j.joca.2007.12.014] [Medline: 18294869]
37. Haverkamp D, Breugem SJ, Sierevelt IN, Blankevoort L, van Dijk CN. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. Acta Orthop 2005 Jun;76(3):347-352. [Medline: 16156462]
38. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
39. Stomp-van den Berg SG, Vlaeyen JW, Ter Kuile MM, Spinhoven P, van Breukelen G, Kole-Snijders AM. Pijn Coping en Cognitie Lijst (PCCL). Meetinstrumenten chronische pijn deel 2 URL: https://docplayer.nl/13110836-Meetinstrumenten-chronische-pijn-deel-2-pijn-coping-en-cognitie-lijst-pccl.html [accessed 2018-02-08]
40. Chan EY, Blyth FM, Nairn L, Fransen M. Acute postoperative pain following hospital discharge after total knee arthroplasty. Osteoarthritis Cartilage 2013 Sep;21(9):1257-1263 [FREE Full text] [doi: 10.1016/j.joca.2013.06.011] [Medline: 23973139]
41. Aasvang EK, Luna IE, Kehlet H. Challenges in postdischarge function and recovery: the case of fast-track hip and knee arthroplasty. Br J Anaesth 2015 Dec;115(6):861-866 [FREE Full text] [doi: 10.1093/bja/aev257] [Medline: 26209853]
42. Bol N, Smets EM, Rutgers MM, Burgers JA, de Haes HC, Loos EF, et al. Do videos improve website satisfaction and recall of online cancer-related information in older lung cancer patients? Patient Educ Couns 2013 Sep;92(3):404-412. [doi: 10.1016/j.pec.2013.06.004] [Medline: 23820196]
43. Selinger CP, Carbery I, Warren V, Rehman AF, Williams CJ, Mumtaz S, et al. The relationship between different information sources and disease-related patient knowledge and anxiety in patients with inflammatory bowel disease. Aliment Pharmacol Ther 2017 Dec;45(1):63-74. [doi: 10.1111/apt.13831] [Medline: 27778366]
44. van Laar MW, van Gestel B, Cruts AA, van der Pol PM, Ketelaars AP, Beenakkers EM, et al. Nationale Drug Monitor - Jaarbericht 2017. Utrecht: Trimbos instituut; 2018.
45. Hernandez NM, Parry JA, Taunton MJ. Patients at Risk: Large Opioid Prescriptions After Total Knee Arthroplasty. J Arthroplasty 2017 Aug;32(8):2395-2398. [doi: 10.1016/j.arth.2017.02.060] [Medline: 28392133]
46. Bedard NA, DeMik DE, Dowdle SB, Callaghan JJ. Trends and risk factors for prolonged opioid use after unicompartmental knee arthroplasty. Bone Joint J 2018 Jan;100-B(1 Supple A):62-67 [FREE Full text] [doi: 10.1302/0301-620X.100B1.BJJ-2017-0547.R1] [Medline: 29292342]
47. Franklin PD, Karbassi JA, Li W, Yang W, Ayers DC. Reduction in narcotic use after primary total knee arthroplasty and association with patient pain relief and satisfaction. J Arthroplasty 2010 Sep;25(6 Suppl):12-16. [doi: 10.1016/j.arth.2010.05.003] [Medline: 20580191]
48. Sinatra RS, Jahr JS, Reynolds LW, Viscusi ER, Groudine SB, Payen-Champenois C. Efficacy and safety of single and repeated administration of 1 gram intravenous acetaminophen injection (paracetamol) for pain management after major orthopedic surgery. Anesthesiology 2005 Apr;102(4):822-831. [doi: 10.1097/00000542-200504000-00019] [Medline: 15791113]
49. YehY, Reddy P. Clinical and economic evidence for intravenous acetaminophen. Pharmacotherapy 2012 Jun;32(6):559-579. [doi: 10.1002/j.1875-9114.2011.01085.x] [Medline: 22570116]
50. Specht K, Agerskov H, Kjaersgaard-Andersen P, Jester R, Pedersen BD. Patients' experiences during the first 12 weeks after discharge in fast-track hip and knee arthroplasty - a qualitative study. Int J Orthop Trauma Nurs 2018 Nov;31:13-19. [doi: 10.1016/j.ijotn.2018.08.002] [Medline: 30297138]
51. Sjøveian AK, Leegaard M. Hip and knee arthroplasty - patient's experiences of pain and rehabilitation after discharge from hospital. Int J Orthop Trauma Nurs 2017 Nov;27:28-35. [doi: 10.1016/j.ijotn.2017.07.001] [Medline: 28781113]
52. LROI. LROI Rapportage 2017. Distribution of ASA score of patients who underwent a primary knee arthroplasty per hospital in The Netherlands in 2016 URL: http://www.lroi-rapportage.nl/knee-primary-knee-arthroplasty-demographics-practice-variation-asa-score [accessed 2019-12-09]
53. LROI. LROI Rapportage 2017. Distribution of BMI of patients who underwent a primary knee arthroplasty per hospital in The Netherlands in 2016 URL: http://www.lroi-rapportage.nl/knee-primary-knee-arthroplasty-demographics-practice-variation-body-mass-index [accessed 2019-12-09]

MULTIMEDIA APPENDIX 1.

Pain management protocol

Table 1 Pain Management Protocol

	Medication	Intake	Dosage
Pre-operative	Gabapentin	2 hours before surgery 2 tablets	600 mg
	Acetaminophen (Paracetamol)	2 hours before surgery 2 tablets	1000 mg
	NSAID[a] (Diclofenac)	2 hours before surgery 1 tablet	50 mg
Per-operative	LIA[b]	***Ropivacaine 2 mg/ml and Adrenalin 1 mg/ml at a total volume of 100 ml*** LIA 1: injection of 2 x 20 ml Ropivacaine 2 mg/ml with Adrenalin in the posterior joint capsule and both collateral ligaments before the prosthesis was placed. LIA 2: After placement of the prosthesis, 2 x 20 ml Ropivacaine 2 mg/ml with Adrenalin injections along the edges of the tibia, in the capsule and in fat and soft tissue around the joint were injected. LIA 3: Inject 20 ml Ropivacaine 2 mg/ml without Adrenalin in subcutaneously layers before the wound was stitched.	
Post-operative during admission	Acetaminophen (Paracetamol)	2 tablets, 4 times per day	500 mg
	NSAID[a] (Diclofenac)	1 tablet, 3 times per day	50 mg
	Gabapentin	Day of surgery 10.00 pm Day after surgery 08.00 am Thereafter at indication	300 mg 300 mg
		If necessary (NRS > 4)	
	Opiate (Oxynorm / oxycodon)	Maximum of 1 tablet, 3 times per day	5 mg
Post-operative at home - *Usually 1 or 2 nights after surgery*	Acetaminophen (Paracetamol)	2 tablets, 4 times per day (until day 14 after surgery)	500 mg
		In presence of pain	
	NSAID[a] (Diclofenac)	1 tablet, 3 times per day (until day 7 after surgery)	50 mg
	Opiate (Oxynorm / oxycodon)	Maximum of 1 tablet, 3 times per day	5 mg
	Gabapentin	At indication on doctor's prescription	

[a]NSAID: non-steroidal anti-inflammatory drug. [b]LIA: local infiltration anesthesia

MULTIMEDIA APPENDIX 2.

Content and use of the PainCoach app

Table 1 Short schematic overview PainCoach app content

Post-operative days	Inputted pain level	Pain medication advice	Other advice
1-2	**No pain**	▪ Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg. ▪ NSAID (Diclofenac)[a]: 1 tablet, 3 times per day. Dosage: 50 mg.	▪ Execute exercises. ▪ Rest between exercises: immobilising the operated leg.
	Bearable pain	▪ Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg. ▪ NSAID (Diclofenac)[a]: 1 tablet, 3 times per day. Dosage: 50 mg.	▪ Execute exercises. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.
	Unbearable pain	▪ Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg. ▪ NSAID (Diclofenac)[a]: 1 tablet, 3 times per day. Dosage: 50 mg. ▪ Opiate (Oxynorm / oxycodon): Maximum of 1 tablet, 3 times per day. Dosage: 5 mg.	▪ Execute exercises without forcing. Contact physiotherapist for advice. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.
	Untenable pain	Call the clinic.	▪ Execute exercises without forcing. Contact physiotherapist for advice. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.
3-7	**No pain**	Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg.	▪ Execute exercises. ▪ Rest between exercises: immobilising the operated leg.
	Bearable pain	▪ Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg. ▪ NSAID (Diclofenac)[a]: 1 tablet, 3 times per day. Dosage: 50 mg.	▪ Execute exercises. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.
	Unbearable pain	▪ Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg. ▪ NSAID (Diclofenac)[a]: 1 tablet, 3 times per day. Dosage: 50 mg. ▪ Opiate (Oxynorm / oxycodon): Maximum of 1 tablet, 3 times per day. Dosage: 5 mg. ▪ Call the clinic for gabapentin at indication on doctor's prescription.	▪ Execute exercises without forcing. Contact physiotherapist for advice. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.
	Untenable pain	Call the clinic.	▪ Execute exercises without forcing. Contact physiotherapist for advice. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.

Table 1 Short schematic overview PainCoach app content (continued)

Post-operative days	Inputted pain level	Pain medication advice	Other advice
8-14	**No pain**	Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg.	▪ Execute exercises. ▪ Rest between exercises: immobilising the operated leg.
	Bearable pain	Acetaminophen (Paracetamol): 2 tablets, 4 times per day. Dosage 500 mg.	▪ Execute exercises. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.
	Unbearable pain	▪ Acetaminophen (Paracetamol): 2 tablets, 4 times per day. ▪ Dosage 500 mg. ▪ NSAID (Diclofenac)[a]: 1 tablet, 3 times per day. Dosage: 50 mg. ▪ Opiate (Oxynorm / oxycodon): Maximum of 1 tablet, 3 times per day. Dosage: 5 mg. ▪ Call the clinic for gabapentin at indication on doctor's prescription.	▪ Execute exercises without forcing. Contact physiotherapist for advice. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.
	Untenable pain	Call the clinic.	▪ Execute exercises without forcing. Contact physiotherapist for advice. ▪ Use ice and heat packs. ▪ Rest between exercises: immobilising the operated leg.

[a]NSAID: non-steroidal anti-inflammatory drug

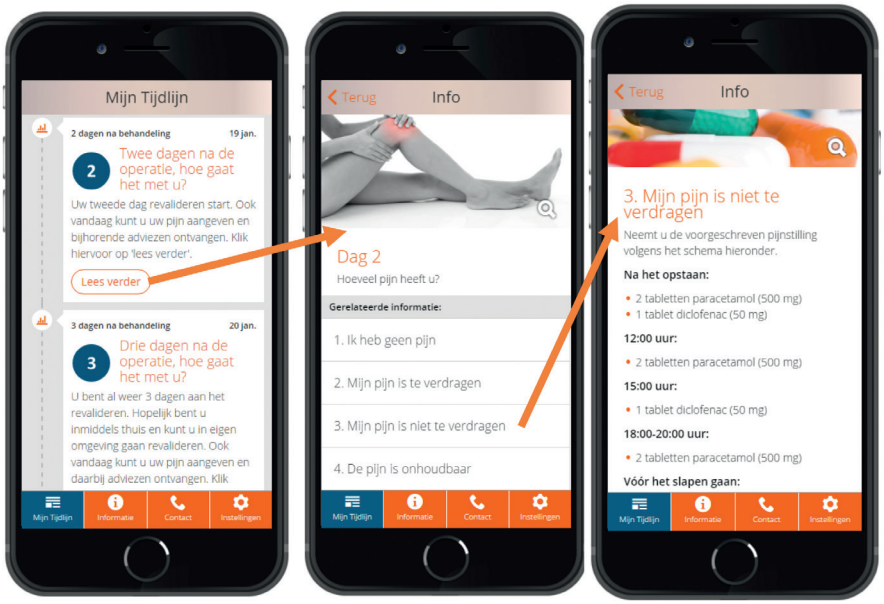


Fig. 1 Screenshots of PainCoach app showing how the PainCoach app works and how the content was presented to the patients

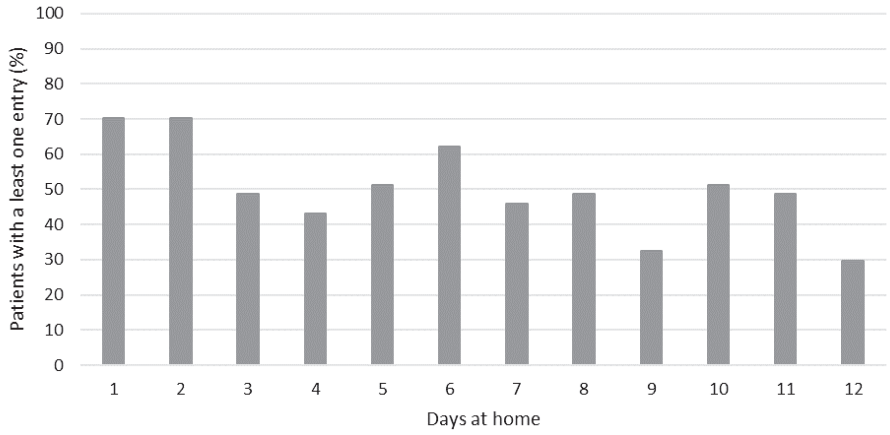


Fig. 2 Patients with at least one entry in the PainCoach app per day at home

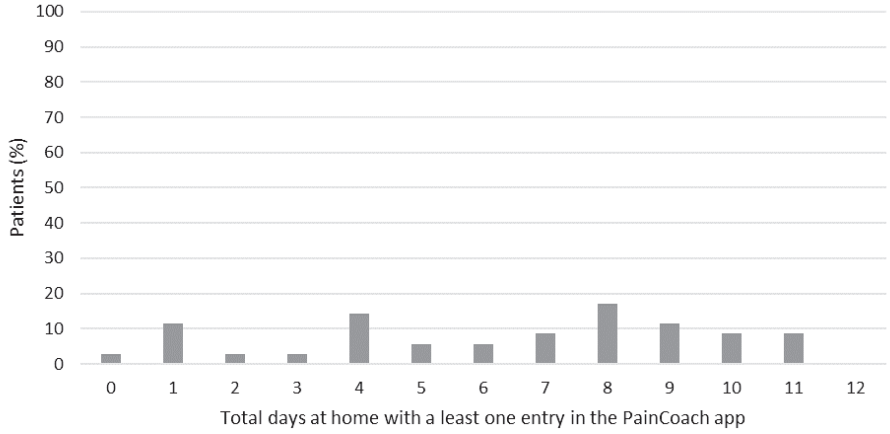
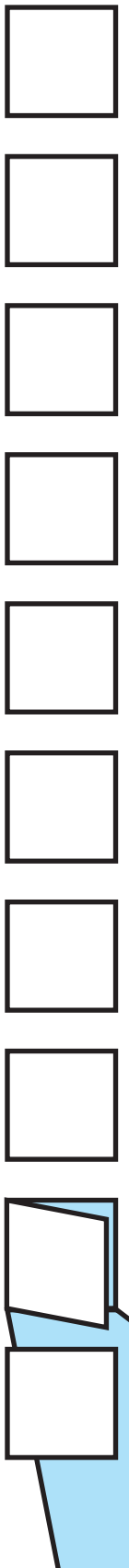


Fig. 3 Total days at home patients entered the PainCoach app



CHAPTER 9

A patient-reported outcome tool to triage total hip arthroplasty patients to hospital or video consultation: pilot study with expert panels and a cohort of 1228 patients

Y. Pronk, P. Pilot, W. van der Weegen, J.M. Brinkman and B.W. Schreurs

JMIR Form Res. 2021 Dec 20;5(12):e31232

ABSTRACT

Background

The digital transformation in health care has been accelerated by the COVID-19 pandemic. Video consultation has become the alternative for hospital consultation. It remains unknown how to select patients suitable for video consultation.

