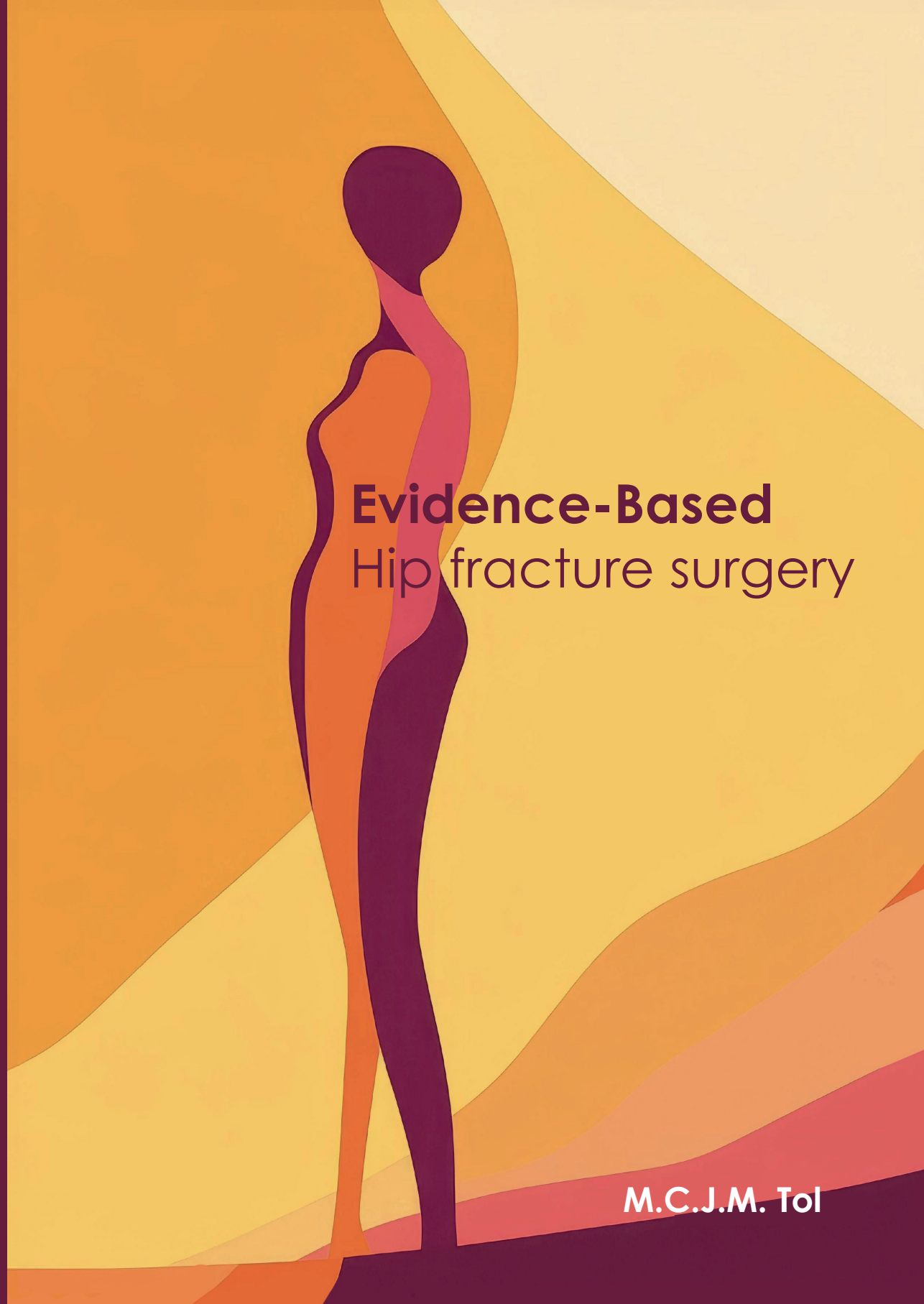

Evidence-Based
Hip fracture surgery
M.C.J.M. Tol

Evidence-based hip fracture surgery

ISBN: 978-94-6534-381-5.
Layout & printing: ProefschriftMaken | proefschriftmaken.nl
Sponsoren: de Nederlandse Orthopaedische Vereniging en Stichting Wetenschap OLVG hebben bijgedragen aan de totstandkoming van dit proefschrift.

Evidence-based hip fracture surgery

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op donderdag 2 juli 2026, te 16.00 uur

door Maria Catharina Johanna Margaretha Tol
geboren te EDAM-VOLENDAM

Promotiecommissie

Promotores:	prof. dr. R.W. Poolman	Universiteit Leiden
	prof. dr. J.C. Goslings	AMC-UvA
Copromotores:	dr. N.W. Willigenburg	OLVG
	dr. H.C. Willems	AMC-UvA
Overige leden:	prof. dr. P. Kloen	Universiteit van Amsterdam
	prof. dr. N. van der Velde	Universiteit van Amsterdam
	prof. dr. C.N. van Dijk	Universiteit van Amsterdam
	prof. dr. I.B. Schipper	Leiden University
	dr. M.G.J. Gademan	Leiden University
	dr. H.A. Formijne Jonkers	OLVG

Faculteit der Geneeskunde

voor mijn ouders

TABLE OF CONTENTS

1

General introduction

Hip fractures in older adults are a worldwide healthcare problem. The global burden of hip fractures continues to rise, with an anticipated increase in disability from 4.5 million to 21 million individuals over the next 40 years (1). Although timely surgery is crucial, functional decline and reduced quality of life often remain challenges (2). Mortality rates within a year after surgery reach over 20% (3), with a significant risk of disability for survivors, even among previously independent individuals (4).

Falls are a significant concern among older, frail adults (5). Each year, about one-third of people aged 65 and older experience a fall, with frequency rising alongside age and frailty (6). In Western Europe, falls account for roughly 70% of all hospital admissions related to injuries, with hip fractures being a major contributor to these admissions (7). A hip fracture has a significant effect on the independence of older adults, their ability to remain living at home, and their overall quality of life. A hip fracture can lead to physical limitations, causing individuals to lose their independence. This can make them dependent on others, which may result in a feeling of loss of control and autonomy and can have negative effects on self-image and mental health. Maintaining mobility reduces the risk of further falls and supports participation in the community, helping to prevent loneliness (8). For frail older adults, preserving independence and quality of life after a hip fracture is essential, not only for their physical health but also for their emotional and social well-being (8).

The treatment options for active older patients facing displaced femoral neck fractures are hemiarthroplasty, or total hip arthroplasty (THA). In THA, both the femoral head and acetabulum are replaced by prostheses, while in hemiarthroplasty, only the femoral head is replaced by inserting a metal femoral prosthesis. Advantages of THA might be a better functional outcome, however it is associated with a higher dislocation rate which often leads to reduction or revision of the prosthesis (9). A recent large randomized study, the HEALTH trial, showed no differences in secondary procedures or quality of life between hemiarthroplasty of THA after 2 years follow-up (10).

The hemiarthroplasty can be placed cemented or uncemented. The WHiTE trial demonstrated a modest but significant improvement in health-related quality of life (HRQoL) and a reduced risk of periprosthetic fractures in favor of cemented over uncemented hemiarthroplasty (11). Therefore, for frail older adults with displaced femoral neck fractures, cemented hemiarthroplasty is the recommended treatment. However, high quality evidence is lacking regarding the optimal surgical approach.

The literature has described various surgical approaches for performing hemiarthroplasty, with the posterolateral (PLA) and direct lateral approach (DLA) being the most used techniques. With the DLA, the fibers of the gluteus medius muscle and vastus lateralis muscle are split parallel to the fibers and during closure, the muscles are sutured(12).

This method can potentially lead to abductor insufficiency, resulting in a positive Trendelenburg sign or post-surgery limping (13, 14). Nonetheless, with the DLA, preserving the posterior capsule helps to prevent dislocation. Contrarily, with a PLA, the surgeon performs a posterior capsulotomy, dividing the short external rotators (12). Using a PLA, the hip abductors are preserved to prevent limping. However, due to insufficient support from the posterior capsule, there might be an increased risk of dislocations (15-17).

Traditionally, a randomized controlled trial (RCT) is considered the highest level of evidence for comparing the effectiveness of the aforementioned treatments or surgical techniques. However, in clinical practice, randomizing patients into treatment groups is not always possible. Additionally, a surgeon or a patient might strongly prefer a particular treatment. In orthopedic trauma care, it is also possible that the necessary expertise for a specific technique is unavailable at the time of admission, necessitating a deviation from the randomization outcome. When dealing with acute trauma scenarios, factors such as the patient's location at the time of injury and the timing of the incident play a pivotal role in deciding the hospital they are transported to and, subsequently, which treatment or technique they get. This pseudo-random treatment allocation is beyond the researchers' control and is regarded as a Natural Experiment. Compared to RCT a Natural Experiment is less costly, faster, and the outcomes potentially exhibit greater generalizability. Therefore, Natural Experiments might be a solution for demonstrating causal treatment effects when randomization is not possible (18). The overarching aim of this thesis is to improve the management and research of femoral neck fractures in fragile older adults.

OUTLINE OF THE THESIS

In the first part of the thesis, we will focus on the interpretation of proxy questionnaires. A significant portion of the hip fracture population comprises frail older patients who frequently experience cognitive impairment. Proxy versions of questionnaires are developed for patients unable to complete the questionnaires themselves. The EQ-5D-5L questionnaire is a widely used patient-reported outcome measure for Health-Related Quality of Life (19), and has two different proxy versions. **Chapter 2** assesses which proxy perspective best represents the orthopedic patient's quality of life. **Chapter 3** explores the relevant barriers and facilitators for implementing the results of this thesis by conducting a qualitative implementation study, providing insights into the perspectives of various stakeholders.

For active older patients who suffered a hip fracture, treatment options include a hemiarthroplasty or a total hip arthroplasty (THA). Literature describes a higher rate of dislocations for THA; however, proponents of THA suggest better patient function and quality of life (20). The HEALTH trial, a large randomized controlled trial (RCT), compared these two

treatments and found no differences in secondary procedures and no clinically significant difference in function and quality of life (10). Nevertheless, due to the limited final follow-up of only two years, long-term outcome differences were not recorded. **Chapter 4** of this thesis presents the 12-year follow-up results of an RCT comparing hemiarthroplasty and THA in active older patients with a hip fracture.

Hemiarthroplasty is the recommended treatment for older, more fragile (i.e. less active) adults following a femoral neck fracture. Considerable variance exists in surgical approaches for hemiarthroplasty worldwide (21). Surgeons typically opt for an approach based on personal preference or hospital agreements. Based on LROI data, the posterolateral approach (PLA) and the direct lateral approach (DLA) are most used in the Netherlands (22). **Chapter 5** systematically reviews and summarizes the available evidence comparing the DLA and PLA for hemiarthroplasty after femoral neck fracture. Due to the lack of high-quality evidence, we initiated an RCT and Natural Experiment comparing the DLA and PLA for cemented hemiarthroplasty. **Chapter 6** outlines the study protocol of our APOLLO trial, detailing the two study designs and an innovative method to fuse the data. **Chapter 7** reports the RCT and Natural Experiment results concerning patient-reported quality of life, function, (fear of) falling, and complications such as dislocation, reoperation, and readmission. A hip fracture can have a significant effect on mobility and balance of older adults. **Chapter 8** compares the PLA and DLA of hemiarthroplasty in terms of physical performance post-treatment in a subgroup of the RCT participants. **Chapter 9** outlines the cost-effectiveness of both hemiarthroplasty approaches.

In the final section, we will reflect on the results and engage in a general discussion to direct future perspectives in **Chapter 10.**

REFERENCES:

1. Cooper C, Campion G, Melton LJ. Hip fractures in the elderly: A world-wide projection. Osteoporos Int. 1992;2:285-9.
2. Johnell O, Kanis JA. An estimate of the worldwide prevalence, mortality and disability associated with hip fracture. Osteoporos Int. 2004;15(11):897-902.
3. Mundi S, Pindiprolu B, Simunovic N, Bhandari M. Similar mortality rates in hip fracture patients over the past 31 years: A systematic review of RCTs. Acta Orthop. 2014;85(1):54–9.
4. VeiligheidNL luidt de noodklok. Valpreventie kan en moet structureel aangepakt worden! [Available from: https://www.veiligheid.nl/actueel/veiligheidnl-luidt-de-noodklok-valpreventie-kan-en-moet-structureel-aangepakt-worden.
5. Organization WH. WHO Global Report on Falls Prevention in Older Age. World Health Orgnanization. 2008.
6. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. NEJM. 1988;319(26):1701-7.
7. Haagsma J, Olij B, Majdan M, al. e. Falls in older aged adults in 22 European countries: incidence, mortality and burden of disease from 1990 to 2017. Inj Prev 2020;26:67-74.
8. Miller RR, Ballew SH, MD S. Repeat falls and the recovery of social participation in the year post-hip fracture. Age Ageing. 2009;38:570-5.
9. Burgers PTPW, Van Geene AR, van den Bekerom MPJ, Van Lieshout EMM, Blom B, Aleem IS, et al. Total hip arthroplasty versus hemiarthroplasty for displaced femoral neck fractures in the healthy elderly: a meta-analysis and systematic review of randomized trials. International Orthopaedics. 2012;36:1549-60.
10. The Health investigators. Total Hip Arthroplasty or Hemiarthroplasty for Hip Fracture. NEJM. 2019;381(23).
11. Fernandez MA, Achten J, Parsons N, Griffin XL, Png ME, Gould J, et al. Cemented or Uncemented Hemiarthroplasty for Intracapsular Hip Fracture New England Journal of Medicine. 2022;386(6).
12. S. Hoppenfeld PdB, R. Buckley Surgical Exposures in Orthopaedics: The Anatomic Approach.
13. Berstock J, Blom A, Beswick A. A systematic review and meta-analysis of complications following the posterior and lateral surgical approaches to total hip arthroplasty. Ann R Coll Surg Engl. 2015;97(1):11-6.
14. Sayed-Noor A, Hanas A, Skoldenberg O, Mukka S. Abductor Muscle Function and Trochanteric Tenderness After Hemiarthroplasty for Femoral Neck Fracture. J Orthop Trauma. 2016;30(6).
15. Rogmark C, Fenstad AM, Leonardsson O, Engesaeter LB, Karrholm J, Furnes O, et al. Posterior approach and uncemented stems increases the risk of reoperation after hemiarthroplasties in elderly hip fracture patients. Acta Orthop. 2014;85(1):18-25.
16. Biber R, Brem M, Singler K, Moellers M, Sieber C, Bail HJ. Dorsal versus transgluteal approach for hip hemiarthroplasty: an analysis of early complications in seven hundred and four consecutive cases. Int Orthop. 2012;36(11):2219-23.
17. Leonardsson O, Rolfson O, Rogmark C. The surgical approach for hemiarthroplasty does not influence patient-reported outcome. The Bone & Joint Journal. 2016;98:542-7.
18. Peter Craig, Cyrus Cooper, David Gunnell, Sally Haw, Lawson K, Sally Macintyre, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. J Epidemiol Community Health. 2012;66:1182-6.
19. Euroqol. [Available from: https://euroqol.org/.

20. Hopley C, Stengel D, Ekkernkamp A, Wich M. Primary total hip arthroplasty versus hemiarthroplasty for displaced intracapsular hip fractures in older patients: systematic review. BMJ. 2010;340:2332.
21. Chechik O, Khashan M, Lador R, Salai M, Amar E. Surgical approach and prosthesis fixation in hip arthroplasty world wide. Arch Orthop Trauma Surg. 2013;133(11):1595-600.
22. LROI. Annual report of Dutch Arthroplasty Register. 2021.

2

Perspectives and Decisions in Orthopaedics

Level of agreement between the self-completed EQ-5D-5L and two proxy perspectives in an orthopaedic population,

A RANDOMIZED AGREEMENT STUDY

Maria C.J.M. Tol[1],MD, Jurrian P. Kuipers[1], MD, Nienke W. Willigenburg[1], PhD,

Hanna C. Willems[2], MD, PhD, prof Rudolf W. Poolman[3], MD, PhD

[1] Department of Orthopaedic Surgery, JointResearch, OLVG, Amsterdam, [2] Department of Internal Medicine and Geriatrics, AMC, Amsterdam, [3] Department of Orthopaedic Surgery, LUMC, Leiden

Health and Quality of Life Outcomes (2021)

ABSTRACT

Objectives: To determine the level of agreement between both proxy versions and the self-completed EQ-5D-5L.

Design: A randomized agreement study

Setting and Participants: We recruited 120 patients (compos mentis) and their proxies at the orthopaedic outpatient clinic. Patients completed the regular EQ-5D-5L and their proxy completed the proxy version of the EQ-5D-5L in and rated the patients' health from their own (proxy-proxy) perspective (i.e. how do you rate the health of the patient), and from the patient's (proxy-patient) perspective (i.e. how do you think the patient would rate their own health if they were able to).

Measures: The primary outcome was the agreement between patients and their proxy, quantified as the intra class correlation coefficient for the EQ-5D-5L Utility score.

Results: Average Utility scores were 0.65 with the self completed EQ-5D-5L, versus 0.60 with the proxy-patient version and 0.58 with the proxy-proxy version. The ICC was 0.66 (95%CI:0.523, 0.753) for the proxy-patient perspective and 0.58 (95%CI:0.411, 0.697) for the proxy-proxy perspective. The mean gold standard score of the VAS-Health was 69.7 whereas the proxy-proxy perspective was 66.5 and the proxy-patient perspective was 66.3.

Conclusion and implications: The proxy-patient perspective yielded substantial agreement with the self completed EQ-5D-5L, while the agreement with the proxy-proxy perspective was moderate. In this study population of patients without cognitive impairment, proxies tended to underestimate the quality of life of their relative.

INTRODUCTION

Due to the risen attention to value-driven care, studies to evaluate existing and new treatments are more common. The patient perspective is frequently at the center of focus during this healthcare evaluation. Health-Related Quality of Life (HRQoL) is an important patient reported outcome when comparing the effectiveness of health care interventions and the value of health care(1).

Scientists widely use the EQ-5D, which is known for its validity, reliability and responsiveness, as a measurement instrument for the HRQoL (2-5). The EQ-5D is a descriptive system of health-related quality of life states consisting of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) as well as a visual analogue scale (EQ VAS) on which patients rate their overall health.

The EQ-5D-5L is a questionnaire completed by the patient. Two by proxy versions were developed for patients who are not capable of completing the questionnaire by themselves due to i.e. cognitive impairment(6). The difference between both proxy versions is the perspective of the answers. The instruction, how to fill in the questionnaire, is different: The first version is from the proxy's perspective (how do you rate the patients health), the second version is from the patients' perspective (how do you think the patient would rate their own health if they were able to do so).

Both versions are used more often, not only due to a growing group of elderly who are incapable of completing their own questionnaires, but also as a result of increased attention to value based health care and cost-efficiency research in which the patient perspective is often used(7-9). Assessing the quality and effectiveness of treatments is challenging when patients, for instance with dementia, are incapable of understanding questionnaires. These developments strengthen the need for outcome assessment by proxy.

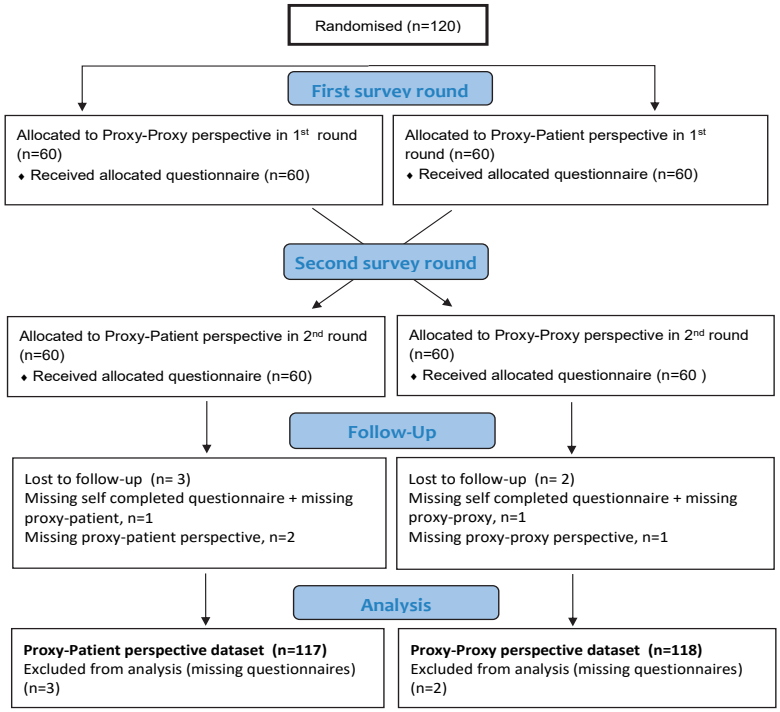
Never has it been investigated which by proxy version of the EQ-5D-5L is most likely to reflect the patient's quality of life best (10-12) We hypothesize that there might be a difference between by proxy versions. Hence, it is important to know which proxy version best reflects quality of life in patients who can not complete the questionnaire themselves since it has a central role in healthcare evaluation research. Therefore, the aim of this study was to determine the agreement of both proxy versions with respect to the self reported quality of life using the EQ-5D-5L. This novel study will investigate the agreement of mentally healthy patients with their proxy.

METHODS

Participants and setting

The study protocol was approved by the medical ethics committee at OLVG (WO 18.059) and conducted according to the principles of the Declaration of Helsinki, as amended in Seoul and Fortaleza (64th WMA General Assembly, October 2013). The trial was registered at The Netherlands National Trial Register (TC 7526). To ensure that self-completed questionnaires could be used as gold standard, the study population consisted of mentally healthy patients. And to accurately reflect a typical clinical setting, their relatives or friends that accompanied them to the outpatient visit were asked as proxy. We asked all consecutive patients of both sexes and all ethnicities who visited the orthopaedic outpatient clinic accompanied with a proxy during the inclusion period, June 2018 – August 2018, to participate in our study. Patients were eligible for study participation when the following inclusion criteria were met: 18 years or older at the time of visiting the outpatient clinic (both patient and proxy giver), compos mentis, Dutch fluency and literacy. We excluded patients when they had signs of cognitive impairment or when they visited the outpatient clinic alone. Any signs of cognitive impairment were objectified by a clinician.

Figure 1.



Study procedures

The EQ-5D-5L descriptive system assesses health in five dimensions (mobility, personal care, usual activities, pain/discomfort, and anxiety/depression), each of which has five levels of response (no problems, slight problems, moderate problems, severe problems, extreme problems/unable to). This part of the EQ-5D questionnaire provides a descriptive profile that can be used to generate a health state profile. Health state index scores generally range from less than 0 (where 0 is the value of a health state equivalent to dead; negative values representing values as worse than dead) to 1 (the value of full health), with higher scores indicating higher health utility. The second part of the questionnaire consists of a visual analogue scale (VAS) on which the patient rates his/her perceived health from 0 (the worst imaginable health) to 100 (the best imaginable health.

There are two proxy perspectives of the EQ-5D-5L . The questionnaire and scoring are the same as the self completed EQ-5D-5L. The instruction, how to fill in the questionnaire, is different: The first version is the proxy-proxy perspective: (how do you rate the patients health) the proxy is asked to rate the patients' health related quality of life in their (the proxy's) opinion. The second version is the proxy-patients perspective (how do you think the patient would rate their own health if they were able to do so) the proxy is asked to rate how he/she (the proxy) thinks the patient would rate his/her own health-related quality of life if the patient were able to communicate it.

All patients received the self-complete version of the EQ-5D-5L, resulting in the gold standard, and all proxy givers completed both proxy versions. Dyads were

randomly assigned to which order they the proxy had to complete the two different questionnaires in a 1:1 allocation ratio, using variable block randomisation in CASTOR EDC (specs) (Figure 1). Prior to participation, the patients and their proxies were not aware of the existence of the different perspectives of the proxy version and were blinded for the purpose and hypothesis of our study. Proxies were only given the second version upon completing the first. Patient and proxy were asked to complete the questionnaire separated from their proxy to assure they answered independently.

Statistical Analysis

We converted the scores on the five dimensions of the EQ-5D-5L into the Utility score, which is reflecting the HRQoL. We used the Euroqol EQ-5D-5L Crosswalk Index Value Calculator for calculating Utility scores derived from the Dutch general population(13). A higher score indicates a higher rated quality of life. To quantify the level of agreement between the (continuous) Utility score of the Self-complete EQ-5D-5L and their proxy versions, we used the Intraclass Correlation Coefficient (ICC), based on absolute agreement in a two-way mixed effects model. We repeated this analysis for the EQ-VAS score. In addition, we performed Weighted Kappa and absolute agreement analysis to quantify the

level of agreement for the (categorical) individual domain scores. For the interpretation of both ICC and Weighted Kappa we used the method of Landis and Koch, with scores >0.81 indicating almost perfect agreement, 0.61-0.8 substantial, 0.41-0.60 moderate, 0.21 – 0.4 fair and <0.20 slight agreement(14). The level of significance was <0.05. To assess whether the health status influences the accuracy in which the proxy giver can assess the patient's health status, we visualized the difference between the self-completed Utility scores and both proxy perspective Utility scores with respect to the self completed EQ-5D-5L Utility scores. We performed all analyses using IBM SPSS Statistics (Version 22.0). Since there was no effect size available to analyse an adequate sample size, the sample size was based on previous studies on this topic (11, 15).

Table 1. Baseline characteristics of patients and proxy givers

Patient gender, n (%)	
Female	74 (61.7)
Proxy gender, n (%)	
Female	62 (51.7)
Patient age group, n (%)	
18 to 45 years	29 (24.2)
45 to 70 years	51 (42.5)
≥70 years	40 (33.3)
Proxy age group, n (%)	
18 to 45 years	23 (19.2)
45 to 70 years	75 (62.5)
≥70 years	22 (18.3)
Relationship to patient, n (%)	
Partner	71 (59)
Sibling	5 (4.2)
Child	15 (12.5)
Parent	15 (12.5)
Friend	9 (7.5)
Neighbour	1 (0.8)
Other	4 (3.3)

Abbreviation: n = number of patients

RESULTS

Population

We enrolled a total of 120 dyads, consisting of 120 patients and their proxies. Between May – July 2018 we received 115 complete and 5 incomplete datasets of the self-completed and their proxy perspectives for data analysis. There was one missing self-completed questionnaire, 1 missing proxy-proxy perspective, 2 missing proxy-patient perspective

and 1 missing questionnaire of both perspectives. The baseline characteristics of the patients and proxy givers are depicted in Table 1.

Absolute scores

The mean Utility score of the patients self completed EQ-5D-5L (gold standard) was 0.65 (95%CI:0.614, 0.686), of the proxy-proxy perspective 0.58 (95%CI:0.538, 0.621) and of the proxy-patient perspective 0.60 (95%CI:0.562, 0.638). The mean gold standard score of the VAS-Health was 69.7 (95%CI:66.4, 73.1), whereas the proxy-proxy perspective was 66.5 (95%CI:63.1, 69.8) and the proxy-patient perspective was 66.3 (95%CI:62.8, 69.8) (Table 2).

Agreement scores

For the Utility score, the proxy-patient perspective had a substantial level of agreement with the gold standard: ICC 0.66 (95%CI:0.523, 0.753). The proxy-proxy perspective had a moderate level of agreement with the gold standard: ICC 0.58 (95%CI:0.411, 0.697) (Table 3). For the overall health status based on the EQ-VAS, the level of agreement was substantial with the proxy-patient perspective: ICC = 0.64 (95%CI: 0.515, 0.737), versus a moderate level of agreement of proxy-proxy perspective: ICC = 0.53 (95%CI: 0.386, 0.65) (Table 3).

Table 2, mean scores EQ-5D and VAS Health

	Patients (Gold Standard) n = 119	**Proxy – proxy** perspective n = 118	**Proxy – patient** perspective n = 117
Utility score EQ-5D-5L mean (95%CI)	0.65 (0.614 - 0.686)	0.58 (0.538 - 0.621)	0.60 (0.562 - 0.638)
	n = 117	n = 118	n = 115
VAS Health mean (95%CI)	69.74 (66.4 - 73.1)	66.48 (63.1 - 69.8)	66.33 (62.8 - 69.8)

Abbreviation: n = number of patients, SD = standard deviation

Table 3, Level of agreement Utility score and VAS-Health

	Proxy – proxy perspective (n =115)	**Proxy – patient** perspective (n =115)
Utility score		
ICC (95% CI)	0.58 (0.411 – 0.697)	0.66 (0.523 – 0.753)
Strength of agreement	Moderate	Substantial
VAS-Health		
ICC (95% CI)	0.530 (0.386 – 0.650)	0.639 (0.515 – 0.737)
Strength of agreement	Moderate	Substantial

The level of agreement, measured with the Weighted kappa, of the individual domain scores are listed in Table 4. No differences in level of agreement in all subdomains were observed between the two proxy versions compared with the gold standard.

Table 4, Level of agreement domain scores EQ-5D-5L

	Proxy-proxy perspective n = 115		**Proxy-Patient perspective** n =115	
	Weighted Kappa (SE)	95 % CI	Weighted Kappa (SE)	95% CI
Mobility	0.596 (0.053) moderate	0.492 – 0.699	0.599 (0.054) moderate	0.494 – 0.705
Selfcare	0.452 (0.065) moderate	0.324 – 0.579	0.417 (0.073) moderate	0.275 – 0.559
ADL	0.337 (0.061) fair	0.216 – 0.457	0.384 (0.061) fair	0.264 – 0.503
Pain	0.392 (0.070) fair	0.255 – 0.529	0.394 (0.068) fair	0.261 – 0.527
Anxiety	0.355 (0.066) fair	0.225 – 0.484	0.406 (0.073) fair	0.262 – 0.549

Abbreviation: n = number of patients, SE = standard error, CI = confidence interval

Absolute agreement for individual domain scores (Table 5) was equal between proxy versions for the domain Selfcare. Absolute agreement was higher for the proxy-patient perspective in Mobility (2.5%) and Anxiety/depression (7.5%). Absolute agreement was higher for the proxy-proxy perspective in ADL (1.7%) and Pain (1.7%).

Table 5, absolute agreement domain scores EQ-5D-5L

Domain	**Proxy-Proxy** (perspective A) N=115	**Proxy-patient** (Perspective B) n=120	**Difference in Agreement between Perspectives**	**95% CI of difference**
Mobility	55.8%	58.3%	A 2.5% < B	-9.887, 14.780
Selfcare	60%	60%	0%	-12.218, 12.218
ADL	39.2%	37.5%	A 1.7% > B	-10.466, 13.796
Pain	52.5%	50.8%	A 1.7% > B	-10.783, 14.108
Anxiety	51.7%	59.2%	A 7.5% < B	-5.012, 19.688

Figure 2 shows that the discrepancy between Utility scores between proxy givers and patients does not depend on the patient's health status (expressed in Utility score).

DISCUSSION

Our study showed substantial agreement between the patient gold standard and the proxy-patient perspective and moderate agreement between the patient gold standard and proxy-proxy perspective for both the Utility score and VAS Health of the EQ-5D-5L. The Utility score is the primary method to interpret health status using the EQ-5D-5L questionnaire. The mean Utility score was higher in the patient gold standard compared with both proxies.

This is the first study which empirically compares the two proxy perspectives of the EQ-5D-5L. Another strength of the study is that both proxy perspectives, given by one proxy, were compared with the same patient self-completed EQ-5D-5L questionnaire, resulting in adequate comparison. Moreover, the study design ruled out recall bias and participants were blinded to the study hypothesis.

Figure 3, Evaluation of effect patients' health status on proxies ability to accurately assess patient's health status (*continued*)

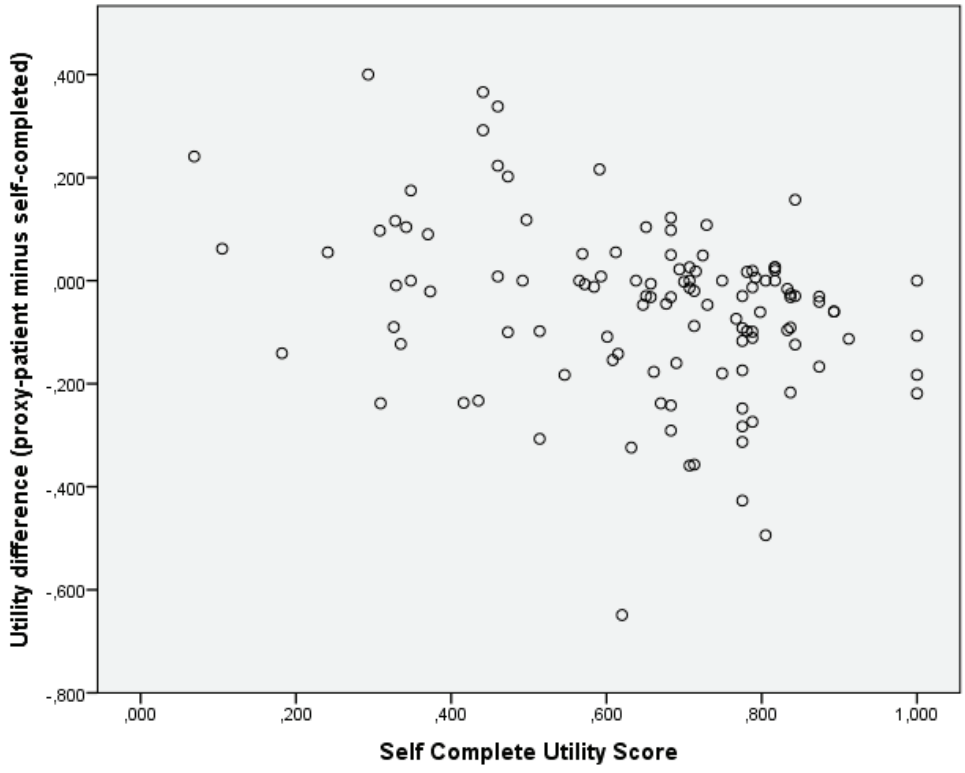
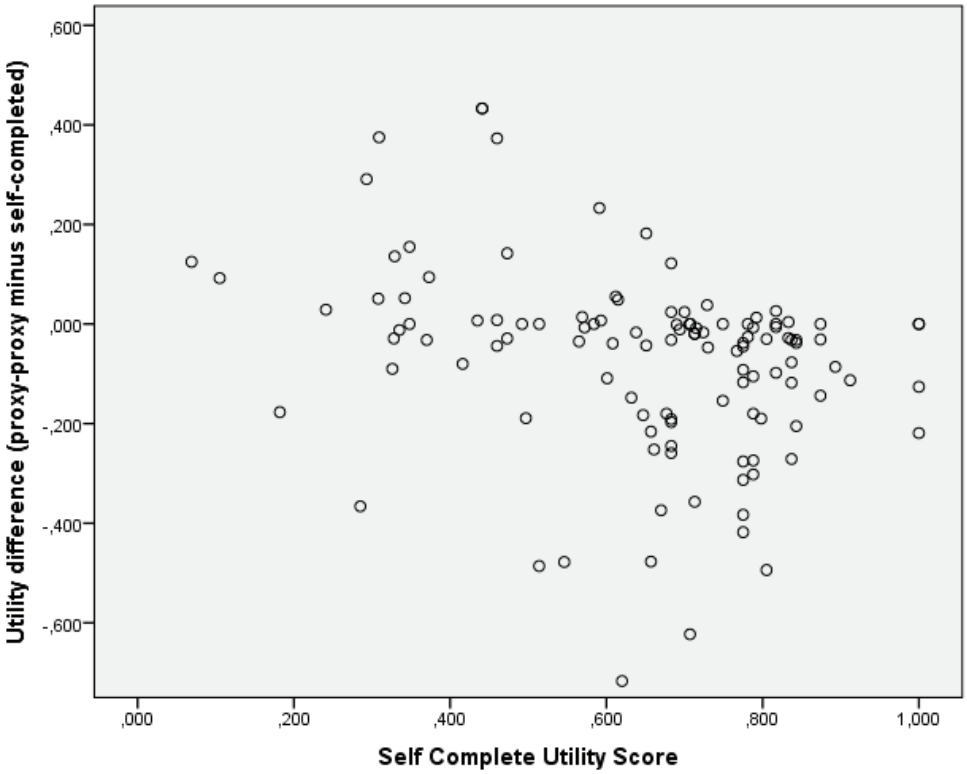


Figure 3, Evaluation of effect patients' health status on proxies ability to accurately assess patient's health status



All participants were Orthopaedic patients, which could be a sub-selection of the population and may be a limitation of the study. However, patient acquisition occurred in the waiting room at the outpatient clinic of all sub specializations of the Orthopaedic surgery where there is a high variance in the patient population. Due to this wide scatter we think the included patients and proxies can be a good representation of the population. The results of our study should however be evaluated in other disciplines. To ensure the self-complete EQ-5D-5L could be used as gold standard, our study population only included mentally healthy patients. Caution should be used when extrapolating the results of this study to a population with cognitive impairments.

We found one previous study that compared the two different proxy perspectives completed by a clinician with the self-completed EQ-5D with three levels (3L)(15). That study reported substantially higher levels of agreement for all domains and utility score compared with our results. They reported Kappa values > 0.7 on all individual domain scores, while our highest kappa was 0.6 (Mobility domain). This is likely explained by the fact that the EQ-5D-3L has only three answer categories per domain, opposed to five categories in

the EQ-5D-5L. Therefore the potential for disagreement is higher for the EQ-5D-5L, resulting in lower levels of agreement for domain scores and eventually Utility score.

All proxy questionnaires where completed by relatives of, or persons close to, the patients in this study. The observed agreement for individual domain scores supports previous findings on higher validity of questionnaires provided by relatives than clinicians for the less observable dimensions ('Anxiety/depression' and 'Pain')(11, 16). However, Bryan and co-investigators showed that data provided by clinicians had higher construct validity regarding more observable dimensions of the EQ-5D-3L instrument (e.g. 'Mobility and Selfcare')(11). Understandably, clinicians are more experienced than relatives at objectively assessing patient functioning as part of routine clinical assessment and have a wide range of knowledge regarding loss of function. Relatives on the other hand are far better acquainted with the patient and therefore more able to empathize with their subjective reality. However, it is remarkable that proxies underestimate the self-related health status of the patient. This phenomenon is observed in more studies, but non of the authors has given an explanation for the observed effect(10-12). One explanation could be that the effect is influenced by the proxies own mood, personal beliefs or expectations(17). Another explanation why elderly overestimate their health status could be that they compare their own health status with peers suffering from worse health, which give them a positive perception of their own function(18).

The effect of the underestimation of the health status by proxies is important for clinicians. This means that treatment decisions based on proxy perspectives could be based on a health status worse than the real health status. On the other hand if patients overestimate their health status, it could be associated with riskier health behavior(19).

Further research could be focused on optimizing the level of agreement between the proxy perspective and the gold standard. For example by completing the observable dimensions of proxy questionnaires through clinicians and the subjective dimensions through relatives. In addition, despite the barriers of cognitive limitations, further research should be done in patients with mild to moderate stage dementia to validate the proxy version for this specific patient group.

CONCLUSION

This study showed a substantial agreement between the patient gold standard and the proxy-patient perspective and moderate agreement between the patient gold standard and proxy-proxy perspective for both the Utility score and VAS Health of the EQ-5D-5L. Therefore we recommend the proxy-patient version of the EQ-5D-5L. Regardless of their perspective, in this study population of patients without cognitive impairment, proxies tended to underestimate the quality of life of their relative.

REFERENCES

1. Porter ME. What Is Value in Health Care? New England Journal of Medicine. 2010;363:2477-81.
2. Hurst NP KP, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). Br J Rheumatol. 1997;36(5):551-9.
3. Haywood KL GA, Lall R, Smith JF, Lamb SE. EuroQol EQ-5D and condition-specific measures of health outcome in women with urinary incontinence: reliability, validity and responsiveness. Qual Life Res. 2008;17(3):475-83.
4. König HH BA, Günther O, Matschinger H, Heinrich S, Riedel-Heller SG, Angermeyer MC, Roick C. Validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with anxiety disorders. Health Qual Life Outcomes. 2010;8:47.
5. Stark RG RP, Leidl R, König HH. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. Inflamm Bowel Dis. 2010;16(1):42-51.
6. Euroqol. [Available from: https://euroqol.org/.
7. Vestjens L CJ, Birnie E, Nieboer AP. Cost-effectiveness of a proactive, integrated primary care approach for community-dwelling frail older persons. Cost Eff Resour Alloc. 2019;17(1):1-15.
8. Nicholson JA SH, MacDonald D, McBirnie J. Cost-effectiveness and satisfaction following arthroscopic rotator cuff repair: does age matter?. Bone Joint J. 2019;101-B(7):860-6.
9. Bailly L, Mosse P, Diagana S, al. FMe. "As du Coeur" study: a randomized controlled trial on quality of life impact and cost effectiveness of a physical activity program in patients with cardiovascular disease. BMC Cardiovasc Disord. 2018;18(1):225.
10. Devine A. TSJC, Spencer A., Diaz-Ordaz K., Eldridge S., Underwood M. The agreement between proxy and self-completed EQ-5D for care home residents was better for index scores individual domains. Journal of Clinical Epidemiology. 2014;67:1035-43.
11. Bryan S. HW, Bentham P., Buckley A., Laight A. . Proxy completion of EQ-5D in patients with dementia. Qual Life Res. 2005;14:107-18.
12. Coucill W. BS, Bentham P., Buckley A., Laight A. EQ-5D in patients with dementia an investigation of inter-rater agreement. Medical care. 2001;39(8):760-71.
13. van Hout B. JMF, Feng Y.S., Kohlmann T., Busschbach J., Golicki D., Lloyd A., Scalone L., Kind P., Pickard A.S. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in health. 2012;15(5):708-15.
14. Landis JR KG. The measurement of observer agreement for categorical data Biometrics 1977;33(1):159-74.
15. McPhail S. BE, Haines T. Two perspecitives of proxy reporting of health-related quality of life using the euroqol-5D, an investigation of agreement. Medical care. 2008;46:1140-8.
16. Pickard S. LH, Knight S., Sharifi R., Wu Z., Hung S., Witt W.P., Chang C., Bennet C.L. Proxy Assessment of Health-Related Quality of Life in African American and White Respondents With Prostate Cancer Perspective Matters. Medical care. 2009;47(2):176-83.
17. Gomez-Gallego M. G-GJ, Ato-Lozano E. . Adressing the bias problem in the assessment of the quality of life of patients with dementia: determinants of the accuracy and precision of the proxy ratings. J Nutr Healt Aging. 2015;19(3):365-72.
18. Henchoz K CS, Girardin M. Health perception and health status in advanced old age: A paradox of association. J Aging Stud. 2008;22(3):282–90.
19. Sakurai R FY, Ishihara M, Higuchi T, Uchida H, Imanaka K. Age-related self-overestimation of step-over ability in healthy older adults and its relationship to fall risk. BMC Geriatr. 2013;13(44):1-9.

3

Perspectives and Decisions in Orthopaedics

Factors influencing the decision between the use of the posterolateral or direct lateral approach approach for hemiarthroplasty in the treatment of hip fractures

MCJM Tol, MD.[1], F van Nassau, MSc, PhD.[2], C Ekhardt, MSc.[2], HC Willems, MD. PhD.[3], RW Poolman, MD. PhD.[4], JR Anema, MSc, PhD.[2]

[1] Department of Orthopedic Surgery, Joint Research, OLVG, Amsterdam, [2] Department of Public and Occupational Health, Amsterdam UMC, Amsterdam, [3] Department of Internal Medicine and Geriatrics, Amsterdam UMC, Amsterdam, [4] Department of Orthopedic Surgery, LUMC, Leiden

(submitted)

ABSTRACT

Introduction: The posterolateral approach (PLA) and direct lateral approach (DLA) are the most used approaches for inserting a hemiarthroplasty after a hip fracture. The decision for which approach is used during the procedure is now mainly based on the surgeon's preference.. The perspectives on implementation of the surgical approaches from different stakeholders are not yet known. The aim of this study is to gain more insight in these perspectives to facilitate the implementation beyond the trial.

Methods: Semi structured interviews were conducted. With the use of purposive and snowball sampling respondents were recruited and interviewed until data saturation. Data were analysed using thematic analysis in open, axial and selective coding.

Results: 12 respondents were interviewed. After analyses of the data, seven themes were identified: 1) Magnitude of the problem, 2) Surgeons' choice for approach and guidelines, 3) Experience, training and attitude of the surgeon, 4) Different roles of stakeholders regarding to implementation, 5) Personal situation of the patient is important, 6) Complications related to hemiarthroplasty, and 7) Strength of the evidence.

Conclusion: Previous experience and training of surgeons were important factors for the decision for which approach was applied. Furthermore, the surgeons emphasized that the robustness of the evidence from the APOLLO trial would significantly influence their future decision-making regarding the approach they adopt. Other stakeholders expressed not to be involved in the decision for the approach but can play a role in the implementation and dissemination of the study results.

INTRODUCTION

Hip fractures are common fractures in the older population and a leading cause of mortality and morbidity in the elderly.(1-4) Hip fractures do not only affect the patient physically but have a considerable impact on the mental and psychosocial aspects of a patient's life(5). For displaced femoral neck fractures in the geriatric population, hemiarthroplasty is a common treatment. There are multiple approaches for this procedure(6). In the Netherlands, the posterolateral (PLA) and the direct lateral approaches (DLA) are the two most frequently used surgical approaches(7). The PLA tends to be less painful, patients are more satisfied, and have a better quality of life. (8, 9) However, some studies reported a higher risk of dislocation after a PLA due to the posterior capsulectomy.(8, 10-13) On the contrary, the DLA preserve the posterior capsule but the hip abductors are split which can lead to loss of muscle strength and limping.(14, 15) Recently, the APOLLO trial, a multicenter randomized controlled trial (RCT) and Natural Experiment comparing DLA and PLA was published. The study showed similar outcomes in quality of life, mobility, and patient-reported outcomes, but there was a higher incidence of dislocation in the PLA group compared to DLA. Despite these results, surgeons are hesitant to change their surgical approach(16) Currently, the perspectives on implementation of both the DLA and PLA from different stakeholders are not yet known. This limits the insight in the barriers and facilitators that could be relevant for implementing one of the surgical techniques after the trial results are known. Therefore, the aim of this study is to determine important factors influencing the decision for the use of the DLA or PLA when inserting a hemiarthroplasty.

METHODS

Design

We conducted a qualitative study design with semi-structured interviews and used a positivistic epistemological approach to require multiple views from different stakeholders. The design was emergent, it was flexible for adjustments if this was beneficial for the study(17). This study is embedded in the APOLLO trial, a RCT and Natural Experiment comparing the PLA and DLA of hemiarthroplasty after femoral neck fractures(18). At time of the interviews, the results of the trial were not yet published and known by the participants. When performing a DLA, the gluteus medius and vastus lateralis muscles are partly dissected from the trochanter mayor. Exposure of the hip joint is reached with an anterior capsulectomy (19). With the PLA the external rotator muscles and piriformis muscle are dissected. Exposure of the hip joint is created by an incision of the posterior capsule.(20)

Theoretical framework

To create the topic list and codebook we used the model of Fleuren et al. as a theoretical framework(21). The model states that multiple levels: socio-political context, organization and innovation, contain different characteristics that are important for implementation. All levels of the theoretical model of Fleuren and the most important determinants were covered in the topic-list (see Appenix)

Setting and participants

We recruited different stakeholders see Table 1. Initially purposive sampling was used to recruit interview participants. All surgeons worked in hospitals that participated in the RCT. Next, snowball sampling techniques were used to recruit more stakeholders. This strategy was continued until data saturation was achieved.

Data collection

Semi-structured interviews were conducted by MT and CE. The interviews were or online conducted by video call (Zoom 2011). The purpose of the interview, rights, use of date and the guarantee of confidentiality were made known to the participant before the interview. The interviews lasted between 15 to 45 minutes.