Objective

This study aimed to develop a tool based on patient-reported outcomes (PROs) to triage total hip arthroplasty (THA) patients to hospital or video consultation.

Methods

A pilot study with expert panels and a retrospective cohort with prospectively collected data from 1228 THA patients was executed. The primary outcome was a PRO triage tool to allocate THA patients to hospital or video consultation 6 weeks postoperatively. Expert panels defined the criteria and selected the patient-reported outcome measure (PROM) questions including thresholds. Data were divided into training and test cohorts. Distribution, floor effect, correlation, responsiveness, PRO patient journey, and homogeneity of the selected questions were investigated in the training cohort. The test cohort was used to provide an unbiased evaluation of the final triage tool.

Results

The expert panels selected moderate or severe pain and using 2 crutches as the triage tool criteria. PROM questions included in the final triage tool were numeric rating scale (NRS) pain during activity, 3-level version of the EuroQol 5 dimensions (EQ-5D-3L) questions 1 and 4, and Oxford Hip Score (OHS) questions 6, 8, and 12. Of the training cohort, 201 (201/703, 28.6%) patients needed a hospital consultation, which was statistically equal to the 150 (150/463, 32.4%) patients in the test cohort who needed a hospital consultation (P=.19).

Conclusions

A PRO triage tool based on moderate or severe pain and using 2 crutches was developed. Around 70% of THA patients could safely have a video consultation, and 30% needed a hospital consultation 6 weeks postoperatively. This tool is promising for selecting patients for video consultation while using an existing PROM infrastructure.

INTRODUCTION

The digital transformation in health care has been accelerated by the COVID-19 pandemic. Health care institutions are challenged by precautionary measures to contain COVID-19 while continuing to provide health care. Especially, managing the physical flow of patients is challenging. After the pandemic, hospitals will begin to eliminate their waiting lists while maintaining a normal patient flow, and they will be challenged again. As a solution, video consultation has become an alternative to the traditional hospital consultation. The number of hospital consultations has dropped by 30%, and the number of telemedicine visits has increased 5-fold [1].

Currently, video consultation provides health care institutions and clinicians the opportunity to increase office efficacy and cost-effectiveness in an era of decreasing reimbursements and increasing time constraints [2-4]. From the patient's perspective, it can also improve care efficacy and patient satisfaction as well as eliminating travel time and expenses [1,2]. However, not all patients might benefit from video consultation, and it is unknown how to select patients suitable for video consultation.

Orthopedic associations in many countries advise hospitals to collect patient-reported outcomes (PROs) of total hip arthroplasty (THA) using selected patient-reported outcome measures (PROMs) to evaluate health care and improve patient care [5,6]. To prevent extra burden in time and costs, it would be efficient to apply these PROs to select which patients need a hospital consultation and who can have a video consultation instead. Therefore, the aim of this study was to develop a tool based on PROs to triage THA patients to hospital or video consultation 6 weeks postoperatively. It was hypothesized that 10% of the THA patients would need a hospital consultation, as around 90% of the performed THAs result in a favorable outcome [7-9].

METHODS

Overview

A pilot study with expert panels and a retrospective cohort with prospectively collected data from THA patients was performed. Regarding the cohort, patients were included in this study if they signed the informed consent form preoperatively to allow further scientific analysis using their anonymized data. Therefore, the institutional review board ruled that formal approval was not required for this study. There were no exclusion criteria.

Outcomes

The primary outcome was a PRO triage tool to allocate THA patients to a hospital consultation or a video consultation for their 6-week postoperative consultation. Hospital consultation was defined as needing a physical examination or other examination, such as an X-ray, for which a patient really needed to be in the hospital. If no hospital consultation was needed, patients were allocated to a video consultation. According to the Dutch guidelines, patients should be seen 6

to 12 weeks after a THA [10], which is mostly held at 6 weeks. As it is advised to collect the first postoperative PROs at 3 months and not at 6 weeks [6], the 3-month PROs were considered the most appropriate for this study. Based on previous studies, the assumption was made that there are limited clinically relevant differences between PROs at 6 weeks and 3 months postoperatively [11,12].

Measurements

Measurements were divided into 3 parts: (1) expert panels defined the criteria and selected the PROM questions, including the thresholds; (2) investigation of the clinimetric qualities of selected questions or triage criteria groups in the retrospective cohort with prospectively collected data; and (3) evaluation of the final triage tool.

Selection by expert panels

Two expert panels were created: clinical expert panel and research expert panel. The clinical expert panel consisted of 4 high-volume THA orthopedic surgeons from 2 different health care institutions. The research expert panel consisted of 3 researchers from 2 different health care institutions. As step one, the clinical expert panel defined the clinical triage criteria for the triage tool. These clinical triage criteria were based on the clinical disabilities for which patients needed to have a physical examination or other examination, such as an X-ray. As the second step, based on the clinical triage criteria, the research expert panel selected the appropriate PROM questions, including the thresholds, based on previous studies. As step three, these questions and their thresholds were presented to the clinical expert panel to discuss if these questions and/or thresholds covered the clinical triage criteria. If no threshold was reported in previous studies, the threshold was set using clinical reasoning by the clinical expert panel. These steps resulted in the PROM questions, including the thresholds, that were determined to be clinically relevant for the triage tool.

Clinimetric qualities of selected questions or triage criteria groups

The retrospective cohort with prospectively collected data consisted of patients who underwent surgery between January 2016 and December 2018 in a medium-sized orthopedic hospital (Kliniek ViaSana, Mill, The Netherlands). Therefore, patients were characterized by an American Society of Anesthesiologists (ASA) score of I-II and BMI ≤35. Four high-volume, experienced orthopedic surgeons performed the primary posterolateral THAs. Length of stay was generally 1 or 2 days.

The data included patient characteristics, PROM response rates, and PROs. Patient characteristics were age on the day of surgery, gender, preoperative BMI, ASA scores, and preoperative Charnley scores collected from the electronic patient records. Response rates were calculated as the number of returned questionnaires that were partially or totally completed divided by the number of THAs minus the number of THAs of patients who were deceased (returned questionnaires / [THAs – THAs of patients who were deceased]) [5]. PROs were primary digitally collected (OnlinePROMs, Rosmalen, The Netherlands). If patients were unable to handle a computer, paper questionnaires were sent. A maximum of 2 reminders to complete the PROMs were sent [13]. PROs

were collected preoperatively and 3 and 12 months postoperatively according to the advice of the Dutch Orthopedic Association. This advice included the following questionnaires: numeric rating scale (NRS) pain at rest, NRS pain during activity, 3-level version of the EuroQol 5 dimensions (EQ-5D-3L), Hip disability and Osteoarthritis Outcome Score – Physical Function Shortform (HOOS-PS), Oxford Hip Score (OHS), and an anchor question about functional improvement [6].

Pain at rest and pain during activity were both measured using an NRS question scored from 0 (no pain) to 10 (severe pain). Quality of life was assessed using the EQ-5D-3L questionnaire consisting of 2 parts: EQ visual analogue scale (EQ VAS; 0-100, with 0 as the worst imaginable health state and 100 as the best imaginable health state) and EQ-5D descriptive system existing of 5 questions about 5 dimensions. These 5 dimensions are mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, scored from 1 (no problems) to 3 (extreme problems) [14]. Furthermore, hip function was measured using the HOOS-PS questionnaire, on a scale from 0 (no difficulty) to 100 (extreme difficulty). This questionnaire consists of 5 questions scored from 0 (no difficulty) to 5 (extreme difficulty) [15,16]. Hip function and pain were assessed using the OHS questionnaire, with scores ranging from 0 (most severe symptoms) to 48 (least symptoms). This questionnaire consists of 12 questions scored from 0 (no difficulty) to 4 (extreme difficulty) [17]. Moreover, functional improvement was inquired on a 7-point Likert scale question ranging from 1 (very much deteriorated) to 7 (very much improved).

Regarding the investigation of the clinimetric qualities of selected questions or triage criteria groups by the expert panels, the cohort was divided into 2 groups: training cohort of patients who underwent surgery in 2016 or 2017 and a test cohort of patients who underwent surgery in 2018. To assess which questions were appropriate for the triage tool, the following clinimetric qualities were investigated per selected question in the training cohort: distribution, floor effect, correlation, responsiveness, and PRO patient journey. PRO patient journey was defined as a change in recovery over time. Regarding distribution, if the question did not show any distinction (median and IQR on the same level), the question was found not to be an appropriate question for the triage tool. For floor effect, if more than 15% of the patients scored the worst score [18], the question had a problem with the floor effect and was not an appropriate triage tool question. Investigating correlation, if the question was correlated ($r \geq .7$) with another selected question(s) [19], this question or one of the other(s) could be chosen instead of all the questions for the triage tool. Regarding responsiveness, if a question was not responsive ($P > .05$) [20,21], it did not distinguish well between clinical relevance and lack of clinical relevance and was not included in the tool. Furthermore, the PRO patient journeys of patients with a worse score and of patients with a better score than the threshold were investigated. If patients with a worse score on a question at 3 months scored well on that question at 12 months, this question was not included in the triage tool. To assess which questions within the selected triage criteria group (for example pain) were appropriate for the triage tool, homogeneity was investigated per triage criteria group in the training cohort. If the homogeneity increased by removing a certain question from this group (Cronbach $\alpha > .7$) [18,19], this question did not fit in this group and could be removed from the triage tool.

Evaluation of the final triage tool

The final triage tool was applied in the test cohort to provide an unbiased evaluation of the final tool fitted on the training dataset. Results in both cohorts were compared to investigate the hypothesis.

Statistical analysis

Results are reported as mean (SD), median (IQR), or n (%) based on the test performed. To investigate if there was any difference in patient characteristics, response rates, and preoperative PROs between the training and test cohorts, continuous variables were first checked for a normal distribution. Second, independent t tests or Mann-Whitney U tests for continuous variables were executed depending on the distribution of the data, and Pearson chi-square or Fisher exact tests were executed for categorical variables.

Distribution was investigated with a boxplot distribution, floor effect was determined by calculating the percentage of patients with a minimum score, and correlation was assessed with Spearman correlation analyses. Responsiveness was evaluated by performing Wilcoxon signed rank tests on the change in preoperative and 3-month scores [20,21]. The PRO patient journey of patients with a worse or better score than the threshold on a question at 3 months was evaluated by boxplot distribution at 12 months. Homogeneity was investigated with a reliability analysis, including "scale if item deleted." Before this analysis was executed, NRS pain and EQ-5D-3L questions were recoded to the same direction as the OHS questions.

Finally, the triage tool was applied for both training and test cohorts. To test the hypothesis, the numbers of hospital and video consultations for both cohorts were compared using Pearson chi-square or Fisher exact tests.

An α of .05 was considered statistically significant. Statistical analyses were performed using SPSS version 25.0 (IBM Corp, Armonk, NY).

RESULTS

Selection by expert panels

"Having moderate or severe pain" and "using 2 crutches" were defined as the triage criteria by the clinical expert panel (step 1). For the criterion of "having moderate or severe pain," the research expert panel selected the following PROM questions: NRS pain at rest, NRS pain during activities, EQ-5D-3L question 4, and OHS questions 1, 8, 10, and 12. For both NRS pain questions, previous studies reported thresholds of $\leq$3 for no or mild pain and >3 for moderate to severe pain [22,23]. For the criterion of "using 2 crutches," EQ-5D-3L question 1 and OHS question 6 were selected (step 2). The clinical expert panel assessed the selected questions, even the NRS pain question thresholds, as appropriate. The other thresholds were discussed and defined (step 3; Table 1).

Table 1 Triage criteria, selected clinical relevant questions, and defined thresholds by expert panels

Triage criteria and selected PROM[a] question	PROM question (score range)	Defined threshold
Having moderate or severe pain		
NRS[b] pain at rest	How much pain from your hip (surgery side) did you experience at rest in the last week? (0-10)	≥4
NRS pain during activity	How much pain from your hip (surgery side) did you experience during activity in the last week? (0-10)	≥4
EQ-5D-3L[c] question 4	Pain/discomfort (1-3)	≥3 (extreme pain)
OHS[d] question 1	During the past 4 weeks… How would you describe the pain you usually had from your hip? (0-4)	≤1 (moderate or severe)
OHS question 8	During the past 4 weeks… After a meal (sitting at a table), how painful has it been for you to stand up from a chair because of your hip? (0-4)	≤2 (moderate, very, unbearable)
OHS question 10	During the past 4 weeks…..... Have you had any sudden, severe pain – "shooting", "stabbing" or "spasms" - from the affected hip? (0-4)	≤1 (most or every)
OHS question 12	During the past 4 weeks… Have you been troubled by pain from your hip in bed at night? (0-4)	≤2 (3 or 4, 5 or 6, all)
Using two crutches		
EQ-5D-3L question 1	Mobility (1-3)	≥3 (confined to bed)
OHS question 6	During the past 4 weeks… For how long have you been able to walk before pain from your hip becomes severe (with or without a stick)? (0-4)	≤2 (5-15 minutes, around the house only, not at all)

[a]PROM: patient-reported outcome measure. [b]NRS: numeric rating scale. [c]EQ-5D-3L: 3-level version of the EuroQol 5 dimensions. [d]OHS: Oxford Hip Score

Clinimetric qualities of selected questions or triage criteria groups

Response rates were statistically significantly equal between both training (n=746) and test (n=482) cohorts preoperatively (745/746, 99.9% versus 482/482, 100%; P=.99) and at 3 months (703/746, 94.2% versus 463/482, 96.1%; P=.24) and 12 months (693/746, 92.9% versus 457/482, 94.8%; P=.29) postoperatively. The training cohort consisted of significantly fewer patients than in the test cohort with an ASA I score (399/746, 53.5% versus 287/482, 59.5%; P=.04), lower Charnely scores (P=.048), higher preoperative HOOS-PS scores (median 46.1, IQR 37.7-55.9 versus median 46.1, IQR 33.9-55.9; P=.01), and lower preoperative OHS scores (median 24.0, IQR 19.0-29.0 versus median 25.0, IQR 19.0-31.0; P=.03; Table 2). The clinical expert panel assessed these differences as not clinically relevant to correct for.

Table 2 Characteristics of training and test cohorts

Characteristics	Training cohort (n=746)	Test cohort (n=482)	*P* value
Response rate, n (%)			
Preoperative	745 (99.9)	482 (100)	.99
3 months postoperative	703 (94.2)	463 (96.1)	.24
12 months postoperative	693 (92.9)	457 (94.8)	.29
Patient characteristics			
Age (years), median (IQR)	66.5 (61.0-72.0)	66.0 (60.0-72.0)	.73
Gender (male), n (%)	295 (39.5)	206 (42.7)	.27
BMI (kg/m^2), median (IQR)	26.1 (24.1-28.4)	26.3 (24.1-28.5)	.48
ASA[a] score (I), n (%)	399 (53.5)	287 (59.5)	.04
Charnley score			.048
One hip affected with OA[b], n (%)	163 (21.8)	108 (22.4)	
Both hips affected with OA, n (%)	309 (41.4)	196 (40.7)	
Contra lateral hip OA, n (%)	163 (21.8)	82 (17.0)	
Multiple joints affected with OA, n (%)	111 (14.9)	96 (19.9)	
Preoperative PROs[c], (median (IQR))			
NRS[d] pain at rest score	6.0 (4.0-7.0)	6.0 (3.8-7.0)	.18
NRS pain during activity score	8.0 (7.0-9.0)	8.0 (7.0-8.0)	.13
HOOS-PS[e] score	46.1 (37.7-55.9)	46.1 (33.9-55.9)	.01
EQ-5D descriptive system[f]	0.693 (0.298-0.775)	0.693 (0.569-0.775)	.16
EQ VAS[g]	76.0 (63.3-89.8)	77.0 (60.0-86.3)	.82
OHS[h]	24.0 (19.0-29.0)	25.0 (19.0-31.0)	.03

[a]ASA: American Society of Anesthesiologists. [b]OA: osteoarthritis. [c]PROs: patient-reported outcomes. [d]NRS: numeric rating scale. [e]HOOS-PS: Hip disability and Osteoarthritis Outcome Score – Physical Function Shortform. [f]EQ-5D descriptive system: EuroQol 5 dimensions descriptive system. [g]EQ VAS: EuroQol visual analogue scale. [h]OHS: Oxford Hip Score

Regarding the questions or triage criteria groups selected by the expert panels (Table 1), OHS question 10 showed no distribution. For floor effect, <15% of patients scored the minimum score on all questions separately. All questions were significantly correlated with each other (P<.001). Regarding correlations ≥0.7, NRS pain during activity correlated with NRS pain at rest (r=0.659, P<.001) and OHS question 1 (r=−0.676, P<.001; Table 3). Furthermore, all questions were shown to be responsive (P<.001; Table 4). Regarding the PRO patient journey, patients with a worse score than the threshold also reported worse scores at 12 months than patients with a better score than the threshold. Only one patient with a better score than the threshold on EQ-5D-3L question 1 at 3 months had a 12-month score (Table 5). The other questions included ≥11 patients below or above the threshold. Regarding homogeneity, a Cronbach α of 0.818 was found for the triage criteria group "pain." When one of the questions in this group was removed, the Cronbach

α was maintained at above 0.7. The triage criteria group "crutches" scored a Cronbach α of 0.628. As there were 2 questions in this group, none of them could be removed to investigate the Cronbach α.

Table 3 Distribution, floor effect, and correlation per selected patient-reported outcome measure (PROM) question

PROM[a] question	Distribution, median (IQR)	Floor effect, n (%)	Correlations		
			Correlations (*r*)[b]	Correlated question	*P* value
NRS[c] pain at rest	0 (0-1)	3[d] (0.4)	0.659	NRS pain during activity	<.001
NRS pain during activity	2 (0-3)	4[d] (0.6)	0.659; -0.676	NRS pain at rest; OHS[e] question 1	Both <.001
EQ-5D-3L[f] question 4	1 (1-2)	13[g] (1.9)	none	N/A[h]	N/A
OHS question 1	3 (3-4)	6[g] (0.9)	-0.675	NRS pain during activity	<.001
OHS question 8	3 (3-4)	0[g] (0.0)	none	N/A	N/A
OHS question 10	4 (4-4)	2[i] (0.3)	none	N/A	N/A
OHS question 12	4 (3-4)	32[i] (4.6)	none	N/A	N/A
EQ-5D-3L question 1	1 (1-2)	2[i] (0.3)	none	N/A	N/A
OHS question 6	4 (3-4)	3[j] (0.4)	none	N/A	N/A

[a]PROM: patient-reported outcome measure. [b]Statistically significant correlations >0.6 or <−0.6 are presented. [c]NRS: numeric rating scale. [d]n=703. [e]OHS: Oxford Hip Score. [f]EQ-5D-3L: 3-level version of the EuroQol 5 dimensions. [g]n=693. [h]N/A: not applicable. [i]n=694. [j]n=690

Table 4 Responsiveness for each patient-reported outcome measure (PROM) question

PROM question	Preoperative, median (IQR)	3 months postoperative, median (IQR)	*P* value
NRS[a] pain at rest	6 (4-7)	0 (0-1)	<.001
NRS pain during activity	8 (7-9)	2 (0-3)	<.001
EQ-5D-3L[b] question 4	2 (2-3)	1 (1-2)	<.001
OHS[c] question 1	1 (0-1)	3 (3-4)	<.001
OHS question 8	2 (2-3)	3 (3-4)	<.001
OHS question 10	2 (1-3)	4 (4-4)	<.001
OHS question 12	2 (0-3)	4 (3-4)	<.001
EQ-5D-3L question 1	2 (2-2)	1 (1-2)	<.001
OHS question 6	2 (2-3)	4 (3-4)	<.001

[a]NRS: numeric rating scale. [b]EQ-5D-3L: 3-level version of the EuroQol 5 dimensions. [c]OHS: Oxford Hip Score

Table 5 Patient journey per patient-reported outcome measure (PROM) question

PROM question	Defined threshold	12 month score of patients with a score below threshold at 3 months, median (IQR)	12 month score of patients with a score above threshold at 3 months, median (IQR)
NRS[a] pain at rest	≥4	0 (0-1)	2 (0-5)
NRS pain during activity	≥4	0 (0-1)	2 (0-4)
EQ-5D-3L[b] question 4	≥3	1 (1-2)	2 (1.5-2)
OHS[c] question 1	≤1	3 (2-4)	4 (3-4)
OHS question 8	≤2	3 (3-4)	4 (4-4)
OHS question 10	≤1	4 (2.5-3)	4 (4-4)
OHS question 12	≤2	3 (1.5-4)	4 (4-4)
EQ-5D-3L question 1	≥3	1 (1-2)	1 (1-1)[d]
OHS question 6	≤2	3 (2-4)	4 (4-4)

[a]NRS: numeric rating scale. [b]EQ-5D-3L: 3-level version of the EuroQol 5 dimensions. [c]OHS: Oxford Hip Score. [d]n=1

Based on the clinimetric qualities of selected questions or triage criteria groups, NRS pain at rest, OHS question 1, and OHS question 10 were removed from the triage tool. The final triage tool consisted of NRS pain during activity; EQ-5D-3L questions 1 and 4; and OHS questions 6, 8, and 12.

Evaluation of the final triage tool

The final triage tool resulted in 201 (201/703, 28.6%) patients in the training cohort needing a hospital consultation, which was statistically equal to the 150 (150/463, 32.4%) patients in the test cohort who needed a hospital consultation (P=.19).

DISCUSSION

This study aimed to develop a tool based on PROs collected using an existing PROM infrastructure to triage THA patients to hospital or video consultation 6 weeks postoperatively. As the main finding, a triage tool based on PROM questions measuring moderate or severe pain and whether the patient used 2 crutches was developed. The included questions were NRS pain during activity; EQ-5D-3L questions 1 and 4; and OHS questions 6, 8, and 12. Applying the final triage tool in both the training and test cohorts resulted in the same outcome: Around 70% of the patients could safely have a video consultation, and 30% needed to have a hospital consultation 6 weeks postoperatively. Therefore, this PRO triage tool is a promising instrument to select patients for video consultation while using an existing PROM infrastructure. The next step is to further investigate this triage tool in daily practice.

This study showed that 70% of the hospital consultations for THA patients 6 weeks postoperatively could safely be done by video. It was hypothesized that 10% of the THA patients would need a

hospital consultation. First, the result that 30% of patients needed a hospital consultation could be explained by the focus of the clinical expert panel. As the experts' beginning point was seeing all patients during a hospital consultation (100%), by developing the triage tool, they desired to see all patients who potentially needed a physical examination or other examination, such as an X-ray, during a hospital consultation. Furthermore, they desired to prevent obtaining more consultations by needing to schedule a hospital consultation after a video consultation. Both implicitly resulted in more liberal criteria for a hospital consultation leading to more patients triaged to a hospital consultation. Second, it could be that specific questions are missing from the triage tool. It was hypothesized that, after an investigation of the triage tool in daily practice, the criteria for the triage tool could be improved, achieving the right health care for each patient and a further reduction in hospital consultations. It would be interesting to investigate how many additional hospital consultations would be needed if the tool triages to video consultation and which PROs are different for patients with an additional hospital consultation.

It is essential to understand that the PRO triage tool is not a tool on its own, but it is the first step in the selection of patients who need a hospital consultation and those who can have a video consultation instead. PROs and clinical judgment produce complementary data and when combined, provide a more accurate description of the patients' symptoms [24]. Therefore, using the current PRO triage tool, clinicians should have the ability to change the outcome of the tool. To further develop the triage tool, it would be interesting to investigate how many times clinicians decide to change the outcome and which PROs are different for the patients whose clinicians decide to change the outcome. Furthermore, it would be interesting to take the patient's preference into account.

Previous studies reported that patients rate their video consultations as excellent or very good (92%-95%) [25,26]. The patient no-show rate has been reported at 2.8%, and their mean estimated saved travel time is 30 minutes [25]. Furthermore, 82% would recommend video consultation to family and friends [26]. Almost all clinicians rate their video consultation experience as very good or excellent (92%). They are comfortable with executing this type of consultation after 1 to 4 sessions (69%) [25]. Therefore, video consultation is a serious alternative for hospital consultation. Numbers could be improved when appropriate patients for video consultation are selected, which makes the developed PRO triage tool a promising instrument.

As a first strength of this study, to the authors' knowledge, this is the first study in which a tool to triage patients to hospital or video consultation was developed. Second, high response rates preoperatively and even postoperatively (above 90%) were achieved, resulting in a representative cohort to execute this study and to generalize the results to the total THA population. A third strength is the application of the training and test cohorts to provide an unbiased evaluation of the final tool.

As a limitation of this study, the triage tool was not investigated in a prospective cohort, and aspects of reliability, validity, sensitivity, and specificity of the triage tool were not investigated yet.

Furthermore, 3-month PROs were used instead of 6-week PROs, as, although based on previous studies, the assumption was made that there is limited clinically relevant difference between PROs at 6 weeks and at 3 months postoperatively [11,12]. Future research should be executed in a prospective cohort, and aspects of reliability, validity, sensitivity, and specificity need to be investigated to further develop the THA PRO triage tool. The triage tool could be improved by investigating which PROs are different for patients with additional hospital consultations after being triaged to video consultation or for patients whose clinician decided to change the outcome of the triage tool. After improving the triage tool, it would be interesting to investigate if and which of the patients triaged to video consultation may not require a consultation at all.

CONCLUSION

A THA PRO triage tool based on moderate or severe pain and using 2 crutches was developed. Around 70% of THA patients could safely have a video consultation, and 30% of patients needed a hospital consultation 6 weeks postoperatively. This tool is promising for selecting patients for video consultation while using an existing PROM infrastructure.

REFERENCES

1. Merkes K, Hagenaars N. Corona: katalysator of struikelblok voor groenere ziekenhuiszorg? Gupta Strategists. 2020 May 28. URL: https://gupta-strategists.nl/studies/corona-katalysator-of-struikelblok-voor-groenere-ziekenhuiszorg [accessed 2021-11-13]
2. Hollander JE, Carr BG. Virtually Perfect? Telemedicine for Covid-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/NEJMp2003539] [Medline: 32160451]
3. O'Connor CM, Anoushiravani AA, DiCaprio MR, Healy WL, Iorio R. Economic recovery after the COVID-19 pandemic: resuming elective orthopedic surgery and total joint arthroplasty. J Arthroplasty 2020 Jul;35(7S):S32-S36 [FREE Full text] [doi: 10.1016/j.arth.2020.04.038] [Medline: 32345566]
4. Kane LT, Thakar O, Jamgochian G, Lazarus MD, Abboud JA, Namdari S, et al. The role of telehealth as a platform for postoperative visits following rotator cuff repair: a prospective, randomized controlled trial. J Shoulder Elbow Surg 2020 Apr;29(4):775-783. [doi: 10.1016/j.jse.2019.12.004] [Medline: 32197766]
5. Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries. Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries Part II. Recommendations for selection, administration, and analysis. Acta Orthop 2016 Jul;87 Suppl 1:9-23 [FREE Full text] [doi: 10.1080/17453674.2016.1181816] [Medline: 27228230]
6. PROMs. Nederlandse Orthopaedische Vereniging. URL: https://www.lroi.nl/invoerders/registreren/proms [accessed 2018-08-01]
7. Palazzo C, Jourdan C, Descamps S, Nizard R, Hamadouche M, Anract P, et al. Determinants of satisfaction 1 year after total hip arthroplasty: the role of expectations fulfilment. BMC Musculoskelet Disord 2014 Feb 24;15:53 [FREE Full text] [doi: 10.1186/1471-2474-15-53] [Medline: 24564856]
8. Anakwe RE, Jenkins PJ, Moran M. Predicting dissatisfaction after total hip arthroplasty: a study of 850 patients. J Arthroplasty 2011 Feb;26(2):209-213. [doi: 10.1016/j.arth.2010.03.013] [Medline: 20462736]
9. Hesseling B, Mathijssen NMC, van Steenbergen LN, Melles M, Vehmeijer SBW, Porsius JT. Fast starters, slow starters, and late dippers: trajectories of patient-reported outcomes after total hip arthroplasty: results from a Dutch nationwide database. J Bone Joint Surg Am 2019 Dec 18;101(24):2175-2186. [doi: 10.2106/JBJS.19.00234] [Medline: 31609887]
10. Routinematige follow-up bij THP - Richtlijnendatabase. Federatie Medisch Specialisten. 2019 Dec 2. URL: https://richtlijnendatabase.nl/richtlijn/totale_heupprothese_thp/postoperatieve_zorg_bij_thp/routinematige_follow-up_bij_thp.html [accessed 2021-11-13]
11. Klapwijk LCM, Mathijssen NMC, Van Egmond JC, Verbeek BM, Vehmeijer SBW. The first 6 weeks of recovery after primary total hip arthroplasty with fast track. Acta Orthop 2017 Apr;88(2):140-144 [FREE Full text] [doi: 10.1080/17453674.2016.1274865] [Medline: 28079428]
12. Hartog YMD, Mathijssen NMC, Vehmeijer SBW. Total hip arthroplasty in an outpatient setting in 27 selected patients. Acta Orthop 2015;86(6):667-670 [FREE Full text] [doi: 10.3109/17453674.2015.1066211] [Medline: 26139431]
13. Pronk Y, Pilot P, Brinkman JM, van Heerwaarden RJ, van der Weegen W. Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection. J Patient Rep Outcomes 2019 Jun 03;3(1):31 [FREE Full text] [doi: 10.1186/s41687-019-0121-6] [Medline: 31155689]
14. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
15. Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert J, 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. Osteoarthritis Cartilage 2009 Jul;17(7):843-847 [FREE Full text] [doi: 10.1016/j.joca.2009.01.005] [Medline: 19215728]
16. Davis AM, Perruccio AV, Canizares M, Tennant A, Hawker GA, Conaghan PG, et al. The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008 May;16(5):551-559 [FREE Full text] [doi: 10.1016/j.joca.2007.12.016] [Medline: 18296074]