Analysis

The transcripts were analysed according to a thematic analysis with the use of the qualitative analysis software ATLAS.ti version 8. Thematic analysis was chosen since this type of analysis is flexible and suits best for identifying and finding patterns in the data (29). The codebook was made by three researchers (MT, FvN, CE). Coding took place first open, second axial and third selective.

Ethical considerations

Study information was provided to the participants, and an informed consent form was signed before the interviews started. All participants were asked for permission and agreed to have interviews recorded for transcription purposes. Withdrawal from the study could occur at any time without any explanation.

RESULTS

Between April 2019 and May 2021 all data were collected. Twelve stakeholders were interviewed. The respondents consisted of three orthopaedic surgeons, four trauma surgeons of which one was also co-author of Dutch national guidelines for hip fractures, two clinical geriatricians, one physiotherapist, one doctor in a nursing home facility, and a senior advisor from the Patients Federation. Five interviews were conducted in-person,

seven interviews took place by video call. Table 1 displays an overview of characteristics of the respondents.

Table 1. Respondents characteristics

Respondent number	Gender	Profession	Years in function
R1	Male	Trauma surgeon	21
R2	Male	Orthopaedic surgeon	12
R3	Male	Trauma surgeon	15
R4	Male	Orthopaedic surgeon	14
R5	Male	Orthopaedic surgeon	17
R6	Female	Clinical geriatrician	11
R7	Female	Physiotherapist	35
R8	Female	Doctor in a nursing home facility	1
R9	Male	Trauma surgeon, author guidelines	20
R10	Male	Trauma surgeon	23
R11	Male	Clinical geriatrician	18
R12	Female	Senior advisor patient federation	31

Seven themes were identified in the coded interviews. 1) Current extent of the problem, 2) Surgeons' choice for approach and guidelines, 3) Experience, training and attitude of the surgeon, 4) Different roles of stakeholders regarding to implementation, 5) Personal situation of the patient is important, 6) Complications related to hemiarthroplasty, 7) Strength of the evidence. An overview of the themes is listed in Table 2. These themes and their corresponding subthemes and related barriers and facilitators will be further described and interpreted below.

Table 2. Overview of themes and subthemes

Themes	Subthemes
1. Current extent of the problem	Magnitude of the problem
2. Harmony on approach and guidelines	Consensus
3. Experience, training and attitude of the surgeon	Surgeons learn multiple approaches during their training
	Experience of the surgeon with both the DLA and PLA
	Attitude of the surgeon
4. Role different stakeholders regarding implementation	Role different stakeholders in the decision for the used approach
5. Personal situation of the patient is important	Personal factors of the patient
	Current way of working

Table 2. Overview of themes and subthemes (*continued*)

Themes	Subthemes
6. Complications related to hemiarthroplasty	Complications related to the DLA
	Complications related to the PLA
	Surgeon is not aware of complications post-surgery
7. Strengths and limitations APOLLO trial	Strengths of the trial
	Limitations of the trial

Theme 1: Magnitude of the problem

Respondents agreed that hemiarthroplasty is a treatment that frequently takes place in many hospitals and therefore it is useful for surgeons to learn a different approach if this is needed according to evidence. If studies demonstrate stronger (cost)effects for one of the approaches, this will lead to a better quality of life, less complications and as a result less costs in healthcare. Considering the large number of these surgeries that take place, this will have positive financial consequences. Respondents mentioned that this can be seen as a factor that can influence the decision for the use of one of the approaches.

'If you think about the fact that most hospitals perform an average of at least 200 surgeries per year, that means they perform around four or five per surgeries per week. So given this fact, you could say that these surgeons should learn the other approach.' (R6)

Theme 2: Surgeons' choice for approach and guidelines

In hospitals throughout the Netherlands, surgeons have a different preferences for which approach they prefer to use. Even within hospitals there is no consensus on which approach works best. One respondent told about the experience he had within different hospitals regarding the approach applied.

'Traditionally in that hospital we all did the total hip surgery with the PLA and the hemiarthroplasty with the DLA. In the other hospital it was very different, because everybody did what they would think was the approach they felt most comfortable with.' (R5)

One other respondent wondered if consensus within hospital on the used approach was important to focus on. He said that the goal always should be the best outcomes for the patient.

'The question is, is it necessary that everybody does the same thing, or that you want to focus on the outcomes. What is important, is that patients are well functioning and do not have major complications post surgery.' (R3)

Multiple respondents stated that when something is written down in guidelines, it does not necessarily mean that surgeons will always adhere to it. This is also expressed by a surgeon involved in guideline development:

'I think that everbody has the right to deviate from the clinical guidelines as long as you can defend it in the first place towards the patient, your colleagues and in the case there is a complaint, against the judge.' (R9)

One potential barrier to implementing the results could be the surgeon's preference and their unwillingness to change the surgical technique because they are not convinced it would lead to better outcomes.

Theme 3: Experience, training and attitude of surgeon

Interviewees expressed that the attitude and opinion of surgeons towards both approaches is shaped by their training and experiences. The opinion of the surgeon about the approaches can both be a barrier or facilitator for the implementation of one of the approaches. This has resulted in the following three subthemes: a) experience as a resident, b) Experience of the surgeon with both the DLA and the PLA and c) Attitude of the surgeon.

A) Experience as a resident

There is a lot of variation in approaches among surgeons. Some surgical residents learn multiple approaches during their training, others only have exposure to one surgical approach. It depends in which hospital the surgeon has his or her surgical training. An orthopedic surgeon said the following:

'I think it is really good in this hospital that residences are trained with all approaches and that they can continue with the one they are comfortable with. This way you give the residences multiple options to choose from. Especially for the ones who would like to continue with hip surgeries, you give them "options out" for the future.' (R2)

There were other variations; the position of the patient during the procedure differed for some surgeons during the surgical training. One respondent gave an overview of all the variations he saw during his training.

'Before my training, we did DLA in supine position, including reoperations. During my training, we did the DLA in side position. Then I went to an academic hospital, where we did the PLA. After that it was the DLA in side position again. So yes, it was a broad training with all different kinds of things.' (R5)

B) Experience of the surgeon with both the DLA and the PLA

All surgeons considered the surgeon' exposure as an important factor in a succesfull surgery outcome. Respondents believed that performing many surgeries using the same approach will lead to less complications.

'I think you measure the volume outcome. Someone who performs many hip surgeries has relatively extensive experience and generally has little complications.' (R3

'Is it someone who performs a hemiarthroplasty weekly? Or is it someone who performs it once in every three to four months, just because it coincidentally occurs during his shift?' (R10)

Surgeons who do not regularly perform a hemiarthroplasty could struggle more in learning a different approach. This could be a barrier for implementation of one of the approaches. However, some surgeons who perform many surgeries and have had a broad training will not perceive these struggles, this could be a facilitator for implementation.

'But I think there are also people who can switch and say, okay, I could do it this approach for a year or another approach for a year.' (R3)

C) Attitude of the surgeon

In general, the surgeons were well disposed to implementation of the results of the trial. Participation in the APOLLO trial(16) comparing the DLA and PLA approach, was considered as a facilitator for implementation. Surgeons had the following to say about this: *'Yes, I do think that if you participate in the trial, that you also have to accept the results.' (R2)*

However, one of the respondents mentioned that the attitude of surgeons that did not want to participate in the APOLLO trial can be opposite. A clinical geriatrician expressed that surgeons in his hospital did not want to participate because they did not think the study was relevant. They placed their own experience above evidence which can be a considerable barrier for implementation of one of the two approaches.

'I was a bit surprised when the orthopedic surgeons in this hospital said that they did not want to participate in the trial, because they already knew which approach was better. I wonder: How can you know for sure?' (R11)

Theme 4: Different roles of stakeholders regarding implementation

The clinical geriatrician, physiotherapist, doctor in a nursing home facility and the Patients Federation are all involved in hemiarthroplasty in a different manner than the interviewed surgeons. These stakeholders have a more indirect role in the implementation process. However, their attitude and opinion are important and can be both a barrier or facilitator for implementation of one of the approaches in practice.

Implementation of one of the surgical approaches in practice would not have a large impact on the working method of the doctor in a nursing home facility. She mentioned that regardless of the approach used, the method for rehabilitation remains the same.

Both clinical geriatricians acknowledged that they are not involved in the decision for which approach is used during surgery. But they did feel like they can play a role in the implementation of one of the approaches. They do not want to tell surgeons what they have to do but they are willing to address relevancy of the problem or take part in spreading information. This can be seen as a facilitator for implementation of one of the approaches in practice.

'I think that as a clinical geriatrician, I am not in a position where I can say to a surgeon 'I think you have to use this approach on the patient.' I think that is too much. I can tell them that there is an interesting discussion going on and I can address the topic, but that is as far as I can go.' (R11)

The senior advisor from the Patients Federation mentioned that they do play a role in policy making regarding to clinical health care. She expressed that the Patient Federation tries to represent patients and propagate values that patients find important. According to her shared discussion making is an important factor and they try to spread this message to other stakeholders.

'I also see that things have changed, so we have really gained our place in the field and are considered serious discussion partners. Of course, we have a different input compared to the hospital association, the medical specialists' association, the nursing association, or the clinics – those are the parties that all participate. But it's not like we can only address the patient experience in the end of the process' (R12)

Theme 5: Personal situation of the patient is important

According to all stakeholders the personal situation of the patient is an important factor. However different stakeholders shared different perspectives in how they cooperate the situation of the patient in their working method. This theme is further divided in the two subthemes personal factors of the patient and current way of working.

Personal factors of the patient

In general the patient population with hip fracture is old and fragile, according to all stakeholders. However there can be considerable differences between patients. Some are already dependent of a wheelchair and are institutionalized while others are more mobile and live independently. The most important values for patients were mobility and autonomy, according to both clinical geriatricians. Also multiple stakeholders confirmed

that quality of life and preventing a next fall are also very important for this patient population.

'These patients have a certain life expectancy. We know that 25 to 30% of them dies in the first year and 80% of them die in the first four years after surgery. So it is all about a way of giving the patients a certain level of comfort and in addition to that you don't want them to fall and hurt themselves again.' (R6)

Shared decision

Patients are currently not involved in the decision for the approach that is used. The main reason is that not all surgeons on call at the time of surgery can perform both approaches.

'It also depends on the surgeon. Is (s)he equally skilled in both approaches? Do you actually want to give the patients the option to choose? Because in some hospitals only one of the approaches is used, in that case there is nothing to choose. If you would ask for the other approach as a patient, that would be an issue, because the risk of complications is higher due to the fact that the surgeon is not used to the approach.' (R6)

Another opinion is that it is perceived to be too complicated and technical for the patient or family, and there is also a short time between the fall and time of surgery. However, the senior advisor of the Patients Federation not completely agreed with this working method. She said:

'In general, patients are not stupid. You can really explain it to them, even if they are old. In general, when the patient is old, they often have a partner or family member who can think along. I think it is important that the patient is able to take the arguments into account to make a decision.' (R12)

Theme 6: Complications related to hemiarthroplasty

Respondents described complications for both approaches. These complications are factors that play a role in the decision for the approach. These findings are described within the following subthemes: Complications related to the direct lateral approach (DLA), Complications related to the posterolateral approach (PLA), Surgeon is not aware of complications post-surgery and Strengths and limitations of the APOLLO trial**.**

Complications related to the posterolateral approach (PLA)

The most seen complications, according to interviewed surgeons, when this approach is used is a dislocation of the prosthesis which can lead to re-operation of the patient. One of the surgeons said the following about this:

'What I noticed, was that we saw some dislocations. And yes, these can be a tragedy for those people, because re-operation of someone that is over eighty years old is of course not very easy for them.' (R3)

Another orthopedic surgeon said the following about this: *'Look, patients with a dislocation after the posterolateral approach… you can at least offer them something and they will eventually more or less walk normally after a re-operation.' (R2)*

Complications related to the direct lateral approach (DLA)

The most common complications that arise when the DLA is used according to the surgeons were abductor weakness, lateral hip pain and an irritating scar.

'There are few patients that we will always keep tracing. Those are the people that remain having problems from the used approach, such as lateral hip pain, not enough muscle strength and those sort of things. So those people, you remember, those that often come back to your outpatient clinic.' (R4)

Multiple surgeons motivated their choice to use the PLA when executing a hemiarthroplasty as complications that can occur using the DLA are worse than complications that arise when using the PLA. The abductor weakness leading to a limping walking pattern was considered as one of the worst complications. One respondent said the following about this:

' The abductor weakness that you can find in the DLA group, I believe that to be a different entity. So, I believe that it also needs a different focus, because it is really mutilating for the patient because they will never be able to walk alright again.' (R2)

Surgeon is not always aware of all postoperative consequences

A clinical geriatrician mentioned that the approach that is used could have consequences in the recovery and quality of life of patients post-surgery. She suspected that one of the approaches has a higher risk of limping and as a consequence of that, a higher risk of falling. It was mentioned that the surgeon might not always be aware of these consequences for the patient. On the other hand certain movements could provoke a dislocation, movements which can occur if patients having sex.

"I think that at this moment there is a delay in knowledge and if you want to change the way of acting than you have to understand the ratio. You first need to feel the urgency, before you want to change it." (R6)

However, it can be a barrier if the different approaches have different advantages. It gives a surgeon an exit if they not want to change their preference of approach. The clinical geriatrician continued:

"I don't know if there is motivation, it could be a barrier to change strategy if patients under a certain age who still are intimate are better off with a DLA or a way out which can be important for surgeons and use it so they do not have to change their surgical approach" (R6)

Theme 7: Strength of the evidence

Strengths of the APOLLO trial

Most of the respondents considered the strength of evidence of the trial as an important factor for implementation of the two approaches in practice. Yet, there needs to be a significant difference in patient outcomes for an approach to change the way surgeons are used to work.

'So, if there is a little difference then it doesn't matter that much. In that case, you don't have to put so much effort in implementing your trial results. So, it really depends on the results of the trial.' (R4)

A facilitator was the generalizability of the study by including patients with dementia. A respondent said: '*I truly believe that including people with dementia in this trial is a huge advantage.' (R4)*

Limitations of the APOLLO trial

On the other hand, surgeons addressed that there are more surgical approaches that can be used when executing a hemiarthroplasty. The antero(lateral) approach is one of these approaches that is often used. This approach has not been studied in the trial, this can be seen as a limitation on the study, hindering implementation.

'If we compare the PLA against the DLA and we were to pick a winner out of those two, then it is important to compare that approach with the anterolateral approach as well. And after that you, really know which approach is better.' (R4)

DISCUSSION

The most important factors for implementation seemed to be related to the experience, training and the willingness to change surgery techniques of the surgeon. After the results of the APOLLO trial are known, the strengths and limitations of this trial are an important factor for making a decision of which approach a surgeon will use. Guidelines

are not always considered as definitive, and they can be deviated from if there is a good explanation. Therefore, the trial's conclusion as a recommendation in the guideline may not substantially improve implementation. Additional measures should be done to successfully implement the study results.

Respondents addressed that the personal situation of the patient was only sometimes taken into account during the decision-making process, but according to the patient federation this factor should play a bigger role in the decision process. Logistically, this would be very challenging because of the availability and experience of the surgeons. Other stakeholders than surgeons had no direct influence on the decision regarding which approach is used when a surgeon performs a hemiarthroplasty. However, they believed to have an indirect role in the decision making of the surgical approach. Clinical geriatricians will not instruct which approach the surgeon should perform, but they can address the relevancy of the outcomes or problems of the approaches and spread this information to increase awareness. At policy level, the Patients Federation can influence the policymaking regarding to clinical health care. For example, health insurers can stop reimbursing a procedure if it is not effective or causes many complications.

Important themes found in this study were influence of the experience, training and attitude of the surgeon on the decision for the approach. Especially the broad training of the surgeon seemed to be important. Previous literature confirmed that surgical education is based on a broad training with a range of technical skills and experts opinions which determine the practice of a young surgeon.(22, 23) Lin et al. emphasizes that successful mentoring in training of surgeons is important in the self-development of a resident(24). To overcome this barrier, the respondents specified implementation strategies of different learning techniques such as cadaveric dissection for surgical training.

Complications related to the hemiarthroplasty are also an important factor when deciding which surgical approach according to this study. One of the subthemes accompanying this theme is that the surgeon is not always aware of all the post-operative complications regarding independence and intimacy, which affects the quality of life. Existing literature emphasizes that the surgery can have great psychosocial consequences for patients.(5) However, in this study none of the surgeons addressed this. This knowledge gap might be a barrier for implementation of the study results of the APOLLO trial, and education of these possible complications is paramount.

If, according to the APOLLO trial, one of the approaches is significantly better and this will be implemented in practice, this also results in de-implementation of the other approach. The experience of the surgeon could be a barrier in this de-implementation process. Surgeons have been doing an approach for many years already and may not feel the need to change it. To add a recommendation in the hip fracture guideline, does not necessarily lead to changes

in clinical practice. To address this barrier, it is important to disseminate the new found evidence by presenting the results at (international) meetings or by active learning from experts.

One of the strengths of this study was its strong credibility, achieved by covering all levels involved in implementation through the use of Fleuren's theoretical model(25). Additionally, the study benefited from interviews conducted by two different researchers, each having a distinct relationship with the respondents. The data analysis was performed by multiple researchers, ensuring a low risk for biased interpretation and enhancing the study's dependability. The iterative process was another strength that contributed to the study's dependability. Furthermore, the inclusion of a surgeon who mainly used the anterolateral approach, differing from the other participating surgeons' working methods, added diversity to the study due to purposive sampling. Finally, the study's transferability was enhanced by interviewing various stakeholders engaged in the hemiarthroplasty process, providing insights from different perspectives. Most of the surgeons we interviewed were involved in the APOLLO trial. This means that these surgeons are more motivated to increase knowledge about potential differences in effectiveness both procedures . This may declines the generalizability of the results and could be seen as a limitation of this study. Another limitation is that the study results of the APOLLO trial were not yet known at the time of the interview. However, we believe it was beneficial to determine the barriers and facilitators of implementation prior to the end of the study, as it allows for anticipation on these factors.

Furthermore, using a hybrid effectiveness-implementation trial design can be beneficial. This approach allows researchers to not only assess the clinical effectiveness of an intervention but also to evaluate the feasibility and practicality of its implementation in real-world settings. Incorporating this dual focus early on helps in understanding how the intervention can be best translated into routine practice, improving both patient outcomes and implementation success(26).

Conclusion and recommendations

The most important factors were training, experience and attitude of the surgeon. Surgeons stated that the outcome for the APOLLO trial will be important for their decision of which approach they will use in the future. Other stakeholders have an indirect influence on the decision for the approach. Awareness can be created by clinical geriatricians by addressing the relevance of the study results on patient outcomes. At policy level, policy-making regarding hip fracture management can be influenced by the Patients Federation, which represents patients and their values. The multiple barriers need different interventions to change surgeons habits, create awareness of the relevance of the outcomes and enhance implementation. The multifaceted intervention should focus on hands-on learning, education, active learning from experts and presenting the evidence at (international) meetings.

APPENDIX I: TOPICLISTS

Topiclist

Chirurg

1. Kenmerken vernieuwing

1. Is er al genoeg informatie bekend/beschreven over voor en nadelen van benaderingen?
2. Moeten er dingen in de huidige werkwijze veranderd worden om de behandeling te implementeren?
3. Welke voordelen voor de zorgverlener zijn er mogelijk? (tijdswinst, dienstbelasting, kostenwinst?
4. Hoeveel patiënten per jaar zijn mogelijk geholpen met de behandeling? Is het het waard?

2. Kenmerken toekomstige gebruiker (zorgverlener)

1. Steun die de zorgverlener ervaart van collega uit de zelfde beroepsgroep om de nieuwe behandelmethode uit te voeren.
2. Steun die de zorgverlener ervaart van collegae uit de andere beroepsgroepen (chirurgie/orthopedie) om de nieuwe behandelmethode uit te voeren
3. Steun die de zorgverlener ervaart van het hoger management met betrekking tot uitvoer van de nieuwe behandeling
4. Mate waarin de zorgverlener vaardigheden heeft die nodig zijn om de vernieuwing uit te kunnen voeren. Het aanleren van een nieuwe benadering.
5. Is er sprake van tegengestelde doelen tussen verschillende beroepsgroepen? Indien beroepsgroep niet de ander benadering zou willen aanleren. Zijn er dan financiële consequenties? Bv de trauma patiënten kunnen/willen niet met ander benadering geopereerd worden.
6. Mate waarin de zorgverlener verwacht dat de patiënt meewerkt/*wil deenemen* aan de vernieuwing. Het gaat niet om daadwerkelijke medewerking, maar om de verwachting van de zorgverlener op voorhand.
7. Mate waarin de zorgverlener verwacht dat de patiënt tevreden is over de vernieuwing. Het gaat niet om daadwerkelijke tevredenheid, maar om de verwachting van de zorgverlener op voorhand.

3. Kenmerken organisatie

1. Plaats waar besluitvorming over invoering van de vernieuwing plaatsvindt: centraal (topmanagement) of decentraal (professionals).
2. Besluitvormingsstructuur: lengte van de beslislijnen waarover de communicatie over de vernieuwing verloopt. (a) Korte lijnen: de beslissing verloopt sneller en er is

weinig verlies van informatie mede door de bijsturingmogelijkheden. Lange lijnen: vertragend. Mogelijk gevolg is demotivatie, verzanden en mislukken vanwege communicatiestoringen in de verschillende schakels. Mogelijke vraag hierbij; Hoe verloopt besluitvorming rondom dit soort vernieuwingen?

3. Relaties met andere afdelingen indien blijkt dan een benadering beter is. Mogelijke vraag hierbij; In welke mate wordt er goed samengewerkt tussen de verschillende betrokken afdelingen?
4. Samenwerking tussen afdelingen die bij de vernieuwing betrokken zijn. Goede samenwerking: men komt makkelijker op een lijn. Slechte samenwerking: omgekeerde redenering.
5. Aanwezige expertise op afdelingsniveau m.b.t. de vernieuwing. Bedoeld wordt expertise in algemene zin. Veel expertise: kan onzekerheid en beginnersfouten voorkomen en vanwege de expertise gaat de invoering sneller. Weinig expertise: omgekeerde redenering.
6. Aantal mensen dat betrokken is in de organisatie dat wil zeggen met de vernieuwing gaat werken. Weinig mensen: vergemakkelijkt directe en persoonlijke communicatie, er is een nauwere samenwerking en mensen zijn meer betrokken. Veel mensen: geen consensus.
7. Hoeveelheid geld die voor het uitvoeren van de vernieuwing beschikbaar is. Veel geld: veranderen betekent investeren in scholing, tijd en materialen. Weinig geld: omgekeerde redenering.
8. Financiële vergoeding voor zorgverleners c.q. organisatie als tegemoetkoming aan extra inspanningen die gepleegd moeten worden bij het werken met de vernieuwing. Extra vergoeding: vernieuwen kost tijd en geld. Een tegemoetkoming is motiverend en geeft financiële armslag om extra middelen en mensen in te zetten. Geen extra vergoeding: omgekeerde redenering. Leidt tot extra werkdruk bij degenen die de vernieuwing moeten toepassen.
9. Coördinator van het project? Is er een duidelijk aanspreekpunt (ook voor implementatie)?

4. Kenmerken sociaal politieke omgeving

1. Bereidheid van patiënt om mee te werken aan de uitvoering van de vernieuwing. Medewerking: vooral bevorderend als vernieuwing invloed heeft op de patiënt. Geen medewerking: kost veel tijd en energie in het overtuigen van de patiënt en kan leiden tot weerstand bij de zorgverlener.
2. Mate waarin patiënt op de hoogte is van de (gezondheids)voordelen voor zichzelf als gevolg van de vernieuwing. Goed op de hoogte: bevordert motivatie en medewerking. Slecht op de hoogte: geen consensus.
3. Twijfel van patiënt aan deskundigheid van de zorgverlener met betrekking tot de uitvoering van de vernieuwing. Geen twijfel: nauwelijks van invloed omdat patiënt er op voorhand vanuit gaat dat de zorgverlener het goed doet, tenzij het tegendeel

blijkt. Twijfel: leidt tot weerstand. Patiënt zal mogelijk een ander alternatief zoeken. Daarnaast moet de zorgverlener meer tijd stoppen in het overtuigen van de patiënt.
4. Extra kosten die patiënt moet maken als met de vernieuwing gewerkt wordt.
5. Extra mentale of fysieke belasting die patiënt ondervindt als met de vernieuwing gewerkt wordt.
6. Wat ziet de patiënt als mogelijk nadelen van de behandeling? (kosten en belasting zijn onderwerpen die daar mogelijk een rol kunnen spelen)

Topiclist

Klinisch geriater

Domein: Huidige werkwijze

Graag wil ik eerst met u bespreken wat u doet voor patiënten die een operatie waarbij een kop-halsprothese wordt geplaatst ondergaan.

- Kunt u mij vertellen wat u als klinisch geriater zoal doet voor deze patiënten?
- Hoe komt het besluit tot stand om de patiënt te opereren met één van de twee technieken? Wordt u hier bij betrokken? En wordt de patiënt hier bij betrokken?
- Welke factoren spelen een rol bij de keuze voor deze benadering(en)? te denken valt aan type patiënt (leeftijd, geslacht, etc), voorkeur voor behandeling van patient voor en nadelen van benaderingen, , kennis en ervaring arts, etc
- Wat zijn uw ervaringen met de twee verschillende benaderingen? Ziet u verschillen in het herstel van de patiënt? (qua voortraject en of nazorg)
- Wat zijn de ervaringen van patiënten wat betreft het herstel van de operatie? En wat zijn hierbij belangrijke uitkomsten voor de patiënt?
- Wordt u als klinisch geriater op dit moment betrokken bij het opstellen van beleid of protocollen omtrent deze ingreep? En op welke manier?

Domein: Nieuwe werkwijze nav resultaten onderzoek

Als uit ons onderzoek blijkt dat een van deze benaderingen een betere patiënten uitkomst blijkt te hebben zoals een hogere mate van kwaliteit van leven, minder complicaties en meer kosten effectief is, dan willen wij dit doorvoeren in de praktijk en dit als standaard aanbevelen.

- Wat zijn uw gedachtes hierover?
- Hoe denkt u dat patiënten hier tegenover staan?
- Wat betekent deze verandering voor patiënten?
- Wat zouden voor patiënten potentiële voor- en nadelen zijn als één van de technieken de standaard wordt?
- Zou u betrokken willen worden bij het proces van implementatie van de onderzoeksresultaten? En op welke manier?

Afsluiting

Dit was mijn laatste vraag. Zijn er nog dingen die we niet besproken hebben, maar waarvan u denkt dat deze wel relevant zijn voor dit project?

- Zijn er nog andere stakeholders waarvan u denkt dat het nuttig is dat we ook met hen spreken?

Topiclist

Patiëntenfederatie

Domein: Huidige werkwijze

- Kunt u mij vertellen wat een patiëntenfederatie zoal doet? En wat doet u specifiek voor patiënten die een operatie met kophalsprothese ondergaan?
- Hoe worden patiënten momenteel voorbereid op deze operatie? En hoe ervaren zij dit?
- Zijn patiënten op de hoogte van de verschillende benaderingen?
- In hoeverre zijn patiënten op de hoogte van de veel voorkomende complicaties bij de verschillende benaderingen?
- Wat zijn de ervaringen van patiënten wat betreft het herstel van de operatie? En wat zijn hierbij belangrijke uitkomsten voor de patiënt?
- Wordt de patiëntenfederatie op dit moment betrokken bij het opstellen van beleid of protocollen omtrent deze ingreep? En op welke manier?
- Wat gaat er volgens u goed in de huidige werkwijze?

Domein: Nieuwe werkwijze nav resultaten onderzoek

Als uit ons onderzoek blijkt dat een van deze benaderingen een betere patiënten uitkomst blijkt te hebben zoals betere kwaliteit van leven, minder complicaties en meer kosten effectief is, dan willen wij dit doorvoeren in de praktijk en dit als standaard aanbevelen.

- Wat zijn uw gedachtes hierover?
- Hoe denkt u dat patiënten hier tegenover staan?
- Wat betekent deze verandering voor patiënten?
- Wat zouden voor patiënten potentiele voor- en nadelen zijn als één van de technieken de standaard wordt?
- Zijn er nog andere stakeholders waarvan u denkt dat het nuttig is dat we ook met hen spreken?

Topiclist

Fysiotherapeut

Domein: Huidige werkwijze

- Kunt u mij vertellen wat u als fysiotherapeut zoal doet voor deze patiënten?
- Bent u op de hoogte van de verschillende benaderingen? En past u hier uw behandelingen op aan?
- Wat zijn uw ervaringen met patiënten die herstellen van het plaatsen van een kophalsprothese?
- Indien op de hoogte van de verschillende benaderingen:
  Ziet u verschillen in het herstel van de patiënt?
- Wat zijn de ervaringen van patiënten wat betreft het herstel van de operatie? En wat zijn hierbij belangrijke uitkomsten voor de patiënt?
- Worden fysiotherapeuten op dit moment betrokken bij het opstellen van beleid of protocollen omtrent deze ingreep? En op welke manier?

Domein: Nieuwe werkwijze nav resultaten onderzoek

Als uit ons onderzoek blijkt dat een van deze benaderingen een betere patiënten uitkomst blijkt te hebben zoals een hogere mate van kwaliteit van leven, minder complicaties en meer kosten effectief is, dan willen wij dit doorvoeren in de praktijk en dit als standaard aanbevelen.

- Wat zijn uw gedachtes hierover?
- Hoe denkt u dat patiënten hier tegenover staan?
- Hoe zouden we fysiotherapeuten kunnen betrekken in het proces van het implementeren van één van de technieken als standaardmethode?
- Wat zouden voor patiënten potentiele voor- en nadelen zijn als één van de technieken de standaard wordt?

Afsluiting

Dit was mijn laatste vraag. Zijn er nog dingen die we niet besproken hebben, maar waarvan u denkt dat deze wel relevant zijn voor dit project?

- Zijn er nog andere stakeholders waarvan u denkt dat het nuttig is dat we ook met hen spreken?

Topiclist

Arts in een verpleeghuis

Domein: Huidige werkwijze

- Kunt u mij vertellen wat u als arts in een verpleeghuis zoal doet? En wat u doet voor deze patiënten?
- Bent u op de hoogte van de verschillende benaderingen? En past u hier de nazorg op aan?
- Wat zijn uw ervaringen met patiënten die herstellen van het plaatsen van een kop-halsprothese?
- Indien op de hoogte van de verschillende benaderingen:
  Ziet u verschillen in het herstel van de patiënt?
- Wat zijn de ervaringen van patiënten wat betreft het herstel van de operatie? En wat zijn hierbij belangrijke uitkomsten voor de patiënt? Zijn hierbij eventueel ook nog verschillen per benadering?
- Worden artsen in verpleeghuizen op dit moment betrokken bij het opstellen van beleid of protocollen omtrent deze ingreep? En op welke manier?

Domein: Nieuwe werkwijze nav resultaten onderzoek

Als uit ons onderzoek blijkt dat een van deze benaderingen een betere patiënten uitkomst blijkt te hebben zoals een hogere mate van kwaliteit van leven, minder complicaties en meer kosten effectief is, dan willen wij dit doorvoeren in de praktijk en dit als standaard aanbevelen.

- Wat zijn uw gedachtes hierover?
- Hoe denkt u dat patiënten hier tegenover staan?
- Hoe zouden we artsen in verpleeghuizen kunnen betrekken in het proces van het implementeren van één van de technieken als standaardmethode?
- Wat zouden voor patiënten potentiele voor- en nadelen zijn als één van de technieken de standaard wordt?

Afsluiting

Dit was mijn laatste vraag. Zijn er nog dingen die we niet besproken hebben, maar waarvan u denkt dat deze wel relevant zijn voor dit project?

- Zijn er nog andere stakeholders waarvan u denkt dat het nuttig is dat we ook met hen spreken?

APPENDIX II: CODEBOOK

Codeboek belemmerende en bevorderende factoren voor implementatie van de chirurgische techniek bij een kop-hals prothese

1 Codes omtrent de interventie

1.1 Complicaties gerelateerd aan de direct laterale benadering (DLA)

1.1.1 Trendelenburg

Factoren die een rol spelen in het ontwikkelen van Trendelenburg (mank lopen) als gevolg van een operatie waarbij een kop-hals prothese is geplaatst middels de DLA.

1.1.2 Verhoogd valrisico

Factoren die een rol spelen bij een verhoogd valrisico als gevolg van een operatie waarbij een kop-hals prothese is geplaatst middels de DLA.

1.1.3 Dislocatie

Factoreen die gerelateerd zijn aan dislocatie van het gewricht na een operatie waarbij een kop-hals prothese is geplaatst middels de DLA.

1.2 Complicaties gerelateerd aan de posterolaterale benadering (PLA)

1.2.1 Dislocatie

Factoren die gerelateerd zijn aan dislocatie van het gewricht na een operatie waarbij een kop-hals prothese is geplaatst middels de PLA.

1.2.2 Re-operatie

Factoren die gerelateerd zijn aan revisie/ re-operatie na een operatie waarbij een kop-hals prothese is geplaatst middels de PLA.

1.3 Langdurig herstel

Factoren die bijdragen aan het herstel en mobiliteit van de patiënt op de lange termijn.

1.4 Overtuigingskracht trial

1.4.1 Gebruik van meerdere uitkomstmaten in de trial (+)

1.4.2 Groot verschil in uitkomst tussen de twee benaderingen (+)

Factoren die bijdragen in het aantonen dat één van de benaderingen overtuigend beter is voor de patiënt.

1.4.3 Klein verschil in uitkomst tussen de twee benaderingen (-)

Factoren die bijdragen in het aantonen dat één van de benaderingen niet overtuigend beter is voor de patiënt.

1.4.4 Includeren kwetsbare ouderen (+)

1.4.5 Omvang studie (+)

1.5 Concurrerende aanpakken (-)

1.5.1 Anterolaterale methode

Factoren gerelateerd aan het gebruik van de anterolaterale methode.

1.6 Variatie in het gebruik binnen beide benaderingen
Factoren gerelateerd aan de variatie die gebruikt wordt binnen de DLA en de PLA

1.7 Overzicht van de chirurg tijdens de operatie
Factoren die gerelateerd zijn aan een goed overzicht op het gewricht tijdens de operatie

2 Codes omtrent de patiënt

2.1 Persoonlijke situatie patiënt
Factoren gerelateerd aan de persoonlijke situatie van de patiënt pre operatief zoals thuissituatie en mobiliteit.

2.2 Waardes van belang voor de patiënt na de operatie
Factoren gerelateerd aan waardevolle zaken voor de patiënt zoals behoud van kwaliteit van leven

2.3 Ervaringen van de patiënt
Factoren gerelateerd aan de ervaringen van de patiënt omtrent de operatie, het keuze en het herstelproces.

2.4 Samen beslissen
Factoren gerelateerd aan het meebeslissen over keuze van benadering.

3 Codes omtrent de implementeerder (chirurg)

3.1 Attitude chirurg
Factoren gerelateerd aan de houding van de chirurg ten opzichte van de twee verschillende benaderingen.

3.1.1 Ziet geen reden voor verandering (-)
Factoren gerelateerd aan dat de chirurg geen reden ziet om zijn werkwijze te veranderen.

3.2 Ervaringen van de chirurg

3.2.1 Skills van de chirurg
Factoren gerelateerd aan de bekwaamheden van de chirurg.

3.2.2 Hoeveelheid operaties voert de chirurg uit
Factoren gerelateerd aan hoe vaak de chirurg de operatie uitvoert.

3.2.3 Leercurve
Factoren gerelateerd aan de leercurve van het leren van een benadering en het effect op de uitkomst van de operatie.

3.3 Routine

3.3.1 Voorkeur van de chirurg
Factoren gerelateerd aan de voorkeur voor een bepaalde benadering van de chirurg

3.3.2 Het gaat altijd al op deze manier

Factoren gerelateerd aan het argument dat de keuze voor de chirurgische benadering wordt gemaakt op basis van dat het altijd al op deze manier heeft plaats gevonden.

3.4 Complicaties buiten het zicht van de chirurg
Factoren gerelateerd aan complicaties die de chirurg niet terug ziet op zijn polikliniek of spoedeisende hulp.

3.4.1 Gebrek aan kennis over gevolgen complicaties
Factoren gerelateerd aan het gebrek van kennis over de gevolgen en impact die complicaties kunnen hebben voor de patiënt.

4 Codes omtrent de setting waar de interventie plaats vindt (organisatie)

4.1 Opleiding van de chirurg
Factoren gerelateerd aan de opleidingen van de chirurg bijvoorbeeld opleidingsplaats.

4.2 Huidige werkwijze binnen de organisatie
Factoren gerelateerd aan de chirurgische benadering die gebruikt wordt binnen de organisatie.

4.3 Rol van de geriater
Factoren gerelateerd aan de rol die de geriater speelt bij patiënten die behandeld worden met een kop hals prothese.

4.4 Consensus

4.4.1 Consensus binnen het team over type benadering
Factoren gerelateerd aan consensus binnen het team over type benadering dat gebruikt wordt bij het plaatsen van een kop-hals prothese.

4.4.2 Consensus binnen het team omtrent de richtlijnen
Factoren gerelateerd aan het bereiken van consensus over het wel of niet gebruiken van richtlijnen.

4.4.3 Consensus tussen verschillende afdelingen
Factoren gerelateerd aan consensus tussen betrokken afdelingen over bijvoorbeeld werkwijze en het gebruik van richtlijnen.

<u>5 Codes omtrent de bredere context (sociaal-politieke context)</u>

5.1 Omvang van het probleem
Factoren gerelateerd aan de omvang van de patiëntenpopulatie die behandeld worden met een kop-hals prothese

5.2 Makers van de richtlijnen
Factoren gerelateerd aan de schrijvers van richtlijnen omtrent de interventie

5.3 Verschillende betrokken stakeholders
Factoren gerelateerd aan de verschillende stakeholders die betrokken zijn omtrent kop-hals protheses bij het proces van beslissing, uitvoering en herstel.

5.4 Financiële gevolgen door complicaties

Factoren gerelateerd aan financiële gevolgen voor de maatschappij.

5.5 Innovatie in de toekomst

5.5.1 Moderne technologie

Factoren gerelateerd aan de ontwikkelingen in moderne technologieën en de uitkomst na een operatie waar een kop-hals prothese wordt geplaatst.

5.5.2 Ontwikkelingen in instrumenten

Factoren gerelateerd aan de ontwikkelingen van instrumenten en de uitvoering van beide chirurgische benaderingen bij een operatie waar een kop-hals prothese wordt geplaatst

6 Codes omtrent Strategieën

- Specialisten schrijven de richtlijnen zelf
- Het is niet moeilijk om een andere chirurg een andere benadering te leren
- De chirurg heeft genoeg tijd nodig om de andere benadering te leren
- Bereid om onderzoeksresultaten te implementeren indien nodig
- Geriater kan en wil een rol spelen in chirurgen overtuigen van de urgentie
- Ervoor zorgen dat onderzoeksresultaten bekend worden onder mensen

REFERENCES

1. Bergstrom U. JH, Gustafson Y., Pettersson U., Stenlund H., Svensson O. The hip fracture incidence curve is shifting to the right. Acta Orthop. 2009;80(5):520-4.
2. Hu F. JC, Shen J., Tang P., Wang Y. Preoperative predictors for mortality following hip fracture surgery: a systematic review and meta-analysis. injury. 2012;43(6):676-85.
3. Pugely AJ MC, Gao Y, Klocke NF, Callaghan JJ, Marsh JL. A risk calculator for short-term morbidity and mortality after hip fracture surgery. J Orthop Trauma. 2014;28(2):63-9.
4. Vidan M SJ, Moreno C, Riquelme G, Ortiz J. Efficacy of a comprehensive geriatric intervention in older patients hospitalized for hip fracture: a randomized, controlled trial. J Am Geriatr Soc. 2005;53(9):1476-82.
5. Alexiou KI. RA, Varitimidis SE., Malizos KN. . Quality of life and psychological consequences in elderly patients after a hip fracture: a review. Clin Interv Aging. 2018;13:143-50.
6. Tol MCJM, van Beers LWAH, Willigenburg NW, Gosens T, Heetveld MJ, Willems HC, Bhandari M, Poolman RW. Posterolateral or direct lateral approach for hemiarthroplasty after femoral neck fractures: a systematic review. Hip int. 2021;31(2):154-65.
7. LROI. Chirurgische benadering kophalsprothese. 2020.
8. Leonardsson O, Rolfson O, Rogmark C. The surgical approach for hemiarthroplasty does not influence patient-reported outcome. The Bone & Joint Journal. 2016;98:542-7.
9. Kristensen TB, Vinje T, Havelin LI, Engesaeter LB, Gjertsen JE. Posterior approach compared to direct lateral approach resulted in better patient-reported outcome after hemiarthroplasty for femoral neck fracture. Acta Orthop. 2017;88(1):29-34.
10. Rogmark C, Fenstad AM, Leonardsson O, Engesaeter LB, Karrholm J, Furnes O, et al. Posterior approach and uncemented stems increases the risk of reoperation after hemiarthroplasties in elderly hip fracture patients. Acta Orthop. 2014;85(1):18-25.
11. Biber R, Brem M, Singler K, Moellers M, Sieber C, Bail HJ. Dorsal versus transgluteal approach for hip hemiarthroplasty: an analysis of early complications in seven hundred and four consecutive cases. Int Orthop. 2012;36(11):2219-23.
12. Unwin AJ, Thomas M. Dislocation after hemiarthroplasty of the hip: a comparison of the dislocation rate after posterior and lateral approaches to the hip. Ann R Coll Surg Engl. 1994;76:327-9.
13. Hongisto MT, Nuotio MS, Luukkaala T, Vaisto O, Pihlajamaki HK. Lateral and posterior approaches in hemiarthroplasty. Scand J Surg. 2018:1457496917748226.
14. Berstock JR BA, Beswick AD. A systematic review and meta-analysis of complications following the posterior and lateral surgical approaches to total hip arthroplasty. . Ann R Coll Surg Engl. 2015;97(1):11-6.
15. Sayed-Noor A, Hanas A, Skoldenberg O, Mukka S. Abductor Muscle Function and Trochanteric Tenderness After Hemiarthroplasty for Femoral Neck Fracture. J Orthop Trauma. 2016;30(6).
16. Tol MCJM, Willigenburg NW, Rasker AJ, Willems HC, Gosens T, Heetveld MJ, et al. Posterolateral or Direct Lateral Surgical Approach for Hemiarthroplasty After a Hip Fracture: A Randomized Clinical Trial Alongside a Natural Experiment. JAMA Network Open. 2024;7(1):e2350765-e.
17. Green J TN. Qualitative methods for health research. 4 ed. London: SAGE; 2018.
18. Tol MCJM, Willigenburg NW, Willems HC, Gosens T, Rasker A, Heetveld MJ, et al. Posterolateral or direct lateral approach for cemented hemiarthroplasty after femoral neck fracture (APOLLO): protocol for a multicenter randomized controlled trial with economic evaluation and natural experiment alongside. Acta Orthop. 2022;12(93):732-8.
19. Hardinge K. The direct lateral approach to the hip. J Bone Joint Surg Br. 1982;64(1):17-9.

20. Moore A. Metal hip joint; a new self-locking vitallium prosthesis. South Med J. 1952;45(11):1015-9.
21. Fleuren M WK, Paulussen T. Determinants of innovation within health care organizations: literature review and Delphi study. Int J Qual Health Care. 2004;16(2):107-23.
22. K B. How Does the Adult Surgeon Learn? Thoracic Surgery Clinics. 2019;29(3):233-8.
23. Verrier ED. The Surgeon as Educator. Thoracic Surgery Clinics. 2019;29(3):227-32.
24. Lin J. RR. Teaching, Mentorship, and Coaching in Surgical Education. Thoracic Surgery Clinics. 2019;29(3):311-20.
25. Frambach JM. vdVC, Durning SJ. Quality criteria in qualitative and quantitative research. Acad Med. 2017;92(3):423-33.
26. Wolfenden L, Foy R, Pressau J, Grimshaw JM, Ivers NM, Powell BJ, et al. Designing and undertaking randomised implementation trials: guide for researchers. BMJ. 2021;371:3721.

4

Treatment Strategies for Femoral Neck Fractures

Hemiarthroplasty or Total Hip Arthroplasty for Femoral Neck Fractures: 12-year Follow-up of Randomised Trial.

[1]M.C.J.M. Tol, MD, [2]M.P.J. van den Bekerom, MD, PhD, [3]I.N. Sierevelt, PhD, [4]E.F. Hilverdink, MD, [4]E.L.F.B. Raaymakers, MD, PhD, [1]J.C. Goslings, MD, PhD, On behalf of the surgeons of the ARTHRO trial group.

[1]Department of Trauma Surgery, Academic Medical Center, Amsterdam, [2]Department of Orthopaedic Surgery, OLVG, Amsterdam, [3]Department of Orthopaedic Surgery & Research, Slotervaart center, [4]Department of Orthopaedic Surgery, Academic Medical Center

Bone and Joint Journal (2017)

ABSTRACT

Introduction The aim of this study was to analyse the long term functional outcome of two different treatments for an intracapsular femoral neck fracture in active patients aged over 70 years old. Treatments were randomised and involved either a hemiarthroplasty (HA) or a total hip replacement (THR).

Patients and methods Initially 252 patients were included of whom 205 (81%) were women, with a mean age of 81.1 years (70.2 to 95.6). They were treated with a cemented HA (137 patients) or cemented THR (115 patients). At the long term follow up of 12 years we analysed the modified Harris Hip Score (HHS), post-operative complications and intra—operative data of the patients who were still alive.

Results After 12-year follow-up 50 patients (20%) (HA n= 32/137, THR n=18/115) were still alive, of which 47 (94%) were female. No significant differences were observed in the modified HHS, mortality, general complications or revision rate of the prosthesis between patients with a HA or a THR.

Conclusion There is no difference in the functional outcome between HA and THR treatments after 12-year follow-up.

INTRODUCTION

Displaced intracapsular fractures of the proximal femur in the elderly population are a worldwide health care problem. Annually, there are 64.000 patients admitted to the hospital suffering from a hip fracture in England, which is associated with impaired mobility and an increased morbidity and mortality, and therefore an impaired quality of life. (1-3)

The treatment options for patients with displaced femoral neck fractures are internal fixation, hemiarthroplasty (HA) or a total hip replacement (THR). While internal fixation is recommended as the treatment of choice in young patients(4) and elderly patients who are medically not able undergoing prosthetic surgery(5), there is still no golden standard in the treatment for active patients aged over 70 years. Burgers et al. showed in a systematic review that due to the higher functional outcome, treatment with a THR might be preferable in active elderly patients with a displaced femoral neck fracture. Although there was a higher frequency of dislocations in this patient group(6). Despite the advantages of HA, such as a quick and highly standardized procedure that allows an early weight bearing and recovery, a revision to a THR is sometimes needed due to developing of symptomatic acetabular erosion or even protrusion.(7) The revision rate reported in literature is very variable but based on recent research lower than previously thought.(8)

Recently a protocol was published of a large multicenter randomized controlled trial, the HEALTH trial, to assess the patient outcome after THR and hemiarthroplasty. However, due to the limited final follow up of only two years, the differences in long term functional outcome will not be observable with this study.(9)

Currently there is a lack of studies that observed the long term results of hip fracture surgery in elderly patients who either received a hemiarthroplasty or a THR. In the past two decades several studies were published with a follow up more than five years.(10-14) However, the availability and so the number of randomized controlled trials is small. Therefore, the ongoing discussion whether the active elderly patient with a hip fracture should be treated with a THR or HA remains the question. Hence a well conducted randomized trial with a large sample size is needed to draw a definitive conclusion.(6) With this study we aim to contribute to an answer in this long lasting controversy in orthopaedic trauma surgery. Therefore, this study was to analyze the long term functional outcome of two different treatments for an intracapsular femoral neck fracture in active patients aged over 70 years old, without having advanced osteoarthritis or rheumatoid arthritis. In a previously study we reported the outcome one and five years post-operatively.(15) In current study the results after twelve-years follow-up are presented.