17. Gosens T, Hoefnagels NHM, de Vet RCW, Dhert WJA, van Langelaan EJ, Bulstra SK, 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 Apr;76(2):204-211. [doi: 10.1080/00016470510030580] [Medline: 16097545]

18. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007 Jan;60(1):34-42. [doi: 10.1016/j.jclinepi.2006.03.012] [Medline: 17161752]

19. Streiner DL, Norman GR, Cairney J. Reliability. In: Streiner DL, Norman GR, Cairney J, editors. Health Measurement Scales: A practical guide to their development and use. Oxford, England: Oxford University Press; 2019. [doi: 10.1093/med/9780199685219.003.0008]

20. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PMM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003 Jun;12(4):349-362. [doi: 10.1023/a:1023499322593] [Medline: 12797708]

21. Moayyedi P, Duffett S, Braunholtz D, Mason S, Richards ID, Dowell AC, et al. The Leeds Dyspepsia Questionnaire: a valid tool for measuring the presence and severity of dyspepsia. Aliment Pharmacol Ther 1998 Dec;12(12):1257-1262 [FREE Full text] [doi: 10.1046/j.1365-2036.1998.00404.x] [Medline: 9882035]

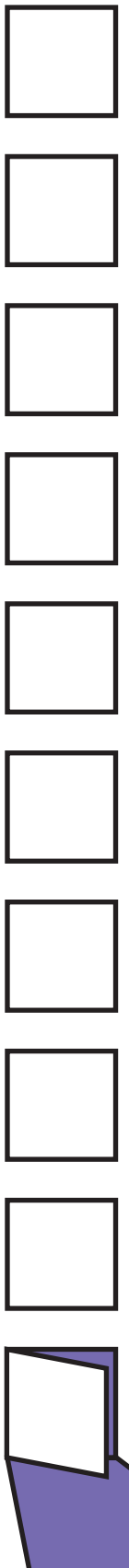
22. Pinto PR, McIntyre T, Ferrero R, Almeida A, Araújo-Soares V. Risk factors for moderate and severe persistent pain in patients undergoing total knee and hip arthroplasty: a prospective predictive study. PLoS One 2013;8(9):e73917 [FREE Full text] [doi: 10.1371/journal.pone.0073917] [Medline: 24058502]

23. Breivik H, Borchgrevink PC, Allen SM, Rosseland LA, Romundstad L, Hals EKB, et al. Assessment of pain. Br J Anaesth 2008 Jul;101(1):17-24 [FREE Full text] [doi: 10.1093/bja/aen103] [Medline: 18487245]

24. Basch E, Bennett A, Pietanza MC. Use of patient-reported outcomes to improve the predictive accuracy of clinician-reported adverse events. J Natl Cancer Inst 2011 Dec 21;103(24):1808-1810. [doi: 10.1093/jnci/djr493] [Medline: 22157639]

25. Tenforde AS, Iaccarino MA, Borgstrom H, Hefner JE, Silver J, Ahmed M, et al. Telemedicine during COVID-19 for outpatient sports and musculoskeletal medicine physicians. PM R 2020 Sep;12(9):926-932 [FREE Full text] [doi: 10.1002/pmrj.12422] [Medline: 32424977]

26. Donelan K, Barreto EA, Sossong S, Michael C, Estrada JJ, Cohen AB, et al. Patient and clinician experiences with telehealth for patient follow-up care. Am J Manag Care 2019 Jan;25(1):40-44 [FREE Full text] [Medline: 30667610]



CHAPTER 10

Summary and General discussion

With the shift to a more patient centred orthopaedic health care, measuring patient-reported outcomes (PROs) of hip and knee arthroplasty patients using selected patient-reported outcome measures (PROMs) became the gold standard. The aim of this thesis was to investigate how routine PRO collection can be optimized (part I) and subsequently how health care can be optimized with routine use of PROs (part II) in hip and knee arthroplasty. In this chapter a short overview of the context of this thesis is presented, the main findings are highlighted and reflected on, and recommendations for future steps are suggested.

CONTEXT

Daily practice shows a large diversity between health care institutions in their success of PRO collection and how they use PROs to improve health care. Wide ranges in response rates (RRs) are observed in both the Dutch and international arthroplasty registries [1, 2]. Routine PRO collection and routine PRO use have been increased, however, multiple PRO-related questions remain unanswered. From that perspective, it is questionable if the current PRO collection and use are justifiable from an ethical and value-based health care perspective. Although examples and recommendations how to collect [3–11] and use PROs exist [12–14], scientific evidence on how to optimize routine PRO collection and how to optimize health care with routine use of PROs in hip and knee arthroplasty is lacking.

What does this thesis add to the scientific evidence regarding PRO(M)s in arthoplasty health care?

- An automated PRO collection system with additional manual steps results in a two times higher PROM RR for patients responding at all time points and adds a cost of €6 per surgical procedure compared to the use of an automated system alone (Chapter 3).
- To draw valid conclusions on PROs, a minimum RR (MRR) of 60% is advised (Chapter 4).
- Only 16% of the Dutch health care institutions reaches this advised 60% on their total hip arthroplasty (THA) PRO collection. Therefore, it remains questionable if the aim 'improving arthroplasty health care' is achieved at the Dutch national level by collecting PROs (Chapter 5).
- When routine PRO collection is sufficient (RR ≥60%), PROs are routinely useful to optimize health care. Scientifically investigated examples:
  - PROs are useful to gain knowledge for shared decision making and for making recommendations to stakeholders on choosing a mobile bearing or a fixed bearing implant design for patients with a medial unicompartmental knee arthroplasty (UKA) performed by a high-volume surgeon. This is an example of health care evaluation (Chapter 6).
  - PROs are not clinically relevant useful to gain knowledge for shared decision making and for making recommendations to stakeholders by preoperatively predicting patient satisfaction after a total knee arthroplasty (TKA) (Chapter 7).
  - A PRO based app, called PainCoach app, is useful to guide TKA patients after surgery in pain control and opiate use (Chapter 8).

- PROs are useful to triage THA patients to whether a traditional hospital consultation is needed or alternatively a video consultation is adequate (Chapter 9).

Based on these results, the next steps in optimizing routine PRO collection in hip and knee arthroplasty are defined: besides an automated system additional manual steps, requiring extra budget, are needed to have a sufficient (RR ≥60%) routine PRO collection (part I of this thesis). Furthermore, when routine PRO collection is sufficient, examples show how health care can be optimized with routine use of PROs in hip and knee arthroplasty (part II of this thesis). A coordinated effort of all stakeholders has to be initiated to improve routine PRO collection and routine PRO use to optimize health care.

REFLECTION ON MAIN FINDINGS

Part I: Optimizing the routine PRO collection

In the last decades, health care has been digitalized to a large extend. Software programmes have been developed to facilitate PRO collection by creating online access for patients and for health care institutions employees, and by automating collection steps. In addition to previous studies [3–11, 15], chapters 3 and 4 show next steps how to optimize routine PRO collection in hip and knee arthroplasty. Chapter 3 is one of the first studies that creates transparency in the effort and costs needed for PRO collection. It shows that a two times higher RR is achievable using an automated system with additional manual steps compared to an automated system alone. This applies to all types of orthopaedic surgeries as well as for only hip and knee arthroplasty procedures separately. However, these additional manual steps add a cost of €6 per surgical procedure on top of the costs for automated PRO collection alone (Chapter 3). Still, collecting PROs routinely of all operated orthopaedic patients is recommend based on three reasons. (1) PRO collection for a larger group of patients reduces the cost per patient because the shared fixed costs (e.g. a licence for the an automated system) (Chapter 3). (2) In addition, incorporating PRO collection in daily health care for all patients to create routine is one of the keys to increase RR [3, 5, 6]. Mainly because, in general, automatic behaviour is dominant over exceptional behaviour. (3) Furthermore, from an ethical perspective, it cannot be justified that surgeons, health care institutions and other stakeholders only should know what the results of arthroplasty patients are instead of all operated orthopaedic patients.

To justify the costs, an important factor to investigate is the PROM RR needed to adequately evaluate hip and knee arthroplasty procedures, called MRR. Based on two of the three pre-determined conditions, a MRR of 60% is advised. When all three conditions should be met, an impossible 100% MRR is required. This MRR is needed on both the preoperative and the 3 month postoperative time points of the Dutch THA PROMs set (Chapter 4). As this is the first study addressing the methodological challenges in investigating the MRR, it can be seen as a starting point to clarify this issue and a first step towards a truly value-based PRO collection. Unfortunately, the Dutch arthroplasty register reported a mean RR of 37% on both the preoperative and the 3

month postoperative time points [16]. To achieve a MRR of 60%, adding manual collection steps to an automated PRO collection system is needed (Chapter 3). The advised MRR of 60% also provides an estimation of what costs are justified for now.

If the MRR of 60% becomes the threshold in the Netherlands, the RR can be monitored via the Dutch arthroplasty register [17, 18]. Although the percentage of health care institutions collecting THA PROs increased over the years, low RRs with large interquartile ranges reveal a large diversity in PRO collection. Only 16% of these institutions achieve the MRR of 60% and, therefore, have sufficient PROs to evaluate THAs from a patients' perspective (Chapter 5). Internationally, most registries reported that the RR was at least 40% but, disappointingly, one-third of the registries do not know their RR [1]. Also the PROMs working group of International Society of Arthroplasty Registries (ISAR) advises a MRR of 60%. However, they mention that this is only based on the external difficulties to collect PROs that may be unrelated to survey logistics and the requirement of ≥60% for a survey study [11, 19], without any further scientific evidence. Based on chapter 4, ISAR has to tighten up its MRR advice and subsequently health care institutions have to increase their RRs to 60%.

Conversely, if multiple health care institutions do not achieve the MRR of 60%, it raises the question whether it is worth to continue PRO collection. Continuing PRO collection in its current form including the effort and costs might not be justifiable from an ethical and value-based health care perspective. Therefore, a coordinated effort of all stakeholders has to be initiated to improve routine PRO collection in daily health care to achieve sufficient data quality.

Apart from the results of one Dutch orthopaedic health care institution (Chapters 3 and 4), more insight into the process of routine PRO collection including effort, costs and data quality is needed to further optimize routine PRO collection.

Implementation and monitoring of PRO collection is beyond the scope of this thesis, but needs to be mentioned as an important part of the recommended coordinated effort.

Recommendations on how to optimize PRO collection in hip and knee arthroplasty

- Incorporate PRO collection in daily health care routine for all operated orthopaedic patients. This will decrease costs per patient and increase the RR (Chapter 3).
- Add manual PRO collection steps to an automated system to achieve a MRR of 60% (Chapters 3 and 4).

Part II: Optimizing health care with routine use of PROs

When routine PRO collection is optimized, PROs at least have to be used for the aim(s) they are collected for. Patients expect data or advice in return for completing their PROMs. This will most likely positively influence their participation and adherence. PROs add the patients' perspective on the outcome in a valid and reliable manner which is needed in addition to all other outcomes,

such as complication rate and implant survival rate, to assess the quality of health care. Therefore, it is expected that, if all stakeholders take PROs into account, health care will improve.

Multiple joint arthroplasty registries have incorporated PROs to improve arthroplasty health care as their main aim. Unfortunately, the patient-reported quality of the received THA health care remained equal in the Netherlands between 2016 and 2019. However, only 16% of the Dutch health care institutions achieved the 60% MRR. It is, therefore, unlikely that the current Dutch THA PROs are sufficiently useful to improve health care on a national level (Chapter 5). Unfortunately, the same conclusion can be stated regarding the international arthroplasty PROs [1]. This means that there is no reply yet on the question whether the main aim of arthroplasty registries by collecting PROs is achieved.

So, in what direction PRO collection and use have to be developed to improve quality of arthroplasty health care at a (inter)national level?

1. Investigate if stakeholders use the collected PROs to evaluate arthroplasty health care. Underlying the movement towards incorporating PRO collection in daily health care is the assumption that if PROs are made available, they will be used. However, studies examining this assumption have found limited use of PROs. Main reasons according to surgeons are a lack of knowledge on how to use PROs in daily health care, the perception that PROs do not provide actionable information, and because gathering and handling of PROs add work to an already busy schedule [20, 21]. In addition, orthopaedic surgeons state that using PROs on an individual patient level is difficult based on logistical barriers (access and display issues, time required) and perceptual barriers (concerns about patients understanding, and validity and reliability of measures). They prefer to talk with patients about personal outcomes. However, they mention that using PROs on an aggregated level is valuable for health care institutions and individual surgeons [22].
2. Support stakeholders to evaluate arthroplasty outcomes from a patients' perspective using the already existing multiple examples and recommendations how to use the PROs [12, 13] (Chapters 6 to 9). Barriers have to be taken away.
3. Investigate how all stakeholders rate the quality of arthroplasty health care provided today. Of course, improvement is always desirable, however, there might be a consensus that the delivered quality is already of such a high level that improvement is unlikely or that the desired improvement is not value-based.
4. Increase the RRs to at least 60% to improve the data quality. Recommendations are provided in part I of this thesis.
5. Evaluate the set aim(s) of PROs. Maybe the goal of improving arthroplasty health care is not achievable or not formulated well. Each aim sets different requirements for the PRO(M)s, time points of collecting PROs and statistical analysis. The primary aim is the basis.

These five points need to be part of the recommended coordinated effort to improve routine PRO collection and to optimize health care with routine use of PROs in hip and knee arthroplasty.

Scientifically investigated examples of routine PRO use in daily orthopaedic health care

The use of PROs as a part of standard health care is justified, but future, adequately powered studies are necessary [23, 24]. In this thesis several examples of routine PRO use, based on a sufficient PRO collection (RR ≥60%), were investigated in daily health care in one Dutch orthopaedic health care institution.

The ***first*** example is a health care evaluation which shows that PROs are useful to gain knowledge for shared decision making and for making recommendations to stakeholders. Even though implant manufacturers claim the superiority of their implant design, for patients with end-stage medial knee osteoarthritis a medial UKA performed by a high-volume surgeon with either a mobile bearing or a fixed bearing implant design is a successful treatment option according to the PROs. Both designs result in excellent patient satisfaction, pain relief, functional improvement and quality of life improvement at 6, 12 and 24 months after surgery (Chapter 6). This means that recommendation and use of one over the other is not justified. This result adds knowledge from a patients' perspective to what is known from previous studies: with strict patient selection and accurate implant positioning, excellent functional outcomes, implant survival rates and complication rates for both implant designs are achievable [25–29].

A ***second*** example of using PROs to gain knowledge useful for shared decision making and for making recommendations to stakeholders is in preoperatively predicting postoperative patient satisfaction. Dissatisfaction after a TKA (up to 20% of the patients [30–37]) remains a difficult problem. Being able to preoperatively predict the outcome might improve patient selection for a TKA. Unfortunately, the degree of patient satisfaction and the chance of being dissatisfied or satisfied at 6, 12 and 24 months after a TKA are predictable by patient characteristics and preoperative PROs but not at a reliability level that is clinically useful (Chapter 7). In future studies, preoperatively set expectations should be included as not fulfilling these expectations is associated with a 10.7 times higher risk of patient dissatisfaction one year after a TKA [32]. It might be possible to influence preoperatively set (possibly) unrealistic expectations using PROs. Currently, a randomized controlled trial has been performed on fulfilling or exceeding of preoperative expectations by reporting patients' own PROs in relation to PROs of patients who underwent a TKA [38].

A main reason why PRO collection and use are limitedly integrated in health care is that most PROMs used in arthroplasty are not validated on an individual patient level [22]. PROs at an individual level could contain much information but caution have to be taken into account when interpreting these PROs. Therefore, as the ***third*** example, it was questioned if PROs could be used to individually guide patients after surgery. TKA patients feel uncertain and left alone after discharge which negatively affect their pain coping and subsequent management [39, 40]. The PainCoach app, a novel application of PROs incorporated in a wearable device, was developed to satisfy the need of individualized guidance after surgery. This app is not based on sufficient routine PRO collection of a health care institution, but on sufficient routine PRO input of the patient self. The randomized controlled trial shows that the use of the PainCoach app contributes

to reduced opiate use while patients reported similar pain levels compared to usual care only. Interestingly, daily use of this app leads to a further reduction in opiate use and to improved pain control (Chapter 8). Apps like this can have a prominent place in the battle of reducing the current wave of increasing opiate use and its undesired health consequences. The app scores high on usability, likelihood of being recommended to others and added value (Chapter 8). This makes the PainCoach app a successful pain management tool showing that PROs are useful to guide patients.

Video consultation, another digital solution, became the alternative for the traditional hospital consultation during the COVID-19 pandemic [41]. As it remained unknown for which patient video consultation was suitable, it was investigated if PROs could be used to triage patients to video or traditional hospital consultation as the ***fourth*** example. A tool for the consultation 6 weeks after a THA including moderate or severe pain and using two crutches was developed. Executing this tool on a retrospective cohort shows that at least 70% of the patients can safely have a video consultation (Chapter 9). The next step is to further investigate this triage tool in daily health care. Interestingly, several Dutch health care insurance companies aim to have at least 25% digital consultations in 2023 or mention 'digital health care when possible, physical health care when needed'. In their opinion digital health care can play a role in shortage of staff; efficiency, effectivity and affordability of health care; reducing waiting listing; and encouraging future-proof and sustainable health care [42]. As the number of patients with osteoarthritis is expected to increase with 40% between 2015 and 2040 [43], digital solutions, such as the PainCoach app and the triage tool, have great potential to ease the burden on future health care demands.

The discussion on PRO(M)s in daily health care is probably more complex then mentioned so far

PROs need to be interpreted in combination with other health care outcomes to have a complete picture of the delivered quality of health care. Scores or change scores are influenced by patient level variables (e.g. patient demographics; social status such as educational level, working status and social network; psychical status and expectations), health care professional level variables (e.g. volume, years of experience and access to materials such as implants), health care institutional level variables (e.g. preoperative assessment arrangements, rehabilitation protocols and access to post-discharge facilities) and (inter)national health care level (e.g. police, reimbursement and involvement of stakeholders) [24, 44, 45]. When possible, these variables were taken into account in this thesis. Caution with use of PROs, the use of a product based on PROs or the degree of generalizable were mentioned when needed.

Moreover, both on an aggregated and individual level, there is a lack of an established standard (minimal clinical important difference (MCIDs) or comparable values) on what change in PROs should be achieved when conducting an arthroplasty [46, 47]. Furthermore, there is no gold standard for the method to determine this standard [48]. In this thesis statistical difference and clinical reasoning were used. Further research has to focus on quantifying this standard as even on the best method to determine this.

Are PROMs the most suitable instruments to measure PROs?

PROMs, the gold standard to measure PROs, have their limitations. Patient-Reported Outcomes Measurement Information System (PROMIS) has the potential to be more valid, reliable and responsive [49, 50]. The generic short forms, a type of PROMIS, are replied digitally and on paper forms. However, computer adaptive testing, another type of PROMIS, can only be performed digitally [51] and is, therefore, not suitable to reach the lower limit of 60% RR (Chapters 3 and 4). Currently, a prospective cohort study in three health care institutions has been executed to investigate which instrument is preferred in hip and knee arthroplasty, separately [52]. If PROMIS is preferred, crosswalks between PROMs and PROMIS need to be developed and be publicly available to maintain the value of already collected PROs with PROMs.

A combination of PROMs, PROMIS and wearable devices could also be the answer. Already today, wearable devices, such as smartphone apps and smart watches, show to be capable of monitoring physical activity and of improving patient engagement following a TKA [53]. Wearable devices will give the subjective PROMs a more objective insight. Previous studies comparing wearable devices to arthroplasty PROMs are limited [54, 55]. Future research is needed to investigate if wearable devices are the most suitable instruments to measure PROs or a part of the PROs.

THE FUTURE BEGINS TODAY

Since the introduction of PRO(M)s it can be concluded that the patients' perspective is increased in hip and knee arthroplasty. Nevertheless, still a considerable amount of work and analysis need to be performed until significant benefits with respect to patient care, outcomes and quality improvement are perceived [56]. To optimize routine PRO collection and to optimize health care with routine use of PROs, a coordinated effort is needed. If all stakeholders work together, improvement can be realized. Based on this thesis, this effort need to included:

- Collect PROs of and use PROs for all operated orthopaedic patients (Chapter 3).
- Increase the RR to the lower limit of 60% (Chapter 4).
- Explore on the MRR (Chapter 4).
- Create more insight into the process of routine PRO collection including effort, costs and data quality (Chapters 3 and 4).
- Incorporate routine PRO collection and use in daily health care.
- Create assess for all stakeholders to collected PROs and examples of PRO use.
- Investigate if stakeholders use PROs to optimize health care (Chapter 5).
- Support stakeholders to use PROs to optimize health care (Chapter 5).
- Investigate what each stakeholder's opinion is on the degree of quality of arthroplasty health care today and if improvement is achievable (Chapter 5).
- Evaluate the set aim(s) of PROs (Chapter 5).
- Support stakeholders to share their examples of PRO use to optimize health care (Chapters 6 to 9).
- Investigate which (combination of) instrument(s) is the most suitable to measure PROs.

Several recommendations are highlighted below.

Coordinated effort in arthroplasty health care

Already incorporated several of aforementioned recommendations, the Dutch orthopaedic association revised their hip and knee arthroplasty PROMs advice in cooperation with orthopaedic surgeons, physiotherapists, patient federation, rehabilitation specialists, scientists and the Dutch arthroplasty register in 2020 [57]. The main goal of PRO collection and use switched from improving quality of health care to improving individual patient care. Short term goals included the incorporation of PRO collection in daily health care. Long term goals were formulated to stimulate research on and with PRO(M)s. Currently, several studies have been performed and funding is received to execute this PROMs advice [58, 59]. This shows a recognition for the value of PRO(M)s in hip and knee arthroplasty and a recognition that effort is needed to optimize PRO collection and to optimize health care with the use of PROs. These are the first steps to improvement.

Starting today, the Dutch orthopaedic association has to included PROs and RRs in both the outlier analysis and conversations with health care institutions [60]. This association mentions 99% completeness and more than 90% validity of the Dutch arthroplasty register [61, 62]. However, these numbers do clearly not include the registered PROs and RRs. All information in the registry has to be used to optimize the quality of arthroplasty health care. Additionally, this can lead to an optimization of the RR to the lower limit of 60%.

Furthermore, starting today, the Dutch arthroplasty register, other registries and stakeholders are recommended to include the MRR of 60% when reporting about quality of arthroplasty health care to improve the quality of their analyses and conclusions.

National coordinated effort

With the focus on the total Dutch health care optimization, the federation of medical specialists together with health care institutions, association of nurses and paramedics, health insurance companies, government and patient federation reached a 4-years agreement (in Dutch: hoofdlijnenakkoord medisch-specialistische zorg 2019 - 2022). They aim to improve quality and efficiency of health care, and to guarantee accessibility and payable health care on the long term [63]. This agreement resulted in several national thematic programmes such as 'Outcome related health care' (in Dutch: Uitkomstgerichte zorg) and 'Health care evaluation and suitable use' (in Dutch: Zorgevaluatie en Gepast Gebruik) which incorporate several aforementioned recommendations [64, 65]. It may be possible that using PROs will be more effective in other medical areas than arthroplasty, because the outcomes of certainly hip and too a less amount knee arthroplasty are relatively exceptional [66].

Incorporating routine PRO collection and use in daily health care is one of the main aforementioned recommendations. This includes understanding of the value of PROs, listing and investigating knowledge gaps, implementing PRO(M)s, and monitoring collection and use. If PROs are a routine

part of health care, health care evaluations can be performed easier (Chapter 6). Difficult to quantify scientifically, half of the executed health care nowadays has no sufficient scientific evidence for its effectiveness [67, 68]. When this main recommendation is achieved, patients' perspective becomes a standard part of these evaluations. The developed generic PRO(M)s set will be a substantial step forwards in optimizing routine PRO collection and use. This set is a small set of PROs and corresponding PROMs suitable for most patients [69]. Specific hip and knee osteoarthritis outcome sets, based on the generic PRO(M)s set, are developed as well [70–72]. As a first benefit, these sets can realize the recommended PRO collection of all operated orthopaedic patients as even of all other patients. Secondly, the patients' perspective of arthroplasty health care can be compared to, for example, the patients' perspective of oncologic health care to quantify the value of arthroplasty health care (inter)nationally. Thirdly, these sets can be a solution for the overload of PROMs which arthroplasty patients can obtain due to visiting other health care professionals besides an orthopaedic surgeon. After investigating the sets in daily health care, barriers have to be solved with all stakeholders to successfully implement, collect and use these sets in daily health care.

Accessibility and visibility of PROs have to be improved to enhance routine PRO use. Although arthroplasty PROs are publicly available for all stakeholders on a governmental website [2] and in limited edition on the Dutch arthroplasty register website [66], it is not easy to use PROs mainly due to low visibility of these locations and how the PROs are presented. Creating access to collected PROs, interpreting PROs in combination with other health care outcomes and supporting stakeholders to share examples of the use of outcomes are imbedded in the programme 'Outcome related health care' [73]. Of course, guiding or teaching stakeholders in how to interpret these outcomes is needed as well.

A coordinated effort of all stakeholders will lead to promising steps in optimizing routine PRO collection and optimizing health care with routine use of PROs.

CONCLUSION

A MRR of 60% is advised to adequately evaluate THA PROs. To achieve this MRR an automated PRO collection system with additional manual steps is needed, adding a cost of €6 per surgical procedure to automated PRO collection alone. Unfortunately, only 16% of the health care institutions achieved this MRR for THA. It remains, therefore, questionable if arthroplasty registries could achieved their aim 'to improve arthroplasty health care on a (inter)national level' by incorporating PROs. When routine PRO collection is sufficient (RR ≥60%), PROs are routinely useful to optimize health care to gain knowledge useful for shared decision making and for making recommendations to stakeholders, in health care evaluation, to guide patients and to triage patients. Although there is still a considerable amount of work to perform before routine PRO collection is optimal and health care is optimal by routinely using PRO in hip and knee arthroplasty, a coordinated effort of all stakeholders will lead to promising next steps.

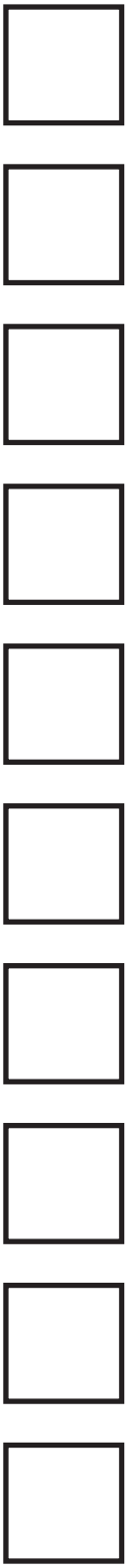
REFERENCES

1. Bohm ER, Kirby S, Trepman E, et al (2021) Collection and Reporting of Patient-reported Outcome Measures in Arthroplasty Registries: Multinational Survey and Recommendations. Clin Orthop Relat Res 479:2151–2166
2. Zorginstituut Nederland. Zorginzicht - Open data Medisch-specialistische zorg. https://www.zorginzicht.nl/openbare-data/open-data-ziekenhuizen-en-zelfstandige-behandelcentra---medisch-specialistische-zorg. Accessed 24 Nov 2021
3. Ho A, Purdie C, Tirosh O, Tran P (2019) Improving the response rate of patient-reported outcome measures in an Australian tertiary metropolitan hospital. Patient Relat Outcome Meas 10:217–226
4. Patel J, Lee JH, Li Z, SooHoo NF, Bozic K, Huddleston JI (2015) Predictors of Low Patient-Reported Outcomes Response Rates in the California Joint Replacement Registry. J Arthroplasty 30:2071–2075
5. Tariq MB, Vega JF, Westermann R, Jones M, Spindler KP (2019) Arthroplasty studies with greater than 1000 participants: analysis of follow-up methods. Arthroplast Today 5:243–250
6. Slover JD, Karia RJ, Hauer C, Gelber Z, Band PA, Graham J (2015) Feasibility of integrating standardized patient-reported outcomes in orthopedic care. Am J Manag Care 21:e494-500
7. Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I, Cooper R, Felix LM, Pratap S (2009) Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2010:2009–2011
8. Hunter J, Corcoran K, Leeder S, Phelps K (2013) Is it time to abandon paper? The use of emails and the Internet for health services research - a cost-effectiveness and qualitative study. J Eval Clin Pract 19:855-61
9. Brook EM, Glerum KM, Higgins LD, Matzkin EG (2017) Implementing Patient-Reported Outcome Measures in Your Practice: Pearls and Pitfalls. Am J Orthop (Belle Mead NJ) 46:273–278
10. Franklin PD, Chenok KE, Lavalee D, Love R, Paxton L, Segal C, Holve E (2017) Framework To Guide The Collection And Use Of Patient-Reported Outcome Measures In The Learning Healthcare System. eGEMs (Generating Evid Methods to Improv patient outcomes) 5:17
11. Rolfson O, Bohm E, Franklin P, et al (2016) Patient-reported outcome measures in arthroplasty registries: Report of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty RegistriesPart II. Recommendations for selection, administration, and analysis. Acta Orthop 87:9–23
12. Boyce MB, Browne JP, Greenhalgh J (2014) The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. BMJ Qual Saf 23:508–518
13. Van Schie P, Van Bodegom-Vos L, Zijdeman TM, Nelissen RGHH, Marang-Van De Mheen PJ (2021) Awareness of performance on outcomes after total hip and knee arthroplasty among Dutch orthopedic surgeons: how to improve feedback from arthroplasty registries. Acta Orthop 92:54–61
14. Baumhauer JF, Bozic KJ (2016) Value-based Healthcare: Patient-reported Outcomes in Clinical Decision Making. Clin Orthop Relat Res 474:1375–1378
15. Pronk Y, Pilot P, Brinkman JM, van Heerwaarden RJ, van der Weegen W (2019) Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection. J Patient-Reported Outcomes 3:31
16. Nederlandse Orthopaedische Vereniging. Three months postoperative PROMs response percentage hip. http://www.lroi-rapportage.nl/hip-2018-proms-hip-2018-proms-response-three-months-postoperative-proms. Accessed 28 Aug 2019
17. LROI. Hip PROMs Response. https://www.lroi-report.nl/hip/proms/response/. Accessed 2 Dec 2021
18. LROI. Knee PROMs Response. https://www.lroi-report.nl/knee/proms/response/. Accessed 2 Dec 2021
19. JAMA. Internal Medicine Instructions for Authors. https://jamanetwork.com/journals/jamainternalmedicine/pages/instructions-for-authors. Accessed 1 Aug 2018
20. Alshehri F, Alarabi A, Alharthi M, Alanazi T, Alohali A, Alsaleem M (2020) Use of patient-reported outcome measures (PROMs) by orthopedic surgeons in Saudi Arabia. J Orthop Surg Res 15:598
21. Mou D, Sisodia RC, Castillo-Angeles M, Ladin K, Bergmark RW, Pusic AL, del Carmen MG, Heng M (2022) The Surgeon's Perceived Value of Patient-reported Outcome Measures (PROMs). Ann Surg 275:500–505