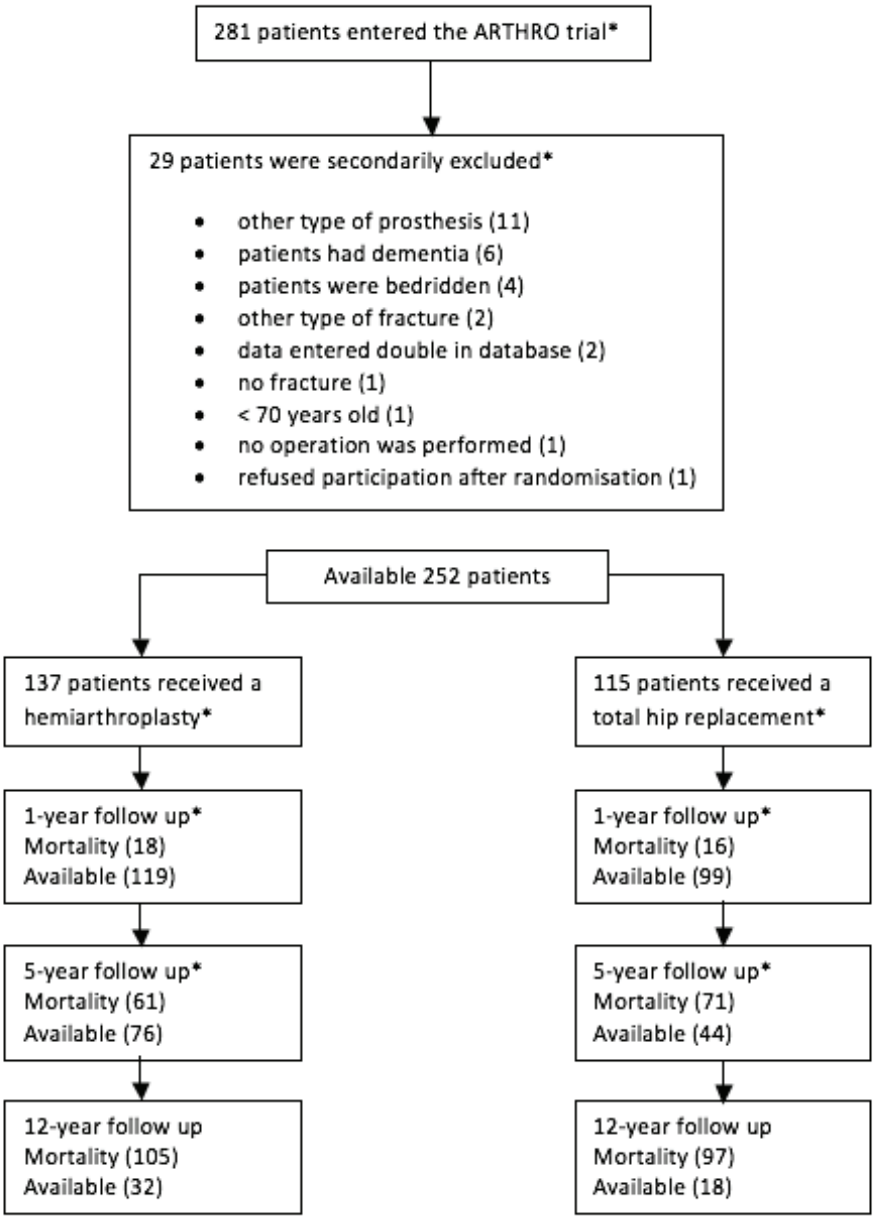
PATIENTS AND METHODS

Between January 1995 and January 2002, all patients who were admitted in a participating hospital with a displaced femoral neck fracture were considered for trial entry and a total of 252 patients provided written informed consent. The patients were randomly assigned to receive either a HA or a THR using either a Weber Rotationsprosthese (Sulzer AG, Winterthur, Switzerland) or a Müller Geradschaftprothese (Protek AG, Münsingen, Switzerland). Blinding of surgeons, patients or outcome investigators was not applied. Details of the study methodology, including surgical technique, have been previously reported.(15)

For the twelve-year follow-up evaluation, patients were interviewed by telephone. Patients completed a modified HHS as primary outcome.(2) In the modified HHS, physical examination score was not included. This outcome score has been used in previous studies. (3,4,15) The modified HHS score was converted to the maximal of 100 points. Secondary outcomes included revision rate, mortality, prosthesis-related and general complications and intra-operative data such as blood loss and duration of surgery.

To complete the demographic and health status details of the patient, information was obtained from first degree family members, general practitioners and other known healthcare employees. At the twelve-year follow-up, radiographic imaging of the pelvis and the axial hip was only indicated in case of hip complaints. The modified HHS of patients with new imaging was assessed during their visit. The institutional Review Board (Medical Ethics Committee) of the participating hospitals approved the protocol and the study was conducted in accordance with the Helsinki declaration, as amended in Tokyo and Seoul.(16)

Fig I, Flowchart of the patients recruited



***Data conducted** from "A comparison of hemiarthroplasty with total hip replacement for displaced intracapsular fracture of the femoral neck"

Statistical analyses

All data were entered by three investigators (MPJB, EMBPR, EH) in a SPSS database v. 15.0 for statistical analyses and calculations. The accuracy of this data was analysed by two other investigators (EFH, INS). Depending on their distribution, statistical significant differences between the treatment groups of both the HHS as the other continuous parameters were tested by either a Mann-Whitney U test or a Student's t-test. Chi-squared tests or Fisher's exact test were applied to analyse dichotomous variables, such as revision rates and displacements of the prosthesis. All analyses were two-sided executed, a p-value <0.05 was considered as statistically significant.

Table I, Baseline characteristics of all patients

	Original population		**12 year follow up**	
	HA† (n= 137)	THR‡ (n= 115)	HA† (n= 32)	THR‡ (n= 18)
Gender, n (%)				
Female	115 (84)	90 (78)	30 (94)	17 (94)
Male	22 (16)	25 (22)	2 (6)	1 (6)
Mean age of trauma (SD)	80.3 (6.2)	82.1 (6.3)	77.2 (4.4)	78.5 (5.5)
Mean years of follow up (SD)			12.4 (2.3)	11.3 (2.1)
Fracture side, n (%)				
Left	78 (57)	74 (64)	15 (47)	11 (61)
Right	59 (43)	41 (36)	17 (53)	7 (39)
Co-morbidities, n (%)				
Malignancy	11 (8)	6 (5)	1 (3)	0 (0)
Cardiovascular	34 (25)	38 (33)	4 (13)	7 (39)
Respiratory	16 (12)	18 (16)	1 (3)	1 (5)
Neurological	26 (19)	33 (29)	2 (6)	3 (17)
Diabetes Mellitus	19 (14)	11 (10)	3 (9)	0 (0)
Musculosketal	22 (16)	31 (27)	4 (13)	7 (39)
ASA classification¥, n (%)				
I	19 (14)	11 (9)	9 (28)	4 (22)
II	77 (56)	48 (42)	19 (59)	9 (50)
III	33 (24)	44 (38)	4 (13)	4 (22)
IV	5 (4)	10 (9)	0 (0)	1 (6)
V	0 (0)	0 (0)	0 (0)	0 (0)
Unknown	3 (2)	2 (2)	0 (0)	0 (0)

†HA, Hemiarthroplasty
‡THR, Total Hip Replacement
¥ASA, American society of Anesthesiologists

RESULTS

Initially a total of 281 patients were considered for study entry. After randomisation 29 patients were excluded since they did not match the inclusion criteria or due to protocol violations. Of these 29 patients, 11 did not receive the prosthesis for which they had been randomised, one patient refused participation after randomisation, six patients suffered from dementia, four patients were bedridden, two patients were diagnosed with intertrochanteric hip fractures, one patient did not have a proximal femur fracture, one patient who did not receive an operation, one patient was only 69 years old at the time of admission and data of two patients were entered double in the database.

Consequently, 252 patients were included and their data were available for analysis, 137 were randomised to receive a HA and 115 patients received a THR (Fig. 1). After 12 years follow-up 50 patients (20%) (HA n= 32/137(23%), THR n=18/115(15%)) were still alive, of which 47 (94%) were female. Mean age at time of trauma was 77.7 (SD 4.4) years in the HA group versus 78.5 (SD 5.5) in the THR group. The side of the fracture was evenly distributed between the two groups, with 24 on the right and 26 on the left. Patient characteristics are depicted in Table 1.

Table 2, outcome measures at 12 year follow-up

	HA† (n=32)	**THR**‡ (n=18)	**p-value**
Mean modified HHS (SD)® ¥	70.3 (16.3)	69.3 (20.0)	P=0.85
Mean HHS pain (SD)¥	39.8 (9.1)	18.3 (7.4)	P=0.44
Mean HHS function (SD)¥	16.4 (8.8)	37.2 (10.0)	P=0.34
Satisfaction of treatment, n (%)*			
Yes	27 (84)	11 (61)	P=0.47
Revision operation, n (%)*	0 (0)	0 (0)	
Dislocation of the prosthesis, n (%)*	0 (0)	0 (0)	

® **The** total modified Harris Hip Score (HHS) was converted to a maximum of 100 points
† **Hemiarthroplasty**
‡ **Total** Hip Replacement
¥ **Student** t-test
* **Chi** Square test

The 12-year outcome measures are shown in Table 2. The mean modified HHS was 70.3 (SD 16.3) in patients who received an HA and 69.3 (SD 20.0) in patients with a THR (P=0.85). There were no dislocations in either the hemiarthroplasty group or THR group after 12-year follow up. Furthermore, at the final follow up none of the patients had un-

dergone a revision operation. Similarly, no significant differences were noted between the interventions HA and THR in overall mortality rate, respectively 105 (77%) patients versus 97 (84%) patients (P=0.013). At the 12-year follow-up 84% of the patients treated with a HA were satisfied with their treatment, versus 61% in the THR group (P=0.47).

However, in a previously study significant differences were observed in intra-operative data between the HA and THR group (Table 3). A blood loss of less than 500 ml during surgery occurred in 28 patients (97%) who were treated with a HA and in eight patients (57%) with a THR. Six patients (43%) with THR surgery had a blood loss of more than 500 ml versus one patient with a hemiarthroplasty (P<0.01). Moreover, the THR group had a longer duration of surgery compared with the HA group (p<0.01)(15).

Table 3, operative details of the patients

	HA† (n=32)	**THR**‡ (n=18)	**p-value***
Time between trauma and surgery			
≤1 day, n (%)	25 (78)	10 (56)	P=0.06
Missing, n (%)	1	0	
Blood loss, n (%)			
< 500 ml	28 (97)	8 (57)	P<0.01
> 500 ml	1 (3)	6 (43)	
Missing	3	4	
Duration of operation, n (%)			
< 1 hour	14 (50)	0 (0)	P<0.01
1 – 1.5 hour	13 (46)	9 (75)	
> 1.5 hour	1 (4)	3 (25)	
Missing	4	6	

† Hemiarthroplasty
‡ Total Hip Replacement
*** Chi** Square test

DISCUSSION

Our findings in this update of our previous 5-year follow-up suggest that the functional outcome of the treatment of active elderly patients with displaced femoral fractures with hemiarthroplasty are similar to THR. These findings relate to patients aged ≥70 years, who are in relatively good physical and mental condition and have neither osteoarthritis or rheumatoid arthritis affecting the hip. However, there seems to be a trend towards treatment with a THR in patients with these conditions without conclusive evidence supporting this direction.(17)

Several trials compared hemiarthroplasty with THR for displaced femoral neck fractures before.(5,6,18) In most of these studies there was a lack of randomization and/or long

term follow-up though. Two other studies are similar and based on computer randomization with a long term follow-up of more than five years.(10,11) Despite a tendency to better functional outcome scores in patients with a THR, Avery et al. came to similar findings with no significant differences in functional outcome or revision rate.(11) However these conclusions are in contrast with the findings of a randomized controlled trial that observed a better functional outcome of THR after a follow up of 13 years.(10) Although, in contrast to our patient population, Ravikumar et al. compared uncemented hemiarthroplasty with cemented THR and a number of studies have suggested that cemented arthroplasty will lead to better mobility and less pain in comparison to an uncemented prosthesis. (19,20)

The possible higher need for revision is the major concern after hemiarthroplasty for intracapsular femoral neck fractures.(21-23) Although the preference of some surgeons to perform a THR to avoid long-term wear of acetabular cartilage and the need for revision of the hemiarthroplasty to a THR, we did not observe revisions in either the hemiarthroplasty or the THR group after 12 years of follow-up.(24) These findings are similar to the results of a systematic review, which enclosed pooled data of a total of 986 patients. (6) Despite hypothetically more revisions after hemiarthroplasty in the small number of patients of follow up, Avery et al. neither showed significant differences in revision rate.

Compared to the original and 5 year follow-up, the population is altered over time. After 12 years follow-up the majority of the patient group were healthy women with less comorbidities and were younger at the time of trauma. Therefore, it is a selection of the original population and in view of the above the results have to be seen.

In our study population of patients who were still alive after 12 years follow-up we observed neither dislocations in the hemiarthroplasty group or the THR group. However, contrary results were found in our five years follow-up study, where there were significant more dislocations in patients treated with a THR.(15) Hypothetically, revision surgery or dislocations might be a risk factor that increases mortality and therefore only patients without these complications are still alive after 12 years of follow-up. However, we did not assess the need of revision or dislocation in patients who died between five and 12 years follow-up. Hence, the question if revision surgery or dislocations influences the mortality remains unknown.

Regarding the overall limited life expectancy of most elderly patients especially in the femoral neck fracture population, the satisfaction of the treatment is paramount. It is debatable if the outcome of this question is reliable in this population. With aging, the number of comorbidities is rising as well. The question is, if these factors influence their well-being and thereby satisfaction of their physical state after hip surgery.

The conclusions of our study have to be drawn in the light of some limitations. Per protocol analysis was a commonly used method to analyze data in 1995 when our study was initiated. We might have introduced a risk of bias due to the assessment of clinical variables by outcome investigators who were not blinded to the type of surgical intervention. Complications after discharge from hospital, rather than dislocations and revision surgery, were not recorded. We used the modified HHS. (3,4,15) Finally, the small sample caused by mortality in a population with already a higher age at time of trauma.

Historically randomized controlled trials have been scarce in orthopaedic trauma literature but should play a major role in determining the best practices for fracture care. (25) This study includes a multicenter prospectively computer generated randomization design with a substantial long term follow up of twelve years, which are the strengths of this study. Thereby due to the strict inclusion criteria, the population to whom the conclusions can be applied was clearly defined.

Conclusion

We conclude that after 12 years of follow-up there is no difference in functional outcome between hemiarthroplasty and THR treatments. Due to the higher rate of dislocations in patients with a THR after 5 years follow-up and the intra-operative data, we recommend hemiarthroplasty as the treatment of choice for patients aged ≥70 years with a fracture of the femoral neck, in the absence of advanced radiological osteoarthritis or rheumatoid arthritis of the hip. Further randomised controlled trials with a larger sample size are required to confirm our findings and investigate the influence of revision surgery and dislocations on mortality.

REFERENCES

1. Griffin XL, Parsons N, Achten J, Fernandez M, Costa ML. Recovery of health-related quality of life in a United Kingdom hip fracture population. The Warwick Hip Trauma Evaluation--a prospective cohort study. Bone Joint J. 2015 Mar ;97-B(3):372–82.
2. Lanting LC , Stam C , Hertog PC den BM. Dutch National Public Health Compass. 2006 Sep
3. Johnell O, Kanis J a. An estimate of the worldwide prevalence, mortality and disability associated with hip fracture. Osteoporos Int [Internet]. 2004 Nov ;15(11):897–902.
4. Bhandari M, Devereaux PJ, Swiontkowski MF, Tornetta P, Obremskey W, Koval KJ, et al. with Arthroplasty for Displaced Fractures of the Femoral Neck: a meta-analysis. J Bone Joint Surg [Am] 2003 ;85-A:1673-81.
5. Parker MJ, Khan RJ, Crawford J PG. Hemiarthroplasty versus internal fixation for displaced intracapsular hip fractures in the elderly. A randomised trial of 455 patients. J Bone Jt Surg Br. 2002 Nov; 84(8):1150–5.
6. Burgers PTPW, Van Geene AR, Van den Bekerom MPJ, Van Lieshout EMM, Blom B, Aleem IS, et al. Total hip arthroplasty versus hemiarthroplasty for displaced femoral neck fractures in the healthy elderly: a meta-analysis and systematic review of randomized trials. Int Orthop. 2012 Aug ;36(8):1549–60.
7. Maxted MJ, Denham RA. Failure of hemiarthroplasty the femur. Injury. 1984 Jan ;15(4):224–6.
8. van den Bekerom MPJ, Sierevelt IN, Bonke H, Raaymakers ELFB. The natural history of the hemiarthroplasty for displaced intracapsular femoral neck fractures. Acta Orthop. 2013 Dec ;84(6):555–60.
9. Bhandari M, Devereaux PJ, Einhorn T a, Thabane L, Schemitsch EH, Koval KJ, et al. Hip fracture evaluation with alternatives of total hip arthroplasty versus hemiarthroplasty (HEALTH): protocol for a multicentre randomised trial. BMJ Open. 2015 Jan ;5(2):e006263.
10. Ravikumar KJ, Marsh G. Internal fixation versus hemiarthroplasty versus total hip arthroplasty for displaced subcapital fractures of femur — 13 year results of a prospective randomised study. Injury. 2000;31:793–7.
11. Avery PP, Baker RP, Walton MJ, Rooker JC, Squires B, Gargan MF BG. Total hip replacement and hemiarthroplasty in mobile, independent patients with a displaced intracapsular fracture of the femoral neck: a seven- to ten-year follow-up report of a prospective randomised controlled trial. J Bone Jt Surg Br. 2011 Aug ;93(8):1045–8.
12. Healy WL, Iorio R. Total hip replacement: optimal treatment for displaced femoral neck fractures in elderly patients. Clin Orthop Relat Res. 2005 Jan ;429:43–8.
13. Schleicher I, Kordelle J, Jürgensen I, Haas H MC. Femoral neck fractures in the elderly -- bipolar hemiarthroplasty in total hip replacement. Unfallchirurg. 2003 Jun ;106(6):467–71.
14. Xu X, Liu Y, Liu J LY. Prosthetic replacement in treatment of subcapital femoral neck fractures in the elderly. Chin J Traumatol. 2002 Feb ;5(1):28–31.
15. van den Bekerom MPJ, Hilverdink EF, Sierevelt IN, Reuling EMBP, Schnater JM, Bonke H, et al. A comparison of hemiarthroplasty with total hip replacement for displaced intracapsular fracture of the femoral neck: a randomised controlled multicentre trial in patients aged 70 years and over. J Bone Joint Surg Br. 2010 Oct ;92(10):1422–8.
16. Declaration Helsinki H. World Medical Association Declaration of Helsinki. 2008.
17. Miller BJ, Callaghan JJ, Cram P, Karam M, Marsh JL, Noiseux NO. Changing trends in the treatment of femoral neck fractures: a review of the american board of orthopaedic surgery database. J Bone Joint Surg [Am] 2014 ;96(17):e149

18. Zhao Y, Fu D, Chen K, Li G, Cai Z, Shi Y, et al. Outcome of Hemiarthroplasty and Total Hip Replacement for Active Elderly Patients with Displaced Femoral Neck Fractures : A Meta-Analysis of 8 Randomized Clinical Trials. PLoS ONE. 2014;9(5):1–7.
19. Emery R, Broughton N, Desal K, Bulstrode C, TL T. Bipolar Hemiarthroplasty for subcapital fracture of the femoral neck: a prospective randomised trial of cemented Thompson and uncemented Moore stems. J Bone Jt Surg Br. 1991;73:322–4.
20. Dorr lawrence d, Glousman R, Sew Hoy A l, Vanis R, Chandler R. Treatment of femoral neck fractures with total hip replacement versus cemented and noncemented hemiarthroplasty. J Arthroplasty. 1986;1(1):21–8.
21. Poignard A, Bouhou M, Pidet O, Flouzat-Lachaniette CH, Hernigou P. High Dislocation Cumulative Risk in THA versus Hemiarthroplasty for Fractures. Clin Orthop Relat Res. 2011;469 (11):3148–53.
22. Baker RP, Squires B, Gargan MF, Bannister GC. Total hip arthroplasty and hemiarthroplasty in mobile, independent patients with a displaced intracapsular fracture of the femoral neck. A randomized, controlled trial. J Bone Joint Surg Am. 2006;88(12):2583–9.
23. Rodriquez-Merchan E. Displaced intracapsular hip fractures: hemiarthroplasty or total arthroplasty? Clin Orthop Relat Res. 2002;399:72–7.
24. Blomfeldt R, Törnkvist H, Eriksson K, Söderqvist a, Ponzer S, Tidermark J. A randomised controlled trial comparing bipolar hemiarthroplasty with total hip replacement for displaced intracapsular fractures of the femoral neck in elderly patients. J Bone Joint Surg Br. 2007;89(2):160–5.
25. Poolman RW, Struijs PAA, Krips R, Sierevelt IN, Lutz KH, Bhandari M. Does a " Level I Evidence " rating imply high quality of reporting in orthopaedic randomised controlled trials ? BMC Med Res Methodol. 2006;8:1–8.

5

Posterolateral or direct lateral approach for hemiarthroplasty after femoral neck fractures: a systematic review

Maria C.J.M. Tol[1], MD; Loes W.A.H. van Beers[1], MSc; Nienke W. Willigenburg[1], PhD; Taco Gosens[2], MD, PhD; Martin J. Heetveld[3], MD, PhD; Hanna C. Willems[4], MD, PhD; Mohit Bhandari[5], MD, PhD, Prof; Rudolf W. Poolman, MD, PhD[1,6]

[1]Department of Orthopaedic Surgery, OLVG, Amsterdam, [2]Department of Orthopaedic Surgery, Elisabeth-TweeSteden Hospital, Tilburg, [3]Department of Trauma Surgery, Spaarne Gasthuis, Haarlem, [4]Department of Internal Medicine and Geriatrics, AMC, Amsterdam, [5]Department of Orthopaedic Surgery, McMaster University, Hamilton, [6]Department of Orthopaedic Surgery, LUMC, Leiden

Hip International (2020)

ABSTRACT

Background: The posterolateral approach (PLA) and direct lateral approach (DLA) are the most commonly used approaches for inserting a hemiarthroplasty in the treatment of femoral neck fractures. A recent review concludes that the routine use of PLA should be questioned, but this conclusion itself can be questioned. The aim of this study is to provide an updated overview and critical appraisal of the available evidence, focussing on outcomes most relevant for patients.

Methods: We conducted a comprehensive search of literature in the MEDLINE and EMBASE databases and Cochrane Library. Studies (till June 2018) to identify hip fracture clinical trials/comparative studies comparing alternative surgical approaches (PLA and DLA). We explored sources of heterogeneity and conducted pooled analyses when appropriate.

Results: 264 potentially eligible studies were identified of which 1 RCT, 3 prospective, 3 registry data and 5 retrospective studies were included. The RCT consisted performance and attrition bias. The mean MINORS score of the prospective/register studies was 17.3 (SD 3.5) and 13.8 (SD 1.9) of the 5 retrospective studies. The GRADE score for all the outcomes were very low . Due to the high and various types of biases across the included studies, we did not pool the data. None of studies assessed the activities of daily living functionality. Six studies reported significantly more dislocations or reoperations due to dislocation in the PLA group, six other studies found no differences. DLA patients were more likely to develop abductor insufficiency leading to limping and more need for walking aids. The PLA patients tended to have better quality of life, less pain and more satisfaction compared to the DLA patients.

Conclusion: Based on low quality studies, PLA may be associated with more dislocations, but patients had less walking problems and a lower tendency to abductor insufficiency compared with DLA. Further clinical trials with methodology rigor are needed to determine which approach is more effective in terms of outcomes relevant to patients.

INTRODUCTION

Hemiarthroplasty is a commonly used procedure in the treatment of displaced proximal femur fractures. One of the first hemiarthroplasties was introduced in 1940 by Moore and a posterolateral approach (PLA) was initially recommended as surgical technique(1, 2). In the subsequent years various operative techniques were described in the literature including the direct lateral approach (DLA) of Hardinge(3). In the Netherlands, 83% of all hemiarthroplasties are inserted through a PLA or DLA(4).

A recent systematic review and meta-analysis concludes that the routine use of PLA for fracture related hemiarthroplasty should be questioned. This conclusion was based on an increased risk of dislocation compared with other approaches, with no evident advantages to counterbalance this disadvantage of the PLA (5). However, the methodological quality of the included studies was low, and combining data of studies with a high and differential risk of bias may result in misleading findings and compromise the conclusion(6). Moreover, potential advantages of the PLA may be less evident, but could still be present. Therefore, the firm conclusion that a PLA would be inferior to DLA should also be questioned.

When inserting a hemiarthroplasty through a DLA the fibers of the gluteus medius and vastus lateralis muscles are split in their own line and when closing the muscles are sutured(7). This approach may lead to abductor insufficiency, resulting in a positive Trendelenburg sign or limping after surgery(8, 9). However, in the DLA the posterior capsule is preserved, preventing dislocation. With a PLA, the surgeon performs a posterior capsulotomy, dividing the short external rotators(7). During a PLA the hip abductors are protected and preserved to prevent limping. However, due to inadequate posterior capsule support there may be an increased risk of dislocations(10-12).

Sustaining a hip fracture increases the risk of falling and often leads to subsequent fractures, which results in high morbidity and increased mortality in these patients(13). While highly relevant for patients, the effect of surgical approach on subsequent fall risk is unknown. It could be speculated that the PLA may lead to faster rehabilitation and better balance due to the scatheless gluteus musculature. Contrarily, the loss of abductor muscle strength as a result of the DLA could cause a loss of balance which may increase the risk of falling and subsequent trauma, further reducing patients' mobility.

In this systematic review we aim to provide an update and critical appraisal of the available evidence of the differences between the most frequently performed surgical approaches for inserting hemiarthroplasty after femoral neck fractures: PLA and DLA. We focus on outcomes most relevant for patients, with the patients' independence in Activi-

ties of Daily Living (ADL) as primary outcome. Secondary outcomes include: postoperative complications, mobility, function, quality of life and falls.

METHODS

We conducted this systematic review using the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines (14). The protocol of this study is registered in the PROSPERO register (CRD42017055162).

Eligibility criteria

For the effects on patient outcome after acute hip fracture, we included all study designs in which the posterolateral approach was directly compared with the direct lateral approach in hemiarthroplasty after hip fractures. We excluded studies that enrolled patients with pathological fractures, fractures due to advanced rheumatoid- or osteoarthrosis, multi-trauma-patients (ISS> 16) or patients who had secondary surgery after failed internal fixation.

Information sources and Literature search

An independent clinical librarian conducted a literature search in the MEDLINE and EMBASE databases and Cochrane Library at the 13th of June 2018. There were no restrictions to publication year and languages. Search terms were: Femoral neck fracture*, Hip fracture*, Proximal femur fracture*, Hemiarthroplasty, Direct lateral, Posterolateral.(Appendix I)

Study selection

Two reviewers (*first and second author, initials blinded for peer review*) independently screened all titles and abstract for eligibility with the criteria mentioned above using Covindence(15). References resulting from the included studies were reviewed for additional relevant articles. When there was any discrepancy between the two reviewers consensus was reached through discussion.

Data collection process

One researcher (*first author, initials blinded for peer review*) extracted all data of the following predefined data items: (1) study information: authors and year of publication, study design, number of patients enrolled and follow-up; (2) baseline characteristics: surgical approach, age, gender, cemented/uncemented prosthesis, cognitive impairment and ASA classification; (3) patient outcomes

Outcomes

In a focus group formed from members of the collaboration of orthopaedics, trauma surgeon and clinical geriatrics we determine the outcomes of this systematic review by their expert opinion. These outcomes were supported by the Dutch Federation of patients.

The primary outcome of this study is the patients' independence in ADL using the KATZ index(16). Secondary outcomes are postoperative complications, mobility, function, Health Related Quality of Life and falls.

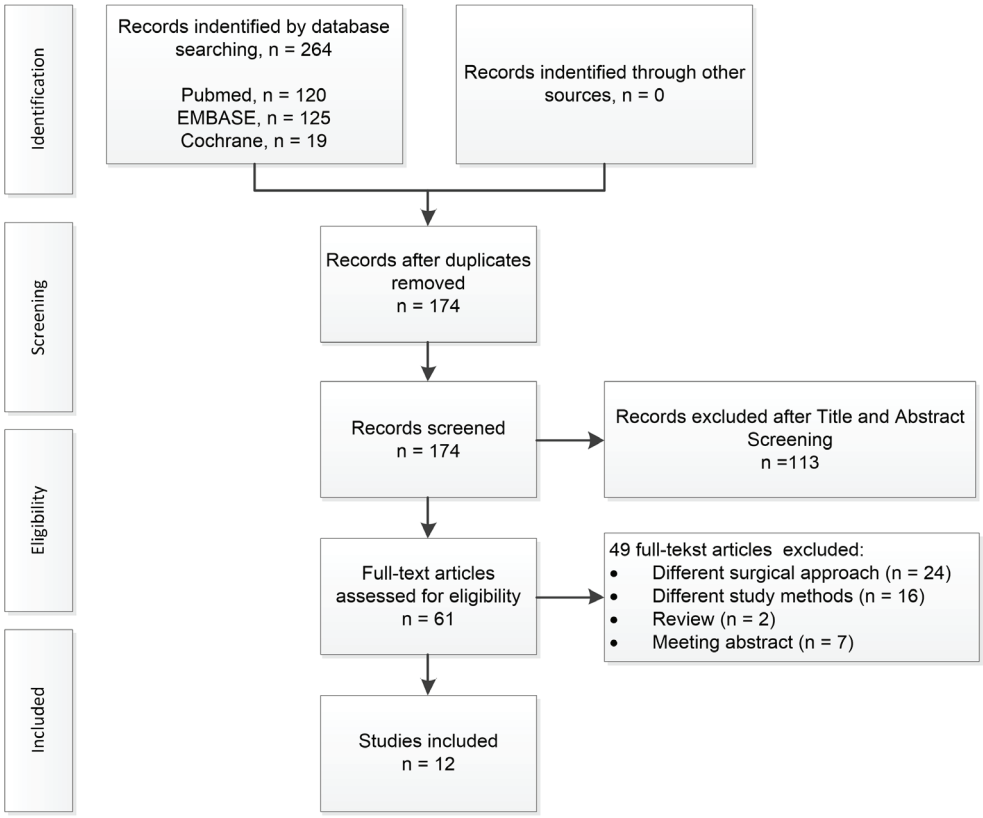
Critical appraisal

To assess the level of evidence we used the Oxford Level of Evidence criteria (17). Two reviewers (*first and second author, initials blinded for peer review*) independently completed the critical appraisal. Consensus was reached after discussion. The quality assessment for Randomised controlled trials was conducted by the Cochrane tool for risk of bias using Review Manager(18). Studies are classified as "low risk", "moderate risk" or "high risk". We assessed the methodological quality for non-randomised controlled trials with the MINORS criteria(19). The MINORs criteria contains 12 items subsequently scored 0-2; 0 indicating that it was not reported, 1 indicating that it was reported but inadequate, and 2 indicating that it was reported and adequate. The maximal score for comparative studies is therefore 24 points.

The type of study design was considered to be of significant importance for the for the interpretation of the study results. Based on observational data causal inferences cannot be made. Causation is defined by a different risk in the study population under two potential exposure values, which only can be reached after randomisation. Therefore, in observational studies only associations could be made(20). Recently, Natural Experiments (NE) are growing interest as a alternative to RCT's(21). A NE is a study design, whereby the dividing of the population into exposed and unexposed to an intervention is not under control of the researcher(22). Natural factors, such as geographical location, leads to exposed and unexposed individuals, which creates as-if randomness. There are several methods to evaluate the selective exposure in NE's. These additional design features can strengthen the causal inference(21).

We also assessed the overall quality of evidence and strength of recommendation by the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach(23). This approach quantifies potential limitations for each outcome in five domains: risk of bias, inconsistency, imprecision, indirectness, and publication bias.

Figure 1.



RESULTS

Search results

The search produced 264 records, of which 90 records were duplicates and were removed. A total of 174 records remained to be screened for title and abstract, after which 113 irrelevant records were eliminated. Sixty-one full-text articles were assessed for eligibility, of which 49 were excluded for the following reasons: different surgical approaches (n=24), different study methods (n=16), review (n=2), meeting abstract (n=7). There were no articles added after screening the references of the included studies. A total of 12 articles were included in this systematic review (Figure 1).

Of the 12 included articles one was a RCT (24), 3 were prospective studies(9, 25, 26), 3 were registry studies(10, 12, 27) and 5 were retrospective studies(11, 28-31)

Table 1

Study	Design	n	Mean Follow-up time, mo	Surgical approach, n (%)	Age, mean (SD)	Female, n (%)	Cemented prothesis, n (%)	ASA 3-5, n (%)	Cognitive impairement, n (%)	Primary outcomes	Main conclusion
Hongisto, 2018	ROS	269	12	PLA: 118 (44) DLA: 151(56)	82.8 (6.3)	PLA: 94 (80) DLA: 118 (78)	PLA: 24 (20) DLA: 12 (8)	PLA: 103 (87) DLA: 129 (85)	PLA: 25 (21) DLA: 40 (27)	Living arrangements, need for mobility aids, mobility level and pain	DLA needs more ambulatory aids 1 year after hip fracture
Svenoy, 2017	POS (NE)	583	12	PLA: 186 (32) DLA: 397 (68)	82.8 (7.8)	PLA: 143 (77) DLA: 291 (73)		PLA: 89 (48) DLA: 235 59)	PLA: 48 (26) DLA: 128 (32)	Early complications	PLA has an increased risk for (recurrent) dislocations
Sayed-Noor, 2016*	POS	48	12	PLA: 24 (50) DLA: 24 (50)		PLA: 19 (79) DLA: 20 (83)	PLA: 24 (100) DLA: 24 (100)	PLA: 9 (38) DLA: 14 (58)	NR	Sign of Trendelenburg	The incidence of Trendelenburg was higher in DLA
Mukka, 2016*	POS	185	12	PLA: 83 (45) DLA: 102 (55)		PLA: 61(74) DLA: 68 (67)	NR	PLA: 42 (51) DLA: 57 (56)	NR	Hip function, pain, complication and reoperation rate	No differences in function or complications
Leonardsson, 2016	RS	2118	12	PLA: 978 (46) DLA: 1140 (54)	85 (8)	PLA: 731 (75) DLA: 838 (74)	PLA: 954 (98) DLA: 1110 (97)	PLA: 349 (44) DLA: 579 (55)	PLA: 173 (22) DLA: 237 (23)	Patient-reported outcomes (PROMs)	After adjusting, no differences in PROMs
Ozan, 2016	ROS	233	17	PLA: 147 (63) DLA: 86 (37)		PLA: 79 (54) DLA: 57 (66)	NR	PLA: 102 (69) DLA: 60 (69)	NR	Postoperatieve complications	No differences in postoperative complications

Table 1 (*continued*)

Study	Design	n	Mean Follow-up time, mo	Surgical approach, n (%)	Age, mean (SD)	Female, n (%)	Cemented prothesis, n (%)	ASA 3-5, n (%)	Cognitive impairement, n (%)	Primary outcomes	Main conclusion
Kristensen, 2016	RS	20908	12	PLA: 1990 (10) DLA: 18918 (90)	83 (7)	PLA: 1424 (72) DLA: 137770 (73)	PLA: 851 (43) DLA: 14283 (75)	PLA: 1242 (63) DLA: 11970 (64)	PLA: 582 (29) DLA: 4809 (25)	PROMs and re-operation rate	PLA results in better PROMs. No differences in reoperation rate.
Parker, 2015	RCT	216	12	PLA: 108 (50) DLA: 108 (50)		PLA: 98 (91) DLA: 100 (93)	NR	PLA: 65 (60) DLA: 64 (59)	NR	Regain of walking ability	No differences in mobility between both approaches.
Rogmark, 2014	RS*	33205	32	PLA: 11999 (36) DLA: 21206 (64)	84 (6.7)	Overall 24059 (72)	Overall 29990 (90)	NR	NR	Reoperation rate	PLA: more reoperation due to dislocation and due to infection in DLA.
Biber, 2012	ROS	704	NR	PLA: 487 (69) DLA: 217 (31)	80.4 (9.8)	Overall 492 (70)	Overall 673 (96)	NR	NR	Early surgical complications	PLA is predisposed to dislocation, DLA is predisposed to haematoma.
Unwin, 1994	ROS	2906	3	PLA: 1656 (57) DLA: 1250 (43)	NR	NR	Overall 628 (22)	NR	NR	Dislocation rate	Higher dislocation rate in PLA
Paton, 1989	ROS	171	6-48	PLA: 93 (54) DLA: 78 (46)	79.3	NR	NR	NR	NR	Dislocation rate	No differences in dislocation rate between approaches

Abbreviations: mo= months, RCT= randomised controlled trial, POS= prospective observational study, ROS= retrospective observational study , NE= Natural Experiment, RS= register study, NR= not reported, SD= standard deviaton, ASA= merican association of anesthesiologists, DLA=Direct lateral approach, PLA= posterolateral approach

***consist the** same patient populatio

Table 2, Quality assessment of non-randomised studies (MINORS)

Quality evaluation criteria*	Hongisto 2018	Svenoy 2017	Kristensen 2017	Ozan 2016	Mukka 2016	Sayed-Noor, 2016	Leonardsson 2016	Rogmark 2014	Biber 2012	Unwin 1994	Paton 1989
Clearly stated aim	2	2	2	1	2	2	2	2	2	2	2
Inclusion of consecutive patients	2	0	2	1	2	2	2	2	2	0	2
Prospective collection of data	1	2	1	1	2	2	1	1	1	1	1
Endpoints appropriate to the study aim	1	1	1	1	1	1	1	1	1	1	1
Unbiased assessment of the study endpoint	0	0	0	0	2	2	0	0	0	0	0
Appropriate follow-up period	2	2	2	2	2	2	2	1	0	2	1
Loss to follow-up less than 5%	1	0	1	0	2	1	1	0	2	1	0
Prospective calculation of the study size	0	0	0	0	2	1	0	0	0	0	0
Adequete control group	2	2	2	2	2	2	2	2	2	2	2
Contemporary groups	2	2	2	2	2	2	2	2	2	2	2
Baseline equivalence	2	2	1	1	2	2	1	0	1	0	0
Adequate statistical analysis	2	2	2	2	2	2	2	2	2	1	1
Total	17	15	16	13	23	21	16	13	15	12	12

***The items** are scored 0: non-reported; 1: reported but inadequate; 2:reported and adequate. The ideal score being 24 for comparative studies.

Table 3, Grade analyse

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**PLA**	**DLA**	**Relative (95% CI)**	**Absolute (95% CI)**		
Dislocation (follow up: mean 12-17 months)												
6	observational studies	very serious [a]	not serious	not serious	serious [b]	none	57/1129 (5.0%)	13/1061 (1.2%)	not pooled	see comment	⨁◯◯◯ VERY LOW	
Infection (follow up: mean 12-17 months)												
6	observational studies	very serious [a]	not serious	not serious	very serious [c]	none	44/1011 (4.4%)	38/910 (4.2%)	not pooled	see comment	⨁◯◯◯ VERY LOW	
Quality of life (follow up: mean 1 years; assessed with: EQ-5D; Scale from: -0.11 to 1.0)												
3	observational studies	very serious [a]	not serious	not serious	serious [d]	none	2992	20082	not pooled	see comment	⨁◯◯◯ VERY LOW	
Pain (follow up: mean 1 years; assessed with: VAS (SD); Scale from: 0 to 100)												

Table 3, Grade analyse (*continued*)

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**PLA**	**DLA**	**Relative (95% CI)**	**Absolute (95% CI)**		
2	observational studies	very serious [e]	not serious	not serious	not serious	none	2968	20.058	not pooled	see comment	⊕◯◯◯ VERY LOW	
Satisfaction (follow up: mean 1 years; assessed with: VAS (SD); Scale from: 0 to 100)												
2	observational studies	very serious [e]	not serious	not serious	not serious	none	2968	20058	not pooled	see comment	⊕◯◯◯ VERY LOW	
Limping (follow up: mean 1 years; assessed with: OR)												
1	observational studies	serious [f]	not serious	not serious	not serious	none	24	24	not pooled	see comment	⊕◯◯◯ VERY LOW	

CI: Confidence interval

Explanations

a. No blinding, no randomisation, follow-up not reported properly in some studies

b. Of 6 of the 7 included studies the sample size is too small

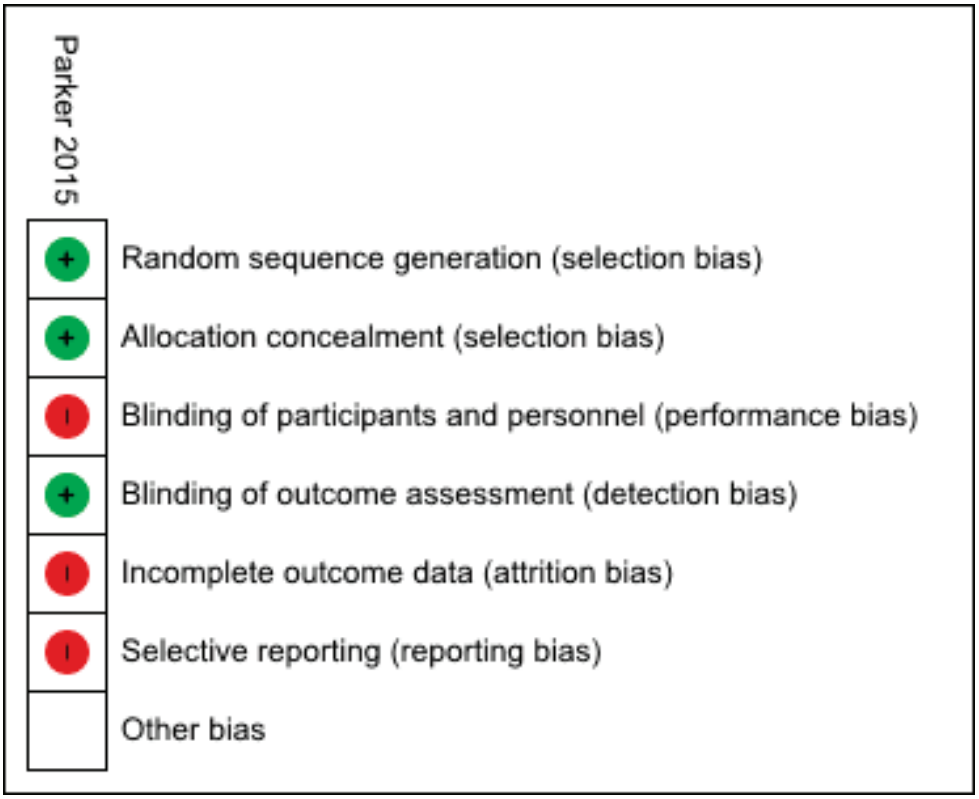
c. Of 5 of the 6 included studies the sample size is too small

d. Of 1 of the 3 included studies the sample size is too small

e. No blinding, no randomisation

f. No randomisation

Figure 2, Risk of bias of randomised controlled trial



Study populations

The study populations ranged from 48 to 33205 patients. In total, more than 60.000 patients participated in the included studies. Baseline characteristics of the patient populations are presented in Table 1. The patient population of Sayed-Noor et al.(9) was a subset of the population that participated in Mukka et al.(25) (confirmed after contact with the corresponding author). Rogmark et al.(10) combined the Norwegian register database, which was also used by Leonardsson (27), with the Swedish registry, which was also used by Kristensen et al. (12).

There was heterogeneity in populations across the included studies (Table 1). Patients treated using the PLA in Kristensen et al.(27) had more cognitive impairment (29%) compared with DLA (25%)(27). They also reported more cemented prostheses in the DLA group (75% vs. PLA 43%)(27). These differences in the use of cemented prostheses were seen across several included studies: the percentage of cemented prostheses in the DLA group was >90% in four studies (9-12), 72% in one registry study (27), 21% in one retrospective study(29) and only 13% of the patients were treated with a cemented prosthesis in the study of Hongisto(31). The latter study reported significantly more patients in the PLA group (20%) treated with a cemented hemiarthroplasty compared with DLA (8%)(31).

Also, Hongisto et al.(31) reported a significant difference in surgeon's expertise. Patients in the DLA group were more frequently treated by a post-registrar (76%) compared to the PLA patients (64%)(31). In five studies the surgical approach was not equally divided in the study population(10, 11, 26, 27, 30). No other significant differences between the PLA and DLA groups were described in the included studies. Other co-morbidities were not reported in the included studies.

Methodological quality

The RCT, all prospective and retrospective studies had level of evidence 2b (individual cohort studies or low quality RCT (<80% follow-up)). The RCT was prematurely ended when half of the sample size was included, because the PLA was found technically more difficult(24). This might have introduced performance bias in the included patients and attrition bias due to incomplete outcome data (Figure 2). The MINORS criteria of the non-randomised studies are listed in Table 2. The mean MINORS score of the prospective cohort and registry studies was 17.3 (SD 3.5) and of the 5 retrospective cohort studies the mean MINORS score was 13.8 (SD 1.9) out of 24 points. Figure 3 presents the overall MINORS score of the individual methodological items of all included studies. In 7 of the 12 items there were considerable differences in methodological quality. Hence, the majority of the included studies showed imperfections in the study design. By design, non-randomised studies lack random allocation of patients to study groups. This increases the risk for selection bias and confounding, especially when factors other than surgical approach also differ between groups (see above). Due to the diverse study designs and the high and differential risk of bias, pooling of the data may result in misleading findings and therefore should be avoided(6).

Grade approach

The GRADE approach was feasible for the following outcomes: postoperative complications, Health Related Quality of Life and function. The summary of findings for each of the included outcomes is listed in the GRADE evidence profile (Table 3). The GRADE quality of all outcomes was very low, in other words, any estimate of effect is very uncertain.

PRIMARY OUTCOME

None of the included studies assessed the Independence in Activities of Daily Living using the KATZ score or anything similar.

Table 4, Postoperative complications

		Follow-up	PLA	DLA	p-value
Single dislocation, n (%)					
Hongisto, 2018		12 mo	4 (3.4)	0 (0)	0.04
Svenoy, 2017		12 mo	15 (8)	4 (1)	< 0.01
Mukka, 2016		12 mo	1 (1.2)	2 (2)	ns
Ozan, 2016		6-39 mo	17 (11.5)	4 (4.6)	ns
Parker, 2015		12 mo	1 (0.9)	2 (1.9)	ns
Biber, 2012		Unknown	19 (3.9)	1 (0,5)	0.01
Unwin, 1994		3 mo	149 (9)	41 (3.3)	< 0.01
Paton, 1989		6-48 mo	8 (8.6)	2 (2.6)	ns
Recurrent dislocation, n (%)					
Svenoy, 2017		12 mo	9 (5)	2 (0.5)	< 0.01
Mukka, 2016		12 mo	6 (7.2)	1 (1)	ns
Reoperation due to dislocation, n (%)					
Mukka, 2016		12 mo	5 (6)	1 (1.2)	ns
Leonardsson, 2016		7-22 mo	20 (2)	10 (0.9)	0.02
Rogmark, 2014	HR	32 mo	2.2 (CI 1.8-2.6)		< 0.01
Kristensen, 2016	RR	96 mo	1.2 (CI 0.92-1.4)		ns
Infection, n (%)					
Svenoy, 2017		12 mo	12 (6)	20 (5)	ns
Mukka, 2016		12 mo	5 (6)	5 (5)	ns
Ozan, 2016		6-39 mo	11 (7.4)	3 (3.4)	ns
Parker, 2015		12 mo	4 (3.7)	3 (2.8)	ns
Biber, 2012		Unknown	12 (2.5)	7 (3.2)	ns
Reoperation due to infection, n (%)					
Mukka, 2016		12 mo	5 (5)	4 (3)	ns
Leonardsson, 2016		7-22 mo	13 (1.3)	12 (1.1)	ns
Rogmark, 2014	HR	32 mo	0.8 (0.7-1.3)		0.05

Abbreviations: PLA = posterolateral approach, DLA = Direct lateral approach, mo = months, ns = non significant, HR = Hazard Ratio, RR = Risk Ratio

SECONDARY OUTCOMES

Postoperative complications

All postoperative complications are listed in Table 4.

Dislocations

Ten out of twelve included studies with in total 59111 included patients reported dislocation (Figure 4a)(10, 11, 24-31). Dislocation rates ranged from 1% to 12% in the PLA group and from 0% to 6% in the DLA group during a follow-up period of 3 to 48 months (11, 24-26, 28-31). The only included RCT found no differences in dislocation rate between the two approaches(24). A large registry study which combined the Norwegian and Swedish registry data, (n= 33205), showed that PLA was a significant risk factor for reoperation due to dislocation (10). The Swedish registry data alone supported this finding, with 2% of all PLA patients undergoing a reoperation due to dislocation versus 0.9% in the DLA group(12). However, the Norwegian registry study (n=20908) found similar risks of reoperation due to dislocation of both approaches in the first 8 year(27). After one year follow-up three retrospective studies described an increased risk of dislocation in patients treated using the PLA (11, 29, 31). Other observational studies with in total 589 patients did not report more dislocations in the PLA group(25, 28, 30)(Table 4).

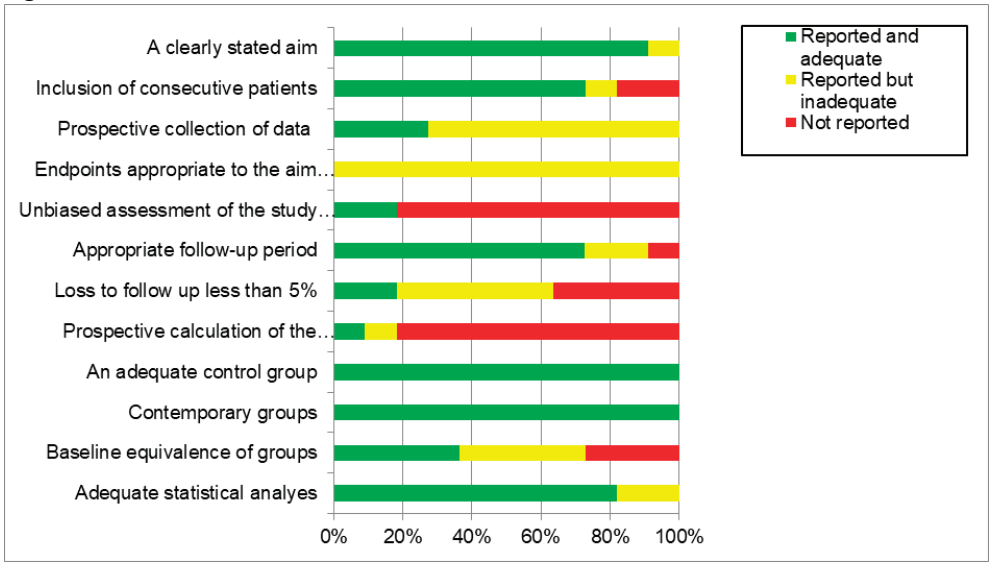
Infections

One registry study (n=20908) reported DLA as a risk factor for reoperation due to infection compared to PLA. No other significant differences in occurrence of infection between PLA and DLA were described in previous studies (Figure 4b, Table 4)(11, 24-26, 30).

Length of stay

Only the RCT assessed the length of hospital stay, and found no significant differences between PLA (18.5 days) and DLA (20.3 days, p=0.4)(24).

Figure 3,



Quality of life, pain and satisfaction

One RCT and three prospective studies with 23924 patients reported quality of life after hip fracture(9, 12, 24, 27). The patient reported quality of life, mobility and function are listed in Table 5. In the RCT (n=216) no differences in modified Charnley pain score(32) were found between PLA and DLA(24). Two studies that prospectively assessed the HRQoL using the EQ-5D reported significant differences after 1-year follow-up in favour of the PLA(12, 27). Specifically, Kristensen et al. (27) observed an adjusted mean difference of 0.03. (n=20908)(27). When assessing the patient's self-rated health on a vertical visual analogue scale (EQ-VAS), in which the endpoints were labelled 'The best health you can imagine' (100 points) and 'The worst health you can imagine' (0 points), PLA patients scored after adjustment 61 points vs. 59 points in DLA patients(27). In addition, when measuring the VAS pain ranged from zero (no pain) to 100 (unbearable pain) and the level of satisfaction ranged from zero (very satisfied) to 100 (dissatisfied), PLA patients had less pain and were more satisfied compared with DLA patients one year after surgery(27). Leonardsson et al.(12) also observed significant differences favouring PLA, however after adjusting for confounders, no significant difference in EQ-5D, pain and satisfaction between the approaches were seen(12).

Figure 4a, dislocation

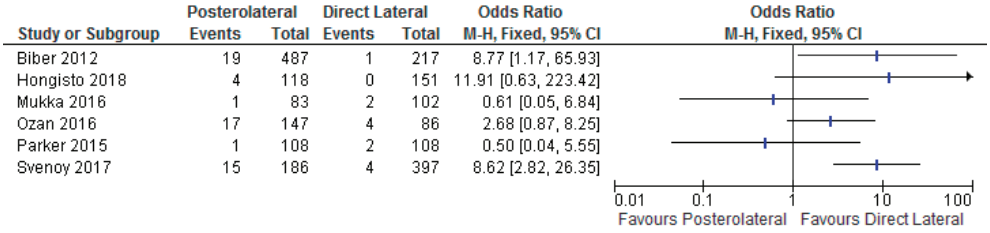
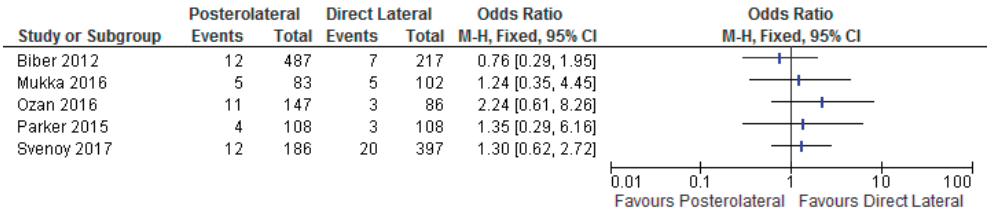


Figure 4b, infection



Mobility, Function

Five studies, of which 1 RCT, 3 prospective and 1 retrospective studies, with a total of 21626 patients reported mobility or function after hip fracture surgery. The RCT (n=216) used a mobility score with a range from 0 to 9 points and found no clinically relevant differences between the groups(24). Patients in a prospective cohort, treated using the DLA had an 18 times higher risk of having a positive Trendelenburg sign and had a 16-

fold higher risk of developing limping(9). Remarkably, no differences in abductor muscle strength, measured with an electronic dynamometer, were found in this same patient population(9). This may be due to the difficulty to follow the instructions when using the dynamometer is this elderly population(9). The Harris Hip Score (HHS) was assessed in one prospective study, and no difference was observed(25)(Table 5). The Norwegian registry data estimated a rate of 42% of the PLA group with no walking problems one year after surgery compared with the estimated rate of 30% in the DLA group ($p<0.01$)(27). A retrospective study reported a significant OR of 2.73 for being mobile without walking aid favouring the PLA(31). However in this retrospective study, both groups were equally fully mobile (with or without walking aids), since there were no differences in mobility level between the groups(31).

Falls

Parker et al. observed the incidence of falling and found no difference between the approaches during the hospital stay(24). No study reported falls or fall incidents during follow up.

DISCUSSION

This review is an update of the available evidence of the differences between the two most common surgical approaches of hemiarthroplasty after femoral neck fractures. We focussed on outcome measures that are most relevant for patients.

Key findings

Since all outcomes were rated as very low quality according to the GRADE approach, any estimated effect could be considered as uncertain. There were major concerns about the quality of the included studies. Eleven out of twelve studies were non-randomised studies(9-12, 25-31), which might have introduced selection bias due to lack of randomisation that could have caused unbalanced confounding factors among patient groups. The MINORS tool identified various types of biases across the included studies. It is unclear which of these biases have had the greatest impact and how they have varied between clinical situations. The only RCT was prematurely ended and was therefore underpowered. Moreover, the RCT terminated because the PLA was felt technically more difficult by the surgeon. Therefore the quality of the study results can be questioned.

Large nationwide registry data could be seen as a NE, such as Kristensen et al.(27) However, in the other included registry studies it was not clearly described whether the surgical approach was chosen based on particular patient characteristics, or because the surgeon was only competent in one single approach. The latter case, as-if randomness could be

reached when the surgical approach is merely determined by chance (which surgeon is on call when the patient is admitted to the hospital).

Table 5

	PLA	DLA	p-value	Adjusted difference	p-value
Quality of life					
EQ-5D (SD)					
Sayed-Noor, 2016*	0.71 (0.33)	0.81 (0.22)	ns		
Leonardsson, 2016	0.52 (0.37)	0.47 (0.37)	<0.01	- 0.01	ns
Kristensen, 2016	0.64	0.61	0.01	- 0.03	0.01
EQ-VAS					
Kristensen, 2016	64	62	0.05	- 2.1	0.05
Pain, VAS (SD)					
Leonardsson, 2016	17 (19)	19 (20)	0.02	1.4	ns
Kristensen, 2016	17	20	<0.01	3.1	< 0.01
Satisfaction (SD)					
Leonardsson, 2016	22 (23)	24 (24)	0.02	1.5	ns
Kristensen, 2016	21	25	<0.01	4.7	< 0.01
Mobility					
No walking problems (%)					
Kristensen, 2016	42	30	<0.01		
Mobile without walking aid (%)				OR	
Hongisto, 2018	22	11.9	0.02	2.73	0.02
Full mobility after surgery (%)					
Hongisto, 2018	46.2	54	ns		
Mean mobility score					
Parker, 2015	5	5.2	ns		
Function					
Trendelenburg (OR)					
Sayed-Noor, 2017	1	17.5	0.02		
Limping (OR)					
Sayed-Noor, 2017	1	16	<0.01		
HHS (SD)					
Mukka, 2017	72 (17)	71 (18)	ns		

*** is** a subset of the population of Mukka, 2017

Abbreviations: PLA: posterolateral approach, DLA: direct lateral approach, SD: standard deviation, OR: odds ratio, HHS: Harris hip score, ns: non

Previous literature

None of the included studies focussed on the independency in ADL and therefore we have no results on the primary outcome. From our point of view, in in this frail population where the life expectancy is not high, the ADL dependency and quality of life is paramount. Especially because hip fracture patients can become more vulnerable and more dependent after subsequent falls.

This current critical appraisal focussed on the outcomes relevant to the patient, which makes our study different compared to Van der Sijp et al. that described a variety of outcome measures(5). Moreover, we performed an additional assessment of the study quality (i.e. GRADE) which could have resulted in a different study interpretation. Furthermore, the study of Van der Sijp et al. mentioned that the Scandinavian registry studies excluded cognitively impaired patients and therefore they introduced sampling bias(5). However, in contrast to this statement, both studies have actually included all consecutive patients including those with cognitive impairment(12, 27). Lastly, Van der Sijp et al also stated that the high prevalence of a Trendelenburg sign and limping does not affect the clinical outcome(5). However, this conclusion was based on a sample of 48 patients, which is underpowered for this patient reported quality of life outcome.

In regards to the secondary outcome measure we found that the majority of the literature reported less dislocations in the DLA group 1 year after surgery(10-12, 31). The only RCT found similar dislocation rates, however this study was ended halfway(24). Contrary results were also found in the Scandinavian registry studies were the Swedish data showed that PLA was an increased risk for reoperation due to dislocation (12) but the Norwegian data observed no differences in reoperation 8 years after surgery between the approaches(27). Recurrent dislocation is a major complication which results in a loss of HRQoL and can lead to a reoperation, but a single dislocation seems to cause only a temporary deterioration of the quality of life (33). This is supported by the Swedish registry study, that reported more reoperation due to dislocation in PLA patients but also a slightly higher quality of life in this group one year postoperatively. The natural experiment by Kristensen et al also showed a better quality of life in PLA patients, which supports the findings of better outcomes regarding pain, satisfaction and quality of life after a total hip arthroplasty using the PLA(34, 35).

The mobility was assessed in the included RCT, which reported no difference. Patients were randomised and treated by one single surgeon. It might have introduced a performance bias due to the fact that the PLA was felt technically more difficult and with no differences of better functional outcomes, the study was terminated prematurely(24). On the contrary, more than 10% less walking problems in the PLA group were found in the natural experiment(27). Furthermore, more patients operated using the DLA needed a walking aid one year postoperatively compared with the PLA, which suggest a worse

functional outcome after hemiarthroplasty using the DLA(31). However, when assessing the Harris Hip Score in a prospective study, this effect was not seen (25).One study with a low risk of bias according to the MINORS criteria, reported a strong association with abductor insufficiency (a positive Trendelenburg sign) leading to limping in DLA patients(9). Whether limping after hemiarthroplasty is associated with a higher tendency to fall has never been investigated. However, it seems likely that the loss of muscle strength leads to a loss of function and to a higher level of frailty, which might intensify the risk of falling and possibly lead to more consecutive injuries and disability during daily activities(36). Only the RCT reported the incidence of falling. They observed no differences in the incidents of falling during a hospital stay(24). However, there is no literature on the tendency to fall after discharge and it remains unknown if the abductor insufficiency after DLA causes to more falls. Too little is currently known about the effect of abductor insufficiency on the patient's daily activities and the risk of falling after both approaches.

Strengths and limitation

There were limitations to our study. First, we were not able to pool the results in a meta-analysis due the heterogeneity of the included studies. Second, in this systematic review the adult population as a whole was described and it is unknown if certain subpopulations (i.e. patients with cognitive impairment) had different outcomes and might benefit from one of the two approaches. Finally, we excluded seven meeting abstracts because there was no data available, this might have introduced publication bias. The strengths of this study are the thorough critical appraisal including the GRADE approach and the comprehensive description of outcomes that matter to patients.

Conclusion

Compared with DLA, PLA might be associated with more dislocations, but patients had less walking problems and a lower tendency to abductor insufficiency. However, based on the current evidence causal inference can not be made. Moreover, little is known on the consequences of the major risks (dislocation vs. abductor insufficiency) in terms of fall risk and independent functioning in ADL.

Implications for future research

Randomised clinical trials with methodological rigor are needed, focusing on outcomes which are important for the patient.

APPENDIX I

("Femoral Neck Fractures"[Mesh] OR "Hip Fractures"[Mesh] OR proximal femur fracture*[tiab] OR femoral neck fracture*[tiab] OR femur neck fracture*[tiab] OR hip fracture*[tiab]) AND ("Hemiarthroplasty"[Mesh] OR hemiarthroplast*[tiab] OR hemi-arthroplast*[tiab]) AND (direct lateral*[tiab] OR lateral*[tiab] OR transgluteal[tiab] OR hardinge[tiab] OR posterolateral[tiab] OR anterior[tiab] OR smith-petersen[tiab] OR anterolateral[tiab] OR watson-jones[tiab] OR posterior[tiab])

REFERENCES

1. Moore A. Metal hip joint; a new self-locking vitallium prosthesis. South Med J. 1952;45(11):1015-9.
2. Moore A, Bohlman H. Metal hip joint, a case report J Bonte Jt Surg Br 1943;25:688.
3. Hardinge K. The direct lateral approach to the hip. The journal of Bone and Joint Surgery. 1982;64.
4. LROI. Surgical approaches of hemiarthroplasty. 2017.
5. Max P.L. van der Sijp DvD, Pieta Krijnen, Arthur H.P. Niggebrugge, Inger, Schipper B. Surgical approaches and hemiarthroplasty outcomes for femoral neck fractures: a meta-analysis. The journal of Arthroplasty. 2018.
6. Joshua D. Harris MD, Jefferson C. Brand, M.D., Mark P. Cote, P.T., D.P.T., M.S.C.T.R., and, Aman Dhawan MD. Research Pearls: The Significance of Statistics and Perils of Pooling. Part 3: Pearls and Pitfalls of Meta-analyses and Systematic Reviews. Arthroscopy: The Journal of Arthroscopic and Related Surgery. 2017;33(8):1594-602.
7. S. Hoppenfeld PdB, R. Buckley Surgical Exposures in Orthopaedics: The Anatomic Approach.
8. Berstock J, Blom A, Beswick A. A systematic review and meta-analysis of complications following the posterior and lateral surgical approaches to total hip arthroplasty. Ann R Coll Surg Engl. 2015;97(1):11-6.
9. Sayed-Noor A, Hanas A, Skoldenberg O, Mukka S. Abductor Muscle Function and Trochanteric Tenderness After Hemiarthroplasty for Femoral Neck Fracture. J Orthop Trauma. 2016;30(6).
10. Rogmark C, Fenstad AM, Leonardsson O, Engesaeter LB, Karrholm J, Furnes O, et al. Posterior approach and uncemented stems increases the risk of reoperation after hemiarthroplasties in elderly hip fracture patients. Acta Orthop. 2014;85(1):18-25.
11. Biber R, Brem M, Singler K, Moellers M, Sieber C, Bail HJ. Dorsal versus transgluteal approach for hip hemiarthroplasty: an analysis of early complications in seven hundred and four consecutive cases. Int Orthop. 2012;36(11):2219-23.
12. Leonardsson OR, C. Rogmark. The surgical approach for hemiarthroplasty does not influence patient-reported outcome. The Bone & Joint Journal. 2016;98:542-7.
13. Paul T. P. W. Burgers SMZ, Adinda K. E. Mailuhu,, Martin J. Heetveld MHJV, Gert R. Roukema, Peter Patka, Rudolf W. Poolman, Esther M. M. Van Lieshout, on behalf of the Dutch femoral neck fracture investigator group. Cumulative incidence and treatment of non-simultaneous bilateral femoral neck fractures in a cohort of one thousand two hundred and fifty patients. International Orthopaedics. 2014;38:2335-42.
14. Moher D, Liberati A, Tetzlaff J, Altman D, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339.
15. Covidence [Available from: www.covidence.org.
16. Katz S. Assessing self-maintenance: Activities of dialy living, mobility and intrumental activities of daily living. JAGS. 1983;31(12):721-6.
17. Gugiu PC WC, Coryn CLS, Hobson KA. An application of a new evidence grading system to research on the chronic care model. Eval Health Prof 2013;36:3-43
18. Higgings J, Altman D, Gotzsche P. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343.
19. K. Slim EN, D. Forestier, F. Kwiatkowski, Y. Panis, J. Chipponi. Methodological index for non-randomized studies (MINORS): development and validation of a new instrumetn. ANZ J Surg. 2003;73:712-6.
20. Hernán MA. A definition of causal effect for epidemiological research. J Epidemiol Community Health. 2004;58:265-71.

21. P. Craig SVK, A. Leyland, F. Popham. Natural Experiments: An overview of methods, approaches, and contributions to public health intervention research. Annual Review of Public Health. 2017;38:39-56.
22. Craig P CC, Gunnell D, Macintyre S, Petticrew M, . Using natural experiments to evaluate population health interventions: new Medical Research Council Guidance. J Epidemiol Community Health. 2012;66:1182-86.
23. Guyatt GH1 OA, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M, Schünemann HJ; GRADE Working Group. GRADE guidelines: 8. Rating the quality of evidence--indirectness. J Clin Epidemiol. 2011;64(12):1303-10.
24. Parker MJ. Lateral versus posterior approach for insertion of hemiarthroplasties for hip fractures: A randomised trial of 216 patients. Injury. 2015;46(6):1023-7.
25. Mukka S, Mahmood S, Kadum B, Skoldenberg O, Sayed-Noor A. Direct lateral vs posterolateral approach to hemiarthroplasty for femoral neck fractures. Orthop Traumatol Surg Res. 2016;102(8):1049-54.
26. Svenoy S, Westberg M, Figved W, Valland H, Brun OC, Wangen H, et al. Posterior versus lateral approach for hemiarthroplasty after femoral neck fracture: Early complications in a prospective cohort of 583 patients. Injury. 2017;48(7):1565-9.
27. Kristensen TB, Vinje T, Havelin LI, Engesaeter LB, Gjertsen JE. Posterior approach compared to direct lateral approach resulted in better patient-reported outcome after hemiarthroplasty for femoral neck fracture. Acta Orthop. 2017;88(1):29-34.
28. Paton RW HP. Hemiarthroplasty of the hip and dislocation. injury. 1989;20:167-9.
29. Unwin AJ TM. Dislocation after hemiarthroplasty of the hip: a comparison of the dislocation rate after posterior and lateral approaches to the hip. Ann R Coll Surg Engl. 1994;76:327-9.
30. F. Ozan ESO, S. Koyuncu, K. Gurbuz, F. Dogar, F. Vatansever, F. Duygulu. Effects of Hardinge versus Moore approach on postoperative outcomes in elderly patients with hip fracture. Int J Clin Exp Med. 2016;9(2):4425-31.
31. Hongisto MT, Nuotio MS, Luukkaala T, Vaisto O, Pihlajamaki HK. Lateral and posterior approaches in hemiarthroplasty. Scand J Surg. 2018:1457496917748226.
32. J. C. The long-term results of low-friciton arthroplasty of the hip performed as a primary intervention. . J Bonte Jt Surg Br. 1972;54:61-76.
33. Enocson A PH, Ponzer S, Törnkvist H, Dalén N, Tidermark J. Quality of life after dislocation of hip arthroplasty: a prospective cohort study on 319 patients with femoral neck fractures with a one-year follow-up. Qual Life Res 2009;18(9):1177-84.
34. Amlie E HLI, Furnes O, Baste V, Nordsletten L, Hovik O. Worse patient-reported outcome after lateral approach than after anterior and posterolateral approach in primary hip arthroplasty. Acta Orthop. 2014;85(5):463-9.
35. Lindgren J V WP, Karrholm J, Garellick G, Rolfson O. Patientreported outcome is infl uenced by surgical approach in total hip replacement: a study of the Swedish Hip Arthroplasty Register including 42,233 patients. Bone Joint J. 2014;96-B(5):590-6.
36. Organization WH. WHO Global Report on Falls Prevention in Older Age. 2007.

6

Posterolateral or Direct Lateral approach for cemented Hemiarthroplasty after Femoral Neck Fracture (APOLLO): protocol for a multicenter Randomized Controlled Trial with economic evaluation and Natural Experiment alongside.

Maria C.J.M. Tol[1]; Nienke W. Willigenburg[1]; Hanna C. Willems[2]; Taco Gosens[3,4]; Ariena Rasker[1]; Martin J. Heetveld[5]; Martijn G.M. Schotanus[6,7]; Johanna M van Dongen[8,9]; Bart Eggen[10]; Mate Kormos[10]; Stéphanie L. van der Pas[11,12]; Aad W. van der Vaart[10]; Rudolf W. Poolman[13]; on behalf of the APOLLO research group.

[1]Department of Orthopedic surgery, Joint Research, OLVG, Amsterdam, [2]Department of Internal Medicine and Geriatrics, Amsterdam UMC, Amsterdam, [3]Department of Orthopedics and Trauma Surgery, ETZ, Tilburg, [4]Department of Medical and Clinical Psychology, Tilburg University, [5]Department of Trauma Surgery, Spaarne Gasthuis, Haarlem, [6]Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen, [7]School of Care and Public Health Research Institute, Faculty of Health, Medicine and Life Science, Maastricht University, [8]Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences Research Institute, [9]Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam Public Health Research Institute, Amsterdam, [10]Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, Mekelweg 4, Delft, [11]Amsterdam UMC location Vrije Universiteit Amsterdam, Epidemiology and Data Science, De Boelelaan 1117, Amsterdam, [12]Amsterdam Public Health, Methodology, Amsterdam, [13]Department of Orthopedic surgery, LUMC, Leiden, the Netherlands

Acta Orthopaedica (2022)

APOLLO research group:
· Frank van Roon, MD. Department of Orthopedics and Trauma Surgery, ETZ, Tilburg,
· Martijn van Dijk, MD, PhD. Department of Orthopedic Surgery, Antonius Ziekenhuis, Utrecht,
· Jort Keizer, MD. Department of Trauma Surgery, Antonius Ziekenhuis, Utrecht,
· Anne J.H. Vochteloo, MD, PhD. Department of Orthopedic Surgery, OCON, Hengelo,
· Pieter Joosse, MD, PhD. Department of Trauma Surgery, Noord-West Ziekenhuis, Alkmaar,
· Bert Boonen, MD, PhD, Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen,
· Jetse Jelsma, MD, PhD, Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen,
· Dieuwertje Theeuwen, MD. Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen,
· Joris J.W. Ploegmakers, MD, PhD. Department of Orthopedic Surgery, UMCG, Groningen,
· Tim Schepers, MD. PhD. Department of Trauma Surgery. Amsterdam UMC, Amsterdam,
· Evelien van der Meij, MD. Department of Orthopedic Surgery, IJsselland, Capelle aan de Ijssel,
· Svenhjalmar H. van Helden, MD, PhD. Department of Orthopedic & Trauma Surgery, Isala, Zwolle,
· Rutger Zuurmond, MD, PhD. Department of Orthopedic & Trauma Surgery, Isala, Zwolle,
· Bart A. van Dijkman, MD. Department of Trauma Surgery, Flevoziekenhuis, Almere,
· Thomas D. Berendes, MD, PhD. Department of Orthopedic Surgery, Meander MC, Amersfoort,
· Hans G.E. Hendriks, MD, PhD. Departement of Orthopedic Surgery & Trauma, Máxima MC, Eindhoven, The Netherlands

ABSTRACT

Background: The posterolateral and direct lateral surgical approach are the 2 most common surgical approaches for performing a hemiarthroplasty in patients with a hip fracture. It is unknown which surgical approach is preferable in terms of (cost-)effectiveness and quality of life.

Methods and analysis: We designed a multicenter randomized controlled trial (RCT) with an economic evaluation and a natural experiment (NE) alongside. We will include 555 patients ≥ 18 years with an acute femoral neck fracture. The primary outcome is patient-reported health-related quality of life assessed with the EQ-5D-5L. Secondary outcomes include healthcare costs, complications, mortality, and balance (including fear of falling, actual falls, and injuries due to falling). An economic evaluation will be performed for Quality Adjusted Life Years (QALYs).We will use variable block randomization stratified for hospital. For continuous outcomes, we will use linear mixed-model analysis. Dichotomous secondary outcome measures will be analyzed using chi-square statistics and logistic regression models. Primary analyses are based on the intention-to-treat principle. Additional as treated analyses will be performed to evaluate the effect of protocol deviations.

INTRODUCTION

The worldwide incidence of hip fractures is expected to rise by 4 fold to 6.26 million in 2050, making it a global important health care problem(1). Hemiarthroplasty is a common treatment for hip fractures. Preferences for surgical approach vary between surgeons, hospitals and countries(2). A systematic review (2) comparing the posterolateral approach (PLA) and direct lateral approach (DLA) reported only 1 randomized study, which was prematurely ended(3).

The PLA is assumed to be beneficial regarding HRQoL and is thought to induce less walking problems postoperatively(4). However, these results may be based on a relatively healthy and cognitive fit group of patients, as another registry study did not report this superior HRQoL(5). Another Scandinavian observational study showed less need for walking aids with PLA 1 year after the hip fracture(6).

The presumably faster rehabilitation and better balance due to the scatheless gluteus musculature in patients treated using the PLA may be counterbalanced by the increased risk of dislocation(5-8). Recurrent dislocations can be devastating and cause a persistent declined HRQoL(9). On the contrary, the loss of abductor muscle strength after hemiarthroplasty through the DLA can lead to limping and reduced mobility (4, 6, 10). Hypothetically, this loss of abductor strength influences the balance, which might increase the risk of falling. Conclusive evidence on which of these 2 approaches results in better patient outcomes is lacking.

Therefore, we will assess the (cost-)effectiveness of the PLA compared to DLA in patients treated with a cemented hemiarthroplasty for a hip fracture. The primary outcome is health-related quality of life and secondary outcomes include physical performance, independency, complications, costs, and tendency to fall.

Table 1, inclusion criteria

Inclusion criteria	Exclusion criteria
≥18 years at the time of trauma	Multi trauma (defined as an Injury Severity Score > 15)
Acute fracture of the femoral neck (< 7 days old)	Secondary surgery after failed internal fixation
Cemented hemiarthroplasty as recommended treatment according to national guidelines	A known metastatic disease and a confirmed pathological fracture of the hip
Dutch or English fluency and literacy	High risk of non-compliance/adherence to study procedures*
Written informed consent (by proxy in patients with mental impairment)	

*** (e.g.** no Dutch residency – such as tourists- during follow-up period, or other factors that impair follow-up data collection such as patients who have a life expectancy of fewer than 6 months)

METHODS AND ANALYSIS

Overview of study design

We will perform a randomized controlled multicenter superiority trial in the Netherlands with a natural experiment (NE) and economic evaluation alongside. Randomization takes place in hospitals where orthopedic surgeons can perform both the DLA and PLA. The NE takes place in hospitals where surgeons solely perform the PLA or the DLA, according to their preference. In the NE, the topographical location where the trauma takes place determines the hospital where the patient is admitted to and thereby determines allocation to the PLA or DLA. This is assumed to resemble random assignment.

Surgeons expertise

Surgeons participating in the RCT have to be competent in both surgical approaches, they have to meet 1 of the following criteria of expertise for the PLA and DLA: 1) performed at least 20 hemiarthroplasties with the PLA and DLA in their career (including residency experience in which they were the primary surgeon during the procedure), 2) performed at least 5 hemiarthroplasties with the PLA and DLA in the last year. Residents may perform the procedure if the attending supervising surgeon meets the above criteria.

Patient selection

Eligibility criteria

In order to be eligible for this study, a subject must meet all of the criteria listed in Table 1. Cognitive impairment, such as dementia, is not an exclusion criteria. We will recruit incapacitated patients for study entry with involvement of their proxy in the informed consent procedure and data collection.

Patient recruitment and screening

We will screen all patients admitted to the hospital with a hip fracture for eligibility. Eligible patients admitted to the hospitals where both surgical techniques are performed are invited to participate in our RCT prior to the surgery. Patients in the NE are invited to participate prior to or at the latest 1 day after surgery. To obtain informed consent, we inform the patients verbally and the patient information letter will be handed out to eligible patients or to their health care proxy. The first patient was included on 6/2/2018 and the last patient on 7/1/2022.

Randomization and blinding

After obtaining informed consent at the emergency department or patient ward, patients will be randomly assigned in a 1:1 allocation ratio to either the PLA or the DLA. Randomization will be done in CASTOR EDC (www.castoredc.com), an online secured study and data management system with built-in randomization (variable block method,

stratified per center). Surgeons, patients or outcome assessors will not be blinded since the different surgical approaches are easily distinguishable (i.e. based on the location of the scar). Data analysts and the Steering Committee will remain blinded throughout the trial. We will first interpret the blinded results of the primary analysis before breaking the randomization code.

Study interventions

Posterolateral approach

In the PLA group, the skin incision starts posterior to the lateral side of the greater trochanter and runs slightly curved towards the femoral axis. Dissection of the insertions of the external rotators and piriformis follows and the surgeon performs a posterior capsulotomy. The gluteus medius and vastus lateralis are preserved. When closing the hip joint the capsule is sutured and the piriformis is reattached. Whether the piriformis was spared or reattached was left to the surgeon's preference.

Direct lateral approach

In the DLA group a longitudinal incision starts 3-5 cm proximal, crossing the greater trochanter and runs over the femoral axis. Release of the anterior insertion of the gluteus medius proximally and splitting fibers of the vastus lateralis distally follows. The surgeon performs an anterior capsulotomy. The stronger posterior capsule is preserved. When closing the hip joint the capsule is sutured as well as the fibers of the vastus lateralis and gluteus medius.

Peri- and postoperative care

All operations will be performed by experienced surgeons or residents under the direct supervision of an experienced surgeon. All implants are inserted with cement. The type and brand of the prosthesis are at the surgeons' discretion. Antibiotic and thromboembolic prophylaxis, suture materials and –techniques, and wound dressing are done according to the surgeons' judgment and local guidelines. Physical therapy and rehabilitation will be administered following the standard protocols and local guidelines. According to the Dutch guideline patients are advised to early weight baring as tolerated. There are no movement restrictions or mandatory use of ADL aids after a hemiarthroplasty. Patients will use the assistive devices when they need it. To improve generalizability to regular clinical practice, we designed a pragmatic trial without substantial restrictions on other clinical care processes which are known to vary between hospitals.

Study outcomes

Table 2 provides an overview of the outcomes at the different measurement moments.

Table 2. Overview of outcomes and measurement moments

Variable	Description	Baseline	4 weeks	3 months	4 months	6 months
Baseline charateristics	Including age, gender, BMI, comorbidities, living status, ASA, prescribed medication	RCT / NE				
Peri- postoperative outcomes	Including length of stay, surgery time, blood loss, discharge destination	RCT / NE				
EQ-5D-5L	Health-related quality of life states consisting of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) that are scored on 5 levels (no problems, slight problems, moderate problems, severe problems and extreme problems)	RCT / NE		RCT / NE		RCT / NE
Katz ADL	Activities of Daily Living (ADL) functionality as introduced by Katz, resulting in a score ranging from 0 (ADL independent) to 6, (ADL dependent)	RCT / NE		RCT / NE		RCT / NE
Mobility score	5 item mobility score(27), ranging from 0, indicating no walking aids to 5, indicating no functionality of lower extremity.	RCT / NE		RCT / NE		RCT / NE
Health-related and societal costs	Cost questionnaires to assess the use of healthcare resources and informal care as well as productivity-losses from unpaid and paid work (i.e. absenteeism and presenteeism). All resource use will be valued in accordance with the "Dutch Manual of Costing".(29)		RCT	RCT		RCT
Pain	Numerical Rating Scale (NRS) ranging from 0 (no pain) to 10 (worst imaginable pain) for mean and maximum pain over the week			RCT / NE		RCT / NE
Fear of falling	Falls Efficacy Scale International (FES-I), resulting in a score of 16, (no concern about falling) to 64 (severe concern about falling)			RCT / NE		RCT / NE

Table 2. Overview of outcomes and measurement moments (*continued*)

Variable	Description	Baseline	4 weeks	3 months	4 months	6 months
Tendency to fall	Number of falls, additional injuries as a result of falling, as reported by patients and/or retrieved from hospital charts			RCT / NE		RCT / NE
Complications	Re-interventions and (surgical) complications as reported by patients, and/or retrieved from hospital charts.			RCT / NE		RCT / NE
Mortality	As reported by patients' contact person and/or retrieved from hospital charts					RCT / NE
Short Physical Performance Battery test (SPPB)	SPPB is a group of measures that combines the results of the gait speed, chair stand and balance tests. Each segment has a maximum of 4 points and the total score has a maximum of 12 points, high scores suggest better physical performance.(28)				RCT*	

*a subgroup of 70 patients will perform the SPPB test

Primary study outcome

The primary outcome is the health related quality of life (EQ-5D-5L) as reported by the patient or proxy at 6 months after surgery(11). Patients' EQ-5D-5L health states will be converted to utility values using the Dutch tariff.(12)

Secondary study outcomes

Secondary outcomes are listed in Table 2.

Study follow-up

The primary outcome EQ-5D-5L and the secondary outcomes will be assessed through questionnaires online, by hardcopy or by phone at baseline, 3 and 6 months postoperatively. The cost questionnaires will be administered at 4 weeks, 3 and 6 months follow-up. One additional measurement moment (SPPB) at 4 months follow up is introduced for a subgroup who participate in the RCT. These patients will be asked to perform the SPPB test during a home visit by a researcher. We will check the patient's medical record 6 months postoperatively if they had any complications or readmissions during the study period.

Sample size

We calculated the sample size for superiority using a 2-sided significance level (α) of 0.05 and a power (ß) of 80%. With a standard deviation (SD) of 0.3 on the EQ-5D and a minimal clinically important difference (MCID) of 0.08 (retrieved from the study of Walters & Brazier(13)) a minimum of 222 subjects are needed in each treatment arm. Taking into account a 20% loss to follow-up after 6 months, a total number of 555 participants will be included in the RCT. A subgroup of 70 randomized patients will perform an additional physical test to assess balance and physical performance. During the inclusion period of the RCT, additional patients will be included in the NE in hospitals who are only comfortable with 1 of both surgical approaches.

Data analysis

Effectiveness analysis

To investigate the difference in the clinical effectiveness of both surgical approaches, we will use linear mixed-model analysis for continuous outcomes. Primary analyses will be based on the intention-to-treat principle. Additional 'as treated' analyses will provide insight into the influence of protocol deviations. Analyses will be done using the Statistical Package for the Social Sciences (SPSS, Version 27.0. Armonk, NY: IBM Corp). For all analyses, a 2-tailed value of $\alpha < 0.05$ is considered to be significant. In the absence of statistical significance, the potential relevance of any differences between groups will be discussed with respect to the study sample size and reported thresholds for clinical relevance. The primary database is the RCT database. For the NE data, a crude analysis will be performed similar to the RCT. Because selection bias may be present in the NE, we will assess balance on the covariates. We will consider the covariates balanced if the absolute standardized mean difference for each covariate is at most 0.1. If the covariates are unbalanced, different matching methods will be used until balance is obtained.(14, 15) The matched data will then be analyzed in a similar way to the RCT.

The goal of combining the randomized controlled trial and NE data is to improve the precision of the estimators for the primary and secondary outcome. Lu et al. developed

a method specifically for this type of study(16). They show that the estimators can be improved if there are no confounders besides the measured baseline covariates and they propose a test for this assumption. If the assumption is satisfied, the data will be combined according to their method. If the proposed test fails, a sensitivity analysis will provide insight in the effect of an unmeasured confounder on the study outcomes. The sensitivity method developed by Dorie et al. will be used, which models the response surface with a linear model and machine learning methods(17). Dorie et al. show empirical validation of the latter method, even in the presence of non-linearities in the true response surface(18).

Primary study parameter

In the primary linear mixed model, EQ-5D-5L utility scores will be analyzed as dependent variable. Treatment allocation (DLA vs. PLA) and baseline EQ5D score will be included as fixed factors. Repeated measures within-subjects and groups of patients within hospitals will be clustered using random effects. Differences between groups over time will be evaluated by incorporating the interaction term of group and time. The primary endpoint is 6 months after trauma.

Secondary parameters

Continuous secondary outcome variables (functionality on the KATZ, physical performance on the SPPB, tendency to fall, fear of falling on the FES-I, number of falls, and pain on the NRS) will be analyzed using similar linear mixed models. Dichotomous secondary outcome measures (additional injuries as a result of falling, re-interventions, discharge destination, and (surgical) complications will be analyzed using chi-square statistics and logistic regression models.

Heterogeneity of treatment effect / Exploratory analyses

Treatment effects can vary across the patients in both intervention groups. Besides the crude analysis of the primary and secondary outcomes, we will also adjust for potential confounders, by adding their baseline values as covariates in a multivariable mixed model (eg. age, gender, living status, dementia, ASA, BMI, Katz ADL, mobility) in an adjusted regression analysis.

Treatment effects can vary across the patients in both intervention groups. We will therefore assess effect modification, by exploring interactions between the treatment group and each of the potential confounders listed above.

Cost-effectiveness analysis

An economic evaluation will be performed for QALYs, from both the societal and healthcare perspective, and in accordance with the intention-to-treat principle. QALYs will be estimated by multiplying the patients' utility values by the duration with which they

experienced a certain health state.(19) Missing data will be imputed using multivariate imputation by chained equations(20). Cost and effect differences will be estimated using linear mixed models. It is very important to use such mixed model analyses and account for the possible clustering of cost and effect data (e.g. at the hospital level), as most economic evaluations fail to do so, whereas ignoring the possible clustering of data might lead to inaccurate levels of uncertainty and inaccurate point estimates(21). Incremental Cost Effectiveness Ratios (ICERs) will be calculated by dividing the difference in costs by that in effects. Bootstrapping techniques will be used to estimate the uncertainty surrounding the cost-effectiveness estimates. Uncertainty will be shown in cost-effectiveness planes and cost-effectiveness acceptability curves. Sensitivity analyses will be performed to test the robustness of the study results(22-24).

Registration, ethics, data sharing plan, funding, and potential conflicts of interests

This trial was registered at clinicaltrials.gov (NCT04438226) prior to the start of inclusion.

The study has been approved by the local and the Medical Ethics Committee (METC) (number NL63378.100.17) an will be conducted according to the principles of the Declaration of Helsinki, as amended in Seoul and Fortaleza (64th WMA General Assembly, October 2013)(25) and in accordance with the Medical Research Involving Human Subjects Act (WMO) and other guidelines, regulations and Acts. In all participating hospitals the study protocol will be submitted for review and approval by their local research ethics board. Any substantial amendments will be notified to the accredited Medical Ethical Committee. The investigator will report all SAEs related to the treatment to the sponsor and report the SAEs through the web portal *ToetsingOnline* to the accredited METC that approved the protocol, within 7 days of first knowledge for SAEs that result in death or are life threatening. All other SAEs will be reported within a period of maximum 15 days after the sponsor has first knowledge of the serious adverse events.

Data will be managed and archived for 15 years at the initiating hospital (OLVG). We intend to facilitate data sharing in line with the FAIR (Findability, Accessibility, Interoperability, and Reuse) principles, taking into account European laws and guidelines for privacy, and upon reasonable request. All included patients receive a trial code, which pseudonyms their personal data. The link between the trial code and the patient personal data is saved in a separate secured file with access only by the coordinating investigator (MCJMT) and research assistant (AR). The outcome data is only accessible for the coordinating investigator (MCJMT), principal investigator (RWP), research assistant (AR), supervisor (NWW) and authorized research personnel of the Joint Research team in OLVG Amsterdam. The handling of personal data will comply with the Dutch Personal Data Protection Act. The results from the study will be submitted for publication in peer-reviewed journals and presented at international conferences. This trial is supported by the Dutch Organisation

for Health Research and Development (ZonMw grant number: 8430041 12). There are no conflicts of interests of all authors.

Steering and data monitoring committee

The steering committee for this study consists of 2 independent orthopedic surgeons (RGHHN, DJFM) and 1 independent trauma surgeon (IBS). The interim analyses will be performed by the research assistant (AR) and blinded for treatment groups. The steering committee will evaluate the interim results to decide whether the study can be continued without compromising patient safety. Data monitoring is conducted by an independent study monitor of the initiating hospital (OLVG).

DISCUSSION

We designed a randomized controlled trial with a natural experiment and economic evaluation alongside, to compare the PLA and DLA for a hemiarthroplasty after a hip fracture. Currently, the choice of surgical approach is mostly based on surgical preference since there is a lack of evidence. This study will be the largest RCT worldwide addressing this subject and may improve the quality of life and health care for patients with hip fractures treated with a hemiarthroplasty. Furthermore, this RCT will be the first to conduct an economic evaluation and provide detailed insight into the healthcare and societal costs of both approaches. Our study outcomes, including quality of life, fear of falling, level of independence, physical performance, and complications are important for the patients.

We aimed to design a study with a good reflection of the elderly population. Therefore we will not exclude patients with dementia. Since the older adults with dementia are well represented in the population of patients with a hip fracture, including them will increase the generalizability of the study results. The primary outcome can be completed by proxy, which is a validated questionnaire in the elderly patient(26).

In the Netherlands the majority of the trauma surgeries are conducted by trauma surgeons, general surgeons specialized in trauma. Approximately 1/3 of all hip hemiarthroplasties in 2020 are placed by general trauma surgeons. Trauma surgeons are mostly competent in one surgical approach.

There are some limitations to our study. First we did not exclude patients with fractures or contusion of the lower / upper limbs with a ISS score of less than 16. Such additional injuries are likely to affect patient outcomes. Given the large sample size and the randomized design, we expect the incidence of additional injuries to be similar between the 2 groups and therefore will not affect the intervention effect. Another limitation is that we did not document whether the piriformis was spared or reattached, which was at the surgeon's

discretion. There is no high-level evidence of the effect of piriformis sparing approaches of hemiarthroplasty. Currently, the HemiSpaire study is comparing the direct lateral approach with a piriformis sparing posterior approach. This will give additional insights to the results of our study.

In addition to the traditional RCT design, we will conduct a NE in hospitals where only one of the surgical approaches is performed. Although we are aware this is not formal randomization on participant level, the NE design has several advantages: 1) prevent surgical expertise bias. 2) facilitate better generalizability of our trial results since more centers are able to participate. 3) help implement our trials results. 4) reduce selection bias, by including patients and hospitals who may not have agreed to randomization.

Currently there is a substantial practice variation due to the absence of high quality evidence reporting which approach is most valuable for the patient. With this study we aim to close the existing knowledge gap about which surgical approach is preferable for the patient outcome.

REFERENCES

1. Cooper C, Campion G, Melton LJ. Hip fractures in the elderly: A world-wide projection. Osteoporos Int. 1992;2:285-9.
2. Tol MCJM, van Beers LWAH, Willigenburg NW, Gosens T, Heetveld MJ, Willems HC, et al. Posterolateral or direct lateral approach for hemiarthroplasty after femoral neck fractures: a systematic review. Hip int. 2021;31(2):154-65.
3. Parker MJ. Lateral versus posterior approach for insertion of hemiarthroplasties for hip fractures: A randomised trial of 216 patients. Injury. 2015;46(6):1023-7.
4. Kristensen TB, Vinje T, Havelin LI, Engesaeter LB, Gjertsen JE. Posterior approach compared to direct lateral approach resulted in better patient-reported outcome after hemiarthroplasty for femoral neck fracture. Acta Orthop. 2017;88(1):29-34.
5. Leonardsson O, Rolfson O, Rogmark C. The surgical approach for hemiarthroplasty does not influence patient-reported outcome. The Bone & Joint Journal. 2016;98:542-7.
6. Hongisto MT, Nuotio MS, Luukkaala T, Vaisto O, Pihlajamaki HK. Lateral and posterior approaches in hemiarthroplasty. Scand J Surg. 2018:1457496917748226.
7. Rogmark C, Fenstad AM, Leonardsson O, Engesaeter LB, Karrholm J, Furnes O, et al. Posterior approach and uncemented stems increases the risk of reoperation after hemiarthroplasties in elderly hip fracture patients. Acta Orthop. 2014;85(1):18-25.
8. Unwin AJ, Thomas M. Dislocation after hemiarthroplasty of the hip: a comparison of the dislocation rate after posterior and lateral approaches to the hip. Ann R Coll Surg Engl. 1994;76:327-9.
9. Enocson A, Petterson H, Ponzer S, Törnkvist H, Dalén N, Tidermark J. Quality of life after dislocation of hip arthroplasty: a prospective cohort study on 319 patients with femoral neck fractures with a one-year follow-up. Qual Life Res. 2009;18(9):1177-84.
10. Sayed-Noor A, Hanas A, Skoldenberg O, Mukka S. Abductor Muscle Function and Trochanteric Tenderness After Hemiarthroplasty for Femoral Neck Fracture. J Orthop Trauma. 2016;30(6).
11. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med. 2001;33(5):337-43.
12. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit AG, Prenger R, Stolk EA. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health. 2016;19(4):343-52.
13. Walter SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Quality of Life Research. 2005;14(6):1523-32.
14. Stuart EA. Matching methods for causal inference: A review and a look forward. Statistical science : a review journal of the Institute of Mathematical Statistic. 2010;25(1):1-21.
15. Gu XS, Rosenbaum PR. Comparison of Multivariate Matching Methods: Structures, Distances, and Algorithms. Journal of Computational and Graphical Statistics, . 1993;2(4):405-20.
16. Lu Y, Scharfstein DO, Brooks MM, Quach K, Kennedy EH. Causal Inference for Comprehensive Cohort Studies. arXiv: Methodology. 2019.
17. Dorie V, Harada M, Carnegie NB, Hill J. A flexible, interpretable framework for assessing sensitivity to unmeasured confounding. Statist Med. 2016(35):3453– 70.
18. Dorie VJ, Hill JL, Shalit U, Scott M, Cervone D. Automated versus Do-It-Yourself Methods for Causal Inference: Lessons Learned from a Data Analysis Competition. . Statistical Science. 2019.
19. Drummond, Michael F, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford university press. 2015.
20. White I, Royston P, Wood A. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011;30:377-99.