22. Whitebird RR, Solberg LI, Ziegenfuss JY, Norton CK, Chrenka EA, Swiontkowski M, Reams M, Grossman ES (2022) What Do Orthopaedists Believe is Needed for Incorporating Patient-reported Outcome Measures into Clinical Care? A Qualitative Study. Clin Orthop Relat Res 480:680–687
23. Ishaque S, Karnon J, Chen G, Nair R, Salter AB (2019) A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). Qual Life Res 28:567–592
24. Browne JP, Cano SJ, Smith S (2017) Using Patient-reported Outcome Measures to Improve Health Care. Med Care 55:901–904
25. Cao Z, Niu C, Gong C, Sun Y, Xie J, Song Y (2019) Comparison of Fixed-Bearing and Mobile-Bearing Unicompartmental Knee Arthroplasty: A Systematic Review and Meta-Analysis. J Arthroplasty 34:3114-3123.e3
26. Huang F, Wu D, Chang J, Zhang C, Qin K, Liao F, Yin Z (2021) A Comparison of Mobile- and Fixed-Bearing Unicompartmental Knee Arthroplasties in the Treatment of Medial Knee Osteoarthritis: A Systematic Review and Meta-analysis of 1,861 Patients. J Knee Surg 34:434–443
27. Peersman G, Stuyts B, van den Langenbergh T, Cartier P, Fennema P (2015) Fixed- versus mobile-bearing UKA: a systematic review and meta-analysis. Knee Surgery, Sport Traumatol Arthrosc 23:3296–3305
28. The New Zealand Joint Registry (2019) The New Zealand Joint Registry Twenty Year Report January 1999 To December 2018. https://nzoa.org.nz/system/files/DH8328_NZJR_2019_Repor t_v4_7Nov19.pdf. Accessed 11 Mar 2020
29. Von Keudell A, Sodha S, Collins J, Minas T, Fitz W, Gomoll A (2014) Patient satisfaction after primary total and unicompartmental knee arthroplasty: An age-dependent analysis. Knee 21:180–184
30. Ali A, Lindstrand A, Sundberg M, Flivik G (2017) Preoperative Anxiety and Depression Correlate With Dissatisfaction After Total Knee Arthroplasty: A Prospective Longitudinal Cohort Study of 186 Patients, With 4-Year Follow-Up. J Arthroplasty 32:767–770
31. Baker PN, Rushton S, Jameson SS, Reed M, Gregg P, Deehan DJ (2013) Patient satisfaction with total knee replacement cannot be predicted from pre-operative variables alone. Bone Joint J 95-B:1359–1365
32. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ (2010) Patient Satisfaction after Total Knee Arthroplasty: Who is Satisfied and Who is Not? Clin Orthop Relat Res 468:57–63
33. Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ (2018) The rate and predictors of patient satisfaction after total knee arthroplasty are influenced by the focus of the question. Bone Joint J 100-B:740–748
34. Noble PC, Conditt MA, Cook KF, Mathis KB (2006) The John Insall Award: Patient Expectations Affect Satisfaction with Total Knee Arthroplasty. Clin Orthop Relat Res 452:35–43
35. Van Onsem S, Van Der Straeten C, Arnout N, Deprez P, Van Damme G, Victor J (2016) A New Prediction Model for Patient Satisfaction After Total Knee Arthroplasty. J Arthroplasty 31:2660-2667.e1
36. Schnurr C, Jarrous M, Güdden I, Eysel P, König DP (2013) Pre-operative arthritis severity as a predictor for total knee arthroplasty patients' satisfaction. Int Orthop 37:1257–1261
37. Scott CEH, Howie CR, MacDonald D, Biant LC (2010) Predicting dissatisfaction following total knee replacement. J Bone Joint Surg Br 92-B:1253–1258
38. Peters MCWM. Personalized Information Before and After Total Knee Arthroplasty: PAK-study (PAK). https://clinicaltrials.gov/ct2/show/NCT05266482?term=NCT05266482&draw=2&rank=1. Accessed 4 Aug 2022
39. Specht K, Agerskov H, Kjaersgaard-Andersen P, Jester R, Pedersen BD (2018) Patients' experiences during the first 12 weeks after discharge in fast-track hip and knee arthroplasty – a qualitative study. Int J Orthop Trauma Nurs 31:13–19
40. Sjøveian AKH, Leegaard M (2017) Hip and knee arthroplasty - patient's experiences of pain and rehabilitation after discharge from hospital. Int J Orthop Trauma Nurs 27:28–35
41. Merkes K, Hagenaars N (2020) Corona: katalysator of struikelblok voor groenere ziekenhuiszorg? https://gupta-strategists.nl/studies/corona-katalysator-of-struikelblok-voor-groenere-ziekenhuiszorg. Accessed 13 Nov 2021
42. BDO (2022) Zorginkoopvergelijker 2023 - Medisch Specialistische Zorg.
43. Nederlandse Orthopaedische Vereniging (2019) Dossier artrose. Zorg voor Beweg 22–27
44. Oldsberg L, Garellick G, Osika Friberg I, Samulowitz A, Rolfson O, Nemes S (2019) Geographical variations in patient-reported outcomes after total hip arthroplasty between 2008 - 2012. BMC Health Serv Res 19:343

45. Greenhalgh J, Gooding K, Gibbons E, Dalkin S, Wright J, Valderas J, Black N (2018) How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. J Patient-Reported Outcomes 2:42
46. Prodinger B, Taylor P (2018) Improving quality of care through patient-reported outcome measures (PROMs): expert interviews using the NHS PROMs Programme and the Swedish quality registers for knee and hip arthroplasty as examples. BMC Health Serv Res 18:87
47. Çelik D, Çoban Ö, Kılıçoğlu Ö (2019) Minimal clinically important difference of commonly used hip-, knee-, foot-, and ankle-specific questionnaires: a systematic review. J Clin Epidemiol 113:44–57
48. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, Croft P, de Vet HCW (2010) Mind the MIC: large variation among populations and methods. J Clin Epidemiol 63:524–534
49. Karhade A V., Bernstein DN, Desai V, Bedair HS, O'Donnell EA, Tanaka MJ, Bono CM, Harris MB, Schwab JH, Tobert DG (2022) What Is the Clinical Benefit of Common Orthopaedic Procedures as Assessed by the PROMIS Versus Other Validated Outcomes Tools? Clin Orthop Relat Res 480:1672–1681
50. Baron JE, Parker EA, Wolf BR, Duchman KR, Westermann RW (2021) PROMIS Versus Legacy Patient-Reported Outcome Measures for Sports Medicine Patients Undergoing Arthroscopic Knee, Shoulder, and Hip Interventions: A Systematic Review. Iowa Orthop J 41:58–71
51. Terwee CB, Roorda LD, de Vet HCW, et al (2014) Dutch–Flemish translation of 17 item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS). Qual Life Res 23:1733–1741
52. Wolterbeek N, Terwee C, Braaksma C, Veen M, Roorda L, Willigenburg N, Poolman R, Pronk Y, Brinkman J. Summary of research proposal LROI. https://www.lroi.nl/downloads/216/abstract-lroi-website-lroi-rg-2021-001-wolterbeek.pdf. Accessed 4 Aug 2022
53. Constantinescu D, Pavlis W, Rizzo M, Vanden Berge D, Barnhill S, Hernandez VH (2022) The role of commercially available smartphone apps and wearable devices in monitoring patients after total knee arthroplasty: a systematic review. EFORT Open Rev 7:481–490
54. Jelsma J, Van Kuijk S, Buil I, Heyligers I, Grimm B, Schotanus M (2021) Only limited correlations between patient-reported outcomes and objectively monitored physical activity 10-years after THA. Acta Orthop Belg 87:593–599
55. Bendich I, Chung C, Hwang K, Patterson J, Mulvihill J, Barry J, Bini S (2019) Changes in prospectively collected longitudinal patient-generated health data are associated with short-term patient-reported outcomes after total joint arthroplasty: a pilot study. Arthroplast Today 5:61–63
56. Wilson I, Bohm E, Lübbeke A, Lyman S, Overgaard S, Rolfson O, W-Dahl A, Wilkinson M, Dunbar M (2019) Orthopaedic registries with patient-reported outcome measures. EFORT Open Rev 4:357–367
57. Nederlandse Orthopaedische Vereniging (2020) NOV PROMs-advies orthopedie 2020. https://www.orthopeden.org/downloads/775/nov-proms-advies.pdf. Accessed 7 Jul 2022
58. ZonMw (2022) Project: Verduurzamen uitkomstgerichte zorg: showcase knieartrose. https://www.zonmw.nl/nl/onderzoek-resultaten/kwaliteit-van-zorg/programmas/project-detail/kwaliteitsgelden/verduurzamen-uitkomstgerichte-zorg-knieartrose/. Accessed 7 Jul 2022
59. LROI. Lopend onderzoek. https://www.lroi.nl/wetenschap/lopend-onderzoek. Accessed 7 Jul 2022
60. Nederlandse Orthopaedische Vereniging. Procedure Uitkomstanalyse. https://www.orthopeden.org/kwaliteit/kwaliteitsbeleid/procedure-uitkomstanalyse. Accessed 20 Dec 2021
61. LROI. Data Quality - Completeness. https://www.lroi-report.nl/data-quality/completeness/. Accessed 20 Dec 2021
62. LROI. Data Quality - Validity. https://www.lroi-report.nl/data-quality/validity/. Accessed 7 Jul 2022
63. Federatie Medisch Specialisten. Hoofdlijnenakkoord. https://demedischspecialist.nl/themas/thema/hoofdlijnenakkoord. Accessed 4 Aug 2022
64. Programma Uitkomstgerichte Zorg. Doelen uitkomstgerichte zorg. https://www.uitkomstgerichtezorg.nl/. Accessed 13 Oct 2022
65. Programma Zorgevaluatie en Gepast Gebruik. Samen voor de bewezen beste zorg voor patiënten. https://zorgevaluatiegepastgebruik.nl/. Accessed 4 Aug 2022
66. LROI. LROI Report. https://www.lroi-report.nl/. Accessed 29 Aug 2022
67. van der Geest M (2019) Sjoerd Repping, de man die alle bewezen onzinzorg uit de ziekenhuizen moet schrappen. de Volkskrant
68. Stuurgroep Zorgevaluatie (2016) Adviesrapport Zorgevaluatie - Van project naar proces.

69. Programma Uitkomstgerichte Zorg. Inzicht in uitkomsten - Generieke PROMs. https://platformuitkomstgerichtezorg.nl/themas/inzicht+in+uitkomsten/default.aspx. Accessed 4 Aug 2022
70. Programma Uitkomstgerichte Zorg (2022) Eindrapport aandoeningswerkgroep knieartrose.
71. Programma Uitkomstgerichte Zorg. Meer inzicht in uitkomsten - stand van zaken (januari 2022). https://www.uitkomstgerichtezorg.nl/themas/meer-inzicht-in-uitkomsten/stand-van-zaken-sept.-2021. Accessed 4 Aug 2022
72. Programma Uitkomstgerichte Zorg. ICT en toegankelijkheid - stand van zaken (september 2021). https://www.uitkomstgerichtezorg.nl/themas/ict-en-toegankelijkheid/stand-van-zaken. Accessed 4 Aug 2022
73. Programma Uitkomstgerichte Zorg. Over Uitkomstgerichte Zorg - aanpak Uitkomstgerichte zorg. https://www.uitkomstgerichtezorg.nl/over-uitkomstgerichte-zorg/aanpak-uitkomstgerichte-zorg. Accessed 4 Aug 2022



APPENDICES

Summary (Dutch – Samenvatting)

Context

Sinds de verschuiving naar een meer patiënt gecentraliseerde orthopedische gezondheidszorg is het meten van patiënt gerapporteerde uitkomsten (patient-reported outcomes; PROs) van heup- en knieprothese patiënten toegenomen. PROs geven een inzicht in het behandelresultaat vanuit het perspectief van de patiënt. Voor het meten van deze PROs zijn patiënt gerapporteerde uitkomstmetingen (patient-reported outcome measures; PROMs) de gouden standaard. Er zijn voorbeelden en aanbevelingen over hoe PROs te verzamelen en te gebruiken. De dagelijkse praktijk laat echter zien dat er een grote diversiteit is tussen zorginstellingen betreffende hun succes in het verzamelen van PROs en hoe zij PROs gebruiken om de gezondheidszorg te verbeteren. Zo is er een grote spreiding in responspercentages op de PROMs zichtbaar. Vele PRO gerelateerde vragen zijn nog onbeantwoord. Het is daarom de vraag of de huidige PRO verzameling en gebruik te rechtvaardigen zijn, zowel ethisch als vanuit waardegedreven zorg gezien. Het doel van dit proefschrift was om te onderzoeken hoe PRO verzameling kan worden geoptimaliseerd (deel I) en vervolgens hoe de gezondheidszorg kan worden geoptimaliseerd met gebruik van PROs (deel II) in de dagelijkse praktijk in heup en knie prothesiologie (**Hoofdstukken 1 en 2**).