21. Gomes. Multiple imputation using chained equations: Issues and guidance for practice. Heal Econ. 2012;21(9):1108-18.
22. Black W. The CE plane: a graphic representation of cost-effectiveness. Med Decis Mak. 1990;10:212-4.
23. Fenwick E. Cost-effectiveness acceptability curves. Heal Econ. 2004;13:405-15.
24. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the Economic Evaluation of Health Care Programmes 3rd Edition. Oxford Univ Press. 2005.
25. 64th WMA General Assembly. Declaration of Helsinki. Fortaleza, Brazil2013.
26. Toh HJ, Yap P, Wee SL, Koh G, Luo N. Feasibility and validity of EQ-5D-5L proxy by nurses in measuring health-related quality of life of nursing home residents. Qual Life Res. 2021;30(3):713-20.
27. Bowers T, Parker M. Assessment of outcome after hip fracture: development of a universal assessment system for hip fractures. SICOT J. 2016;2(27).
28. Guralnik J, Simonsick E, Ferrucci L, Glynn R, Berkman L, Blazer D. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol. 1994;49(2):85-94.
29. Kanters TA, Bouwmans CAM, van der Linden N, Swan Tan S, Hakkaart-van Roijen L. Update of the Dutch manual for costing studies in health care. PLoS One. 2017;12(11).

7

Posterolateral or Direct Lateral Surgical Approach for Hemiarthroplasty after a Hip Fracture A Randomized Controlled trial alongside a Natural Experiment

Maria C.J.M. Tol[1], M.D., Nienke W. Willigenburg[1], Ph.D., Ariena J. Rasker[1], M.Sc., Hanna C. Willems[2], M.D., Ph.D., Taco Gosens[3,4], M.D., Ph.D., Martin J. Heetveld M.D.[5], Ph.D., Martijn G.M. Schotanus[6,7], Ing., Ph.D., Bart Eggen[8], M.Sc., Mate Kormos[8], M.Sc., Stéphanie L. van der Pas[9,10], Ph.D., Aad W. van der Vaart[8], Ph.D., J.Carel Goslings[11], M.D., Ph.D., and Rudolf W. Poolman[1,12], M.D., Ph.D., on behalf of the APOLLO research group (Supplementary Appendix 1)

[1]Department of Orthopedic Surgery, Joint Research, OLVG, Amsterdam, [2]Department of Internal Medicine and Geriatrics, Amsterdam UMC, Amsterdam, [3]Department of Orthopedics and Trauma Surgery, ETZ, Tilburg, [4]Department of Medical and Clinical Psychology, Tilburg University, Tilburg, [5]Department of Trauma Surgery, Spaarne Gasthuis, Haarlem, [6]Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen, [7]School of Care and Public Health Research Institute, Faculty of Health, Medicine and Life Science, Maastricht University, [8]Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, Mekelweg 4, Delft, [9]Amsterdam UMC location Vrije Universiteit Amsterdam, Epidemiology and Data Science, De Boelelaan 1117, Amsterdam, [10]Amsterdam Public Health, Methodology, Amsterdam, [11]Department of Trauma Surgery, OLVG, Amsterdam, [12]Department of Orthopedic Surgery, LUMC, Leiden

JAMA Network Open (2024)

KEY POINTS

Question: Is there a difference in patient outcomes between the posterolateral approach (PLA) and direct lateral approach (DLA) for cemented hemiarthroplasty after acute femoral neck fracture?

Findings: Quality of life six months after trauma did not differ between surgical approaches in 843 patients who participated in a Randomized Clinical Trial (RCT, n=555) or Natural Experiment (NE, n=288). PLA resulted in significantly more dislocations and reoperations than DLA.

Meaning: This first well-powered randomized trial showed no difference in patient-reported quality of life between PLA and DLA for cemented hemiarthroplasty, despite higher rates of dislocation and reoperation after PLA. Randomized (RCT) and pseudo-randomized (NE) data yield similar outcomes, which strengthens these findings.

ABSTRACT

Importance: Hip fractures in older adults are serious injuries that result in disability, higher rates of illness and death, and a significant strain on healthcare resources. High-quality evidence to improve hip fracture care regarding the surgical approach of hemiarthroplasty is lacking.

Objective: To compare six-month outcomes of the posterolateral approach (PLA) and direct lateral approach (DLA) for hemiarthroplasty in patients with acute femoral neck fracture.

Design, Setting and Participants: We performed a multicenter, randomized clinical trial (RCT) comparing DLA and PLA, with a natural experiment (NE) alongside in 14 centers in the Netherlands. In total 843 patients aged 18 years or older were included, with or without dementia and an acute femoral neck fracture. Secondary surgery of the hip, pathological fractures or multi trauma patients were excluded. Recruitment took place between February 2018 and January 2022. Treatment allocation was random (n=555) or pseudorandom (n=288) based on geographical location and surgeon preference.

Main outcome: The primary outcome was health-related quality of life six months after surgery, quantified with the EuroQol Group 5-Dimension questionnaire (EQ-5D-5L). Secondary outcomes included dislocations, (fear of) falling, activity of daily living, pain, and reoperations. To improve generalizability, we performed a novel technique for data fusion of the RCT and NE.

Results: A total of 843 patients participated; 555 in the RCT (283 DLA, 272 PLA) and 288 in the NE (172 DLA and 116 PLA). In the RCT, mean EQ-5D-5L utility scores at six months were 0.50 (95%CI 0.45–0.55) after DLA and 0.49 (95%CI 0.44–0.54) after PLA, with 77% completeness. The between-group difference was not statistically nor clinically significant (-0.04, 95%CI -0.11–0.04). Most secondary outcomes were comparable between groups, but PLA resulted in more dislocations (5.5% in RCT, 5.2% in NE) than DLA (0.4% in RCT, 0.6% in NE). Data fusion resulted in an effect size of 0.00 (95%CI -0.04–0.05) for the EQ-5D-5L and an odds ratio of 12.31 (95%CI 2.77–54.7) for suffering a dislocation after PLA.

Conclusion and Relevance: Among patients treated with a cemented hemiarthroplasty after an acute femoral neck fracture, PLA does not result in a better quality of life than DLA. Rates of dislocation and reoperation are higher after PLA. Randomized (RCT) and pseudo-randomized (NE) data yield similar outcomes, strengthening these findings.

Trial registration number: NCT04438226 (clinicaltrails.gov)

INTRODUCTION

Hip fractures in older adults are disabling injuries that increase morbidity and mortality(1), and cause excessive utilization of healthcare resources(2). High-quality evidence to improve hip fracture care is emerging. The HEALTH trial showed no clinically meaningful differences between hemiarthroplasty and total hip arthroplasty in the treatment of hip fractures in older adults(3). The WHiTE trial showed a modestly but significantly better health-related quality of life (HRQoL) and a lower risk of periprosthetic fractures in favor of cemented compared with uncemented hemiarthroplasty(4). Such evidence is lacking for the choice of surgical approach.

Surgical approaches for hemiarthroplasty vary widely around the globe. The choice of approach is usually determined by the surgeon's preference or by the agreements made within a hospital. The posterolateral approach (PLA) and the direct lateral approach (DLA) are currently most used(5, 6). A systematic review on outcomes most relevant for patients suggests that the PLA might have advantages compared to DLA in HRQoL, abductor insufficiency, and walking problems(7). A meta-analysis concluded that the PLA had no advantages that counterbalanced its increased risk of dislocation and reoperation compared to DLA and the direct anterior approach(8). However, both systematic reviews were based primarily on observational studies and concluded that high-quality clinical trials were needed to confirm or refute their conclusions.

To compare the HRQoL and other relevant patient outcomes between DLA and PLA in adult patients with a displaced femoral neck fracture, we conducted the 'surgical Approaches of cemented hemiarthroplasty after hip fractures; POsteroLateral versus direct Lateral apprOach' (APOLLO) trial. We hypothesized better HRQoL in patients treated with the PLA.

METHODS

Trial design

The APOLLO trial was a multicenter, randomized superiority trial in the Netherlands with a natural experiment (NE) and economic evaluation alongside. Details on the trial objectives, design, procedures, and statistical analysis plan can be found in the published protocol(9) and at clinicaltrials.gov (NCT04438226).

Trial oversight

Fourteen centers participated, of which nine only in the NE (Supplementary Appendix 2). Randomization at the individual patient level occurred when orthopedic surgeons could perform the DLA and the PLA.

The NE was conducted in specific hospitals where surgeons were specialized in either the PLA or DLA techniques, and the group of surgeons opted to one of these approaches as their standard of care. Given their specialization, these surgeons did not have the flexibility to randomize between the two methods. A patient's proximity to a hospital at the time of the incident (and thus where they were brought to) dictated the surgical approach they would receive, as hospitals were implicitly designated as either 'DLA' or 'PLA' based on the expertise of their surgeons. This geographical-based allocation was outside of research parameters and control. We postulated that this setup mirrored a pseudorandom allocation mechanism, and we rigorously verified this presumption prior to integrating our data.

The trial was funded by the Dutch Organization for Health Research and Development (ZonMw grant number: 8430041-12). We received approval from the Medical research Ethics Committees United (MEC-U, NL63378.100.17) and the local institutional review boards of all participating centers. The steering committee consisted of three independent (orthopedic) trauma surgeons, who evaluated the interim analysis.

Participants

Inclusion criteria were: adult patients (≥18 years) with an acute femoral neck fracture (<7 days), cemented hemiarthroplasty as recommended treatment according to the national guidelines, Dutch or English fluency, and literacy, and written informed consent. Exclusion criteria were: multi-trauma (ISS >15), secondary surgery after failed internal fixation, pathological fracture, and high risk of non-compliance (i.e., no Dutch residency, such as tourists, or patients with a life expectancy <six months). Cognitive impairment, such as dementia, was not an exclusion criterion.

Intervention, Randomization and Blinding

Experienced surgeons or residents under direct supervision of an experienced surgeon performed all operations. Surgical details (i.e., whether the m. piriformis was spared or reattached with the PLA, and how the m. gluteus medius was closed with the DLA) were left to the surgeon's discretion.

CASTOR EDC (Amsterdam, the Netherlands), an online secured data management system with built-in randomization (variable block method, stratified per center), randomly assigned patients in a 1:1 allocation ratio to either PLA or DLA. Surgeons, patients, and outcome assessors were aware of the assignment group since the different surgical approaches were easily distinguishable (i.e., based on the scar's location). Data analysts and the Steering Committee were blinded. We interpreted the blinded results before breaking the randomization code.

Outcomes

The primary outcome was the health-related quality of life (EQ-5D-5L) as reported by the patient or proxy at six months. Proxies of incapacitated patients were asked to rate how they thought the patient would rate their health-related quality of life(10). The patients' EQ-5D-5L health states were converted to utility values (ranging from -0.446 to 1, with higher scores indicating higher health utility) using the Dutch tariff(11). Deceased patients obtained a score of 0(4, 12).

Secondary outcomes are listed in Supplementary Table S1. We assessed EQ-VAS (higher scores indicating a better health status), activities of daily living functionality (Katz ADL) ranging from 0 (ADL independent) to 6 (ADL dependent), and a 5-item mobility score ranging from 0 (no walking aids) to 5 (no functionality of lower extremity). We also considered dichotomous Katz (ADL independent: 0-1 point; ADL dependent: 2 to 6 points) and mobility scores (good: with or without 1 crutch; impaired: more walking aids). Patients also scored their mean and maximum pain during the week using the numeric rating scale (NRS-pain) ranging from 0-10, with higher scores indicating worse pain. We assessed the fear of falling with the Falls Efficacy Scale International (FES-I), ranging from 16 (no concern about falling) to 64 (severe concern about falling). We also recorded the actual fall incidents and additional injuries resulting from falling.

Assessments

We obtained the outcomes through questionnaires online, by hardcopy, or by phone at three and six months. In addition, we checked all patients' medical records up to six months postoperatively to record relevant baseline and surgical characteristics, as well as any complications, re-admissions, or reoperations during the study period.

Sample size

To detect a minimally important difference (MCID) of 0.08 in the primary outcome (EQ-5D-5L utility score)(13) with a standard deviation of 0.3, we needed 555 patients based on a two-sided significance level (α) of 0.05 with 80% power and a loss to follow-up of 20% after six months(9).

Statistical analysis

We used linear mixed-model analysis to investigate the difference in the primary outcome (EQ-5D-5L utility score) between both surgical approaches. The primary analysis was based on the intention-to-treat principle, with additional as treated analyses to quantify the effect of protocol deviations. For the crude analyses, the fixed factors were treatment allocation (DLA vs. PLA) and the EQ-5D-5L utility score at baseline. We evaluated differences between groups over time by adding time and a time by treatment interaction. Repeated measures within subjects and groups of patients within hospitals were clustered using random intercepts. For the adjusted analyses, we added the potential confounders

age, sex, living status, dementia, ASA-classification, BMI, mobility (good versus impaired), and Katz ADL (dependent versus independent) to the model as fixed factors.

The continuous secondary outcomes, fear of falling on the FES-I, functionality on the Katz, and pain on the NRS were analyzed using similar linear mixed models. Categorical and dichotomous secondary outcomes (i.e., mobility, discharge destination, complications, ≥1 fall incident(s), additional injuries as a result of falling, or reoperations) were compared using a chi-square or Fisher's Exact test. We used Statistical Package for the Social Sciences (SPSS, Version 27.0. Armonk, NY) and a two-tailed value of $\alpha < 0.05$ was considered statistically significant for all analyses.

Data fusion

Besides unconfoundedness, which was assumed, for data fusion it is necessary that there are no significant difference between the patients in the NE and the RCT in any of the treatment groups when correcting for the confounders. We tested this condition on one separately imputed combined dataset similarly to Lu et al.(14). We fitted a linear model for the EQ-5D-5L index at 6 months as a function of an experiment indicator and the confounding variables, separately in both groups of surgical approaches. The experiment indicator was 0 for RCT and 1 for NE. If the estimated coefficient belonging to the experiment indicator was close to zero with a p-value larger than 0.05, indicating no statistically significant effect of the experiment indicator on the outcome while keeping all other confounders constant.

We used the augmented inverse probability weighting estimator (AIPW) to determine the average treatment effect between the two surgical approaches. We corrected for possible confounders and performed a sensitivity analysis of our chosen methods. Multiple imputation was used on all confounders (15). Estimates of the average treatment effect were pooled according to Rubin's rule to form a single estimate and a confidence interval. For more information, see the published protocol and Supplemental Methods.

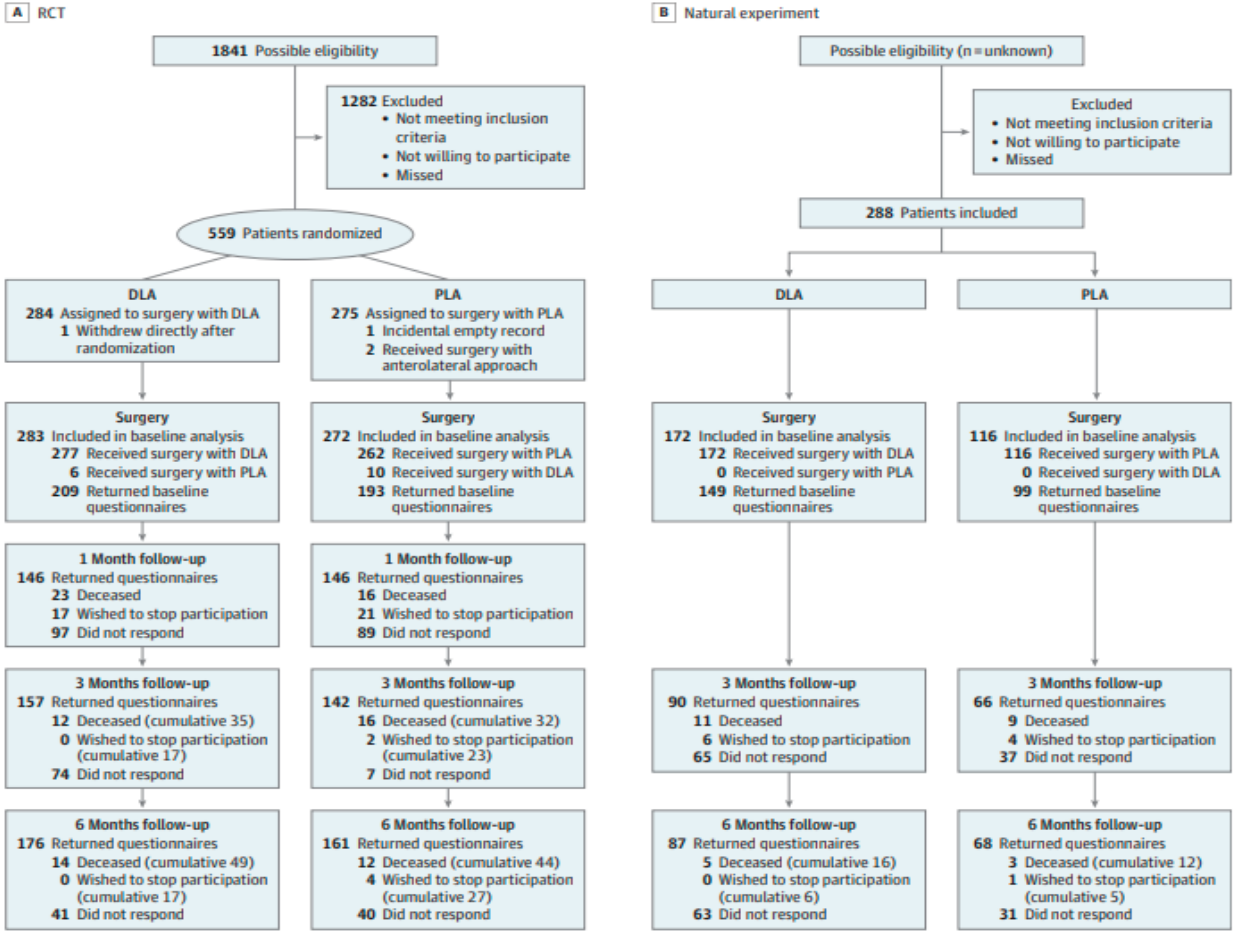


Figure 1.

DLA indicates direct lateral approach; PLA, posterolateral approach; and RCT, randomized clinical trial.

RESULTS

Between February 2018 and January 2022, 843 patients consented to participate. Five hundred fifty-five patients participated in the RCT (283 DLA and 272 PLA) and 288 in the NE (172 DLA and 116 PLA). The final six months follow-up was completed in July 2022, with 77% data completeness for the primary outcome in the RCT. Figure 1 depicts the patient flow and reasons for exclusion. The percentages of missing data are provided in Supplementary Fig. S1. Non-responders had a similar age as responders, but were less mobile. Patients with dementia were overrepresented in the DLA non-responder group (Supplementary Table S2).

Baseline characteristics (Table 1) were comparable between treatment groups (DLA vs. PLA) and between study designs (RCT vs. NE), but patients with arrhythmia were overrepresented in the NE-DLA group. Most (62% RCT, 69% NE) patients were female, the mean age was 83, and the majority (78% RCT, 82% NE) lived independently (with or without ADL help) before the hip fracture. Signs of dementia were present in 137 patients (25%) in the RCT and 49 (17%) in the NE.

Adherence to the assigned intervention

Sixteen protocol deviations occurred: six patients assigned to DLA (2.1%) had surgery with PLA and 10 patients assigned to PLA (3.7%) had surgery with DLA (Fig. 1).

RCT Primary endpoint

Mean EQ-5D-5L utility scores at six months follow-up were 0.50 (95%CI 0.45-0.55) in the DLA group and 0.49 (95%CI 0.44–0.54) in the PLA group (Fig. 2). Intention to treat analyses showed no statistically significant between-group difference at 6 months: crude -0.05 (95%CI, -0.14–0.04), adjusted: -0.04 (95%CI -0.11–0.04); the negative sign means in favor of PLA (Supplementary Table S3). As-treated analyses yielded similar results: crude -0.06 (95%CI -0.15-0.03) and adjusted -0.04 (95%CI -0.12-0.03) between-group difference.

Table 1. Baseline characteristics

	RCT DLA (n=283)	RCT PLA (n=272)	NE DLA (n=172)	NE PLA (n=116)	P value of NE
Sex					
Female, n (%)	172 (61)	172 (63)	116 (67)	82 (71)	0.42
Age, years (SD)	82 (8)	82 (8)	82 (7)	84 (6)	0.08
Dementia, n (%)					
Evident	56 (20)	44 (16)	20 (12)	11 (9)	0.60
Possibly	18 (6)	19 (7)	9 (5)	9 (8)	
ASA, n (%)					
I	7 (2)	4 (1)	6 (3)	6 (5)	0.78
II	101 (36)	80 (29)	64 (37)	41 (35)	
III	152 (54)	160 (59)	85 (49)	56 (48)	
IV	10 (4)	11 (4)	15 (9)	7 (6)	
BMI, mean (SD)	24.2 (4)	24.6 (4)	24.3 (4)	24.9 (4)	0.64
Katz, mean (SD)	1.4 (1.7)	1.3 (1.7)	0.9 (1.4)	0.9 (1.6)	0.69
Prescribed medicines, n (SD)	7 (5)	7 (4)	7 (4)	6 (4)	0.67
Comorbidities, n (%)					
Cardiac	29 (10)	30 (11)	22 (13)	8 (7)	0.11
MCI	20 (7)	20 (7)	25 (15)	9 (8)	0.08
Heart failure	49 (17)	56 (21)	58 (34)	21 (18)	**<0.01**
Arrhythmia	14 (5)	8 (3)	8 (5)	5 (4)	1.0
Neurological	26 (9)	25 (9)	18 (10)	6 (5)	0.11
Hemiparalysis	20 (7)	15 (6)	3 (2)	3 (3)	0.69
CVA	3 (1)	2 (1)	1 (1)	1 (1)	1.0
Parkinson	39 (14)	46 (17)	20 (12)	9 (8)	0.29
Epilepsy	10 (4)	10 (4)	6 (3)	4 (3)	0.99
Pulmonal					
COPD					
Asthma					
Pre-fracture mobility, n (%)					
Without aids	94 (33)	99 (36)	70 (41)	43 (37)	0.35
1 crutch	26 (9)	27 (10)	18 (10)	18 (16)	
Walker,	106 (37)	92 (34)	53 (31)	40 (34)	
Outside with help	29 (10)	20 (7)	18 (10)	8 (7)	
No mobility	2 (1)	5 (2)	3 (2)	0	
Living status, n (%)					
Independent	153 (54)	149 (55)	115 (67)	80 (69)	0.40
Independent with help	61 (22)	69 (25)	36 (21)	15 (13)	
Residential care	29 (10)	29 (11)	9 (5)	5 (4)	
Nursing home	32 (11)	20 (7)	7 (4)	9 (8)	
Rehabilitation unit	3 (1)	3 (1)	1 (1)	2 (2)	
Other	0	0	3 (2)	2 (2)	

Abbreviations: MCI = Myocardial infarct, CVA = cerebral vascular accident, BMI: Body Mass Index, COPD: Chronic Obstructive Pulmonary Disease.

RCT secondary endpoints

At six months follow-up, the EQ-VAS was 68 points (95%CI 65.12–71.01) after DLA and 66 (95%CI 62.78–68.38) after PLA (Fig. 2). The QALY did not differ between the groups with 0.255 in PLA and 0.243 in DLA (mean difference 0.012, CI95% mean difference 0.012, CI95% -0.482, 0.025). Figure 3 and Figure S2 summarises the other secondary outcomes. There were no statistically significant differences in ADL Katz score between the groups, nor in the dichotomized Katz scores (65.4% ADL independent patients in the DLA group and 55.8% in the PLA group, p=0.11). Mean and maximum pain scores were similar after DLA and PLA. At six months follow-up, 21.5% of the DLA patients had good mobility compared with 13.8% of the PLA patients (p=0.11). The frequency of falls did not differ between the surgical approaches; 51 patients (18%) reported at least 1 fall incident in the DLA group versus 44 (16%) in the PLA group (p=0.77). There were no differences in the rate of emergency department admissions or additional injuries due to the falls. The fear of falling measured with the FES-I questionnaire was not significantly different between the groups: 37 points with DLA (95%CI 34.21–39.14) and 34 with PLA (95%CI 31.48–36.18) (Fig. 3). We observed similar outcomes in both treatment groups regarding the length of stay, surgery time, postoperative complications, and discharge destinations (Supplementary Table S4).

The rate of prosthesis dislocation was significantly different between groups, both statistically and clinically (Supplementary Table S5). Dislocation occurred in 15 of the 272 patients (5.5%) in the PLA group compared with 0.4% (1 out of 283 patients) in the DLA group, resulting in an odds ratio of 16.46 (95%CI 2.16 – 125.48, p<0.01). Six (2.2%) patients in the PLA group suffered recurrent dislocations, and 8 (2.9%) underwent revision surgery. The overall number of reoperations was 35 in the PLA group, and 18 in the DLA group, see Supplementary Table S6.

During the six month follow-up, death occurred in 93 of the 555 patients (17%). Mortality was similar in both groups; 17% after DLA and 16% after PLA (p=0.69). Between-group differences at three and six months follow-up are detailed in Supplementary Table S3.

Figure 2, EQ-5D-5L and EQ-VAS outcomes

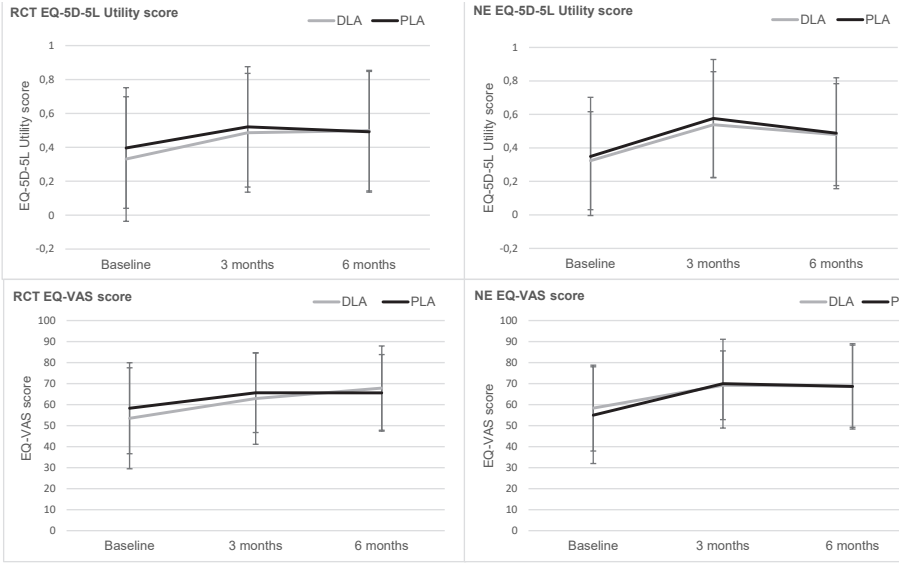


Figure 3. Secondary outcomes at 6-months.

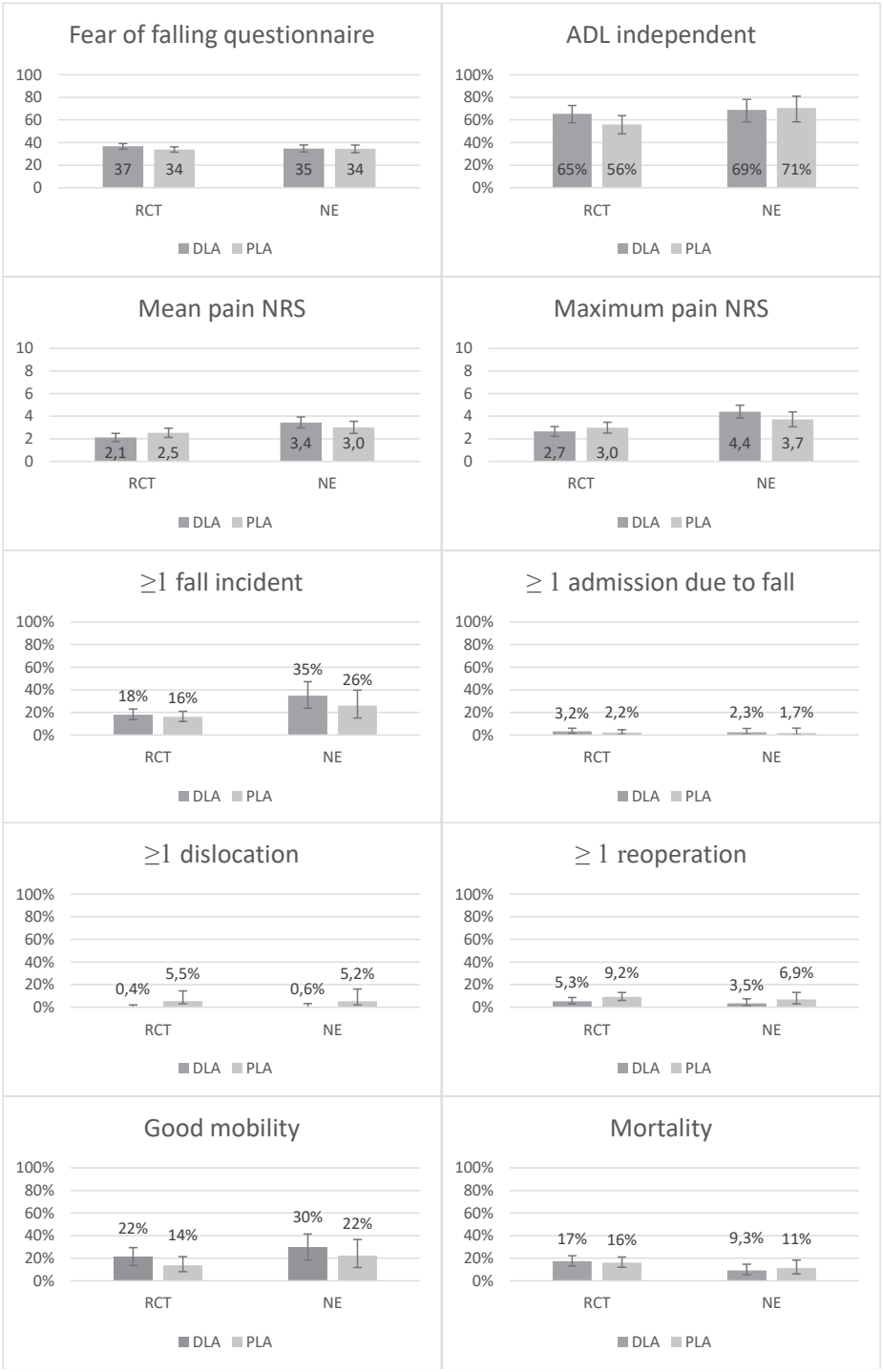
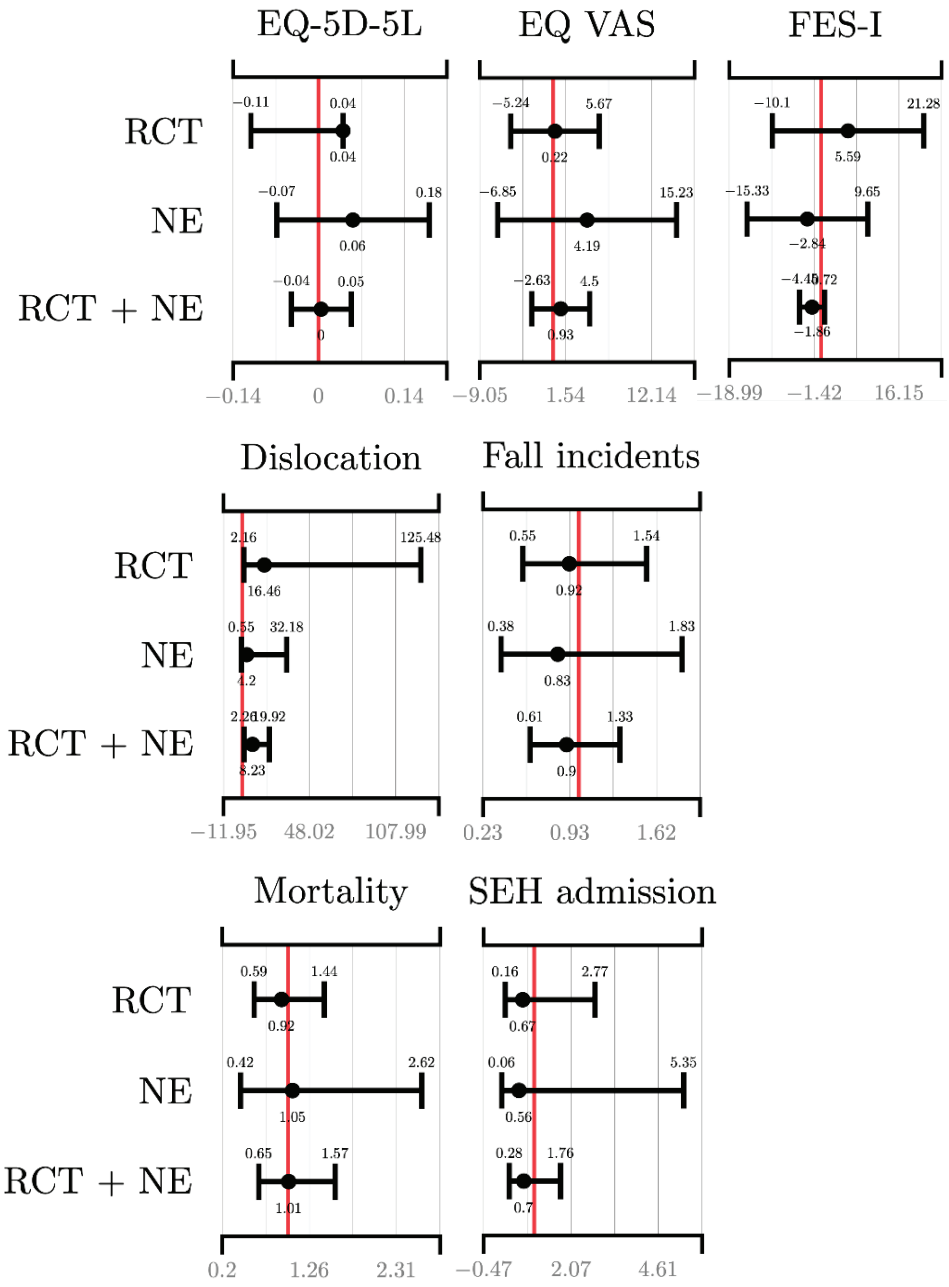


Figure 4, Fused results of primary and secondary outcomes

*effect sizes: EQ-5D-5L, EQ VAS, FES-I, Odds Ratio: Dislocation, Fall incidents, Mortality, SEH admission



NE primary and secondary endpoints

The EQ-5D-5L utility score at six months did not differ between the treatment groups: 0.53 (95%CI 0.47–0.60) for DLA and 0.57 (95%CI, 0.50–0.64) for PLA (group effect 0.06, 95%CI -0.07–0.18) (Fig. 2, Supplementary Table S3). We found a higher risk of dislocation in the PLA group (5.3%) compared with DLA (1.1%) (Table S4). Figure 3 presents all secondary outcomes of the RCT and NE.

Data fusion outcomes

For the primary outcome, there was no significant difference between the groups in the different experiments when all other confounders were included in the model, so no evidence was found against the assumptions for data fusion (see Supplemental Methods and Results). Fusion tests were rejected for the Katz ADL score, mobility, pain mean and pain max variables. All estimates, including 95% confidence intervals, are presented in Figure 4.

*effect sizes: EQ-5D-5L, EQ VAS, FES-I, Odds Ratio: Dislocation, Fall incidents, Mortality, SEH admission

DISCUSSION

The surgical approach for a cemented hemiarthroplasty after a femoral neck fracture did not influence the health-related quality of life in adult patients with an acute hip fracture six months after surgery. The observed mean difference in EQ-5D-5L utility score between DLA and PLA was neither statistically significant nor clinically relevant. In addition, the secondary outcomes in function, pain, mobility, and tendency to fall did not differ between the groups. However, patients in the PLA group had a higher risk of prosthesis dislocation and reoperation due to dislocation, which was statistically significant and is also clinically relevant.

Our study has several strengths. First, this is the largest RCT worldwide addressing this topic. Second, the NE that we designed alongside the RCT showed strong potential as a feasible alternative to randomization at the individual patient level. Treatment allocation in the NE was purely based on the geographical location (and thereby the hospital) and the timing of the fracture (and subsequently the surgeon who would perform the hemiarthroplasty). Our study provides important insights in the similarities between RCT and NE study populations and outcomes, which is highly relevant given the known difficulties with surgical RCTs(16). The addition of a NE not only raised the number of participants but also increased the generalizability of our trial results. Since we were not restricted to surgeons with expertise in both approaches, more hospitals could participate. Another strength of our study is the inclusion of patients with cognitive impairments such as dementia. Older adults with dementia are well represented in the population of patients

with hip fractures; however, they are often excluded from RCTs(17). Including them further raised the generalizability of the results.

There were some limitations to our study. First, missing data were prevalent in the RCT and NE datasets. This was mainly caused by non-response to the follow-up questionnaires, which was associated with dementia, living status, and mobility. However, it is essential to include these frail patients since it is a substantial part of the hip fracture population. Sensitivity analyses with multiple imputation analyses for confounders found no differences between the approaches. There are no missing data/loss to follow up for the outcomes dislocation, reoperation, re-admissions and mortality. For these variables, we reviewed all medical reports 6 months postoperatively, and used registry data for possible adverse events in other hospitals. Second, a screening log was not maintained, so we do not formally know why not all eligible patients were randomized. We know that some patients declined participation, and that not all patients received information about the study due to logistical and time constraints. Our overview of potentially eligible patients based on the Dutch arthroplasty registry(6) shows that their characteristics were similar to the randomized group. Third, we did not gather information on whether the m. piriformis was spared or reattached. Currently, there is no high-level evidence for the effect of piriformis-sparing approaches of hemiarthroplasty. While we acknowledge the potential variability in surgeon expertise and specialty, our study's outcomes remain consistent across the board. Future studies might benefit from a deeper dive into the influence of surgeon subspecialties on hemiarthroplasty outcomes.

No RCT to date evaluating the surgical approach for hemiarthroplasty reported the EQ-5D-5L. A large Norwegian observational study showed a significant difference of 0.03 in favor of PLA(18). We observed a non-significant difference of 0.04 in favor of the PLA, but both differences are under the threshold of the MCID of 0.08(13). Two other prospective observational studies did not report any differences in the EQ-5D-5L related to the surgical approach after adjusting for confounders(19, 20).

Prosthesis dislocation was more common in the PLA group, with 5.5% vs. 0.4% in the DLA group, which was similar in the NE. Thereby, the risk of reoperation due to dislocation was statistically significantly higher in the PLA group (2.9% vs. 0). These findings support the existing evidence of observational studies(21-24). While most clinical trials do not include patients with dementia, we did include this fragile population, which may contribute to a relatively high dislocation rate compared to the existing literature. The higher risk of dislocation rate was seen in both study designs. The potential bias of performance bias in the RCT was not seen in this trial. This increased risk of dislocation did not result in a substantially lower quality of life in the PLA group, which raises the question whether the EQ-5D-5L is sufficiently sensitive to quantify the effectiveness of orthopaedic interventions. On the other hand, the more specific patient-reported secondary outcomes, includ-

ing mobility, (fear of) falling, and Katz ADL, were also not significantly different between surgical approaches.

In conclusion, PLA does not result in a better quality of life than DLA in adult patients treated with a cemented hemiarthroplasty after an acute femoral neck fracture. Most secondary outcomes are similar between groups, but PLA is associated with more dislocation and reoperation compared to DLA. Pseudo-randomization in our natural experiment resulted in similar outcomes and could be a valid and more feasible alternative to the traditional RCT.

The implications of our results to improve patient care are not straightforward. The increased risk for dislocation and reoperation after PLA without clear benefits of that approach could justify a recommendation for DLA. Alongside this trial we conducted an economic evaluation and we thoroughly quantified physical performance and balance in a subgroup of patients. These additional outcomes will help to weigh a broader spectrum of costs and benefits associated with DLA and PLA, to better inform evidence based decisions on the surgical approach for older patients with hemiarthroplasty after a hip fracture.

Contributers

Writing Committee: MCJM Tol (Department of Orthopaedic Surgery, Joint Research, OLVG, Amsterdam), NW Willigenburg (Department of Orthopaedic Surgery, Joint Research, OLVG, Amsterdam), A Rasker (Department of Orthopaedic Surgery, Joint Research, OLVG, Amsterdam) , HC Willems (Department of Internal Medicine and Geriatrics, Amsterdam UMC, Amsterdam), T Gosens (Department of Orthopaedics and Trauma Surgery, ETZ, Tilburg; Department of Medical and Clinical Psychology, Tilburg University, Tilburg), M Heetveld (Department of Trauma Surgery, Spaarne Gasthuis, Haarlem), B Eggen (Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, Delft), M Kormos (Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, Delft) SL van der Pas (Amsterdam UMC location Vrije Universiteit Amsterdam, Epidemiology and Data Science, Amsterdam; Amsterdam Public Health, Methodology, Amsterdam), AW van der Vaart (Delft University of Technology, Electrical Engineering, Mathematics and Computer Science, Delft), MGM Schotanus (Department of Orthopaedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen), JC Goslings (Department of Trauma Surgery, OLVG, Amsterdam) and RW Poolman (Department of Orthopaedic Surgery, Joint Research, OLVG, Amsterdam; Department of Orthopaedic Surgery, LUMC, Leiden)

Data acquisition: AR, MT, TG, RP, MH, MS, F van Roon (Department of Orthopaedics and Trauma Surgery, ETZ, Tilburg), M van Dijk (Department of Orthopaedic Surgery, Antonius Ziekenhuis, Utrecht), J Keizer (Department of Trauma Surgery, Antonius Ziekenhuis,

Utrecht), AJH Vochteloo (Department of Orthopaedic Surgery, OCON, Hengelo), P Joosse (Department of Trauma Surgery, Noord-West Ziekenhuis, Alkmaar), B Boonen (Department of Orthopaedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen), J Jelsma (Department of Orthopaedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen), D Theeuwen (Department of Orthopaedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen), JJW Ploegmakers (Department of Orthopaedic Surgery, UMCG, Groningen), T Schepers (Department of Trauma Surgery, Amsterdam UMC, Amsterdam), E van der Meij (Department of Orthopaedic Surgery, IJsselland, Capelle aan den Ijssel), SH van Helden (Department of Orthopaedic & Trauma Surgery, Isala, Zwolle), R Zuurmond (Department of Orthopaedic & Trauma Surgery, Isala, Zwolle), BA van Dijkman (Department of Trauma Surgery, Flevoziekenhuis, Almere), TD Berendes (Department of Orthopaedic Surgery, Meander MC, Amersfoort), HE Hendrinks (Departement of Orthopaedic Surgery & Trauma, Máxima MC, Eindhoven).

Statistical support: NW, BE, MK, SP, AvdV, RP

Steering Committee: RGHH Nelissen (Department of Orthopaedic Surgery, LUMC, Leiden), DJF Moojen ((Department of Orthopaedic Surgery, OLVG, Amsterdam), IB Schipper (Department of Trauma Surgery, LUMC, Leiden)

Study-centre coordination, management of the database, and, data analyses was organised by the Joint Research group at OLVG Amsterdam. The study design was developed by MT, NW, HW, TG, MH, RP. The study protocol and statistical analysis plan was created by the Writing Committee. MT, NW, AR, RP were responsible for data completeness, and accuracy. Data analyses are done MT, AR, NW, BE, MK, SP, AvdV, RP. The first author, MT, wrote the first draft of the manuscript; the writing committee made revisions. All authors had full access to the data and read and, approved the final manuscript.

Acknowledgements

This study was funded by grants from the Dutch Organization for Health Research and Development (ZonMw grant number: 8430041-12 & 10330112010006).

Declaration of interests: RP reports grants from the Dutch Organization for Health Research and Development (ZonMw grant number: 8430041-12 & 10330112010006) for conducting this study, grants from foundation of medical research OLVG, Amsterdam, the Netherlands, van Rens Foundation, ReumaNederland, personal fees from NOV-Dutch Orthopaedic Society, Waldemar LINK and LIMA corporate outside the submitted work

MT, NW, AR, MS, TG, reports grants from the Dutch Organization for Health Research and Development during the conduct of this study. TG was Member of the Board of the

Dutch Orthopaedic Society till 2022. AV reports a grant from the Dutch Research Council (Spinoza prize) and is president elect of the International Society of Bayesian Analysis, outside the submitted work. MK, BG reports a grant from the Dutch Research Council (Spinoza prize). SP reports grants from the Dutch Research Council (Veni 192.087, OCENW. M20.190) and personal fees from Dutch Arthroplasty Register, outside the submitted work. HW reports speakersfee for lectures on fractures in elderly persons from AMGEN, UCB and is Director of the Foundation for Health care evaluation and chair of the working group Leading the Change, outside the submitted word. MH, JG declare no competing interests.

Data sharing

Data will be available upon reasonably request and after approval of the writing committee. Requests can be sent to JointResearch@olvg.nl. The study protocol and statistical analysis plan are available in the published protocol(11).

REFERENCES

1. Johnell O, Kanis JA. An estimate of the worldwide prevalence, mortality and disability associated with hip fracture. Osteoporos Int. 2004;15(11):897-902.
2. Bhandari M, Swiontkowski M. Management of Acute Hip Fracture. New England Journal of Medicine. 2017;377:2053-62.
3. The Health investigators. Total Hip Arthroplasty or Hemiarthroplasty for Hip Fracture. NEJM. 2019;381(23).
4. Fernandez MA, Achten J, Parsons N et al. Cemented or Uncemented Hemiarthroplasty for Intracapsular Hip Fracture. New England Journal of Medicine. 2022;386(6).
5. Chechik O, Khashan M, Lador R, Salai M, Amar E. Surgical approach and prosthesis fixation in hip arthroplasty world wide. Arch Orthop Trauma Surg. 2013;133(11):1595-600.
6. LROI. Annual report of Dutch Arthroplasty Register. 2021.
7. Tol MCJM, van Beers LWAH, Willigenburg NW et al. Posterolateral or direct lateral approach for hemiarthroplasty after femoral neck fractures: a systematic review. Hip int. 2021;31(2):154-65.
8. van der Sijp MPL, van Delft D, Krijnen P, Niggebrugge AHP, Schipper IB. Surgical approaches and hemiarthroplasty outcomes for femoral neck fractures: a meta-analysis. The journal of Arthroplasty. 2018; 33(5):1617-1627
9. Tol MCJM, Willigenburg NW, Willems HC et al. Posterolateral or direct lateral approach for cemented hemiarthroplasty after femoral neck fracture (APOLLO): protocol for a multicenter randomized controlled trial with economic evaluation and natural experiment alongside. Acta Orthop. 2022;12(93):732-8.
10. Tol MCJM, Kuipers JP, Willigenburg NW, Willems HC, Poolman RW. How are you doing in the eyes of your spouse? Level of agreement between the self-completed EQ-5D-5L and two proxy perspectives in an orthopaedic population: a randomized agreement study. Health Qual Life Outcomes. 2021;19(1):35.
11. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit AG, Prenger R, Stolk EA. Dutch Tariff for the Five-Level Version of EQ-5D. Value Health. 2016;19(4):343-52.
12. Euroqol. [Available from: https://euroqol.org/.
13. Walter SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Quality of Life Research. 2005;14(6):1523-32.
14. Lu Y, Scharfstein DO, Brooks MM, Quach K, Kennedy EH. Causal Inference for Comprehensive Cohort Studies. arXiv: Methodology. 2019.
15. Rubin DB. Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons, Inc, . 1987;June.
16. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. BMJ. 2002;324(7351):1448-51.
17. Hebert-Davies J, Laflamme G-Y, Rouleau D, investigators HaF. Bias towards dementia: Are hip fracture trials excluding too many patients? A systematic review. injury. 2012;43(12):1978-84.
18. Kristensen TB, Vinje T, Havelin LI, Engesaeter LB, Gjertsen JE. Posterior approach compared to direct lateral approach resulted in better patient-reported outcome after hemiarthroplasty for femoral neck fracture. Acta Orthop. 2017;88(1):29-34.
19. Leonardsson O, Rolfson O, Rogmark C. The surgical approach for hemiarthroplasty does not influence patient-reported outcome. The Bone & Joint Journal. 2016;98:542-7.
20. Sayed-Noor A, Hanas A, Skoldenberg O, Mukka S. Abductor Muscle Function and Trochanteric Tenderness After Hemiarthroplasty for Femoral Neck Fracture. J Orthop Trauma. 2016;30(6).

21. Rogmark C, Fenstad AM, Leonardsson O et al. Posterior approach and uncemented stems increases the risk of reoperation after hemiarthroplasties in elderly hip fracture patients. Acta Orthop. 2014;85(1):18-25.
22. Leonardsson OR, C. Rogmark. The surgical approach for hemiarthroplasty does not influence patient-reported outcome. The Bone & Joint Journal. 2016;98:542-7.
23. Biber R, Brem M, Singler K, Moellers M, Sieber C, Bail HJ. Dorsal versus transgluteal approach for hip hemiarthroplasty: an analysis of early complications in seven hundred and four consecutive cases. Int Orthop. 2012;36(11):2219-23.
24. Hongisto MT, Nuotio MS, Luukkaala T, Vaisto O, Pihlajamaki HK. Lateral and posterior approaches in hemiarthroplasty. Scand J Surg. 2018:1457496917748226.

SUPPLEMENTARY APPENDIX

Posterolateral or Direct Lateral Surgical Approach

for Hemiarthroplasty after a Hip Fracture

Investigators of the APOLLO research group (all located in the Netherlands):

- Frank van Roon, M.D. Department of Orthopedics and Trauma Surgery, ETZ, Tilburg
- Martijn van Dijk, M.D., Ph.D. Department of Orthopedic Surgery, Antonius Ziekenhuis, Utrecht
- Jort Keizer, M.D. Department of Trauma Surgery, Antonius Ziekenhuis, Utrecht
- Anne J.H. Vochteloo, M.D., Ph.D. Department of Orthopedic Surgery, OCON, Hengelo
- Pieter Joosse, M.D., Ph.D. Department of Trauma Surgery, Noord-West Ziekenhuis, Alkmaar
- Bert Boonen, M.D., Ph.D., Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen
- Jetse Jelsma, M.D., Ph.D., Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen
- Dieuwertje Theeuwen, M.D. Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen
- Joris J.W. Ploegmakers, M.D., Ph.D. Department of Orthopedic Surgery, UMCG, Groningen
- Tim Schepers, M.D. Ph.D. Department of Trauma Surgery, Amsterdam UMC, Amsterdam
- Evelien van der Meij, M.D. Department of Orthopedic Surgery, IJsselland, Capelle aan den IJssel
- Svenhjalmar H. van Helden, M.D., Ph.D. Department of Orthopedic & Trauma Surgery, Isala, Zwolle
- Rutger Zuurmond, M.D., Ph.D. Department of Orthopedic & Trauma Surgery, Isala, Zwolle
- Bart A. van Dijkman, M.D. Department of Trauma Surgery, Flevoziekenhuis, Almere
- Thomas D. Berendes, M.D., Ph.D. Department of Orthopedic Surgery, Meander MC, Amersfoort
- Hans G.E. Hendriks, M.D., Ph.D. Department of Orthopedic Surgery & Trauma, Máxima MC, Eindhoven

Participating hospitals in the RCT and NE.

- Department of Orthopedic and Trauma Surgery, ETZ, Tilburg
- Department of Orthopedic and Trauma Surgery , Zuyderland Medical Center, Heerlen, Sittard-Geleen
- Department of Orthopedic and Trauma Surgery, OLVG, Amsterdam
- Department of Orthopedic Surgery, OCON, Hengelo
- Department of Orthopedic Surgery, IJsselland, Capelle aan de Ijssel

Participating hospitals in the NE only:

- Department of Trauma Surgery, Spaarne Gasthuis, Haarlem
- Department of Orthopedic and Trauma Surgery, Antonius Ziekenhuis, Utrecht
- Department of Trauma Surgery, Noord-West Ziekenhuis, Alkmaar
- Department of Orthopedic Surgery, UMCG, Groningen
- Department of Trauma Surgery, Amsterdam UMC, Amsterdam
- Department of Orthopedic and Trauma Surgery, Isala, Zwolle
- Department of Trauma Surgery, Flevoziekenhuis, Almere
- Department of Orthopedic Surgery, Meander MC, Amersfoort
- Department of Orthopedic and Trauma Surgery, Máxima MC, Eindhoven

SUPPLEMENTAL METHODS: DATA FUSION

Missing data

Missing data were prevalent in both the RCT and the NE data sets, concerning mainly the primary outcome and the covariate BMI. For example, in the NE data set, 35% of patients' primary outcome and 11% of patients' BMI were missing. See Fig. S1 for detailed missingness patterns. Missingness was negligible in the other covariates.

The outcomes were mainly thought to be missing by non-response to the post trial questionnaire, which was attributed to dementia, living status and mobility. Indeed, a logistic regression of the missingness indicator of the primary outcome on the covariates resulted in a statistically significant positive coefficient on mobility (p-value 0.002) in the RCT and a statistically significant positive coefficient on dementia (p-value 0.004) in the NE. Therefore, the outcomes were considered to be missing at random given the confounders, thus indicating that no bias will be introduced to the average treatment effect when missing outcomes are ignored in the analysis.

All missing data in the confounders were imputed with Classification And Regression Trees[1] using the R package MICE[2]. The outcome and treatment variables were not used as predictors for imputing the missing confounder values. The statistical analysis was then performed only on the subjects with an observed outcome.

Estimator

Augmented Inverse Probability Weighting (AIPW) was deployed to estimate the average treatment effect. AIPW requires model specifications for the (i) propensity score, i.e. the conditional probability of receiving PLA given the covariates age, gender, BMI, living status, baseline ASA scores, mobility and dementia; and (ii) conditional mean of the outcome variable given the same covariates. Flexible modeling strategies were chosen: the propensity score and outcome regression were modelled using an ensemble of methods, including elastic net regression[3], generalized linear models[4], generalized additive models[5], generalized boosted models[67] and multivariate adaptive regression splines[8]. These were implemented in the R package SuperLearner[9]. All this was performed using 10-fold cross-validation.

Data fusion

Besides unconfoundedness, which was assumed, another condition is necessary for data fusion. Namely, there should be no significant difference between the patients in the NE and the RCT in any of the treatment groups when correcting for the confounders. This condition was tested on one separately imputed combined dataset similarly to Lu et. Al. (2019)[10]. Specifically, a linear model was fitted for the EQ-5D-5L index at 6 months as a function of an experiment indicator and the confounding variables, separately in both groups of surgical approaches. The experiment indicator was equal to zero for patients in the RCT and one for patients in the NE. For both surgical approaches, the estimated coefficient belonging to the experiment indicator was close to zero with a p-value larger than 0.05, indicating no statistically significant effect of the experiment indicator on the outcome while keeping all other confounders constant. The test suggested that there was no evidence against data fusion, as there was no significant difference between the groups in the different experiments when all other confounders were included in the model. Accordingly, the data sets were combined, and the average treatment effect was estimated with AIPW as indicated in the protocol paper.[11]

Secondary outcomes

For the secondary outcomes, the estimates should only be viewed as exploratory because of multiple testing errors and because the outcomes had small variance (for instance, only 8 patients had a fall incident out of 281 in the NE). For AIPW, only an ensemble of generalized linear models, generalized additive models and multivariate adaptive regression splines were used to estimate the propensity score and outcome regression to limit runtime of the code. When the outcome was dichotomous, tests against data fusion were performed using logistic regression. All fusion tests indicated no evidence against data fusion, except Katz score, mobility, pain mean and pain max variables. For all these variables, the fused data sets were still analyzed, but only as a sensitivity analysis to check robustness of the study conclusions.

Sensitivity analysis

For predicting missing values by multiple imputation, CART was used as a predictive model for all confounders[12]. The findings were robust to the choice of predictive models (predictive mean matching and classification trees for binary, and linear regression and regression trees for continuous variables). The conclusions remained the same when the outcome was also imputed. These robustness findings held irrespective of whether BMI, the covariate with the highest missing percentage (11% in the NE data set), was included in the analysis. Using an ensemble of methods to model the components of the AIPW

estimator rendered the results robust to modelling choices. Fitting individual, instead of ensemble, models led to no changes in the conclusions.

Software

Analysis was performed in R. Versions of R and the packages can be found below

SUPPLEMENTAL RESULTS: DATA FUSION TEST FOR THE PRIMARY OUTCOME EQ-5D-5L UTILITY SCORE

Package	Version
R	4.1.3
foreign	0.8.83
Rcpp	1.0.9
tidyverse	1.3.2
encryptr	0.1.3
withr	2.5.0
mice	3.14.0
AIPW	0.6.3.2
SuperLearner	2.0.28
gbm	2.1.8.1
gam	1.20.2
earth	5.3.1
glmnet	4.1.4
foreach	1.5.2
doParallel	1.0.17
doSNOW	1.0.20

```
[1] "========================================================================"
[1] "Direct Lateral Approach"
[1] "========================================================================"
                   term      estimate   std.error  statistic       df      p.value
1           (Intercept)  0.6604398920 0.218359719  3.0245500 290.7823 2.712350e-03
2       NERE_indicatorRE  0.0103088600 0.035277330  0.2922234 309.4809 7.703117e-01
3    BSL_Mobility_dicho1 -0.2296335543 0.037806397 -6.0739338 229.0922 5.146488e-09
4 BSL_LivingStatus_dicho1 -0.0955672207 0.050561892 -1.8901037 289.3977 5.974322e-02
5            ASA_dicho1 -0.1084115302 0.036087107 -3.0041624 276.6259 2.907060e-03
6        Dementia_dicho1 -0.2237989547 0.048229038 -4.6403363 305.4754 5.168386e-06
7            BSL_Gender1 -0.0043565757 0.033843039 -0.1287289 308.1923 8.976562e-01
8                BSL_Age  0.0010336238 0.002236909  0.4620767 311.3296 6.443490e-01
9                BSL_BMI  0.0009187454 0.004311558  0.2130889 218.1529 8.314566e-01
[1] "========================================================================"
[1] "Postero Lateral Approach"
[1] "========================================================================"
                   term     estimate   std.error  statistic       df      p.value
1           (Intercept)  1.255263275 0.274458984  4.5735915 251.4863 7.530240e-06
2       NERE_indicatorRE -0.062543620 0.042030732 -1.4880450 273.6287 1.378905e-01
3    BSL_Mobility_dicho1 -0.142878167 0.044295464 -3.2255711 234.0806 1.436706e-03
4 BSL_LivingStatus_dicho1 -0.119554908 0.059403907 -2.0125765 269.5273 4.515373e-02
5            ASA_dicho1 -0.063187017 0.043275243 -1.4601193 239.5928 1.455673e-01
6        Dementia_dicho1 -0.173657071 0.054822653 -3.1676152 269.7027 1.713662e-03
7            BSL_Gender1 -0.018725917 0.040444395 -0.4630040 270.7629 6.437334e-01
8                BSL_Age -0.005677388 0.002641269 -2.1494922 271.7434 3.247835e-02
9                BSL_BMI -0.001608018 0.005202790 -0.3090683 198.3985 7.575939e-01
```

Figure S1. Missing values for relevant patient characteristics and outcomes.

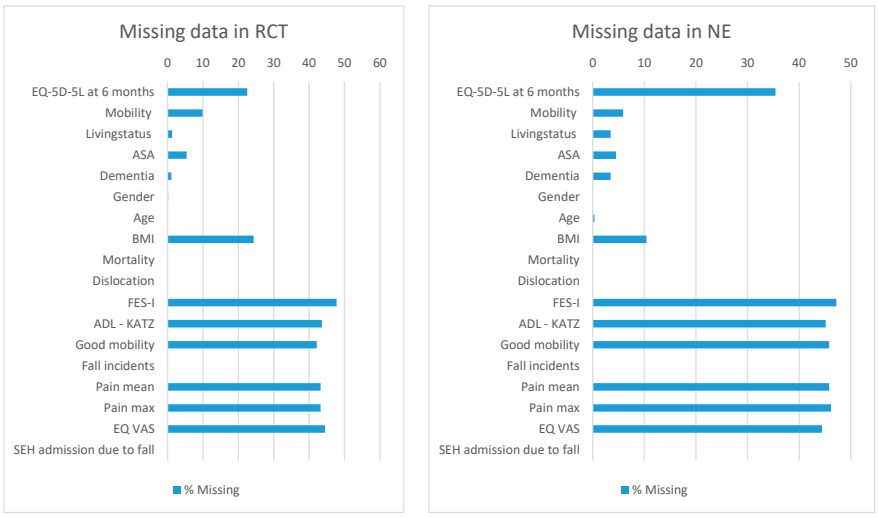


Figure S1. Missing values for relevant patient characteristics and outcomes.

Table S1. Overview of outcomes and measurement moments

Variable	Description	Baseline	4 weeks	3 months	4 months	6 months
Baseline characteristics	Including age, gender, BMI, comorbidities, living status, ASA, prescribed medication	RCT / NE				
Peri- postoperative outcomes	Including length of stay, surgery time, blood loss, discharge destination	RCT / NE				
EQ-5D-5L(1)	Health-related quality of life states consisting of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) that are scored on 5 levels (no problems, slight problems, moderate problems, severe problems and extreme problems)	RCT / NE		RCT / NE		RCT / NE
Katz ADL(2)	Activities of Daily Living (ADL) functionality as introduced by Katz, resulting in a score ranging from 0 (ADL independent) to 6, (ADL dependent)	RCT / NE		RCT / NE		RCT / NE
Mobility score(3)	5 item mobility score(4), ranging from 0, indicating no walking aids to 5, indicating no functionality of lower extremity.	RCT / NE		RCT / NE		RCT / NE
Health-related and societal costs	Cost questionnaires to assess the use of healthcare resources and informal care as well as productivity-losses from unpaid and paid work (i.e. absenteeism and presenteeism). All resource use will be valued in accordance with the "Dutch Manual of Costing".(5)		RCT	RCT		RCT
Pain(6)	Numerical Rating Scale (NRS) ranging from 0 (no pain) to 10 (worst imaginable pain) for mean and maximum pain over the week			RCT / NE		RCT / NE
Fear of falling(7)	Falls Efficacy Scale International (FES-I), resulting in a score of 16, (no concern about falling) to 64 (severe concern about falling)			RCT / NE		RCT / NE

Table S1. Overview of outcomes and measurement moments (*continued*)

Variable	Description	Baseline	4 weeks	3 months	4 months	6 months
Tendency to fall	Number of falls, additional injuries as a result of falling, as reported by patients and/or retrieved from hospital charts			RCT / NE		RCT / NE
Complications	Re-interventions and (surgical) complications as reported by patients, and/or retrieved from hospital charts.			RCT / NE		RCT / NE
Mortality	As reported by patients' contact person and/or retrieved from hospital charts					RCT / NE
Short Physical Performance Battery test (SPPB)(8)	SPPB is a group of measures that combines the results of the gait speed, chair stand and balance tests. Each segment has a maximum of 4 points and the total score has a maximum of 12 points, high scores suggest better physical performance.(28)				RCT*	

*a subgroup of 70 patients will perform the SPPB test

Table S2. Baseline characteristics of lost to follow up patients for primary outcome at end-point

	DLA n = 58	PLA n = 67
Sex, n (%)		
Female	36 (62)	46 (68)
Age, years	82 (8)	82 (7)
Dementia		
Evident + possibly, n/N (%)	23/56 (41)	17/66 (26)
ASA, n (%)	n = 57	n = 64
I	1 (2)	2 (3)
II	24 (42)	24 (38)
III	30 (53)	38 (59)
IV	2 (4)	0
BMI, mean (SD)	23.5 (4)	24.7 (4)
Katz, mean (SD)	1.9 (1.9)	1.1 (1.2)
Comorbidities, n (%)		
Cardial,	5 (9)	10 (15)
MCI	3 (5)	6 (9)
DC	5 (9)	8 (12)
Arrhythmia	1 (2)	3 (4)
Neurological	7 (12)	10 (15)
Hemiparalyse	5 (9)	4 (6)
CVA	2 (3)	0
Parkinson	7 (12)	15 (22)
Epilepsy	4 (7)	4 (6)
Pulmonal		
COPD		
Astma		
Pre-fracture mobility, n (%)	n = 54	n = 57
Without aids	15 (28)	17 (30)
1 crutch	6 (11)	7 (12)
Walker	28 (52)	28 (49)
Outside with help	5 (9)	4 (7)
No mobility	0	1 (2)
Living status, n (%)	n = 57	n = 66
Independent	24 (42)	31 (47)
Independent with help	20 (35)	21 (32)
Residential care	5 (9)	8 (12)
Nursing home	8 (14)	4 (6)
Rehabilitation unit	0	2 (3)

Table S3. Detailed between-group effects at 3 and 6 months for primary and secondary outcomes. Adjusted effects are based on mixed models for the RCT groups and on AIPW models for the NE group and the fused results.

3 months	DLA		PLA		Crude mixed model		Adjusted model			
	n	Mean	n	Mean	Group effect	95%CI	Group effect	95%CI		
EQ-5D-5L										
RCT	192	0.49	172	0.52	-0.08	-0.17 – 0.01	-0.07	-0.14 – 0.01		
NE	101	0.54	75	0.58	0.05	-0.04 – 0.15				
EQ VAS										
RCT	162	62.9	146	65.6	-1.69	-7.52 – 4.13	-1.72	-7.24 – 3.80		
NE	90	66.6	67	69.3	-3.10	-9.28 – 3.07				
FES-I										
RCT	146	38.3	128	33.7	5.08	0.90 – 9.27	4.34	1.15 – 2.83		
NE	82	35.9	65	34.5	-0.26	-4.99 – 4.48				
Pain mean										
RCT	155	2.50	138	2.43	0.14	-0.61 – 0.88	0.23	-0.51 – 0.96		
NE	90	3.42	67	2.87	0.25	-0.53 – 1.03	0.11	-0.79 – 0.55		
Pain max	155	3.06	139	3.20	0.03	-0.82 – 0.89				
RCT	90	4.18	67	3.63	0.23	-0.76 – 1.22				
NE										
ADL – Katz										
RCT	155	1.72	142	1.67	0.23	-0.15 – 0.60	0.18	-0.16 – 0.53		
NE	91	1.36	67	1.61	-0.70	-0.69 – -0.25				
6 months	DLA		PLA		Crude mixed model		Adjusted model		Fused data	
	n	Mean	n	Mean	Group effect	95%CI	Group effect	95%CI	Group effect	95%CI
EQ-5D-5L									0.00	-0.04 – 0.05
RCT	225	0.50	205	0.49	-0.05	-0.14 – 0.04	-0.04	-0.11 – 0.04		
NE	103	0.53	80	0.57	0.02	-0.08 – 0.11	0.06	-0.07 – 0.18		
EQ VAS									0.93	-2.63 – 4.50
RCT	179	67.9	165	65.6	0.30	-5.46 – 6.06	0.22	-5.24 – 5.67		
NE	89	66.6	69	67.6	1.62	-3.05 – 6.29	4.19	-6.85 – 15.23		

Table S3. Detailed between-group effects at 3 and 6 months for primary and secondary outcomes. Adjusted effects are based on mixed models for the RCT groups and on AIPW models for the NE group and the fused results. (*continued*)

3 months	DLA		PLA		Crude mixed model		Adjusted model			
FES-I									-1.86	-4.45 – 0.72
RCT	152	36.8	138	34.0	5.59	-10.10 – 21.28	8.41	-6.65 – 23.48		
NE	86	34.5	66	34.7	-1.89	-6.64 – 2.85	-2.84	-15.33 – 9.65		
Pain mean										
RCT	163	2.14	153	2.53	-0.39	-1.13 – 0.35	-0.30	-1.03 – 0.44		
NE	89	3.42	67	3.04	0.49	-0.17 – 1.16	-0.89	-1.99 - 0.20		
Pain max	165	2.69	151	2.98	0.03	-0.81 – 0.87	0.12	-0.72 – 0.96		
RCT	88	4.36	67	3.76	0.72	-0.11 – 1.55	-1.16	-2.54 – 0.22		
NE										
ADL – Katz										
RCT	160	1.38	154	1.81	-0.02	-0.39 – 0.35	-0.08	-0.42 – 0.26		
NE	91	1.18	67	1.34	-0.44	-0.90 – 0.18	-0.21	-1.32 – 0.90		

Figure S2. Secondary outcomes at 6-months.

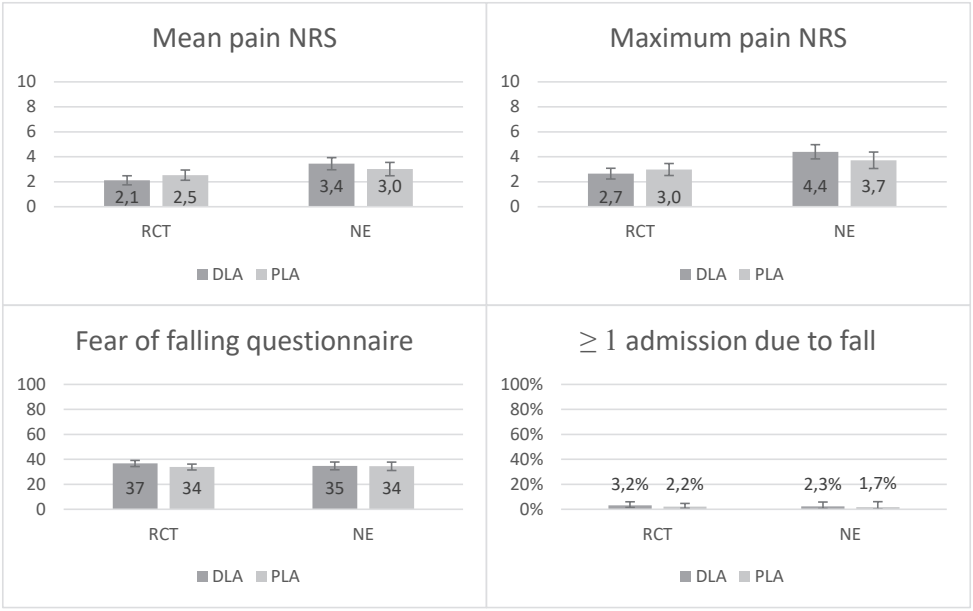


Table S4. Peri- and post-operative outcomes of the RCT

	DLA (n = 283)	PLA (n = 272)	P value
Primary surgeon, n (%)			
Consultant	147 (52)	133 (49)	0.415
Resident	129 (46)	129 (47)	
Missing	7 (2)	10 (4)	
Time to surgery			
Days, mean (SD)	0.95 (1)	1.04 (1)	0.31
Surgery time, min, mean (SD)	57 (16)	59 (14)	0.29
Blood loss, cc, mean (SD)	259 (168)	253 (127)	0.72
Postoperative complication, n (%)			
Pulmonary embolism	4 (1)	2 (1)	0.69
Neuropraxie	2 (1)	1 (0)	1
Anemia with transfusion	18 (6)	20 (7)	0.74
Pneumonia	15 (5)	18 (7)	0.59
DAIR due to infection	3 (1)	4 (1)	0.72
DAIR due to hematoma	1 (0)	0	1
Superficial infection	1 (0)	0	1
Dislocation	0	2 (1)	0.24
Urinary tract infection	12 (4)	7 (3)	0.28
Bladder retention	9 (3)	2 (1)	0.06
Delirium	32 (11)	36 (13)	0.52
In hospital mortality, n (%)	13 (5)	9 (3)	0.52

Table S4. Peri- and post-operative outcomes of the RCT (*continued*)

	DLA (n = 283)	PLA (n = 272)	P value
Fall incident during admission, n (%)	7 (2)	6 (2)	0.85
Additional injury, n (%)	1 (0)	1 (0)	1
Length of hospital stay, days (SD)	8.3 (6)	8.1 (6)	0.76
Discharge destination, n (%)			
Independent	12 (4)	9 (3)	0.95
Independent with help	59 (21)	62 (23)	
Retirement home	21 (7)	18 (7)	
Nursing home	35 (12)	37 (14)	
Rehabilitation home	142 (50)	134 (49)	
Other	13 (5)	10 (4)	
Missing	1 (0)	2 (1)	
Mobility at discharge, n (%)			
Without aids	1 (0)	2 (1)	0.05
1 crutch	2 (1)	0	
Walker	73 (26)	97 (36)	
Outside with help	177 (63)	149 (55)	
No mobility	19 (7)	17 (6)	
Missing	11 (4)	4 (1)	

Table S5, Dislocation rates

Dislocation	**DLA**	**PLA**	**p-value**
Dislocation	1/284 (0.4)	15/275 (5.5)	<0.01
RCT, number/n (%)	2/175 (1.1)	6/113 (5.3)	0.06
NE, number/n (%)	3/459 (0.7)	21/388 (5.4)	<0.01
Total, number/n (%)			
Recurrent dislocation			
RCT, number/n (%)	0	6/272 (2.2)	0.01
NE, number/n (%)	1/175 (0.6)	3/113 (2.7)	0.31
Total, number/n (%)	1/459 (0.2)	9/388 (2.3)	<0.01
Revision due to dislocation			
RCT, number/n (%)	0	8/272 (2.9)	<0.01
NE, number/n (%)	1/114 (0.9)	3/174 (1.7)	0.30
Total, number/n (%)	1/459 (0.2)	11/388 (2.8)	<0.01

Table S6, Reoperations by indication

Indication for reoperation	RCT DLA (n=283)	RCT PLA (n=272)	NE DLA (n=172)	NE PLA (n=116)
Dislocation (revision to THA)	0	8	0	3
Infection	10	14	5	2
Any type of fracture	5	7	0	0
Other	3	6	2	4
Total number of reoperations (1-3 per patient)	18	35	7	9

Abbreviations: THA = total hip arthroplasty.

Table S7, Subgroup analyses

	Dementia	
	Yes	**No**
EQ-5D-5L		
DLA	0.27 (0.18-0.37)	0.56 (0.50-0.61)
PLA	0.27 (0.19-0.35)	0.57 (0.51-0.62)
	Mixed model results	
Between groups	**Effect (95% CI)**	**p-value**
Surgical approach (DLA vs. PLA)	-0.02 (-0.231, 0.181) *	0.81
(Suspected) dementia (yes/no)	0.29 (0.119, 0.465) ^{T}	<0.01
Interaction approach X dementia	-0.01 (-0.236, 0.211)	0.91

***in favor** of PLA

^{T} **in** favor of no dementia

Table S7, Subgroup analyses (*continued*)

	Experience	
	Surgeon	**Resident**
EQ-5D-5L		
DLA	0.48 (0.41-0.55)	0.54 (0.47-0.60)
PLA	0.47 (0.40-0.54)	0.51 (0.44-0.58)
	Mixed model results	
Between groups	**Effect (95% CI)**	**p-value**
Surgical approach (DLA vs. PLA)	-0.08 (-0.205, 0.05)*	0.23
Experience (yes/no)	-0.02 (-0.09, 0.04)^{T}	0.47
Interaction approach X Experience	0.04 (-0.045, 0.135)	0.33

*** in** favor of PLA

^{T} **in** favor of resident

Table S7, Subgroup analyses (*continued*)

	Dislocation, n (%)		
	PLA	**DLA**	P value
Dementia			
Yes	5 out of 63 (8%)	0 out of 74 (0%)	0.02
No	10 out of 205 (4.8%)	1 out of 207 (0.5%)	<0.01
Experience			
Surgeon	6 out of 132 (4.5%)	0 out of 147 (0%)	0.01
Resident	8 out of 129 (6.2%)	0 out of 129 (0%)	<0.05

REFERENCES

1. Breiman, L. (1984). Classification and Regression Trees (1st ed.). Routledge. https://doi.org/10.1201/9781315139470
2. Stef van Buuren and Karin Groothuis-Oudshoorn. mice: Multivariate imputation by chained equations in r. Journal of Statistical Software, 45(3), 2011. https://doi.org/10.18637/jss.v045.i03.
3. Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2):301–320, 2005. doi: https://doi.org/10.1111/j.1467-9868.2005.
4. J. A. Nelder and R. W. M. Wedderburn. Generalized linear models. Journal of the Royal Statistical Society. Series A (General), 135(3):370–384, 1972. ISSN00359238. http://www.jstor.org/stable/2344614.
5. Trevor Hastie and Robert Tibshirani. Generalized Additive Models. Statistical Science, 1(3):297 – 310, 1986. doi: 10.1214/ss/1177013604. URL https://doi.org/10.1214/ss/1177013604.
6. Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 29(5):1189 – 1232, 2001. doi: 10.1214/aos/1013203451. URL https://doi.org/10.1214/aos/1013203451.
7. Jerome H. Friedman. Stochastic gradient boosting. Computational Statistics Data Analysis, 38(4):367–378, 2002. ISSN 0167-9473. doi: https://doi.org/10.1016/S0167-9473(01)00065-2. URL https://www.sciencedirect.com/
8. Jerome H. Friedman. Multivariate Adaptive Regression Splines. The Annals of Statistics, 19(1):1 – 67, 1991. doi: 10.1214/aos/1176347963. URL https: //doi.org/10.1214/aos/1176347963.
9. Mark J. van der Laan, Eric C Polley, and Alan E. Hubbard. Super learner. Statistical Applications in Genetics and Molecular Biology, 6(1), 2007. doi:doi:10.2202/1544-6115.1309. URL https://doi.org/10.2202/1544-6115.1309.
10. Yi Lu, Daniel O. Scharfstein, Maria M. Brooks, Kevin Quach, and Edward H. Kennedy. Causal inference for comprehensive cohort studies, 2019. URL https://arxiv.org/abs/1910.03531.
11. Maria C J M Tol, Nienke W Willigenburg, Hanna C Willems et al. Postero-lateral or direct lateral approach for cemented hemiarthroplasty after femoral neck fracture (APOLLO): protocol for a multicenter randomized controlled trial with economic evaluation and natural experiment alongside. Acta Orthop, 93:732–738, September 2022.
12. Dekkers Olaf M. le Cessie Saskia Choi, Jungyeon. A comparison of different methods to handle missing data in the context of propensity score analysis. European Journal of Epidemiology, 34(1):23 – 36, 2019. doi: 10.1007/ s10654-018-0447-z. URL https://doi.org/10.1007/s10654-018-0447-z.

8

Is the physical performance of patients influenced by the surgical approach of hemiarthroplasty after a femoral neck fracture?

Maria C.J.M. Tol[1], Ariena J. Rasker[1], Nienke W. Willigenburg[1], Taco Gosens[2,3], Martin J. Heetveld[4], Martijn G.M. Schotanus[5,6], J. Carel Goslings[7], Rudolf W. Poolman[8], Hanna C. Willems[9], On Behalf Of The APOLLO Trial Research Group

[1]Department of Orthopedic surgery, Joint Research, OLVG hospital, Amsterdam, [2]Department of Orthopedics and Trauma Surgery, ETZ, Tilburg, [3]Department of Medical and Clinical Psychology, Tilburg University, [4]Department of Trauma Surgery, Spaarne Gasthuis, Haarlem, [5]Department of Orthopedic surgery and Traumatology, Zuyderland MC, Sittard-Geleen-Heerlen, the Netherlands, [6]School of Care and Public Health Research Institute (CAPHRI), Faculty of Health, Medicine and Life Sciences, Maastricht, the Netherlands, [7]Department of Trauma Surgery, OLVG Hospital, Amsterdam, [8]Department of Orthopedic surgery, LUMC, Leiden, [9]Department of Internal Medicine and Geriatrics, Amsterdam UMC, Amsterdam, the Netherlands

(submitted)

ABSTRACT

Background: After an initial hip fracture, patients are at increased risk for recurrent falls and prone for additional injuries. Postoperative physical performance is a key factor contributing to this risk. This study investigated the influence of surgical approach on the postoperative physical performance of patients treated with a hemiarthroplasty after a femoral neck fracture.

Design and methods: The study population was a subsample of a prospective nationwide multicentre RCT comparing the direct lateral approach (DLA) and the posterolateral approach (PLA) after cemented hemiarthroplasty. Variable block randomisation was used. The primary outcome was the physical performance using the Short Physical Performance Battery (SPPB) at 3-4 months after surgery. Secondary outcomes were fear of falling (FoF) and emergency room (ER) visits for falls. Trial registration number: NCT04438226 (clinicaltrials.gov).

Results: A random sample of 70 patients (n=35 DLA and n=35 PLA) was enrolled in this study between 2018 and 2021. The mean SPPB score was 6.6 (SD 2.9) in DLA and 7.6 (SD 3.0) in PLA, mean difference 1.06 (95%CI -2.48–0.37). In the PLA group were significant more good performers (moderate (7-9) + high (10-12) SPPB scores) 25 out of 35 (71%) compared with DLA patients 17 out of 35 (49%).

Conclusion: The effect size of 1.06 out of 12 suggests a clinically relevant difference between the 2 surgical approaches, favoring PLA. However, the lack of statistical significance and the study's limitations warrant cautious interpretation and highlight the need for further research with a larger sample size to confirm these findings.

INTRODUCTION

Falling is a major problem in the older fragile people. A 3rd of this population aged 65 years and over falls each year, and this frequency is increasing with age and frailty(1). In Western Europe, approximately 70% of all injury-related hospital admissions were due to falls(2) and hip fractures are a substantial cause of these admissions. Hip fracture patients are at increased risk for recurrent falls and prone for additional injuries(3), the burden is high for a patient and costly for society(4).

Falling is usually multifactorial, biological factors such as gait performance and balance are identified as consistent predictors of future falls(5). After hip fracture surgery there is a deterioration of physical performance including the gait speed and balance. Some patients develop a Trendelenburg gait or a limping walking pattern after surgery(6). Another major factor in the risk of falling is fear of falling(7). Fear of falling is a predictor of important outcomes in the rehabilitation process after hip fractures. For older adults and frail patients it is important to stay independent (8) and therefore to stay as mobile as possible.

The direct lateral approach (DLA) and the posterolateral approach (PLA) are the most common surgical approaches for inserting a hemiarthroplasty. With the DLA the abductor muscles are split and/or released which can affect the leg press and abduction strength of the hip and might be associated with abductor insufficiency (9). In the PLA these muscles are preserved, however, with this approach, the piriformis is released, and the posterior capsule is incised, which increases the risk of dislocation. There is some literature about the effect of the surgical approach on gait in total hip arthroplasty(10). However, little is known about the effect of the surgical approach of a hemiarthroplasty for a hip fracture on physical performance.

The aim of the present study was to compare the physical performance and fear of falling between patients with DLA and PLA of a cemented hemiarthroplasty after a hip fracture.

METHODS

Trial design and participants

This study was a pre-planned subgroup analysis of the APOLLO trial, a nationwide multicentre randomized controlled trial (RCT)(11). The study compared the DLA and PLA of cemented hemiarthroplasty after a hip fracture. We previously published a detailed description of the study protocol(11). The protocol describes that we invited all eligible patients admitted to the hospitals to participate in the RCT prior to the surgery. Inclusion criteria for the RCT were: < 7 days after initial hip fracture, > 18 years of age. Exclusion

criteria were: pathological fracture, multi-trauma, and secondary surgery after failed internal fixation.

For this subgroup analysis dementia (determined by physicians) was an exclusion criterion. Patients were included in this subgroup analysis at 3 Dutch hospitals among the 14 centers participating in the APOLLO Trial. 2 independent research assistants asked eligible patients after randomization if they wanted to participate in the extra test for this subgroup analysis. They recruited patients until the necessary sample size were achieved in both groups. Selection bias was prevented by inviting all subsequent patients who met the inclusion criteria. The extra tests took place during home visits.

Interventions

Both surgical approaches are extensively described in description of the study protocol(12). All prosthesis were cemented, the type and brand of the prosthesis was left to the surgeons' preference. According to the local guidelines, antibiotic and thromboembolic prophylaxis and wound dressing were chosen. Early weight bearing was advised to the patients, and there was no mandatory use of ADL aids, nor were there movement restrictions. Physical therapy and rehabilitation were conducted following the standard protocols and local guidelines.

Outcomes

The primary outcome of this subgroup study was physical performance assessed with the Short Physical Performance Battery (SPPB) at 3-4 months after surgery. The SPPB test is a predictive tool for possible disability and monitoring of function in older fragile adults(12). It consists of 3 segments of testing: balance, gait speed, and strength of the lower extremity measured with a chair stand. Each segment has a maximum of 4 points and the total score has a maximum of 12 points, high scores suggest better physical performance. The SPPB total score ranges from 0–12; "Very Low" (0-3), "Low" (4-6) "Moderate" (7-9) and "High" (10-12)(13). The groups were divided in good performers with 7 points or more and bad performers with 6 or less points(14). The SPPB test is listed in appendix I. Secondary outcomes were fear of falling (FoF) measured with the Falls Efficacy Scale International (FES-I) score ranging from 16, (no concern about falling) to 64 (severe concern about falling). Other secondary outcomes were mobility, and complications as fall incidents or ER admissions due to falls reported by patients or retrieved from hospital charts within 6 months follow-up.

Sample size and statistics

The sample size was calculated using a 2-sided significance level (α) of 0.05 and a power (ß) of 80% with a standard deviation (SD) of 1.48 and a substantial meaningful change of 1 point on the SPPB (retrieved from the study of Perera et al.)(15). A total of 70 subjects were needed. We analyzed the continuous outcomes using an independent samples

T-test when normally distributed and a Mann-Whitney U-test when not normally distributed. We used a multivariate linear regression analysis to test for potential confounding. Primary outcome of the SPPB were the total scores. Secondary outcomes of the SPPB were the separate scores of the 3 segments and the time it took to complete the 3 segments. The proportion of 'good performers' (i.e. score 7-12) versus 'bad performers' (i.e. score 0-6) per treatment group was compared using chi-square statistics. A value of p ≤0.05 was considered statistically significant.

Table 1. Baseline characteristics

	DLA (n = 35)	PLA (n = 35)	P value
Sex, n (%)			
Female,	22 (63)	17 (49)	0.23
Age, years (SD)	81 (7)	80 (6.2)	0.72
BMI (SD)	24.7 (3.1)	25.3 (3.9)	0.31
KATZ score (SD)	0.87 (1.6)	0.70 (1.1)	0.25
Number of descriptive medicines(SD)	7.7 (6)	7.8 (5)	0.17
Living status prior to trauma, n (%)			0.74
Independent	0	1 (3)	
Independent with help	28 (80)	24 (69)	
Retirement home	5 (14)	7 (20)	
Nursing home	1 (3)	2 (6)	
Rehabilitation nursing home,	1 (3)	1 (3)	
ASA classisfication, n (%)			0.28
I	4 (12)	1 (3)	
II	16 (49)	12 (36)	
III	11 (33)	17 (52)	
IV	2 (6)	3 (9)	
Comorbidity, n (%)			
Cardiac	12 (34)	23 (66)	<0.01
Neurological	4 (11)	8 (23)	0.21
Pulmonary	10 (29)	9 (26)	0.78
Blind	0	1 (3)	0.36
Mobility prior to trauma, n (%)			0.68
No walking aids	19 (58)	22 (67)	
Outdoors with one aid	4 (12)	2 (6)	
Outdoors with 2 aids or frame	7 (21)	8 (24)	
Some indoor mobility	2 (6)	1 (3)	
No functionality of the legs	1 (2)	0	

Table 2. Short Physical Performance Battery (SPPB) outcomes

	DLA (n = 35)	PLA (n = 35)	Mean diff [95%CI]
SPPB, (SD)			
Total	6.6 (2.9)	7.6 (3.0)	1.1 [-2.48-0.37]
Gait speed			
0 points	2	2	
1 point	12	5	
2 points	4	7	
3 points	10	8	
4 points	8	12	
Chair stand test			
0 points	13	7	
1 point	10	10	
2 points	4	6	
3 points	7	7	
4 points	2	4	
Balance			
0 points	1	1	
1 point	3	2	
2 points	2	9	
3 points	14	3	
4 points	16	19	
SPPB, median time in sec (IQR)			
3 meter walking time	4.53 (3.9)	4.06 (2.4)	
Longer than 10 seconds, n (%)	6	2	
Not capable to perform the test, n (%)	0	1	
Time missing, only points available, n (%)	11	8	
5 times sit to stand	18.79 (17.1)	16.13 (5.4)	
Longer than 25 seconds, n (%)	5	2	
Not capable to perform test, n (%)	12	4	
Time missing, only points available, n (%)	6	8	

Additional analysis

The outcome of the primary outcome measure of the SPPB was more widely distributed than we had expected. To gain a better interpretation of the results, we conducted an additional analysis based on progressive insight, in which we categorized the patients into "good" and "bad" performers, using an SPPB score of 7 as the cut-off point. The proportion of 'good performers' (score 7-12) versus 'bad performers' (score 0-6) per treatment group was compared using chi-square statistics. A value of $p \leq 0.05$ was considered statistically significant.

Ethics, registration, data sharing plan, funding, and potential conflicts of interest

The study has been supported by the Dutch Organisation for Health Research and Development (ZonMw grant number: 8430041 12) and was registered before the start of inclusion at clinicaltrials.gov (NCT04438226). The local and the Medical Ethics Committee (METC) of OLVG hospital (number NL63378.100.17) approved the study, which is performed according to the principles of the Declaration of Helsinki(16) and in accordance with the Medical Research Involving Human Subjects Act (WMO) and other guidelines, regulations and Acts. There are no conflicts of interests.

RESULTS

All included patients of the RCT who were eligible for study entry were asked for participation between 2018 and 2021. Patients could not participate when they had dementia or had died within 3-4 months after surgery, so this represents the healthier part of the overall RCT population. In the period of 2020 until march 2021 no patients were included due to restrictions during the COVID pandemic. Because the above affected both groups similarly, this resulted in a 'pseudorandom' sample of 70 patients: 35 with DLA and 35 with PLA. The baseline characteristics are depicted in Table 1. Although there were slightly more females in de DLA (63% vs 49%; p=0.23), linear regression showed that sex was not a confounder. More PLA patients had a cardiac co-morbidity. All other baseline characteristics were not different between both groups. The baseline characteristics of the patients who were eligible for study participation but were not included in the subgroup analysis were comparable (Appendix II). The patients included in the SPPB test subgroup had a slightly higher ASA classification, more were living in a retirement home, but they had less need for walking aids. Timing of the SPPB test was equally distributed between the groups, 124 days (95%CI 118.97 – 127.78) from surgery to test in DLA and 135 days in PLA (95%CI 114.09 – 156.09). Linear regression showed that the timing of the assessment of the test did not influence the primary outcome between the approaches.

SPPB test

The total mean SPPB score at 3-4 months was 6.6 (SD 2.9) in the DLA group and 7.6 (SD 3.0) in the PLA group, with a mean difference of 1.06 (95%CI -2.48–0.37) (table 2). When testing the segments' gait, strength and balance, we found no statistically significant differences in sub-scores.

Since the time dependent measurements of the SPPB were not normally distributed, we analysed these variables using the Mann-Whitney U-test. The median time of gait speed was significantly different between the groups (4.53 sec (IQR 3.9) in DLA and 4.06 sec (IQR 2.4) in PLA, P=0.035). The median time of the chair stand test did not significantly differ between the groups (18.79 sec (IQR 17.1) in DLA vs 16.13 sec (IQR 5.4) in PLA, P=0.265).

However, 12 patients in the DLA group were not able to perform the test against 4 patients in the PLA group. In the PLA group were significant more good performers 25 out of 35 (71%) compared with DLA patients 17 out of 35 (49%).

Fear of falling and fall incidents, and mobility:

Table 3 describes the parameters of falling. DLA patients had statistically significant more FoF 3 months postoperative compared to PLA patients (35.87 points in DLA vs. 26.85 in PLA, (95%CI 2,60 – 15,46)). At 6 months there was still a trend to more FoF in DLA patients (34.33 points in DLA vs. 28.29 points in PLA, (95%CI -0,15 – 12,24)).

In the 6 months follow-up 9 DLA patients (29%) had at least 1 fall incident compared with 13 PLA patients (43%). Only 1 patient needed consultation at the ER due to this fall, this patient was treated with the PLA. At the final follow-up in 16 DLA patients (45.7%) the mobility was worse than prior to the trauma, compared with 11 (31.4%) patients in the PLA group.

Table 3. Fear of falling and fall incidence

	DLA (n =32)	PLA (n = 31)	p-value
FES-I score (SD)			
3 months	35.87 (14.7)	26.85 (10.4)	<0.05
6 months	34.33 (13.9)	28.29 (10)	0.06
Fall incident in 6 months	**DLA (n =35)**	**PLA (n=35)**	
yes, n (%)	9 (29)	13 (43)	0.33
ER admission due to fall			
yes, n (%)	0 (0)	1 (3)	0.61

DISCUSSION

This study with older adult fragile patients suffering a proximal hip fracture treated with a hemiarthroplasty, showed a difference of 1 point SPPB in favor of the PLA. This was not statistically significant, but may be clinically relevant. At 3 months postoperative we observed higher fear of falling in the DLA group compared with the PLA patients, and there were more good performers in the PLA-compared to the DLA group.

The SPPB test consist of 3 elements to measure physical performance: gait speed, balance and strength. Over the years, the simple test of gait speed has become increasingly common for assessing health and functional status. In older adults, gait speed has been suggested as a valuable clinical indicator(17). Studenski et al. found that a lower gait was associated with a reduced life expectancy(18). A systematic review and experts surveys have concluded that gait speed consistently represents a risk factor for disability,

institutionalization, falls, and/or mortality in community-dwelling older adults(19). Lower muscle strength is also a predictor for worse outcomes. A prospective study in older men showed higher risk of suffering a hip fracture when they were unable to complete the chair stand(20). The risk of hip fractures was significantly associated with the repeated chair stand performance(20). In the present study, we observed a slightly lower median gait speed in DLA, and more DLA patients were unable to perform the chair stand test. Compared with PLA, more DLA patients did not return to their pre-fracture mobility. Although this did not result in any differences in mortality, institutionalization, or adverse events in the present study population, probably due to the small sample size and the relatively short follow up. The present follow-up is 6 months, which is probably too short to detect all relevant adverse effect described above.

A prospective longitudinal study of surgically treated hip fracture patients aged 60 years or older found that high FoF at 3 months postoperative predicted poorer functional outcome at 1 year postoperative. This effect was seen in patients with high pre fracture function(8, 21-23). Previous studies found that a higher FoF was also associated with frailty, more institutionalization and greater fall risk. In the present study we found no association of higher risk of falling in patients with a high FoF.

Our findings align with a proposed hypothesis suggesting that patients with lower mobility may experience fewer falls due to limited physical activity, whereas those with moderate mobility engage in more movement but are simultaneously still at risk to falling. However, the literature presents a heterogeneous picture; while some studies corroborate this association(24), others do not observe a consistent relationship(25). In our cohort, the relatively small sample size and wide confidence interval limit the ability to draw definitive conclusions. A strength of this study are the home visits; we made it easy for patients to participate with minimal effort. Another strength is the methodology of the study, it was a pre-planned sub-study, nested in an RCT. There were some limitations. While we reached out to all subsequent RCT patients in 3 hospitals, we may have introduced selection bias because not all patients agreed to participate. Additionally, due to COVID, we couldn't include patients for several months. In this subset, patients had less need for walking aids prior to the trauma compared with the original study population where this subset is obtained from. The results should be seen in this light. Another limitation is the sample size. Although we powered the present study on available literature, our SD was larger than previous studies on the SPPB test. Due to the high variation between patients within treatment groups, some of the observed differences are clinically relevant, but did not reach statistical significance. With a larger sample size, we might be able to enhance the power of these results. With our results in mind, a new study should include at least 284 patients to find a statistical difference. In future studies, we recommend administering the SPPB test to all enrolled patients, rather than a subset, with standardized instructions clearly outlined in the study protocol.