Dit proefschrift

Deel I van dit proefschrift richt zich op hoe PRO verzameling kan worden geoptimaliseerd in de dagelijkse praktijk in heup en knie prothesiologie. Een retrospectieve cohort studie beschrijft het responspercentage en de kosten voor een PRO verzameling met een automatisch verzamelsysteem vergeleken met een automatisch PRO verzamelsysteem met additionele handmatige verzameling. Dit werd vergeleken voor alle orthopedische operaties samen en specifiek voor totale heup prothesiologie, knie prothesiologie en voorste kruisband reconstructies (**Hoofdstuk 3**). Vervolgens geeft een retrospectieve cohort studie inzicht in het minimaal benodigde responspercentage op PROMs om totale heupprotheses adequaat te kunnen evalueren (**Hoofdstuk 4**).

Deel II van dit proefschrift focust zich op hoe de gezondheidszorg kan worden geoptimaliseerd met gebruik van PROs in de dagelijkse praktijk in heup en knie prothesiologie. Een longitudinale studie met openbaar beschikbare, nationale, totale heupprothese indicator datasets laat zien of het doel 'verbeteren van heup prothesiologie gezondheidszorg' middels PRO verzameling is bereikt in Nederland (**Hoofdstuk 5**). Daarna worden wetenschappelijke voorbeelden van PRO gebruik gegeven. In een retrospectieve cohort studie worden PROs gebruikt om twee veel gebruikte unicondylaire knieprothese implantaten designs (mobiel en gefixeerd) met elkaar te vergelijken. Dit eerste voorbeeld geeft inzicht in welk implantaat design volgens patiënten dient te worden gebruikt in de dagelijkse praktijk (**Hoofdstuk 6**). Een tweede voorbeeld focust zich op het preoperatief voorspellen van patiënttevredenheid na een totale knieprothese. Dit werd onderzocht in een retrospectieve cohort studie (**Hoofdstuk 7**). Een gerandomiseerde studie onderzocht het effect van een op PRO gebaseerde eHealth app (PijnCoach app) op pijnbeleving en opiaat gebruik tijdens de eerste periode thuis na een totale knieprothese als derde voorbeeld

(**Hoofdstuk 8**). Een vierde voorbeeld, een vooronderzoek met een panel van deskundigen en een retrospectieve cohort, beschrijft het gebruik van PROs voor het ontwikkelen van een hulpmiddel om totale heupprothese patiënten te selecteren voor of een traditioneel ziekenhuisbezoek nodig is of dat een video consult als alternatief adequaat is (**Hoofdstuk 9**).

In alle retrospectieve cohort studies werden prospectief verzamelde data gebruikt.

Wat voegt dit proefschrift toe aan het huidige wetenschappelijke bewijs over PRO(M)s in de prothesiologie gezondheidszorg?

- Een automatisch PRO verzamelsysteem met een additionele handmatige verzameling resulteert in een twee keer hogere responspercentage op de PROMs op alle tijdsmomenten vergeleken met het gebruik van alleen een automatisch systeem. Deze additionele handmatige stappen voegen extra kosten van €6 per operatie toe (**Hoofdstuk 3**).
- Om een valide conclusie te kunnen trekken op PROs wordt een minimaal responspercentage van 60% geadviseerd (**Hoofdstuk 4**).
- Slechts 16% van de Nederlandse zorginstellingen bereikt deze 60% met hun totale heupprothese PRO verzameling. Daarom blijft het de vraag of het doel 'verbeteren van prothesiologie gezondheidszorg' met PRO verzameling is bereikt in Nederland (**Hoofdstuk 5**).
- Wanneer de PRO verzameling in de dagelijkse praktijk adequaat is (responspercentage ≥60%), zijn PROs bruikbaar in de dagelijkse praktijk voor het optimaliseren van de gezondheidszorg. Wetenschappelijk onderzochte voorbeelden:
  - PROs zijn bruikbaar voor het verkrijgen van kennis voor samen beslissen en het doen van aanbevelingen richting belanghebbenden voor het kiezen van een implantaat design (mobiel of gefixeerd) bij patiënten met mediale unicondylaire knieprothese. Dit is een voorbeeld van zorgevaluatie (**Hoofdstuk 6**).
  - PROs zijn niet klinisch relevant bruikbaar voor het verkrijgen van kennis voor samen beslissen en het doen van aanbevelingen richting belanghebbenden bij het preoperatief voorspellen van patiënttevredenheid na een totale knieprothese (**Hoofdstuk 7**).
  - Een op PRO gebaseerde app, genaamd de PijnCoach app, is bruikbaar voor postoperatieve totale knieprothese patiënten in pijnbegeleiding inclusief opiaten gebruik (**Hoofdstuk 8**).
  - PROs zijn bruikbaar voor het selecteren van totale heupprothese patiënten in of een traditioneel ziekenhuisbezoek nodig is of dat een video consult als alternatief adequaat is (**Hoofdstuk 9**).

Gebaseerd op deze resultaten zijn de volgende stappen gedefinieerd om PRO verzameling in de dagelijkse praktijk in heup en knie prothesiologie te optimaliseren: naast een automatisch verzamelsysteem is handmatige inspanning, en daarmee extra budget, nodig voor een adequate (responspercentage ≥60%) PRO verzameling (deel I van dit proefschrift). Daarnaast, wanneer de PRO verzameling adequaat is, laten de voorbeelden zien hoe de gezondheidszorg geoptimaliseerd kan worden met het gebruik van PROs in de dagelijkse praktijk in heup en knie prothesiologie (deel II van dit proefschrift). Een gecoördineerde inspanning van alle belanghebbenden moet worden geïnitieerd om PRO verzameling én gebruik in de dagelijkse praktijk te verbeteren om de gezondheidszorg te optimaliseren (**Hoofdstuk 10**).

Data management

This thesis is based on the results of human studies, which were conducted in accordance with the principles of the Declaration of Helsinki. Data used within this thesis was collected and stored according to the Findable, Accessible, Interoperable and Reusable (FAIR) principles.

Data obtained during this PhD have been captured and stored on OnlinePROMs (Interactive Studios, Rosmalen, the Netherlands), a digital and online patient-reported outcome measure tool. Completed and uncompleted paper questionnaires were incorporated in OnlinePROMs and stored in a locked room at Kliniek ViaSana. Data management and monitoring were also performed within OnlinePROMs. Furthermore, data have been captured and stored on Chipsoft HIX (ChipSoft B.V., Amsterdam, the Netherlands), an electronic patient record. Datasets from both OnlinePROMs and Chipsoft HIX were combined. These datasets were stored, including daily backed-up, on the local Kliniek ViaSana server. The privacy of the patients in this thesis was warranted by use of encrypted and unique individual subject codes. All data archives are accessible by only those who need to.

Chapters 3 up to and including 7, and chapter 9 applied the Dutch orthopaedic association PROMs advice. The retrospective cohort studies described in chapters 3 (N18.156) and 4 (N18.156) were approved by the medical ethics committee of Maxima Medisch Centrum (Eindhoven, the Netherlands). For the study described in chapter 5, a publically available dataset was used. For the retrospective cohort studies described in chapters 6, 7 and 9, patients were included in this study if they signed the informed consent form preoperatively to allow further scientific analysis using their anonymised data. Therefore, the institutional review board of Kliniek ViaSana ruled that formal approval was not required for these studies. The medical ethics committee of St. Anna Hospital (Geldrop, the Netherlands, 5.12) approved the randomised controlled trial study described in chapter 8. This study was registered at Clinicaltrials.gov retrospectively (NCT03961152).

The data collected for this thesis will be available for further analyses for 15 years if patients signed the informed consent form preoperatively to allow further scientific analysis using their anonymised data. The datasets generated and analysed for this thesis are available from the corresponding author on reasonable request.

Scientific publications

1. Hoorntje A, **Pronk Y**, Brinkman JM, van Geenen RCI, van Heerwaarden RJ. High tibial osteotomy versus unicompartmental knee arthroplasty for Kellgren-Lawrence grade 3-4 knee osteoarthritis in younger patients: comparable improvements in patient-reported outcomes, adjusted for osteoarthritis grade and sex. Knee Surg Sports Traumatol Arthrosc. 2023 Aug 12. doi: 10.1007/s00167-023-07526-5. Online ahead of print. PMID: 37572139

2. **Pronk Y**, van der Weegen W, Schreurs BW, Pilot P. Quality of total hip arthroplasty health care based on four years of patient-reported outcomes in the Netherlands. Health Qual Life Outcomes. 2023 Mar 14;21(1):24. doi: 10.1186/s12955-023-02104-2. PMID: 36918879

3. van Eck J, Liu WY, Goosen JHM, Rijnen WHC, van der Zwaard BC, Heesterbeek P, van der Weegen W and **further members of Regional Prosthetic Joint Infection Group**. Higher 1-year risk of implant removal for culture-positive than for culture-negative DAIR patients following 359 primary hip or knee arthroplasties. J. Bone Joint Infect. 2022 Jul;7:143–149. doi: 10.5194/jbji-7-143-2022. PMID: 35937090

4. van Melick N, **Pronk Y**, Nijhuis-van der Sanden M, Rutten S, van Tienen T, Hoogeboom T. Meeting movement quantity or quality return to sport criteria is associated with reduced second ACL injury rate. J Orthop Res. 2022 Jan;40(1):117-128. doi: 10.1002/jor.25017. Epub 2021 Mar 11. PMID: 33650704

5. **Pronk Y**, Pilot P, van der Weegen W, Brinkman JM, Schreurs BW. A Patient-Reported Outcome Tool to Triage Total Hip Arthroplasty Patients to Hospital or Video Consultation: Pilot Study With Expert Panels and a Cohort of 1228 Patients. JMIR Form Res. 2021 Dec 20;5(12):e31232. doi: 10.2196/31232. PMID: 34931989

6. **Pronk Y**, Peters MCWM, Brinkman JM. Is Patient Satisfaction After Total Knee Arthroplasty Predictable Using Patient Characteristics and Preoperative Patient-Reported Outcomes? J Arthroplasty. 2021 Jul;36(7):2458-2465. doi: 10.1016/j.arth.2021.02.064. Epub 2021 Mar 2. PMID: 33741243

7. **Pronk Y**, Paters AAM, Brinkman JM. No difference in patient satisfaction after mobile bearing or fixed bearing medial unicompartmental knee arthroplasty. Knee Surg Sports Traumatol Arthrosc. 2021 Mar;29(3):947-954. doi: 10.1007/s00167-020-06053-x. Epub 2020 May 12. PMID: 32399616

8. Willems SJ, Coppieters MW, **Pronk Y**, Diks MJF, van der Heijden KWAP, Rooker S, Scholten-Peeters GGM. A Clinical Journey Mobile Health App for Perioperative Patients: Cross-sectional Study. JMIR Hum Factors. 2021 Feb 8;8(1):e20694. doi: 10.2196/20694. PMID: 33555262

9. **Pronk Y**, van der Weegen W, Vos R, Brinkman JM, van Heerwaarden RJ, Pilot P. What is the minimum response rate on patient-reported outcome measures needed to adequately evaluate total hip arthroplasties? Health Qual Life Outcomes. 2020 Dec 2;18(1):379. doi: 10.1186/s12955-020-01628-1. PMID: 33267842

10. van Melick N, Hoogeboom T, **Pronk Y**, Rutten S, van Tienen T, Nijhuis-van der Sanden M, van Cingel REH. Less than half of ACL-reconstructed athletes are cleared for return to play based on practice guideline criteria: results from a prospective cohort study. Int J Sports Phys Ther. 2020 Dec;15(6):1006-1018. doi: 10.26603/ijspt20201006. PMID: 33344017

11. Roelofsen EGJ, van Cingel R, **Pronk Y**, Staal JB, Nijhuis-van der Sanden MWG, Meulenbroek RGJ. Leg-amplitude differentiation guided by haptic and visual feedback to detect alterations in motor flexibility due to Total Knee Replacement. Hum Mov Sci. 2020 Jun;71:102623. doi: 10.1016/j.humov.2020.102623. Epub 2020 Apr 30. PMID: 32452440

12. **Pronk Y**, Peters MCWM, Sheombar A, Brinkman JM. Effectiveness of a Mobile eHealth App in Guiding Patients in Pain Control and Opiate Use After Total Knee Replacement: Randomized Controlled Trial. JMIR Mhealth Uhealth. 2020 Mar 13;8(3):e16415. doi: 10.2196/16415. PMID: 32167483

13. **Pronk Y**, Pilot P, Brinkman JM, van Heerwaarden RJ, van der Weegen W. Response rate and costs for automated patient-reported outcomes collection alone compared to combined automated and manual collection. J Patient Rep Outcomes. 2019 Jun 3;3(1):31. doi: 10.1186/s41687-019-0121-6. PMID: 31155689

14. Kamp MC, van Kempen RWTM, Janssen L, van der Steen MCM, **on behalf of the Regional Prosthetic Joint Infection working group**. First results of a uniform regional treatment protocol and registration for acute prosthetic join infection in the South-East of the Netherlands. J Bone Jt Infect. 2019 May 21;4(3):133-139. doi: 10.7150/jbji.33039. eCollection 2019. PMID: 31192113

15. van Lankveld W, van Melick N, Habets B, **Pronk Y**, Staal JB, van Cingel R. Cross-cultural adaptation and measurement properties of the Dutch knee self efficacy scale (K-SES). BMC Sports Sci Med Rehabil. 2019 Mar 7;11:3. doi: 10.1186/s13102-019-0115-y. eCollection 2019. PMID: 30899517

16. Timmers T, Janssen L, **Pronk Y**, van der Zwaard BC, Koëter S, van Oostveen D, de Boer S, Kremers K, Rutten S, Das D, van Geenen RC, Koenraadt KL, Kusters R, van der Weegen W. 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 Dec 21;6(12):e10742. doi: 10.2196/10742. PMID: 30578185

17. Löwik CAM, Wagenaar FBM, van der Weegen W, Poolman RW, Nelissen RGHH, Bulstra SK, **Pronk Y**, Vermeulen KM, Wouthuyzen-Bakker M, van den Akker-Scheek I, Stevens M, Jutte PC, on behalf of the LEAK study group. Higher randomization rates needed to successfully complete the LEAK study: progress report and solutions for current issues. NTvO. 2018 Jun 25(2):40-45.

18. Hofstede SN, Gademan MGJ, Stijnen T, Nelissen RGHH, Marang-van de Mheen PJ, **on behalf of the ARGON-OPTIMA study group**. The influence of preoperative determinants on quality of life, functioning and pain after total knee and hip replacement: a pooled analysis of Dutch cohorts. BMC Musculoskelet Disord. 2018 Mar 2;19(1):68. doi: 10.1186/s12891-018-1991-0. PMID: 29499692

19. Löwik CAM, Wagenaar FC, van der Weegen W, Poolman RW, Nelissen RGHH, Bulstra SK, **Pronk Y,** Vermeulen KM, Wouthuyzen-Bakker M, van den Akker-Scheek I, Stevens M, Jutte PC; LEAK study group. LEAK study: design of a nationwide randomised controlled trial to find the best way to treat wound leakage after primary hip and knee arthroplasty. BMJ Open. 2017 Dec 28;7(12):e018673. doi: 10.1136/bmjopen-2017-018673. PMID: 29288184

20. Gademan MGJ, Hofstede SN, Vliet Vlieland TPM, Cannegieter SC, Nelissen RGHH, Marang-van de Mheen PJ, **on behalf of the ARGON-OPTIMA study group**. Toward optimal timing of total knee and hip replacement in osteoarthritis patients: the ARGON-OPTIMA study. NTvO. 2017 Dec 24(4):106-110.