To conclude, this study focused on the effect of surgical approaches on outcomes that reflect or predict the physical performance and fear of falling of hip fracture patients, which is important to stay independent. This study found a mean difference of 1.06 in the SPPB scores between the PLA and DLA surgical approach. This was not statistically significant, but seems clinically relevant, especially in combination with other outcomes. Specifically, 71% of patients in the PLA group were classified as good performers compared to 49% in the DLA group.

On the other hand there were mtore dislocations and reoperations after hemiarthroplasty in PLA patients(26). Therefore, it is difficult to make strong recommendations for either surgical approach in these patients. These findings highlight the need for further investigation with a larger sample size to more accurately assess the clinical impact of the surgical approach on postoperative physical performance.

REFERENCES

1. Organization WH. WHO Global Report on Falls Prevention in Older Age. World Health Orgnaization. 2008.
2. Haagsma J, Olij B, Majdan M, al. e. Falls in older aged adults in 22 European countries: incidence, mortality and burden of disease from 1990 to 2017. Inj Prev 2020;26:67-74.
3. Burgers PTPW, Zielinski SM, Mailuhu AKE, Heetveld MJ, Verhofstad MHJ, Roukema GR, et al. Cumulative incidence and treatment of non-simultaneous bilateral femoral neck fractures in a cohort of one thousand two hundred and fifty patients. Int Orthop. 2014;38(11):2335-42.
4. Kanters TA, Bouwmans CAM, van der Linden N, Swan Tan S, Hakkaart-van Roijen L. Update of the Dutch manual for costing studies in health care. PLoS One. 2017;12(11).
5. Ganz DA, Bao Y, Shekelle PG, Rubenstein LZ. Will my patient fall? JAMA. 2007;297(1):77-86.
6. Sayed-Noor A, Hanas A, Skoldenberg O, Mukka S. Abductor Muscle Function and Trochanteric Tenderness After Hemiarthroplasty for Femoral Neck Fracture. J Orthop Trauma. 2016;30(6).
7. Visschedijk J, Achterberg WP, Van Balen R, Hertogh C. Fear of falling after hip fracture: a systematic review of measurement instruments, prevalence, interventions, and related factors. J Am Geriatr Soc. 2010;58(9):1739-48.
8. Friedman S, Munoz B, West S, al. e. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. J Am Geriatr Soc. 2002;50:1329-35.
9. Winther SB, Husby VS, Foss AO, Wik TS, Svenningsen S, Engdal M, et al. Muscular strength after total hip arthroplasty A prospective comparison of 3 surgical approaches. Acta Orthop. 2016;87(1):22-8.
10. Ulivi M, Orlandini L, Vitale JA, Meroni V, Prandoni L, Mangiavini L, et al. Direct superior approach versus posterolateral approach in total hip arthroplasty: a randomized controlled trial on early outcomes on gait, risk of fall, clinical and self-reported measurements Acta Orthop. 2021;92(3).
11. Tol MCJM, Willigenburg NW, Willems HC, Gosens T, Rasker A, Heetveld MJ, et al. Posterolateral or direct lateral approach for cemented hemiarthroplasty after femoral neck fracture (APOLLO): protocol for a multicenter randomized controlled trial with economic evaluation and natural experiment alongside. Acta Orthop. 2022;12(93):732-8.
12. Kameniar K, Mackintosh S, van Kessel G, Kumar S. The Psychometric Properties of the Short Physical Performance Battery to Assess Physical Performance in Older Adults: A Systematic Review. J Geriatr Psych Ther. 2024;47(1):43-54.
13. Guralnik J, Simonsick E, Ferrucci L, Glynn R, Berkman L, Blazer D. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol. 1994;49(2):85-94.
14. Phu S, Kirk B, Hassan EB, Vogrin S, Zanker J, Bernardo S, et al. The diagnostic value of the Short Physical Performance Battery for sarcopenia. BMC Geriatr. 2020;20(1):242.
15. Perera S. MSH, Woodman R.C.,, S.A. S. Meaningful Change and Responsiveness in Common Physical Performance Measures in Older Adults. The American Geriatrics Society. 2006;54:743-9.
16. 64th WMA General Assembly. Declaration of Helsinki. Fortaleza, Brazil2013.
17. Hall W. Update in geriatrics. Ann Intern Med. 2006;145(7):538-43.
18. Studenski SA, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait Speed and Survival in Older Adults. JAMA. 2011;305(1):50-8.
19. Abellan van Kan G, Rolland Y, Andrieu S, al. e. Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people. J Nutr Healt Aging. 2009;13(10):881-9.

20. Cawthon P, Fullman R, Marshall L, Mackey D, Fink H, Cauley J, et al. Physical Performance and Risk of Hip Fractures in Older Men. J Bone Miner Res. 2008;23(7):1037-44.
21. Bower ES, Loebach Wetherell J, Petkus AJ, Rawson KA, Lenze EJ. Fear of Falling after Hip Fracture: Prevalence, Course, and Relationship with One-Year Functional Recovery. Am J Geriatr Psychiatry. 2016;24(11):1228-36.
22. Arfken CL, Lach HW, Birge SJ, al. e. The prevalence and correlates of fear of falling in elderly persons living in the community. Am J Public Health. 1994;84:565-70.
23. Austin N, Devine A, Dick I, al. e. Fear of falling in older women: a longitudinal study of incidence, persistence, and predictors. J Am Geriatr Soc. 2007;55(1598-1603).
24. Chan BKS, Marshall LM, Winters KM, Faulkner K, Schwartz AV, Orwoll E. Incident Fall Risk and Physical Activity and Physical Performance among Older Men The Osteoporotic Fractures in Men Study. American Journal of Epidemiology. 2005;165(6):696-703.
25. Peeters GMEE, van Schoor NM, Pluijm SMF, Deeg DJH, Lips P. Is there a U-shaped association between physical activity and falling in older persons. Osteoporos Int. 2010;21:1189-95.
26. Tol MCJM, Willigenburg NW, Rasker AJ, Willems HC, Gosens T, Heetveld MJ, et al. Posterolateral or Direct Lateral Surgical Approach for Hemiarthroplasty After a Hip Fracture: A Randomized Clinical Trial Alongside a Natural Experiment. JAMA Network Open. 2024;7(1):e2350765-e.

9

Comparing the posterolateral and the direct lateral approach for cemented hemiarthroplasty after femoral neck fracture: a cost-effectiveness analysis

Jonas L. Esser[1], Maria C.J.M. Tol[2], Nienke W. Willigenburg[2], Ariena J. Rasker[2], Taco Gosens[3,4], Martijn G. M. Schotanus[5,6], Hanna C. Willems[7], Martin J. Heetveld[8], J. Carel Goslings[9], Johanna M. van Dongen[1,] Rudolf W. Poolman[2,10]

[1]Department of Health Sciences , Vrije Universiteit Amsterdam, Netherlands, [2]Department of Orthopedic Surgery, Joint Research, OLVG Hospital, Amsterdam, the Netherlands, [3]Department of Orthopedics and Trauma Surgery, ETZ, Tilburg, the Netherlands, [4]Department of Medical and Clinical Psychology, Tilburg University, Tilburg, the Netherlands, [5]Department of Orthopedic Surgery & Traumatology, Zuyderland Medical Center, Heerlen, Sittard-Geleen, the Netherlands, [6]School of Care and Public Health Research Institute, Faculty of Health, Medicine and Life Science, Maastricht University, the Netherlands, [7]Department of Internal Medicine and Geriatrics, Amsterdam UMC, Amsterdam, the Netherlands, [8]Department of Trauma Surgery, Spaarne Gasthuis, Haarlem, the Netherlands, [9]Department of Trauma Surgery, OLVG Hospital, Amsterdam, the Netherlands, [10]Department of Orthopedic Surgery, LUMC, Leiden, the Netherlands

Acta Orthopaedica (2025)

ABSTRACT

Background: The two most common hemiarthroplasty surgical approaches in hip fracture patients are the posterolateral and the direct lateral approach. This study aimed to estimate the relative cost-effectiveness of these two surgical approaches.

Methods: We conducted an economic evaluation alongside a randomized controlled trial for 6 months. The trial included 555 patients over 18 years of age with an acute femoral neck fracture. The effectiveness outcome used was quality-adjusted life years (QALYs), assessed using the EQ-5D-5L. Costs were measured through self-reported questionnaires administered at baseline, after 3 months, and after 6 months. We dealt with missing data through multiple imputation and analyzed the imputed datasets by comparing group means in costs and QALYs. We also include a secondary analysis, where we adjusted for baseline imbalances through linear regression.

Results: The estimated average treatment effect on the QALYs was 0.02 (95%CI: [-0.006, 0.046]). From the healthcare and societal perspective, we found a non-significant average treatment effect on costs of 1508 [-1744, 4760] and 1583 [-1972, 5137], respectively. The probability of cost-effectiveness was 10% at a willingness-to-pay of zero, and then slowly increased to around 50% for higher willingness-to-pay values.

Conclusions: We found no conclusive evidence of any differences between the surgical approaches with respect to costs, QALYs and cost-effectiveness. We therefore suggest that, from an economic viewpoint, the two surgical approaches should be treated as interchangeable.

Trial registration:
The trial was registered at clinicaltrials.gov (NCT04438226).

INTRODUCTION

Hip fractures are common in the older population, and their incidence is expected to rise substantially in the coming decades (1). Given their significant impact on health-related quality of life (HRQoL) and healthcare costs (2-4); effective and cost-effective treatment is essential. Hemiarthroplasty is a commonly used treatment for femoral neck fractures. Different surgical approaches are available for performing a hemiarthroplasty and vary wide between surgeons. The most common approaches are the direct lateral approach (DLA) and the posterolateral approach (PLA)(5, 6).

Until recently, the evidence regarding the comparative effectiveness of DLA versus PLA was derived mainly from observational studies (7, 8). A 2021 systematic review, also largely based on observational data, suggested that PLA may provide advantages over DLA with respect to HRQoL, abductor insufficiency, and gait-related impairments (9). However, a subsequent meta-analysis indicated that these potential benefits of PLA were outweighed by a higher risk of dislocation and reoperation compared with both DLA and the direct anterior approach (10). More recently, the large multicenter randomized controlled APOLLO trial found no differences in HRQoL, pain, activities of daily living (ADL) independence, or mobility between the approaches, but did report a significantly higher rate of dislocation and reoperation after PLA (11).

Evidence on the cost-effectiveness of PLA versus DLA is currently lacking, even though such information is crucial for healthcare decision-making. Cost-effectiveness analyses show whether health gains are achieved in proportion to the resources used, thereby informing policy, reimbursement, and prioritization within constrained healthcare budgets (12). Even when initial treatment costs are similar, as in the case of PLA and DLA, total healthcare and societal costs may diverge considerably if recovery is quicker or more complete, reducing downstream productivity losses and healthcare use. In order to assess the relative cost-effectiveness of PLA and DLA, we therefore present in this paper an economic evaluation based on the APOLLO trial (13). Our effectiveness outcome is the quality-adjusted life year (QALY), while costs are considered from both a healthcare and a societal perspective.

METHODS

Study design and procedures

This economic evaluation was conducted alongside a multicenter randomized controlled trial (RCT) with a superiority design. We recruited patients in five Dutch hospitals, where the local surgeons could perform both PLA and DLA. The trial was registered at clinical-

trials.gov (NCT04438226) and approved by the local and the Medical Ethics Committee (METC) (number NL63378.100.17).

We screened all patients admitted to the recruited hospitals for eligibility and invited them to participate in the RCT before the surgery. Inclusion criteria were: ≥18 years, acute femoral neck fracture (≤7 days), cemented hemiarthroplasty as recommended treatment, and written informed consent. Multi-trauma patients (Injury Severity Score >15), patients with secondary surgery of the hip or pathological fractures were excluded. After informed consent, we randomly assigned each patient to either the PLA or DLA group using CASTOR EDC (www.castoredc.com), with equal probabilities. Patients, surgeons, and other medical personnel were not blinded. For more detailed information about the study design, see the published protocol(13).

Interventions

Posterolateral approach (PLA)
The external rotators and piriformis are dissected in the PLA group, and a posterior capsulotomy is performed. The gluteus medius and vastus lateralis muscles are preserved. The surgeon's preference determined whether the piriformis was spared or reattached.

Direct lateral approach (DLA)
In the DLA group, the anterior insertion of the gluteus medius is released proximally, and the fibers of the vastus lateralis are divided. An anterior capsulotomy is performed while preserving the stronger posterior capsule.

Measurements

The effectiveness outcome was the quality-adjusted life year (QALY), measured with the EQ-5D-5L at baseline, 3 months, and 6 months after the surgery (14). It should be noted that at the time of recruitment all patients had already sustained a femoral neck fracture; "baseline" thus does not represent the pre-fracture health state. The patients' EQ-5D-5L health states were converted into utility scores ranging from -0.446 (worse than dead) to 1 (optimal health), with a score of 0 indicating death, based on the Dutch utility tariff(15). QALYs were calculated as a weighted average of the reported utility scores (see appendix).

Figure 1: patient flowchart

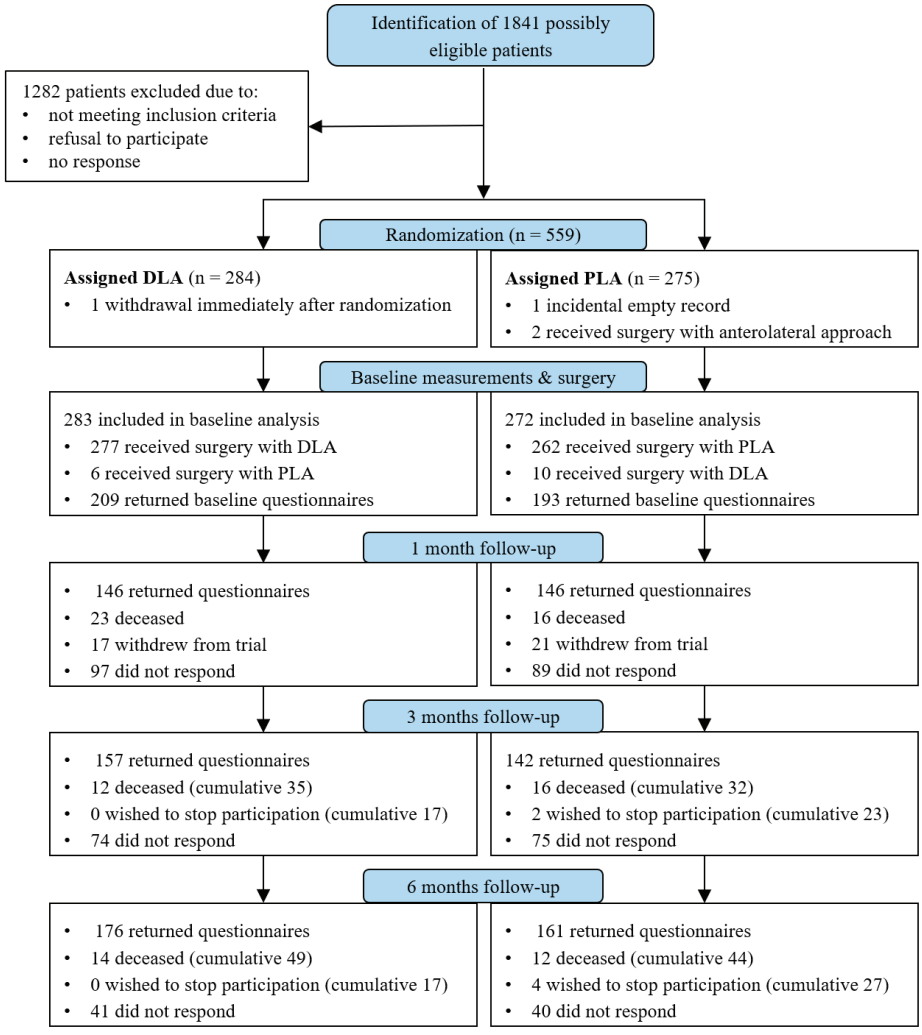


Figure 1: patient flowchart

Resource use information was obtained by questionnaires administered after 1, 3, and 6 months with the following categories :

initial surgery;

follow-up surgeries due to complications;

primary healthcare use (e.g. general practitioner);

secondary healthcare use (e.g. specialists, hospital expenses);

medication use (over-the-counter and prescription-only drugs);

unpaid productivity losses (i.e. volunteer work)

informal care (i.e. care by family members)

All resource use was valued in accordance with the Dutch Manual of Costing (16), with all costs being expressed in Euros (2021). Total costs were estimated from the healthcare perspective (only including surgery costs, primary healthcare costs, secondary healthcare costs, and medication costs), and the societal perspective (including all of the cost categories listed above).

Sample size

To detect a minimally clinically important difference (MCID) of 0.08 in the EQ-5D-5L utility scores (17), from which QALYs were derived, we required a sample size of 555 patients. Based on a two-sided significance level (α) of 0.05 with 80% power, a standard deviation of 0.3, and a 20% loss to follow-up after 6 months (13).

Statistical analysis

Average treatment effects on costs and QALYs were estimated in two ways:
Crude analysis, comparing mean costs and QALYs between both surgical approaches.

Adjusted analysis, where we used two separate linear regressions to estimate the effects on costs and QALYs, respectively. Here, we adjusted for the baseline cost and utility measurements.

Analyses were performed from the healthcare and societal perspective. We used bootstrapping to estimate the sampling variance of all statistics of interest.

Given that the data is entirely composed of questionnaires filled out by elderly patients (or their proxies), we expected to encounter large numbers of partially missing observations. Therefore, we used multivariate imputation by chained equations (MICE) to deal with missing data. Imputation works by generating artificial values which then replace the missing values. We specifically used the predictive mean matching (PMM) method (18) as implemented in the mice software package (19). We imputed a total of 100 datasets. The imputation model included the baseline variables shown in Table 1, as well as all variables which enter into the calculation of the outcome variables. For each imputed dataset, we performed the analyses described below, after which we pooled the point estimates and variances using Rubin's rules(20). Confidence intervals were computed based on a normal approximation, using the estimated variances, since this approach has been shown to perform well in a recent simulation study(21).

Cost-effectiveness analysis

The most important parameters for the cost-effectiveness analysis are the average treatment effects on costs and QALYs, which we estimate based on the observed data. These are then combined into a utility function, called the Net Monetary Benefit (NMB), which informs the implementation decision: if the INB is positive, PLA is considered cost-effective and should be implemented. Vice versa if the INB is negative. Incremental Cost-Effectiveness Ratios (ICERs) were calculated by dividing the estimated effect on costs by the estimated effect on the QALYs. We plotted bootstrapped cost-effect pairs on a cost-effectiveness plane to visually inspect the uncertainty surrounding the estimates (22). We further provided a cost-effectiveness acceptability curve to illustrate the probability of PLA being cost-effective at different levels of willingness-to-pay. In the Netherlands,

decision-makers usually apply thresholds of €20,000, €50,000 and €80,000 per QALY, depending on the severity of the disease (16). More detailed information is provided in the appendix.

RESULTS

Between February 2018 and January 2022, 555 patients were included (272 PLA and 283 DLA). Figure 1 shows a flowchart, adapted from Tol et al, which illustrates the flow of patients through the trial (14). Patient's baseline characteristics per treatment group are summarized in Table 1. Overall, the groups were well-balanced, with no stark differences.

Missing data and imputation

Missing observations were prevalent among the recorded data, with only 72 out of 555 observations being complete, with respect to all variables of interest for this paper. All patients had age and gender baseline measurements recorded. For all other baseline variables, missing values were present; the highest missingness proportion here was the BMI score, which was missing for 24% of patients. For answers to the EQ-5D-5L questionnaire, the missingness proportions at baseline, 3 months, and 6 months after the surgery were 27%, 46% and 39%, respectively. For the cost questionnaires, the missingness proportions at 1 month, 3 months, and 6 months after the surgery were 40%, 32% and 21%, respectively. Hence, all the following tables and figures are based on imputed data.

Table 1: Baseline characteristics in APOLLO trial (SD = standard deviation)

	PLA group (n = 272)	DLA group (n = 283)
Age (years), mean (SD)	82 (8)	82 (7)
Gender (female), %	172 (63%)	172 (61%)
Body mass index, kg/m^2, mean (SD)	24.7 (4.2)	24.2 (4.1)
ASA I	4 (1.5%)	8 (2.8%)
ASA II	86 (31.6%)	107 (37.8%)
ASA III	171 (62.9%)	158 (55.8%)
ASA IV	11 (4%)	10 (3.5%)
Impaired mobility, n (%)	166 (61%)	175 (62%)
Dependent living status (e.g. nursery home)	52 (19%)	64 (23%)
Quality of life (EQ-5D utility score), mean (SD)	0.389 (0.358)	0.333 (0.366)

Utility scores and cost variables

Table 2 shows the patients' EQ-5D-5L utility scores, QALYs, and cost variables. For both treatment groups, the utility scores at 3 months and 6 months were much higher than at baseline, indicating that both treatments effectively improve the patients' quality of life. The utilities and QALYs were slightly higher in the PLA group. Note, however, that the baseline utility was also higher in the PLA group, and the differences in Table 2 are not adjusted for that.

Table 2: group means and differences of disaggregated outcome variables

Variable	PLA group (n = 272), Mean (SE)	DLA group (n=283), Mean (SE)	Difference [95% CI]
EQ-5D utility score – three months	0.530 (0.024)	0.482 (0.023)	0.047 [-0.019, 0.078]
EQ-5D utility score – six months	0.500 (0.024)	0.489 (0.024)	0.012 [-0.054, 0.078]
QALY	0.244 (0.009)	0.224 (0.010)	0.020 [-0.006, 0.046]
Initial surgery costs	3300 (0)	3300 (0)	-
Follow-up surgery costs	895 (188)	358 (99)	537 [121, 954]
Primary health care costs (other than surgery)	5395 (999)	4476 (811)	920 [-1603, 3441]
Secondary health care costs	3369 (854)	3099 (881)	270 [-2135, 2676]
Medication costs	4 (1)	5 (1)	-1 [-2, 1]
Unpaid productivity costs	64 (15)	57 (13)	7 [-32, 46]
Informal care costs	1799 (532)	1688 (407)	109 [-1204, 1422]
Total costs – healthcare perspective	13195 (1381)	11490 (1327)	1706 [-2166, 5458]
Total costs – societal perspective	15056 (1488)	13235 (1387)	1822 [-2109, 5753]

SE = standard error, CI = confidence interval.

There were no significant differences in utility scores and QALYs between both surgical approaches in the crude analyses. The adjusted analyses are presented in Table 3 (column 3). There were no differences in QALYs between PLA and DLA. Costs in all categories were similar in both treatment groups, with the exception of the surgery costs, which were significantly higher in the PLA group. Total healthcare and societal costs were higher in the PLA group than the DLA group; however, these differences were non-significant.

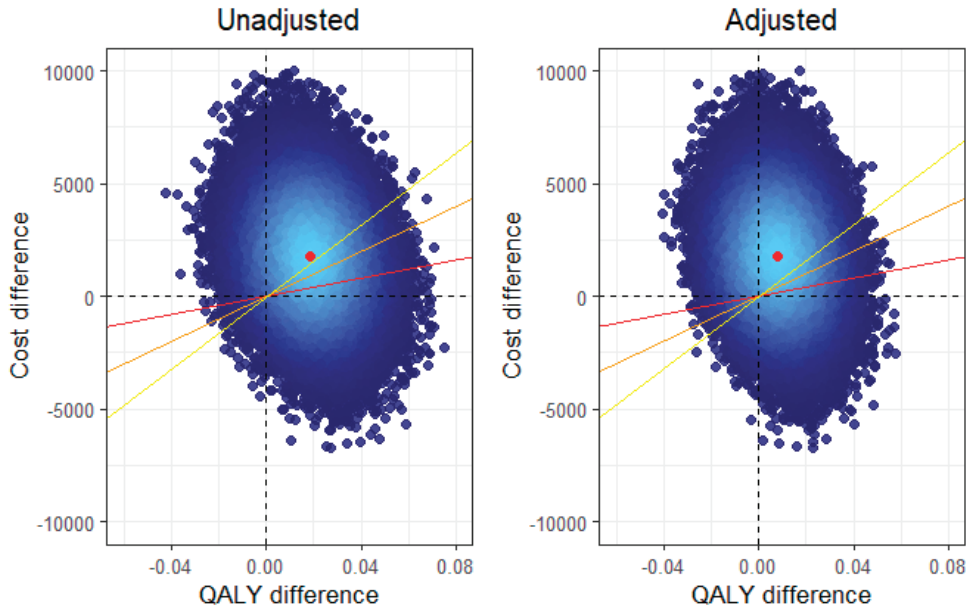


Figure 2: cost-effectiveness plane for unadjusted and adjusted analysis, societal perspective

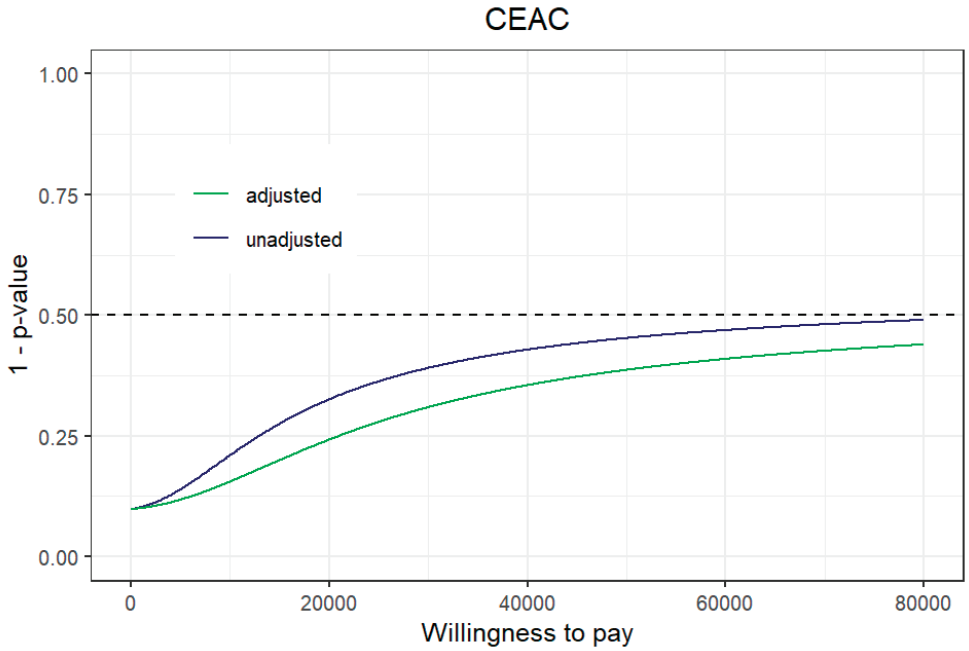
Cost-effectiveness

From both the healthcare and societal perspective, the ICERs showed that PLA was - on average - 'more costly' and 'more effective' than DLA. At willingness-to-pay thresholds of €20,000, €50,000, and €80,000 per QALY, the point estimates for the NMBs were negative, although none were statistically significant (16). Figure 3 shows that, at a willingness-to-pay of €0 per QALY, PLA had a 0.17 probability of being cost-effective compared with DLA. This means that if decision-makers are not willing to pay anything per QALY gained, the probability of PLA being cost-effective compared with DLA is only 17%. This probability increased with higher willingness-to-pay thresholds, but remained below 0.50 across the full range.

Table 3: results of regression analyses (adjusted for baseline measurements)

Analysis	ΔCosts [95% CI]	ΔQALY [95% CI]	NMB(20000) [95% CI]	NMB(50000) [95% CI]	NMB(80000) [95% CI]	ICER
Healthcare perspective	1508 [-1744, 4760]	0.009 [-0.014, 0.032]	-1330.28 [-5589.68, 2929.11]	-1064.17 [-10588.73, 8460.38]	-798.06 [-15791.43, 14195.31]	169970
Societal perspective	1583 [-1972, 5137]	0.009 [-0.014, 0.032]	-1405 [-5756, 2945]	-1139 [-10710.40, 8431.72]	-873.23 [--15899.56, 14153.09]	178444

Figure 3: cost-effectiveness acceptability curves for both analyses, societal perspective



DISCUSSION

The results suggest that, in the treatment of cemented hemiarthroplasty in adults suffering an acute femoral neck fracture, PLA is not cost-effective compared with DLA. Irrespective of the cost-effectiveness threshold used, the probability of cost-effectiveness is smaller than 0.50.

We found no difference in HRQoL between the two approaches, which is line with the previously published main results of the RCT (see Tol et al (11). There were significantly higher follow-up surgery costs in the PLA group, which can be attributed to a much higher rate of dislocations, which was 5.5% and 0.4%, respectively (11). In line with this finding, we found PLA to be associated with higher secondary healthcare costs (e.g., specialists, hospital expenses, such as an emergency visit) compared with DLA; however, this difference was not significant. In the Netherlands, the reduction of a dislocated hip is frequently performed non-surgically at the emergency room with the use of Procedural Sedation and Analgesia (PSA). The costs of this procedure with PSA are not clear since the costs are not adequately documented in Dutch hospitals. For an emergency room consultation, a standard fee is charged, independent of which treatments and anesthesia and team were needed. Therefore, in this study, the secondary healthcare costs for PLA may have been somewhat underestimated.

A direct comparison of our results to the literature is challenging due to the lack of research on the comparative costs and cost-effectiveness of PLA and DLA in the context of hemiarthroplasty. It is noteworthy that the average healthcare costs and utility scores we found for hip fracture patients in the Netherlands were somewhat lower than those estimated in two recent studies (2, 4). These differences may have resulted from variations in methodology, patient population, and the element of chance due to the limited sample sizes in both the aforementioned studies and ours. The latter factor is not quantifiable due to the substantial uncertainty in our estimates. Therefore, we refrain from speculating about differences in methodology or patient population.

Strengths and limitations

This study had several strengths. To date, this is the first economic evaluation of a randomized controlled trial comparing the two most commonly used surgical approaches for hemiarthroplasty for femoral neck fractures. Furthermore, this study is part of a trial which is the largest randomized controlled trial worldwide that addresses this subject. Another strength of the study is that we included patients suffering dementia. Dementia is often an exclusion criteria in clinical trials, even though patients with dementia present a substantial part of the population of patients with a hip fracture (23). We increased the generalizability of the results by including them. One limitation of this study is the substantial proportion of missing data. As a large part of the data is sourced from self-assessments through questionnaires, the observations were partially missing for many patients, as is commonly the case in trial-based cost-effectiveness studies (24). There was no missing data regarding dislocations, reoperations, and admission to the ER, which was used for the follow-up surgery and secondary healthcare costs. We extracted these data from medical reports and registry data, which we reviewed 6 months after surgery. Although multiple imputation can mitigate the bias caused by informative missingness to some extent, we still encountered very large standard errors in the estimates, making it difficult to draw strong conclusions from the results of our study. A further limitation concerns the generalizability of our findings, as not all eligible patients were randomized. Some patients chose not to participate in the trial, while others were excluded due to logistical challenges or time constraints. We lack detailed information on the specific reasons for non-participation, which limits our ability to confirm that the data represents an unbiased sample of the target patient population. Nevertheless, a comparison with the Dutch Arthroplasty Register indicates that the baseline characteristics in our sample are comparable to those of the broader patient population. Another limitation is the relatively short follow-up duration of only 6 months, which restricts the ability to capture the longer-term effects of the intervention on healthcare utilization and costs (12). As joint replacement is expected to provide benefits well beyond this period, future studies should evaluate cost-effectiveness over longer follow-up, ideally combining trial-based data with model-based extrapolations. Please also note that the sample size calculation was based on an MDC value (0.08) derived from the EQ-5D-3L, as evidence for the EQ-5D-

5L was not yet available at the study's initiation. The EQ-5D-5L generally shows improved measurement properties, and more recent evidence suggests a slightly higher MID for improved health states (0.11); see Cheng et al (25). However, this does not affect our interpretation, as the observed difference in HRQoL (0.009) was well below both thresholds.

CONCLUSION

In conclusion, we found no evidence of a difference in cost-effectiveness between PLA and DLA for hemiarthroplasty following acute femoral neck fractures in adult patients. We therefore suggest that, from an economic viewpoint, the two surgical approaches should be treated as interchangeable.

REFERENCES

1. Sing CW, Lin TC, Bartholomew S, Bell JS, Bennett C, Beyene K, et al. Global epidemiology of hip fractures: secular trends in incidence rate, post-fracture treatment, and all-cause mortality. Journal of Bone and Mineral Research. 2023;38(8):1064-75.
2. Kanters T, Van De Ree C, De Jongh M, Gosens T, Hakkaart-van Roijen L. Burden of illness of hip fractures in elderly Dutch patients. Archives of osteoporosis. 2020;15(1):1-7.
3. Williamson S, Landeiro F, McConnell T, Fulford-Smith L, Javaid MK, Judge A, Leal J. Costs of fragility hip fractures globally: a systematic review and meta-regression analysis. Osteoporosis international. 2017;28(10):2791-800.
4. Burgers P, Hoogendoorn M, Van Woensel E, Poolman R, Bhandari M, Patka P, Van Lieshout E. Total medical costs of treating femoral neck fracture patients with hemi-or total hip arthroplasty: a cost analysis of a multicenter prospective study. Osteoporosis international. 2016;27(6):1999-2008.
5. Chechik O, Khashan M, Lador R, Salai M, Amar E. Surgical approach and prosthesis fixation in hip arthroplasty world wide. Arch Orthop Trauma Surg. 2013;133(11):1595-600.
6. LROI. Annual report of Dutch Arthroplasty Register. 2021.
7. Kristensen TB, Vinje T, Havelin LI, Engesæter LB, Gjertsen J-E. Posterior approach compared to direct lateral approach resulted in better patient-reported outcome after hemiarthroplasty for femoral neck fracture: 20,908 patients from the Norwegian hip fracture register. Acta orthopaedica. 2017;88(1):29-34.
8. Leonardsson O, Rolfson O, Rogmark C. The surgical approach for hemiarthroplasty does not influence patient-reported outcome: a national survey of 2118 patients with one-year follow-up. The Bone & Joint Journal. 2016;98(4):542-7.
9. Tol MC, van Beers LW, Willigenburg NW, Gosens T, Heetveld MJ, Willems HC, et al. Posterolateral or direct lateral approach for hemiarthroplasty after femoral neck fractures: a systematic review. Hip International. 2021;31(2):154-65.
10. van der Sijp MPL, van Delft D, Krijnen P, Niggebrugge AHP, Schipper IB. Surgical approaches and hemiarthroplasty outcomes for femoral neck fractures: a meta-analysis. The journal of Arthroplasty. 2018.
11. Tol MC, Willigenburg NW, Rasker AJ, Willems HC, Gosens T, Heetveld MJ, et al. Posterolateral or Direct Lateral Surgical Approach for Hemiarthroplasty After a Hip Fracture: A Randomized Clinical Trial Alongside a Natural Experiment. JAMA Network Open. 2024;7(1):e2350765-e.
12. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford university press; 2015.
13. Tol MC, Willigenburg NW, Willems HC, Gosens T, Rasker A, Heetveld MJ, et al. Posterolateral or direct lateral approach for cemented hemiarthroplasty after femoral neck fracture (APOLLO): protocol for a multicenter randomized controlled trial with economic evaluation and natural experiment alongside. Acta orthopaedica. 2022;93:732-8.
14. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research. 2011;20(10):1727-36.
15. Versteegh MM, Vermeulen KM, Evers SM, De Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. Value in health. 2016;19(4):343-52.
16. Hakkaart-van Roijen L, Van der Linden N, Bouwmans C, Kanters T, Tan SS. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg. 2016.
17. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Quality of life research. 2005;14:1523-32.

18. Little RJ. Missing-data adjustments in large surveys. Journal of Business & Economic Statistics. 1988;6(3):287-96.
19. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of statistical software. 2011;45:1-67.
20. Harel O, Zhou XH. Multiple imputation: review of theory, implementation and software. Statistics in medicine. 2007;26(16):3057-77.
21. Brand J, van Buuren S, le Cessie S, van den Hout W. Combining multiple imputation and bootstrap in the analysis of cost-effectiveness trial data. Statistics in Medicine. 2019;38(2):210-20.
22. Willan AR, Briggs AH. Statistical analysis of cost-effectiveness data: John Wiley & Sons; 2006.
23. Hebert-Davies J, Laflamme G-Y, Rouleau D. Bias towards dementia: are hip fracture trials excluding too many patients? A systematic review. Injury. 2012;43(12):1978-84.
24. Ling X, Gabrio A, Mason A, Baio G. A scoping review of item-level missing data in within-trial cost-effectiveness analysis. Value in Health. 2022;25(9):1654-62.
25. Cheng LJ, Chen LA, Cheng JY, Herdman M, Luo N. Systematic review reveals that EQ-5D minimally important differences vary with treatment type and may decrease with increasing baseline score. Journal of Clinical Epidemiology. 2024;174:111487.
26. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health economics. 2005;14(5):487-96.
27. Löthgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health economics. 2000;9(7):623-30.
28. O'Brien BJ, Briggs AH. Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods. Statistical methods in medical research. 2002;11(6):455-68.

APPENDIX

Calculation of QALYs

As is customary in cost-effectiveness analyses [22], we calculate the total QALY as a weighted average of the measured QoL scores. Let $QALY_i$ be the QALY of individual i, $U_{i,k}$ the measured utility at the follow up point k, with $U_{i,0}$ the baseline utility. Let D_k the time that has passed at follow up point k since the beginning of the study (expressed in years), with $D_0 = 0$. Then
$QALY_i = \sum_{k=1}^{K} \frac{1}{2}(U_{i,k-1} + U_{i,k})(D_k - D_{k-1})$

In our study, patient's quality of life scores were measured at baseline, after three months and finally after six months. So the above formula reduces to
$QALY_i = \frac{1}{2}\left[(U_{i,0} + U_{i,1})\left(\frac{3}{12} - 0\right) + (U_{i,1} + U_{i,2})\left(\frac{6}{12} - \frac{3}{12}\right)\right] = \frac{1}{8}(U_{i,0} + 2U_{i,1} + U_{i,2})$

Calculation and interpretation of the cost-effectiveness acceptability curve (CEAC)

Let Δ_c and Δ_q be the expected differences in costs and QALYs between the two treatment options. Let λ be the maximum price that society is willing to pay per additional QALY gained in exchange for that price. Then the intervention is deemed to be cost-effective if and only if $\lambda\Delta_q - \Delta_c > 0$. The left side of this inequality is known as the net monetary benefit (NMB) and can be computed as a function of λ and the mean differences in costs and QALYs.[23] $\widehat{V} = \begin{bmatrix} \widehat{Var}\left(\widehat{\Delta_c}\right) & \widehat{Cov}\left(\widehat{\Delta_c}, \widehat{\Delta_q}\right) \\ \widehat{Cov}\left(\widehat{\Delta_c}, \widehat{\Delta_q}\right) & \widehat{Var}\left(\widehat{\Delta_q}\right) \end{bmatrix}$
Let $\widehat{\Delta_c}$ and $\widehat{\Delta_q}$ be the point estimates for the treatment effects; let be the estimated covariance matrix. Then the point estimate and variance estimate of the NMB are given as
$\widehat{NMB}(\lambda) = \lambda\widehat{\Delta_q} - \widehat{\Delta_c}$
$\widehat{Var}\left(\widehat{NMB}(\lambda)\right) = \lambda^2\widehat{Var}\left(\widehat{\Delta_q}\right) + \widehat{Var}\left(\widehat{\Delta_c}\right) - 2\lambda\widehat{Cov}\left(\widehat{\Delta_c}, \widehat{\Delta_q}\right)$
Based on the point estimates and the estimated covariance matrix of the treatment effects, we compute the quantity $A(\lambda) = \phi\left(\frac{\widehat{NMB}(\lambda)}{\left[\widehat{Var}\left(\widehat{NMB}(\lambda)\right)\right]^{1/2}}\right)$

where $\phi(.)$ is the cumulative distribution function of standard normal random variable. In the classical statistical paradigm, $A(\lambda)$ is 1 minus the p-value for the one-sided hypothesis test $H_0 : \lambda\Delta_q - \Delta_c \leq 0$ versus $H_1 : \lambda\Delta_q - \Delta_c > 0$ [19]. In contemporary practise, $A(\lambda)$ is often instead interpreted as the "probability of cost-effectiveness", so $P(\lambda\Delta_q - \Delta_c > 0)$. Even though it is often left implicit, such an interpretation is only possible under a Bayesian paradigm.[24] We do not prescribe either framework; the readers may interpret the estimates as they wish.

Regardless of interpretation, it is customary in economic evaluations to not commit to a single willingness-to-pay λ, but to instead compute $A(\lambda)$ for a range of possible values of λ and plot the two against each other. This plot is known as the cost-effectiveness acceptability curve (CEAC).

General discussion

Efficient management of hip fracture care is becoming increasingly critical in light of the projected demographic shift toward an ageing population and the consequent rise in healthcare demand. Hip fractures are particularly prevalent among older adults and are associated with significant morbidity, mortality, and healthcare costs. As the proportion of elderly individuals increases, the absolute number of hip fractures is expected to grow substantially, placing additional strain on already overburdened healthcare systems. Optimizing care pathways—through evidence-based surgical approaches, standardized postoperative protocols, and effective rehabilitation—is essential to improve patient outcomes, reduce complication rates, and contain costs. Moreover, aligning treatment strategies with long-term functional recovery and quality of life is vital to support the autonomy and well-being of older adults. In this context, delivering cost-effective, high-quality care for hip fracture patients is not only a clinical imperative but also a key component of sustainable healthcare planning in ageing societies.

Surgical Strategies for Displaced Femoral Neck Fractures: Existing Knowledge

Prior to this thesis, treatment options for active older adults with displaced femoral neck fractures included hemiarthroplasty and total hip arthroplasty (THA). THA, which replaces both the femoral head and acetabulum, may offer better functional outcomes but carries a higher risk of dislocation. The HEALTH trial (1) found no significant differences in secondary procedures or quality of life between THA and hemiarthroplasty after two years. Furthermore, hemiarthroplasty can be performed with cemented or uncemented fixation; the WHiTE trial(2) showed that cemented implants are associated with improved health-related quality of life and fewer peri-prosthetic fractures, making it the preferred option for frail elderly patients. However, evidence on the optimal surgical approach remained limited. The two most common approaches in the Netherlands are the direct lateral (DLA) and posterolateral (PLA) (3). DLA involves splitting the gluteus medius and vastus lateralis, which may impair abductor function but preserves the posterior capsule, reducing dislocation risk. PLA spares the abductors but compromises posterior stability, potentially increasing dislocation rates.

Surgical approaches in hemiarthroplasty: clinical Outcomes, patient experience, and costs

This thesis specifically addressed the question which surgical approach is most effective in hemiarthroplasty following a femoral neck fracture. Rather than focusing solely on traditional surgical outcomes such as complication or revision rates, we examined outcomes that are directly relevant to the patient's lived experience. For older adults, particularly those facing the challenges of advanced age, preserving independence, maintaining mobility, and achieving the best possible quality of life in the final stages of life are often more meaningful than purely clinical endpoints. By incorporating these patient-centered

outcomes, this thesis contributes to a more comprehensive understanding of surgical success. In doing so, it supports a shift in focus from solely technical measures toward outcomes that truly reflect what matters most to patients—an essential step in designing sustainable, person-centered care pathways in an ageing society.

Our findings underscore the complexity of determining a universally superior surgical approach for hemiarthroplasty in older adults with femoral neck fractures. The results indicate that this is not a binary issue; although the risk of dislocation and reoperation was indeed higher in patients treated via the PLA, patient-reported outcomes did not uniformly favor one technique over the other. Specifically, despite differences in complication rates, overall quality of life was comparable between groups. Moreover, no significant differences were observed in mobility, pain, independence in activities of daily living (ADL), or fear of falling—outcomes that are particularly relevant for this population. Interestingly, in a smaller subgroup of RCT participants, we did observe a clinically relevant difference in physical performance favoring DLA. However, due to the limited sample size and the potential for selection bias, we were unable to establish a robust causal relationship. These findings suggest that while certain surgical risks may differ between approaches, these do not necessarily translate into differences in patient-perceived recovery or autonomy. This highlights the importance of integrating both clinical and patient-centered outcomes when considering treatment strategies.

Importantly, this thesis is the first to investigate whether one of the two commonly used surgical approaches—PLA or DLA—is more cost-effective. Our cost-effectiveness analysis revealed no significant differences between the approaches. This suggests that from an economic standpoint, neither technique holds a clear advantage, further supporting the view that surgical decision-making should be individualized based on patient characteristics and clinical judgment, rather than cost considerations.

Since it shows no differences in patient outcomes, costs, or most clinical measures—despite a higher risk of dislocation—the choice of surgical approach may be well-suited for shared decision-making. Although shared decision-making is a valuable goal, it requires that patients are well-informed and that the surgeon is trained in both surgical approaches. This presents practical challenges. How effectively can a patient process complex medical information immediately after sustaining a hip fracture, particularly when in significant pain and distress? Moreover, shared decision-making assumes the logistical feasibility of offering both options. In reality, if a surgeon is not trained in both approaches, this may limit the availability of choice. Should a patient then wait until a suitably skilled surgeon is available? Such delays may introduce new risks and disrupt the flow of care within the hospital system. In essence, while the ideal of shared decision-making in selecting a surgical approach for hemiarthroplasty promotes autonomy and

respect for patient preferences, there is a risk that it may overshoot its goal if these practical limitations are not acknowledged and addressed.

Dementia and clinical trials: advancing inclusive research

Another important aspect of this thesis is the deliberate inclusion of patients with dementia in the randomized clinical trial—a group often underrepresented in clinical research. Individuals with cognitive impairment constitute a substantial proportion of the hip fracture population, yet they are frequently excluded from clinical trials due to the methodological and ethical challenges their inclusion poses. These challenges include lower response rates to questionnaires, higher rates of missing data, and difficulties in assessing subjective outcomes such as quality of life or functional recovery. Indeed, our own study observed increased missing values among participants with dementia, underscoring the practical difficulties of involving this population. Nonetheless, excluding this vulnerable group limits the generalizability of research findings and risks marginalizing those who may have the greatest need for evidence-based care. Prior to this thesis, it remained unclear which proxy perspective was most appropriate for capturing patient-reported outcomes in individuals with cognitive impairment. Through our agreement study (Chapter 2), we addressed this knowledge gap and provided empirical guidance on the use of proxy respondents in future research. We therefore strongly advocate for the inclusion of patients with dementia in future clinical studies on hip fractures or other conditions with a high prevalence of dementia. Doing so will yield more representative and clinically applicable findings, and contribute to the development of care strategies that are both equitable and relevant to the full spectrum of patients affected by hip fractures.

Future research and recommendations

An important question that remained insufficiently addressed in this thesis is the effect of the surgical approach on physical performance. Although the physical performance study demonstrated a clinically relevant difference, this was not statistically significant. The study was limited by an inadequate sample size, largely attributable to greater variability in outcomes than anticipated from prior literature. Moreover, the inability to perform assessments over several months due to the COVID-19 pandemic introduced a potential selection bias. This could potentially have been solved by conducting balance assessments on all patients, but we lacked the capacity and budget to implement such an approach. Therefore, the methodological design of the study complicated the interpretation of our findings. In hindsight, we would have preferred to make different design choices, though these were difficult to implement given the constraints at the time. Moreover, introducing such a balance assessment in all participants could have introduced a selection bias in the whole trial, because patients who are hesitant to commit to additional physical testing would not agree to participate. Future research supported by greater resources could address this question more rigorously, for instance using more sophisticated (i.e. biomechanical) assessments. This thesis focused on the two most commonly performed

surgical approaches for hip hemiarthroplasty: the direct lateral approach (DLA) and the posterolateral approach (PLA). At the time of study initiation, the direct anterior approach (DAA) was used in only approximately 3% of cases in our setting. As is often the case in scientific research, developments have since evolved, and the DAA has gained considerable popularity in clinical practice. Given this shift, future research should aim to compare the traditional approaches (DLA and PLA) with the increasingly adopted DAA, to determine whether its growing use is supported by differences in clinical outcomes.

Another important recommendation for future research is the long-term follow-up of large cohorts such as the HEALTH trial. In our own 12-year follow-up of the ARTHRO trial, which compared THA and hemiarthroplasty in patients with femoral neck fractures, we found no difference in long-term quality of life between the two groups. However, THA was associated with a higher incidence of dislocations. It is important to note that the sample size at follow-up was limited due to the high mortality rate inherent to this elderly population. A long-term analysis of a larger cohort such as the HEALTH trial could provide more definitive evidence regarding the durability and complications of THA versus hemiarthroplasty. Should these findings confirm our results, it could have a significant impact on clinical practice. After all, if hemiarthroplasty yields equivalent long-term quality of life without increased revision rates, while involving a less extensive surgical procedure, lower costs, and a reduced risk of dislocation—an event that can cause pain, anxiety, and unplanned hospital visits—then the rationale for routine use of THA becomes less compelling. Such evidence would support a shift toward more efficient, value-based hip fracture care, better aligned with patient safety, healthcare sustainability, and resource optimization.

In the light of increasing healthcare costs and scarcity of health care professionals, it is important that we keep evaluating healthcare processes that are considered routine. The focus of cost-effectiveness research should not be limited to the evaluation of new interventions, but must also critically examine existing clinical practices. Clinical routines that have become habitual over decades may no longer represent best practice, particularly if newer evidence challenges their relevance or benefit.

Even minor gains in health outcomes or cost savings, when applied across large patient populations, can yield substantial system-level improvements. Therefore, revisiting established treatment protocols is not just a theoretical exercise—it is a practical necessity for creating high-value care. We must actively question whether ingrained routines remain justified or whether they persist merely out of habit or outdated assumptions. Promoting meaningful, evidence-based care begins with the willingness to challenge the familiar.

Beyond the RCT: embracing natural experiments for smarter, real-world research

Fortunately**,** in recent years there has been more focus on making healthcare more efficient. But it is not only clinical healthcare that demands greater efficiency—our approach to research must also evolve. Important in the quest to reduce research waste is the full utilization of patient data already being generated in routine care. Research waste refers to the loss of value in medical research due to poorly chosen questions, weak study design, non-publication, or incomplete reporting, leading to missed opportunities for better care. A powerful way to reduce research waste is by using the wealth of patient data already available in routine care. This allows us to learn much faster from treatments we are already providing. That is the core idea behind *care evaluation*: embedding research into everyday healthcare, so that evaluating what we do becomes part of how we work. This approach improves relevance, reduces duplication, and accelerates meaningful learning. Traditional clinical trial designs, while rigorous, are often time-consuming, expensive, and poorly suited to the fast-changing realities of healthcare. Especially in the context of an ageing population and limited resources, it is essential to adopt more pragmatic, scalable research methodologies that reflect real-world clinical settings. In this thesis, we demonstrated that a natural experiment design can yield results comparable in reliability and causal inference to those of a traditional RCT. Crucially, natural experiments can be conducted more quickly, with fewer resources, and better reflect real-world clinical decision-making. This makes them particularly suitable for evaluating standard procedures or practices that are already embedded in clinical routines. Moreover, from an ethical and patient-centered perspective, natural experiments can offer a less intrusive alternative. Some patients may feel discomfort or resistance when being randomly assigned to a treatment they would not otherwise have received. By observing the consequences of routine choices already being made by clinicians, natural experiments preserve autonomy and can enhance participation. This thesis introduces an innovative step forward in clinical research methodology: the fusion of findings from both the RCT and the natural experiment. This approach represents a methodological evolution, combining the internal validity of the RCT with the external validity and efficiency of the natural experiment. By integrating these complementary sources of evidence, we can enhance the robustness and applicability of study conclusions. This not only advances methodological thinking in this field but also opens the door for broader application.

Future research can adopt and refine this approach of Natural Experiment (with or without data fusion) across various clinical questions and study designs, especially when aiming to draw causal conclusions using faster and more practical methods than an RCT. A trauma setting is particularly well-suited for this type of research, as treatment or hospital choice is typically not predetermined by the patient or surgeon, but rather influenced by the timing and location of the traumatic event.

By incorporating Natural Experiments, researchers can produce more comprehensive and generalizable evidence while reducing costs, shortening timelines, and respecting patient preferences. As healthcare systems continue to evolve under increasing demographic and economic pressure, the way we generate knowledge must evolve accordingly—and our work contributes a practical tool for that future.

Conclusion

In conclusion, this thesis provides both clinical and methodological contributions to the optimization of hip fracture care and research in an ageing society. By evaluating surgical approaches not only on traditional outcomes such as complications and dislocations but also on patient-centered measures such as quality of life and functional recovery, we offer a more comprehensive perspective on what truly matters in treatment decisions. Our inclusion of patients with dementia, and our work on proxy perspectives, further strengthen the relevance and inclusivity of our findings. Methodologically, we have shown that natural experiments—particularly when combined with RCT data through innovative data fusion—can provide not only efficient and ethically acceptable alternatives to traditional trials, but also retain the crucial ability to draw causal inferences. Being able to establish causal relationships remains a cornerstone of high-quality clinical research. These contributions underscore the need for a more reflective and adaptive approach to both care delivery and research design—one that meets the growing demands of healthcare while remaining centered on the needs of the patient.

REFERENCES:

1. investigators TH. Total Hip Arthroplasty or Hemiarthroplasty for Hip Fracture. NEJM. 2019;381(23).
2. Fernandez MA, Achten J, Parsons N, Griffin XL, Png ME, Gould J, et al. Cemented or Uncemented Hemiarthroplasty for Intracapsular Hip Fracture New England Journal of Medicine. 2022;386(6).
3. LROI. Annual report of Dutch Arthroplasty Register. 2021.

SUMMARY

Chapter 2 evaluated the agreement between self-reported health status (patient gold standard) and proxy assessments using the EQ-5D-5L questionnaire among orthopedic patients. The findings indicate substantial agreement between the patient self-reports and the proxy-patient perspectives, and moderate agreement between the patient and proxy-proxy perspectives for both the Utility score and Visual Analog Scale (VAS) Health. Notably, the Utility scores reported by patients were higher than those provided by proxies, suggesting that proxies generally underestimate the patients' self-rated health status.

This research represents the first empirical comparison of two proxy perspectives within the context of the EQ-5D-5L. A key strength of the study is its design, which mitigated recall bias and ensured that participants were blinded to the study's hypothesis. All subjects were orthopedic patients, reflecting a diverse range of conditions, from sports injuries to hip fractures, which may enhance the representativeness of the sample. However, the findings may not be applicable to populations with cognitive impairments, as the study included only mentally healthy patients. The study also contrasts its results with a previous study utilizing the EQ-5D-3L, which reported higher agreement levels for proxy assessments. The difference in results is attributed to the fewer response categories in the 3L version, which potentially limits disagreement compared to the 5L version.

Interestingly, the agreement in individual domain scores supports existing literature that suggests proxies, particularly relatives, provide more valid assessments for less observable dimensions such as anxiety and pain. Conversely, clinicians tend to assess observable dimensions more effectively. The underestimation of health status by proxies observed in this study aligns with previous findings, but the reasons for this discrepancy remain speculative. Factors such as proxies' mood or personal expectations might influence their assessments.

The implications of these findings are significant for clinical decision-making, as reliance on proxy perspectives could lead to misinformed treatment choices. Moreover, patient overestimation of health status may encourage riskier health behaviors.

In **chapter 3** we investigated the factors influencing the implementation of surgical approaches for hemiarthroplasty in hip fracture management. The study concludes that the training, experience, and attitudes of surgeons are critical factors in determining the surgical approach for hemiarthroplasty. The findings suggest that the outcomes of the APOLLO trial will significantly inform surgeons' choices of surgical approach, although guidelines may not always be strictly adhered to unless accompanied by compelling evidence.

Key themes identified include the importance of broad surgical training and the influence of personal attitudes towards technique adaptation. Respondents noted that while the personal situation of patients is sometimes considered in decision-making, it should be prioritized further, despite logistical challenges related to surgeon availability and experience. Stakeholders other than surgeons, like clinical geriatricians, can impact decision-making indirectly by highlighting the importance of patient-relevant outcomes. Additionally, the Patients Federation can influence healthcare policy, including insurance reimbursement decisions. Strengths of the study include its comprehensive approach, involving multiple stakeholder perspectives and employing a theoretical framework for implementation analysis. However, limitations exist, such as potential biases due to the specific involvement of surgeons in the APOLLO trial and the lack of access to trial results at the time of interviews.

The study recommends a hybrid effectiveness-implementation trial design to assess both clinical outcomes and the feasibility of translating findings into practice. A multifaceted approach, incorporating hands-on learning, education, and dissemination of evidence at conferences, is essential to enhance awareness and ultimately improve surgical practice.

Chapter 4 was the update of a 12-year follow-up of a RCT evaluates functional outcomes in active elderly patients aged ≥70 years with displaced femoral neck fractures treated with hemiarthroplasty (HA) compared to total hip arthroplasty (THA). The findings indicate no significant difference in functional outcomes between the two procedures for patients in good physical and mental condition, and without hip osteoarthritis or rheumatoid arthritis. While there is a trend toward THA for patients with these conditions, conclusive evidence to support this shift was lacking at the time when the study was conducted. A significant concern regarding HA is the potential need for revision due to wear of the acetabular cartilage. Despite theoretical risks, the current study reported no revisions in either group over the 12-year follow-up, consistent with systematic reviews indicating no significant differences in revision rates between HA and THR. The patient population in the study evolved over the follow-up period, with a greater proportion of healthier, younger women, suggesting a selection bias. Notably, there were no dislocations recorded in either group after 12 years, contrasting with earlier findings of increased dislocation rates in THA patients. This raises questions about the relationship between complications like dislocations and long-term mortality, though data on patients who died between follow-ups were not assessed. The study's limitations include potential bias from non-blinded outcome assessments and incomplete recording of complications post-discharge. The sample size was affected by mortality, common in this age group, limiting the generalizability of the findings. In conclusion, after 12 years, HA and THA demonstrate comparable functional outcomes for elderly patients with displaced femoral neck fractures. Given the higher dislocation rates associated with THA observed in prior

follow-ups, the authors recommend HA as the preferred treatment for this demographic, provided there is no significant pre-existing hip pathology.

Chapter 5 presents our systematic review which updates the evidence regarding the two most common surgical approaches—direct lateral approach (DLA) and posterior lateral approach (PLA)—for hemiarthroplasty following femoral neck fractures, emphasizing patient-relevant outcomes. Key findings indicate that all included studies were rated as very low quality according to the GRADE criteria, raising concerns about selection bias and confounding factors, particularly since eleven out of twelve studies were non-randomized. The only randomized controlled trial (RCT) was terminated prematurely due to the technical challenges associated with PLA, which further compromises the reliability of its results. The review also highlights the importance of assessing outcomes like quality of life and ADL dependency, which are often overlooked in existing literature. While secondary outcomes suggested that the DLA group had fewer dislocations one year post-surgery, this finding is tempered by conflicting results from registry studies, which noted an increased risk of reoperation due to dislocation with PLA. The review also mentions that while some studies reported functional advantages for PLA, inconsistencies in results and the presence of biases across studies limit definitive conclusions. The strengths of this review include a thorough critical appraisal using GRADE and a focus on patient-centered outcomes, contrasting with previous reviews that covered a broader range of measures. Limitations include the inability to perform a meta-analysis due to study heterogeneity and the exclusion of some meeting abstracts, which could introduce publication bias. In conclusion, while PLA may be associated with fewer walking problems and lower rates of abductor insufficiency, the evidence does not allow for strong causal inferences between the two approaches regarding their effects on dislocation rates, functional outcomes, and fall risk. Future research should prioritize rigorous randomized clinical trials that focus on outcomes meaningful to patients, particularly those involving cognitive impairment and functional independence.

Chapter 7 describes the results of the RCT alongside a NE investigating the impact of surgical approach on health-related quality of life (HRQoL) in adult patients undergoing cemented hemiarthroplasty following an acute femoral neck fracture. The research finds that the surgical method—direct lateral approach (DLA) versus posterior lateral approach (PLA)—does not significantly influence HRQoL, as measured by the EQ-5D-5L utility score, six months post-surgery.

Secondary outcomes related to function, pain, mobility, and fall risk showed no significant differences between the two groups. However, the PLA group exhibited a statistically significant increase in the risk of prosthesis dislocation and reoperation due to dislocation (5,5% in PLA vs. 0% in DLA), which is clinically relevant.

The study is notable for being the largest randomized controlled trial (RCT) on this topic to date and for its innovative use of a natural experiment (NE) design. This NE, which allocated treatment based on geographical location and timing of the fracture, enhanced participant numbers and generalizability. Additionally, the inclusion of patients with cognitive impairments, often excluded from RCTs, further bolstered the generalizability of the findings.

However, limitations include prevalent missing data from follow-up questionnaires, particularly among patients with dementia, and the absence of a screening log to track eligible patients not randomized. The study did not assess whether the m. piriformis was spared or reattached, highlighting a gap in current high-level evidence regarding its impact.

While previous observational studies suggested a small HRQoL advantage for PLA, this study did not confirm those findings. The increased dislocation risk associated with PLA raises questions about the sensitivity of the EQ-5D-5L to capture the effectiveness of orthopedic interventions. Overall, the findings indicate that PLA does not yield better HRQoL than DLA in this patient population but has a higher risk of dislocation, and the pseudo-randomization approach used in the NE offers a viable alternative to traditional RCT methods.

In **chapter 8** we investigate the influence of surgical approach on the physical performance using the Short Physical Performance Battery (SPPB) score. While no significant difference was found in the (SPPB) total scores between the groups, a clinically relevant difference of 1 point favored the PLA group. Notably, a higher percentage of patients in the PLA group (71%) were classified as good performers compared to 49% in the DLA group. Additionally, the DLA group exhibited significantly higher fear of falling (FoF) at three months postoperatively. The SPPB, which assesses physical performance through gait speed, balance, and strength, indicated that lower gait speed and chair stand performance were more common in the DLA group, which also had more patients unable to return to pre-fracture mobility. Despite these differences, there were no statistically significant impacts on mortality, institutionalization, or adverse events, likely due to the small sample size and short follow-up period of six months.

Previous literature has established associations between higher FoF and poorer functional outcomes, but this study did not find a correlation with fall risk, again potentially limited by sample size. Strengths of the study include comprehensive participation through home visits and its foundation as a pre-planned subgroup analysis of a larger RCT. However, limitations include potential selection bias, reduced patient enrollment due to COVID-19, and a small sample size that may have affected the power of the study. In conclusion, while the study found no statistically significant difference in postoperative

physical performance between PLA and DLA, the PLA group demonstrated a potential clinical advantage. The increased dislocation and reoperation rates in PLA complicate recommendations for either approach. Future research with a larger sample size is necessary to better evaluate the clinical impacts of these surgical methods on postoperative outcomes.

In **chapter 9** we presented the economic evaluation compares the cost-effectiveness of the direct lateral approach (DLA) and the posterior lateral approach (PLA) for cemented hemiarthroplasty in adults with acute femoral neck fractures. The results indicate that DLA is not cost-effective compared to PLA. Both surgical approaches demonstrated no significant differences in quality of life outcomes, aligning with findings from the associated RCT. However, the PLA group incurred higher follow-up surgery costs due to a greater dislocation rate. Although secondary healthcare costs were also higher for PLA, this difference was not statistically significant. The costs related to emergency room consultations and procedural sedation for managing dislocations are not clearly defined, which may lead to an underestimation of PLA costs.

Strengths of the study include its position as the first economic evaluation of an RCT addressing this topic and the inclusion of patients with dementia, enhancing generalizability. Limitations include a significant amount of missing data from self-reported questionnaires, leading to uncertainty in estimates. However, comprehensive documentation of dislocations, reoperations, and emergency admissions reduces bias in those areas.

DUTCH SUMMARY

Hoofdstuk 2 beoordeelt de overeenstemming tussen zelfgerapporteerde gezondheidstoestand (gouden standaard van de patiënt) en proxy-beoordelingen bij orthopedische patiënten, gebruikmakend van de EQ-5D-5L vragenlijst. De bevindingen wijzen op een substantiële overeenstemming tussen de zelfrapportages van patiënten en de perspectieven van proxy's, en op een matige overeenstemming tussen patiënt- en proxy-proxy perspectieven, zowel voor de utiliteitsscore als de Visueel Analoge Schaal (VAS) Gezondheid. Opvallend is dat de door patiënten gerapporteerde utiliteitsscores hoger waren dan die van proxy's, wat suggereert dat proxy's de door patiënten zelfgerapporteerde gezondheid doorgaans onderschatten. Dit onderzoek is de eerste empirische vergelijking van twee proxy-perspectieven in de context van de EQ-5D-5L. Een belangrijk voordeel van de studie is het ontwerp, dat recall-bias minimaliseerde en ervoor zorgde dat deelnemers geblindeerd waren voor de hypothese van de studie. Alle proefpersonen waren orthopedische patiënten, met een breed scala aan aandoeningen, van sportblessures tot heupfracturen, wat de representativiteit van de steekproef kan verhogen. De resultaten zijn echter mogelijk niet van toepassing op populaties met cognitieve beperkingen, omdat de studie alleen mentaal gezonde patiënten omvatte. De studie vergelijkt ook de resultaten met een eerdere studie die de EQ-5D-3L gebruikte, waarin hogere overeenstemmingsniveaus voor proxy-beoordelingen werden gerapporteerd. Het verschil in resultaten wordt toegeschreven aan de beperktere antwoordcategorieën in de 3L-versie, wat mogelijk meningsverschillen beperkt in vergelijking met de 5L-versie.

Opmerkelijk is dat de overeenstemming in individuele domeinscores de bestaande literatuur ondersteunt, die suggereert dat met name familieleden meer valide beoordelingen geven voor minder waarneembare dimensies zoals angst en pijn. Omgekeerd beoordelen clinici eerder waarneembare dimensies nauwkeuriger. De in deze studie waargenomen onderschatting van de gezondheidstoestand door proxy's komt overeen met eerdere bevindingen, maar de redenen voor deze discrepantie blijven speculatief. Factoren zoals de stemming of persoonlijke verwachtingen van proxy's kunnen hun beoordelingen beïnvloeden.

De implicaties van deze bevindingen zijn aanzienlijk voor klinische besluitvorming, aangezien vertrouwen op proxy-perspectieven kan leiden tot verkeerde behandelingskeuzes. Daarnaast kan de neiging van patiënten om hun gezondheid te overschatten risicovoller gezondheidsgedrag bevorderen.

In **hoofdstuk 3** onderzochten we de factoren die van invloed zijn op de implementatie van chirurgische benaderingen voor hemiartroplastiek bij heupfracturen. De studie concludeert dat de opleiding, ervaring en attitudes van chirurgen cruciale factoren zijn bij het bepalen van de chirurgische benadering voor hemiartroplastiek. De resultaten

suggereren dat de uitkomsten van de APOLLO-trial een significante rol zullen spelen in de keuze van de chirurgische benadering, hoewel richtlijnen niet altijd strikt worden gevolgd zonder overtuigend bewijs. Belangrijke thema's zijn onder andere het belang van brede chirurgische training en de invloed van persoonlijke attitudes ten aanzien van techniekadaptatie. Respondenten merkten op dat de persoonlijke situatie van patiënten soms wordt overwogen in de besluitvorming, maar dat dit meer prioriteit zou moeten krijgen ondanks logistieke uitdagingen zoals de beschikbaarheid en ervaring van chirurgen. Andere belanghebbenden, zoals klinisch geriaters, kunnen indirect invloed uitoefenen door het belang van patiëntrelevante uitkomsten te benadrukken. Daarnaast kan de Patiëntenfederatie invloed uitoefenen op het zorgbeleid, inclusief beslissingen over vergoedingen. Sterke punten van de studie zijn de uitgebreide benadering, waarbij meerdere perspectieven van belanghebbenden zijn betrokken en een theoretisch raamwerk is toegepast voor implementatieanalyse. Beperkingen zijn onder andere mogelijke biases vanwege de specifieke betrokkenheid van chirurgen bij de APOLLO-trial en het gebrek aan toegang tot de resultaten van de trial tijdens de interviews.

De studie beveelt een hybride effectiviteits-implementatie onderzoeksopzet aan om zowel de klinische resultaten als de uitvoerbaarheid van het vertalen van bevindingen naar de praktijk te beoordelen. Een veelzijdige aanpak, inclusief praktijkgerichte scholing, educatie en de verspreiding van bewijs op congressen, is essentieel om het bewustzijn te vergroten en uiteindelijk de chirurgische praktijk te verbeteren.

Hoofdstuk 4 betreft de update van een 12-jarige follow-up van een RCT waarin functionele uitkomsten bij actieve oudere patiënten (≥70 jaar) met verplaatste femurhalsfracturen behandeld met hemiartroplastiek (HA) werden vergeleken met totale heupartroplastiek (THA). De bevindingen laten geen significant verschil zien in functionele uitkomsten tussen de twee procedures voor patiënten in goede fysieke en mentale conditie, zonder heupartrose of reumatoïde artritis. Er is een trend waarneembaar naar het gebruik van totale heupprothese (THP) voor patiënten met specifieke aandoeningen, maar overtuigend bewijs om deze verschuiving te ondersteunen ontbrak op het moment van de studie. Een belangrijk aandachtspunt bij hemiarthroplastiek (HA) is de potentiële noodzaak voor revisie als gevolg van slijtage van het acetabulaire kraakbeen. Ondanks theoretische risico's rapporteerde de huidige studie geen revisies in beide groepen gedurende de 12-jarige follow-up, wat consistent is met systematische reviews die geen significante verschillen in revisiepercentages tussen HA en THP aantoonden. De patiëntenpopulatie evolueerde tijdens de follow-up, met een toename van jongere, gezondere vrouwen, wat op een selectiebias kan wijzen. Opmerkelijk is dat er na 12 jaar geen dislocaties zijn geregistreerd in beide groepen, in contrast met eerdere bevindingen die een verhoogd dislocatierisico voor THP-patiënten rapporteerden. Dit roept vragen op over de relatie tussen complicaties, zoals dislocaties, en lange termijn mortaliteit, hoewel gegevens over overleden patiënten tussen follow-upmomenten niet zijn beoordeeld. Beperkingen van de studie

zijn onder meer mogelijke bias door niet-geblindeerde uitkomstbeoordelingen en onvolledige registratie van complicaties na ontslag. De steekproefgrootte werd beïnvloed door mortaliteit, wat gebruikelijk is bij deze leeftijdsgroep, en beperkt de generaliseerbaarheid van de bevindingen. Concluderend tonen HA en THP na 12 jaar vergelijkbare functionele uitkomsten voor ouderen met dislocaties van de femurhals. Gezien het hogere dislocatierisico bij THP, zoals waargenomen in eerdere follow-ups, bevelen de auteurs HA aan als de voorkeurstherapie voor deze demografie, mits er geen significante vooraf bestaande heupziekte is.

Hoofdstuk 5: In dit hoofdstuk wordt ons systematisch literatuuronderzoek gepresenteerd waarin het bewijs wordt geactualiseerd over de twee meest gebruikte chirurgische benaderingen—de directe laterale benadering (DLA) en de posterolaterale benadering (PLA)—voor hemiarthroplastiek bij femurhalsfracturen, met een nadruk op patiëntrelevante uitkomsten. Belangrijke bevindingen tonen aan dat alle geïncludeerde studies volgens de GRADE-criteria als zeer laag werden beoordeeld, wat bezorgdheid opwerpt over selectiebias en confounders, aangezien elf van de twaalf studies niet-gerandomiseerd waren. De enige gerandomiseerde gecontroleerde trial (RCT) werd vroegtijdig stopgezet vanwege technische uitdagingen geassocieerd met PLA, wat de betrouwbaarheid van de resultaten verder ondermijnt. Het onderzoek benadrukt ook het belang van het beoordelen van uitkomsten zoals kwaliteit van leven en afhankelijkheid in ADL, die vaak over het hoofd worden gezien in de bestaande literatuur. Hoewel secundaire uitkomsten suggereren dat de DLA-groep minder dislocaties had na één jaar, wordt deze bevinding getemperd door tegenstrijdige resultaten van registerstudies, die een verhoogd risico op reoperatie wegens dislocatie bij PLA rapporteerden. De beoordeling noemt dat, hoewel sommige studies functionele voordelen voor PLA meldden, inconsistenties en biases beperkingen opleveren voor definitieve conclusies. De sterke punten van deze review omvatten een grondige kritische beoordeling met GRADE en een focus op patiëntgerichte uitkomsten, in tegenstelling tot eerdere reviews die een breder scala aan metingen dekken. Beperkingen zijn onder meer de onmogelijkheid om een meta-analyse uit te voeren vanwege heterogeniteit tussen studies en de uitsluiting van sommige congresabstracts, wat publicatiebias kan introduceren. Concluderend kan PLA geassocieerd worden met minder loopproblemen en lagere percentages abductortekort, maar het bewijs staat geen sterke causale inferenties toe tussen de twee benaderingen met betrekking tot dislocaties, functionele uitkomsten, en valrisico. Toekomstig onderzoek moet zich richten op robuuste gerandomiseerde klinische trials met focus op uitkomsten die belangrijk zijn voor patiënten, met name voor patiënten met cognitieve stoornissen en functionele onafhankelijkheid.

Hoofdstuk 7: In dit hoofdstuk worden de resultaten van de RCT beschreven, samen met een natuurlijke experiment (NE) dat de invloed van de chirurgische benadering op de kwaliteit van leven gerelateerd aan gezondheid (HRQoL) bij volwassen patiënten on-

dezoekt na gecementeerde hemiarthroplastiek bij een acute femurhalsfractuur. Uit het onderzoek blijkt dat de chirurgische methode—directe laterale benadering (DLA) versus posterolaterale benadering (PLA)—geen significante invloed heeft op HRQoL, gemeten aan de hand van de EQ-5D-5L utility score, zes maanden na de operatie. Secundaire uitkomsten met betrekking tot functie, pijn, mobiliteit en valrisico vertoonden geen significante verschillen tussen de twee groepen. De PLA-groep vertoonde echter een statistisch significant verhoogd risico op prothesedislocatie en reoperatie vanwege dislocatie (5,5% bij PLA vs. 0% bij DLA), wat klinisch relevant is. Het onderzoek is opvallend omdat het de grootste RCT op dit gebied is en vanwege het innovatieve gebruik van een NE-ontwerp. Dit NE, dat behandeling toekende op basis van geografische locatie en tijdstip van de fractuur, vergrootte het aantal deelnemers en de generaliseerbaarheid. Daarnaast versterkte de inclusie van patiënten met cognitieve stoornissen, vaak uitgesloten van RCT's, de generaliseerbaarheid van de bevindingen. Beperkingen omvatten ontbrekende gegevens uit follow-up vragenlijsten, vooral bij patiënten met dementie, en het ontbreken van een screeningslog om in aanmerking komende niet-gerandomiseerde patiënten bij te houden. De studie onderzocht niet of de m. piriformis gespaard of opnieuw bevestigd werd, wat een leemte vormt in het huidige hoogstaande bewijs met betrekking tot de impact hiervan. Eerdere observationele studies suggereerden een klein HRQoL-voordeel voor PLA, maar deze studie bevestigde die bevindingen niet. Het verhoogde dislocatierisico bij PLA roept vragen op over de gevoeligheid van de EQ-5D-5L om de effectiviteit van orthopedische interventies te meten. Samenvattend geven de bevindingen aan dat PLA geen betere HRQoL oplevert dan DLA in deze patiëntenpopulatie maar een hoger risico op dislocatie heeft, en de pseudo-randomisatiemethode gebruikt in de NE biedt een levensvatbaar alternatief voor traditionele RCT-methoden.

Hoofdstuk 8: In dit hoofdstuk wordt onderzocht hoe de chirurgische benadering de fysieke prestaties beïnvloedt, beoordeeld met de Short Physical Performance Battery (SPPB)-score. Hoewel er geen significant verschil werd gevonden in de totale SPPB-scores tussen de groepen, was er een klinisch relevant verschil van 1 punt in het voordeel van de PLA-groep. Opmerkelijk was dat een hoger percentage patiënten in de PLA-groep (71%) werd geclassificeerd als goede performers vergeleken met 49% in de DLA-groep. Daarnaast vertoonde de DLA-groep een significant hogere angst voor vallen (FoF) drie maanden postoperatief. De SPPB, die fysieke prestaties beoordeelt door loopsnelheid, balans en kracht, toonde aan dat lagere loopsnelheid en stoel-opstaan-prestaties vaker voorkwamen in de DLA-groep, die ook meer patiënten omvatte die niet konden terugkeren naar pre-fractuurmobiliteit. Ondanks deze verschillen waren er geen statistisch significante effecten op mortaliteit, institutionalisering of bijwerkingen, waarschijnlijk door de kleine steekproefgrootte en de korte follow-up periode van zes maanden. Eerdere literatuur heeft verbanden aangetoond tussen een hogere angst voor vallen (FoF) en slechtere functionele uitkomsten, maar deze studie vond geen correlatie met valrisico, wat mogelijk te wijten is aan de beperkte steekproefgrootte. Sterke punten

van de studie zijn onder andere de uitgebreide participatie door huisbezoeken en de opzet als een vooraf geplande subgroepanalyse van een grotere RCT. Beperkingen zijn echter onder meer een mogelijke selectiebias, verminderde inclusie van patiënten als gevolg van COVID-19 en een kleine steekproefgrootte, wat de kracht van de studie kan hebben beïnvloed. Concluderend vond de studie geen statistisch significant verschil in postoperatieve fysieke prestaties tussen PLA en DLA, maar de PLA-groep toonde een mogelijk klinisch voordeel. De verhoogde dislocatie- en reoperatiepercentages bij PLA bemoeilijken aanbevelingen voor een van beide benaderingen. Toekomstig onderzoek met een grotere steekproefomvang is noodzakelijk om de klinische effecten van deze chirurgische methoden op postoperatieve uitkomsten beter te kunnen beoordelen.

In **hoofdstuk 9** hebben we de economische evaluatie gepresenteerd waarin de kosten-effectiviteit van de directe laterale benadering (DLA) en de posterolaterale benadering (PLA) voor gecementeerde hemiarthroplastiek bij volwassenen met acute femurhals-fracturen wordt vergeleken. De resultaten wijzen uit dat DLA niet kosteneffectief is in vergelijking met PLA. Beide chirurgische benaderingen toonden geen significante verschillen in kwaliteit van leven, in lijn met de bevindingen van de bijbehorende RCT. De PLA-groep maakte echter hogere kosten voor vervolgoperaties door een verhoogd dislocatierisico. Hoewel de secundaire gezondheidszorgkosten voor PLA ook hoger waren, was dit verschil niet statistisch significant. De kosten met betrekking tot spoedeisende hulpconsultaties en procedurele sedatie voor het behandelen van dislocaties zijn niet duidelijk gedefinieerd, wat kan leiden tot een onderschatting van de kosten van PLA.

Sterke punten van deze studie zijn onder meer dat het de eerste economische evaluatie is van een RCT die dit onderwerp behandelt en dat patiënten met dementie zijn opgenomen, wat de generaliseerbaarheid vergroot. Beperkingen zijn een aanzienlijke hoeveelheid ontbrekende gegevens uit zelfgerapporteerde vragenlijsten, wat leidt tot onzekerheid in de schattingen. Echter, een uitgebreide documentatie van dislocaties, reoperaties en spoedopnames vermindert bias in deze gebieden.

Name PhD student: Maria C.J.M. Tol
PhD period: 1-12-2017 – 1-7-2026

Names of PhD supervisor(s) & co-supervisor(s):
prof. dr. R.W. Poolman
prof. dr. J.C. Goslings
dr. N.W. Willigenburg
dr. H.C. Willems

1. PhD training

	Year	ECTS
General courses		
- Research Data Management Course (AMC, Graduate School)	2018	0.9
- GCP certificate	2018	1
- Basic Methods and Reasoning in Biostatistics (LUMC)	2022	1.5
- Scientific Conduct for PhDs (LUMC)	2022	0.5
- PhD meeting – Taking charge (LUMC)	2025	0.5
Presentations		
Diverse More than 30 educational and introductory lectures on the surgical treatment of proximal femur fractures and the APOLLO trial		2.1
NOV/NOF meeting 2024, Rotterdam ACTA session, one of the speakers: How to avoid rejections		0.6
Traumadagen, 2023, Amsterdam The results of the APOLLO trial: Posterolateral or Direct Lateral Surgical Approach or Hemiarthroplasty after a Hip Fracture (oral)		0.6
OTA 2023, Seattle, United States The results of the APOLLO trial: Posterolateral or Direct Lateral Surgical Approach or Hemiarthroplasty after a Hip Fracture (oral)		0.6
NOV Annual meeting, 2023, Utrecht RCT versus Natural Experiment (oral) The results of the APOLLO trial: Posterolateral or Direct Lateral Surgical Approach or Hemiarthroplasty after a Hip Fracture (oral)		0.6
EFORT 2022, Lisbon, Portugal MCJM Tol, A.J. Rasker, H.C. Willems, N.W. Willigenburg, T. Gosens, M.J. Heetveld, R.W. Poolman, on behalf of the APOLLO trial research group. "Is the physical performance of patients influenced by the surgical approach of hemiarthroplasty after a femoral neck fracture?" (oral)		0.6
NOV Annual meeting 2022, Utrecht MCJM Tol, A.J. Rasker, H.C. Willems, N.W. Willigenburg, T. Gosens, M.J. Heetveld, R.W. Poolman, on behalf of the APOLLO trial research group. "Is the physical performance of patients influenced by the surgical approach of hemiarthroplasty after a femoral neck fracture?" (oral).		0.6
EFORT 2019, Lisbon, Portugal M.C.J.M. Tol, J.P. Kuipers, N.W. Willigenburg, H.C. Willems, R.W. Poolman, MD, PhD. "Level of agreement between the self-completed EQ-5D-5L and two proxy perspectives in an orthopaedic population, A RANDOMISED AGREEMENT STUDY" (oral)		0.6
EHS congress 2018, The Hague M.C.J.M. Tol, T. Gosens, H.C. Willems, M.J. Heetveld, R.W. Poolman. "Surgical Approach of Hemiarthroplasty after Femoral Neck Fracture: Posterolateral or Direct Lateral, a Systematic Review" (oral)		0.6
Traumaplatform Young Generation Symposium 2018, Schiphol M.C.J.M. Tol, T. Gosens, H.C. Willems, M.J. Heetveld, R.W. Poolman. "Surgical Approach of Hemiarthroplasty after Femoral Neck Fracture: Posterolateral or Direct Lateral, a Systematic Review" (oral)		

NOV Autumn meeting, 2017, Den Bosch M.C.J.M. Tol, T. Gosens, H.C. Willems, M.J. Heetveld, R.W. Poolman. "Introduction of a RCT: Surgical Approach of Hemiarthroplasty after Femoral Neck Fracture: Posterolateral or Direct Lateral (APOLLO trial)" (oral)		0.6
AAOS Annual meeting 2017, San Diego, United States M.C.J.M. Tol, M.P.J. van den Bekerom, I.N. Sierevelt, E.F. Hilverdink, E.L.F.B. Raaymakers, J.C. Goslings. "Hemiarthroplasty or Total Hip Arthroplasty for Femoral Neck Fractures: 12-year Follow-up of Randomised Trial" (poster/oral)		0.6
Traumadagen 2016, Amsterdam M.C.J.M. Tol, M.P.J. van den Bekerom, I.N. Sierevelt, E.F. Hilverdink, E.L.F.B. Raaymakers, J.C. Goslings. "Kophals prothese of totale heupprothese bij mediale collumfracturen, 12-jaar follow-up van een gerandomiseerde trial" (oral)		0.6
M.C.J.M. Tol, M.P.J. van den Bekerom, I.N. Sierevelt, E.F. Hilverdink, E.L.F.B. Raaymakers, J.C. Goslings "Voorspellers van mortaliteit in de behandeling van mediale collumfracturen" (poster)		0.6
NOV Annual meeting 2016, Den Bosch M.C.J.M. Tol, M.P.J. van den Bekerom, I.N. Sierevelt, E.F. Hilverdink, E.L.F.B. Raaymakers, J.C. Goslings. "Kophals prothese of totale heupprothese bij mediale collumfracturen, 12-jaar follow-up van een gerandomiseerde trial" (oral)		0.6
(Inter)national conferences		
- NOV Annual meeting 2016, 2017, 2018, 2019, 2021, 2022, 2023, 2024, 2025		2.6
- Traumadagen 2016, 2017, 2019, 2023		1.1
- AAOS Annual meeting, San Diego, United States 2017		0.5
- 13th Congress of the European Hip Society, The Hague 2018		0.5
- EFORT Annual meeting, Lisbon, Portugal 2019, 2022		1
- OTA Annual meeting, Seattle, United States 2023		0.5
- NVA Annual meeting, Noordwijk 2025		0.5

2. Teaching		
	Year	**ECTC**
Lecturing		
- Orthopedic course for Theatre nurses	2021, 2022 2025	1.5
- Orthopedic course for summa master students UMCU	2023	1.5
Tutoring, Mentoring		
- Mentoring residents not in training	2020-2025	1
- Tutor research internship master in medicine of Jurrian Kuipers	2018	1.5

3. Parameters of Esteem	
	Year
Grants	
- **The Netherlands Organisation for Health Research and Development, ZonMw Grant – Efficiency Research (€324.697,-)** Surgical Approach of Hemiarthroplasty after Femroal Neck Fracture: Posterolateral or Direct Lateral (APOLLO trial)	2017
- **The Netherlands Organisation for Health Research and Development, ZonMw Grant – Inclusion accelerator (€49.690,-)** Surgical Approach of Hemiarthroplasty after Femroal Neck Fracture: Posterolateral or Direct Lateral (APOLLO trial)	2022
Awards and Prizes	
- Dr. Keeman Award, Best Publication of 2024, OLVG 2025 – Won (€3.000,-)	2025
- Science and Innovation Award, Dutch Federation of Medical Specialists 2025 – Nominated	2025

4. Publications

	Year
Peer reviewed	
J.L. Esser, M.C.J.M. Tol, J.M. van Dongen,, N.W. Willigenburg, T. Gosens, H.C. Willems, M.J. Heetveld, R.W. Poolman. Cost-effectiveness of Posterolateral versus Direct Lateral Approach of Hemiarthroplasty after Femoral Neck Fracture: a randomised controlled trial. *Acta Orthop*. 2025 Dec 18:96:914-919. doi:10.2340/17453674.2025.45056.	2025
Maria C.J.M. Tol; Nienke W. Willigenburg; Hanna C. Willems; Taco Gosens; Ariena Rasker; Martin J. Heetveld; Martijn G.M. Schotanus; Johanna M van Dongen; B. Eggen; M. Kormos; S.L. van der Pas; A.W. van der Vaart; Rudolf W. Poolman; on behalf of the APOLLO research group. Posterolateral or Direct Lateral Surgical Approach for Hemiarthroplasty After a Hip Fracture - A Randomized Clinical Trial Alongside a Natural Experiment. *JAMA Netw Open.* 2024;7(1):e2350765. doi:10.1001/jamanetworkopen.2023.50765	2024
Maria C.J.M. Tol; Nienke W. Willigenburg; Hanna C. Willems; Taco Gosens; Ariena Rasker; Martin J. Heetveld; Martijn G.M. Schotanus; Johanna M van Dongen; B. Eggen; M. Kormos; S.L. van der Pas; A.W. van der Vaart; Rudolf W. Poolman; on behalf of the APOLLO research group. Posterolateral or Direct Lateral approach of cemented Hemiarthroplasty after Femoral Neck Fracture (APOLLO): protocol for a multicenter Randomized Controlled Trial with economic evaluation and Natural Experiment alongside. *Acta Orthop.* 2022 Sep 12:93:732-738. doi: 10.2340/17453674.2022.4547.	2022
M.C.J.M. Tol, J.P. Kuipers, N.W. Willigenburg, H.C. Willems, R.W. Poolman, Level of agreement between the self-completed EQ-5D-5L and two proxy perspectives in an orthopaedic population, A Randomised agreement study. *Health Qual Life Outcomes*. 2021 Jan 27;19(1):35 doi: 10.1186/s12955-021-01679-y.	2021
M.C.J.M. Tol, T. Gosens, H.C. Willems, M.J. Heetveld, R.W. Poolman. Surgical Approach of Hemiarthroplasty after Femoral Neck Fracture: Posterolateral or Direct Lateral, a Systematic Review. *Hip Int.* 2021 Mar;31(2):154-165. doi: 10.1177/1120700020931766.	2021
M.C.J.M. Tol, M.P.J. van den Bekerom, I.N. Sierevelt, E.F. Hilverdink, E.L.F.B. Raaymakers, J.C. Goslings. Hemiarthroplasty or Total Hip Arthroplasty for Femoral Neck Fractures: 12-year Follow-up of Randomised Trial. *Bone Joint J.* 2017 Feb;99-B(2):250-254 doi: 10.1302/0301-620X.99B2.BJJ-2016-0479.R1.	2017
K. Treskes, S.C. Voeten* M.C.J.M. Tol*, W.P. Zuidema, J. Vermeulen, J.C. Goslings, N.W.L Schep; on behalvf of Study group on certification of trauma proximal femoral fractures. Trauma surgery by general surgeons: Still an option for proximal femoral fractures? (*equally contributed to this article) *Injury.* 2017 Feb;48(2):339-344 doi: 10.1016/j.injury.2016.11.020.	2017

ABOUT THE AUTHOR

Maria Catharina Johanna Margaretha (Mathilde) Tol was born on the 14th of April 1990 in Volendam. After completing her secondary school at Don Bosco College in Volendam, she went on to study medicine at the University of Amsterdam.

Mathilde obtained her Master of Medicine degree in 2017 at the University of Amsterdam, where she completed internships in General Surgery and Orthopaedics. Her master's thesis, focusing on trauma surgery outcomes, reflected an early commitment to evidence-based surgical care. Concurrently, she began research activities within the Trauma Surgery Unit at the Academic Medical Center, laying the foundation for her PhD trajectory.

In 2017, Mathilde formally commenced her PhD at the Department of Orthopaedics at OLVG in Amsterdam, under the supervision of Prof. Dr. R.W. Poolman and Prof. Dr. J.C. Goslings. Her doctoral research focused on surgical approaches for hemiarthroplasty following femoral neck fractures, culminating in the APOLLO trial. Her work contributed to international literature and has been presented at numerous national and international conferences, including the Orthopaedic Trauma Association (OTA), EFORT, and the American Academy of Orthopaedic Surgeons (AAOS).

In parallel with her academic work, Mathilde pursued her residency in Orthopaedic surgery within the ROGO Midden-West training program, with clinical rotations at OLVG Amsterdam, UMC Utrecht, and the St. Antonius Hospital. She is expected to complete this residency in July 2026. She received multiple grants, including two ZonMw research awards, and was nominated for the Federation of Medical Specialists Science and Innovation Prize in 2025 and won the Dr. Keeman Publication Award.

Following the completion of her residency, Mathilde will further specialize in lower limb surgery. In 2026, she will begin a combined clinical and research fellowship in lower limb sports knee surgery in Perth, Australia. This fellowship will allow her to deepen her expertise in complex knee pathology and sports-related procedures, while continuing to contribute to scientific research in Orthopaedics. Mathilde lives in Amsterdam with her partner, Jetse, and their son, Karel.

ACKNOWLEDGEMENTS

Allereerst dank aan alle patiënten die deel wilden nemen aan ons onderzoek en aan alle arts-assistenten, PA/VS en chirurgen die deze patiënten wilden includeren en opereren. Zonder jullie geen proefschrift.

Professor Poolman, beste Rudolf, het was een voorrecht om jouw promovenda te zijn. Je gaf me de vrijheid om mijn eigen tempo te bepalen en stuurde me in de juiste richting om zelf tot inzichten te komen. Je creatieve en innovatieve blik op de wetenschap werkt inspirerend, en als mentor brachten je filosofische lessen alles weer in perspectief.

Professor Goslings, beste Carel, als student onder jouw begeleiding is het allemaal begonnen. Je was de katalysator voor de eindsprint en hebt ervoor gezorgd dat we als compleet team de eindstreep konden halen. Dank voor de begeleiding als promotor en als arts.

Nienke Willigenburg, dank voor je scherpe analyses en de 'never settle for less' mentaliteit. Jouw kritische blik hield mij scherp en bracht dit proefschrift naar een hoger niveau.

Hanna Willems, ik heb veel van je geleerd om vanuit een andere invalshoek naar de zorg te kijken. Je nam altijd ruim de tijd om te sparren en nieuwe ideeën uit te denken, dat waardeer ik enorm.

Prof. Taco Gosens, Martijn Schotanus, Martin Heetveld, jullie hebben een aanzienlijke bijdrage geleverd aan de totstandkoming van dit proefschrift. Dank voor alle begeleiding en het was een genoegen om het zuiden te bezoeken tijdens de APOLLO trial.

Prof. Michel van den Bekerom, dank voor het begeleiden van de eerste stappen in de wetenschap en voor het connecten met de juiste mensen, waardoor mijn eerste artikel kon uitgroeien tot dit promotietraject.

Prof. dr. C.N. van Dijk, prof. dr. P. Kloen, prof. dr. N. van der Velde, prof. dr. I.B. Schipper, dr. M.G.J. Gademan en dr. H.A. Formijne Jonkers, hartelijk dank voor het beoordelen van dit proefschrift en uw bereidheid om zitting te nemen in mijn promotiecommissie.

Pim, wat is het bijzonder om met je beste vriendin dezelfde passie te delen en dat ook nog eens tegelijkertijd te doen. Dank voor de talloze uren van ventileren en optimisme. Je weet me altijd weer de juiste richting op te laten kijken. Op naar Tol & van Dijk!

Shanna, lieve zus, ik kan bij jou altijd terecht. Zorgzaam, scherp en als ik het even niet meer wist, zorgde jij dat ik er weer vertrouwen in had. Bouke, dank voor het aanhoren van al onze verhalen terwijl je ons in de watten legde met je briljante kookkunsten.

Ariena, SPSS queen, zo fijn om met jou samen te werken. Dirk, dank voor de mentale support en het sparren over hoe je naast de opleiding ook een RCT runt.

Opleiders, dank voor de fantastische opleiding die ik in jullie ROGO heb mogen volgen.

Alle vrienden uit Amsterdam en Volendam, jullie hebben precies niets bijgedragen aan dit proefschrift, maar zonder jullie had dit proefschrift er niet geweest. In de tijden van hard werken was juist de ontspanning des te belangrijker. Dank voor alle gezelligheid en ik ben blij dat jullie nooit teleurgesteld waren als ik ergens niet bij kon zijn.

Ootje, helaas al een tijdje niet meer onder ons, maar jouw wijsheid stuurt mij nog steeds.

Leon en Anna, altijd even enthousiast en geïnteresseerd, dank voor jullie steun.

Ouders, het begin van de weg hiernaartoe was verre van vanzelfsprekend. Zonder jullie onvoorwaardelijke support had ik hier nooit gestaan en daarvoor ben ik jullie voor altijd dankbaar.

Jetse, de afgelopen maanden waren best uitdagend met alles wat tegelijkertijd op m'n pad kwam, maar dankzij jou is het ons gelukt. Ik kan altijd op jou aan, je bent m'n veilige haven. Ik kan niet wachten op het avontuur met ons kleine mannetje in Australië.