21. Wagenaar FC, Löwik CAM, Stevens M, Bulstra SK, **Pronk Y**, van den Akker-Scheek I, Wouthuyzen-Bakker M, Nelissen RGHH, Poolman RW, van der Weegen W, Jutte PC. Managing persistent wound leakage after total knee and hip arthroplasty. Results of a nationwide survey among Dutch orthopaedic surgeons. J Bone Jt Infect. 2017 Nov 3;2(4):202-207. doi: 10.7150/jbji.22327. eCollection 2017. PMID: 29188171

22. van Heerwaarden R, Brinkman JM, **Pronk Y**. Correction of Femoral Valgus Deformity. J Knee Surg. 2017 Oct;30(8):746-755. doi: 10.1055/s-0037-1602138. Epub 2017 May 5. PMID: 28476063

Acknowledgements (Dutch – Dankwoord)

Een avontuur dat ergens in 2014 begon. Hetzelfde avontuur dat in 2017 meer vorm kreeg. Dit avontuur kreeg in 2020 officieel de naam 'externe promovendus'. Het avontuur is vandaag te lezen in dit proefschrift.

Beste promotie team, 'Hoi heren' startte mijn email meestal. Wat zijn jullie een fijn team! Ik weet dat het mijn taak was om dit traject te voltooien, maar zo heb ik dat nooit gevoeld. Jullie hebben allemaal jullie expertise en jullie combinatie zorgde voor mij voor een top team. Dank voor jullie tijd, energie, gezelligheid én expertise.

Wim, ons eerste gesprek was in Den Bosch op het NOV bureau. Na een x tijd gesproken te hebben, volgde 'ooh ja, je wilde me graag spreken over ..'. Voor mij maakte het stiekem al niet meer uit wat je me kon bieden en hoe we het verder gingen inrichten. Ik had mijn promotor gevonden. Bedankt dat je me mijn eigen weg liet uitstippelen, bedankt dat je me meenam in het oerwoud van regels rondom het promoveren, bedankt voor het toevoegen van je kennis en kunde aan dit proefschrift en bedankt voor het gezellig een kopje koffie drinken op het terras in Lissabon.

Walter, jij hebt de gave om alles wat zwaar/moeilijk/lastig/ingewikkeld is altijd weer luchtig en vrolijk te maken. Ik citeer je even: 'Op de m-schijf voor mijn proefschrift heb ik de submappen … en …, maar de belangrijkste submap is 'feest'. Die moet je vast even aanmaken.' In mijn beginfase bij ViaSana hebben jij en Peter mij onder jullie hoede genomen. Je bent altijd bereikbaar; altijd in om even te overleggen, te brainstormen of om gewoon even bij te kletsen. Het EFORT congres is onze vaste trip, en behalve het geven van (poster)presentaties, bestaat deze uit veel gezelligheid en een ronde hardlopen door desbetreffende stad. Ik waardeer je als mens en hoop nog veel met je te mogen samenwerken. Eerst gaan we een feestje vieren. Bedankt.

Martijn, wat heb ik van jou veel mogelijk leren betreffende hoe nu een wetenschappelijk stuk te schrijven. Het is fijn dat ik in ViaSana altijd even bij je kan langslopen. Je zit vol goede onderzoeksideeën en bent op de hoogte van de nieuwste bevindingen uit de literatuur. Ik waardeer je efficiënte en duidelijke werkwijze welke ook altijd goed beargumenteerd is. Dank en tot morgen weer in Mill, zeer gewaardeerde collega.

Peter, onze telefoongesprekken duurde altijd langer dan gepland. We kwamen absoluut toe aan hetgeen we voor ogen hadden, maar tussendoor volgde een wijze les van jou of een interessant weetje over orthopedie, eHealth of big data. Je ben geïnteresseerd en ziet overal kansen. Dat waardeer ik zeer. Dank ook dat je me in contact hebt gebracht met Rein. Bedankt.

Het avontuur dat ergens in 2014 begon, begon bij jou, Klaas. Je wilde graag wetenschappelijk onderzoek een plek geven in Kliniek ViaSana en zocht daar iemand voor. Blijkbaar zag je iets in mij, want ik blijf het tot op de dag van vandaag stoer van je vinden dat je iemand, die net uit de schoolbanken kwam rollen, de kans gaf wetenschappelijk onderzoek binnen ViaSana op te zetten.

Wat heb ik een ruimte van je gekregen om – inmiddels – de afdeling en mezelf te ontwikkelen, en wat is ViaSana een fijne plek om te mogen werken. Jij was ook de eerste die opperde of een promotie traject niet iets voor mij was. Dat het onderwerp uiteindelijk 'de PROMs' is geworden, is natuurlijk niet verrassend. In 2017 schreef ik in een brief aan je: 'Mocht ik ooit gaan promoveren, laat ik dit je zeker weten.'. Klaas, bedankt.

Lieve Maud, we leerden elkaar kennen toen jij bij mij kwam solliciteren voor een onderzoeksstage. Je bent niet meer weggegaan en bent ondertussen mijn zeer gewaardeerde collega. Samen met onze stagiairs vormen wij de wetenschapsafdeling. Jouw bijdrage is in veel hoofdstukken van die proefschrift zichtbaar: je dagelijkse focus op het invoeren en versturen van vragenlijsten, je kennis en kunde bij statistische analyses en de uren 'achterste voren' of via MS Teams samen brainstormen. Dat jij dan vandaag de dag naast mij staat als paranimf is behalve erg gezellig ook zeer vertrouwd. Bedankt.

Ronald, eigenlijk ben jij onofficieel ook onderdeel van mijn promotie team. Veel vergaderingen heb je bijgewoond. Jij hebt er voor gezorgd dat ik ruimte kreeg binnen mijn werkzaamheden in ViaSana om dit promotie traject te kunnen doen. Ik kan altijd even bij je langs lopen. Je straalt uit dat je alle vertrouwen hebt in hetgeen ik doe. Een betere basis voor mijn promotie traject kon je me niet geven. Dank.

Beste Rein, wat een simpel project leek, bleek een ingewikkeld statistisch project. Op de uren in Maastricht kijk ik met veel plezier terug. Naast dat we een zo goed mogelijk antwoord op onze toch niet zo simpele vraag hebben gekregen, heb je mij ook veel andere statistische zaken geleerd. Een betere statistiek cursus had ik niet kunnen krijgen. Hartelijk dank voor de samenwerking.

Team K&V, lieve Bregje, Hilde, Klaartje, Maud en Ronald. Wat vormen wij samen een goed en gezellig team. Het was even schakelen toen we i.v.m. corona thuis gingen werken, maar in de chat is het nooit stil. Bregje, je bent altijd bereikbaar, meedenkend en vrolijk. Dank. Hilde, een speciaal dankjewel aan jou vanwege jouw design en lay-out van dit proefschrift, en hoogstwaarschijnlijk weer prachtige foto's van de promotie ceremonie. Klaartje, dankjewel voor je jarenlange inzet bij het verzamelen van PROs. Successen moet je vieren, hoe groot of klein ze ook zijn. Tijd voor bitterballen?

Alle stagiaires, Angela, Bas, Daisy, Jason, Rens, Renz, Rik, Vince, Wilma en Maud, dank voor jullie inzet en bijdrage. Ik heb met jullie allemaal een goede tijd gehad. Veel succes met jullie carrière.

Mike, Herman en Maud, ik gun iedereen kamergenoten zoals jullie.

Fijne collega's van ViaSana, DiaSana en FysioSupport, dank allemaal voor jullie inzet en bijdrage aan de wetenschappelijk onderzoeken. Van een goed gesprek bij het koffiezetapparaat, een snelle vraag tussen de patiëntenzorg door tot vergaderingen in de avond, altijd die intrinsieke motivatie

om samen kwalitatief hoogwaardige zorg te willen leveren gecombineerd met gezelligheid. Dit proefschrift is ook zeker een bekroning op jullie werk.

Alle patiënten van ViaSana, dank voor uw bijdrage aan de wetenschappelijke onderzoeken, zoals het invullen van de vragenlijsten. De resultaten uit dit proefschrift toepassen in de kliniek is één van de vele dingen die wij er mee doen om uw zorg verder te optimaliseren.

Nicky, mijn masterstage op jouw voorste kruisband studie was mijn introductie in ViaSana. Ondertussen kom je op de koffie terwijl er drie kinderen rondlopen en ben je een gewaardeerde collega binnen de orthopedie. Dank.

Beste Sjoerd, beste vertegenwoordigers van alle hoofdlijnenakkoord-partijen en beste collega's van het landelijke programma Zorgevaluatie en Gepast Gebruik (ZE&GG). Wat heb ik een fantastisch jaar met jullie mogen hebben in het letterlijk en figuurlijk samen de hoogte bereiken in ons gezamenlijke doel: de bewezen beste zorg voor patiënten. Jullie energie en ambitie spat er vanaf. Dank.

Beste Tom, Marloes, Marian en alle andere collega's van de brancheorganisatie Zelfstandige Klinieken Nederland (ZKN), heel hartelijk dank voor elke keer weer jullie hartelijke ontvangst in Zoetermeer en jullie inzet voor alle klinieken in Nederland.

Thomas en alle collega's van Interactive Studios, dank voor jullie fijne product OnlinePROMs, jullie service in het meedenken hoe de PRO verzameling te verbeteren en jullie open houding voor nieuwe ideeën. OnlinePROMs is een constante factor in mijn proefschrift en daar mogen jullie trots op zijn.

Bart, ik twijfel of ik je in het Nederlands of het Engels moet bedanken, want meestal maken we er een combinatie van. Thank you very much voor je feedback op mijn eerste manuscripten.

Lieve Inge (R), een betere vriendin dan jij kan ik me niet wensen. Altijd geïnteresseerd hoe mijn promotie ervoor stond, waar ik nu mee bezig was en altijd goede raad. Een lach, een traan, een knuffel. We hebben alles al samen meegemaakt, behalve ruzie. Je bent een topper! En een paranimf.

Lieve pap, mam, Denise en Lisan, wat zijn jullie een fijn gezin. Mijn eerste kilometers naar Mill maakte ik vanuit Kaatsheuvel. Gezelligheid, humor, belangstellende gesprekken, motiverende woorden en altijd een dikke knuffel. Ik weet dat jullie trots op mij zijn, wat ik ook doe. Joey, dank dat je zo'n fijne schoonbroer bent. Mam, jij specifiek bedankt voor je feedback op diverse stukken. Daar zitten wat uurtje werk in (sorry), maar elke keer gaf je weer aan met hoeveel liefde en plezier je het gelezen had. Dikke knuffel.

Lieve schoonfamilie, lieve Hans, Mirjam, Marc, Miriam, Paul, Mandy en de kids, je hebt je schoonfamilie niet voor het uitkiezen, maar ik ben bij jullie echt in een warm bad beland. Klussen met 40 graden, weekendjes weg, gezellig samen borrelen/eten/lunchen of gewoon even op de koffie. Het is altijd geslaagd.

Lieve Kim en Astrid, vaak hebben we het er over gehad dat we zo trots op elkaar zijn dat we startend als studentassistentes bij Maastricht University nu staan waar we staan. Jullie ambitie niveau is inspirerend met het gezin altijd op de eerste plaats. Dank voor jullie waardering en goede raad.

Lieve Inge (Z), jij als spits bij de aartsrivaal en ik als voorstopper bij mijn hockeyclub. Het leidde tot een fantastische vriendschap. Dank voor wie je bent.

Lieve Esther, Eva, Fabiola, Lisa, Maite, Malon, Marit en Tessa, de leukste pauze is en blijft met Sublime. Ontelbaar aantal jaarclubavonden, weekendjes weg, borrels, koffietjes en vakanties naar Casa Cologna. Met jullie is het altijd gezellig en ik heb nu al weer zin in onze volgende date.

Lieve Daan, lieve schat, nadat mijn eerste artikel als eerste auteur was gepubliceerd, nam je me mee uit eten. Zo nuchter als je bent, zo trots was je op me. Goede en leuke discussies hebben we gehad over diverse studies en je gaf me vaak weer een nieuw idee of andere kijk op iets. Natuurlijk wist je het vaak ook beter, of ja, heel soms. Je hebt me ook de kunst om 'even niks te doen' bijgebracht. We hebben al vele avonturen meegemaakt welke ik niet zonder jou had willen en kunnen doen. Je laat me zijn wie ik ben en trekt op het juiste moment even aan het juiste touwtje. Je bent een fantastische papa voor Lynn en Eva. Duizendmaal dank.

Lieve Lynn en Eva, dank voor elke lach die jullie me geven. Het is genieten met jullie! Vannacht kom ik jullie stiekem nog even een kus geven.

Lieve familie, vrienden, collega's en iedereen die direct of indirect een bijdrage heeft geleverd aan dit proefschrift, zonder jullie was het nooit zo'n mooi avontuur geweest. Mijn naam staat dan wel op de omslag, maar onderzoek doen, doen we samen. Bedankt!

Curriculum Vitae

Yvette Pronk was born on June 1st 1991 in 's-Hertogenbosch, the Netherlands. In 2009 she finished her secondary education at Willem van Oranje college in Waalwijk, the Netherlands. Thereafter, she started studying Health Sciences at Maastricht University in Maastricht, the Netherlands. She focused her study on human movement sciences and after her Bachelor of Science degree, she started her Master Biology of Human Performance and Health (Human Movement Sciences) at Maastricht University. She completed her internship at Sport Medisch Centrum Papendal at Papendal, the Netherlands. Her subject was which autograft for anterior cruciate ligament reconstruction is preferred for return to sports. During her bachelor and master, she worked at the marketing and communication department of Maastricht University. She was responsible for developing, organising and coordinating counselling activities for potential new students and/or their parents. She obtained her Master of Science degree in 2014. She started to work as a research coordinator at Kliniek ViaSana in Mill, the Netherlands. With increasing interest in patient-reported outcome (measures), her PhD begun to take shape in 2017. In 2020 she became an official PhD candidate at Radboudumc in Nijmegen, the Netherlands. Her PhD was focused on how routine PRO collection can be optimized and subsequently how health care can be optimized with routine use of PROs in hip and knee arthroplasty. In addition to her PhD and her work as a research coordinator at Kliniek ViaSana, she was a member of the committee who modernised the 'Nederlandse Orthopaedische Vereniging PROMs advies' in 2019. Furthermore, she represented the branch organisation Zelfstandige Klinieken Nederland in the national programme 'Zorgevaluatie en Gepast Gebruik' in 2019 and 2020.

Yvette Mathijssen-Pronk is married to Daan Mathijssen in 2022. Together they have two daughters, Lynn (2021) and Eva (2023), and they live in 's-Hertogenbosch, the Netherlands.