Challenges and opportunities in trauma research: study designs and patient-reported outcome measures

Ochen, Y.

CHALLENGES AND OPPORTUNITIES IN TRAUMA RESEARCH

STUDY DESIGNS AND PATIENT-REPORTED OUTCOME MEASURES

YASSINE OCHEN

CHALLENGES AND OPPORTUNITIES IN TRAUMA RESEARCH

STUDY DESIGNS AND PATIENT-REPORTED OUTCOME MEASURES

YASSINE OCHEN

Challenges and opportunities in trauma research
Study designs and patient-reported outcome measures

PhD thesis, Leiden University, the Netherlands

ISBN: 9789464191523
Printed by: Gildeprint, Enschede, the Netherlands

The research presented in this thesis was financially supported by the Stichting Fundatie van de Vrijvrouwe van Renswoude, Traumazorgnetwerk Midden-Nederland, and Prins Bernhard Cultuurfonds.

Printing of this thesis was financially supported by the St. Antonius Hospital, Nieuwegein, the Netherlands.

CHALLENGES AND OPPORTUNITIES IN TRAUMA RESEARCH

STUDY DESIGNS AND PATIENT-REPORTED OUTCOME MEASURES

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op
woensdag 31 maart 2021
klokke 11:15 uur

door

YASSINE OCHEN

geboren op 6 december 1990
te Utrecht

Promotor
Prof. dr. R.H.H. Groenwold

Copromotoren
Dr. R.M. Houwert (Universitair Medisch Centrum Utrecht)
Dr. D. van der Velde (St. Antonius Ziekenhuis Nieuwegein)

Leden promotiecommissie
Prof. dr. I.B. Schipper
Prof. dr. O.M. Dekkers
Prof. dr. Y. van der Graaf (Universitair Medisch Centrum Utrecht)
Prof. dr. F.C. Öner (Universitair Medisch Centrum Utrecht)

Voor mijn oma, Aïcha

TABLE OF CONTENTS

CHAPTER 1

General introduction and outline of this thesis

Challenges in trauma research

Evidence-based medicine, using the best available research evidence to guide clinical decision making, offers both opportunities and challenges. In the field of orthopedic trauma surgery, great progress has been made in the understanding of fracture healing, immune response, timing of interventions, and development of implants.[1] However, musculoskeletal injuries are still a major global health problem, contributing a large burden of disability and suffering.[2] Although injury-related mortality rates have declined over the last decades in many high-income countries, in 2013 injuries still accounted for 10% of the global burden of disease, with a loss of 248 million disability-adjusted life-years.[3]

Improved patient care has enhanced the likelihood of surviving serious injury. As a result, focus has shifted to improving quality of survival and reducing the burden of nonfatal injury.[3] However, research-based advances that improve these outcomes for patients with major orthopedic injuries have been constrained by two important factors; inadequate high-quality studies in the field of orthopedic trauma surgery, and insufficient attention to patient-reported outcome measures (PROMs).[1]

Study designs in trauma research

The first challenge is the lack of high-quality studies of surgical and non-surgical interventions in orthopedic trauma patients. Although randomized controlled trials (RCTs) are considered the highest level of evidence, this design might not always be ethical, feasible, or necessary to address a specific surgical research question. These challenges are especially apparent in the surgical field, with acute and urgent life-threatening situations such as trauma surgery. In this field variation in surgical practice can lead to practical and methodological difficulties in terms of patient recruitment and randomization.[4] In fact, a literature review revealed 21% of surgical RCTs are discontinued and 33% of trials remain unpublished after a median of 4.9 years. These challenges have led to a growing debate on the need of RCTs for the evaluation of surgical interventions, and whether well-designed observational studies might complement and add valuable information to results from RCTs.[5,6]

Importantly, more than 80% of orthopedic surgical trials are methodologically limited by small sample sizes, inadequate allocation of concealment, and no independent assessment of outcomes. The small sample sizes and low quality of studies have restricted the translation of studies to

routine patient care.[1] Furthermore, RCTs can have strong internal validity, however, their external validity (generalizability of results) tends to be low. The patient populations encountered in daily clinical practice can differ from the often highly selected patient populations enrolled in RCTs.[7] Moreover, RCTs might not have sufficient follow-up or sample size to assess infrequent outcome measures or long-term treatment effects.[8,9] Finally, RCTs are accompanied by the need for substantial resources and are therefore not always justified for questions regarding small modifications to a treatment or technique, which are typical for the surgical field.[10] Given the limited added value of surgical trials (no concealment of allocation and blinding of outcome assessor) and need for substantial recourses, observational studies may provide an alternative to assess effects of surgical and non-surgical interventions in orthopedic trauma patients, provided such studies are of sufficient quality.[4] The potential added value of observational studies might differ between different comparisons and research questions.[4] In a health care system with growing financial burden, the relative low cost and feasibility underline the possible added value of observational studies.

Meta-analyses are valuable tools for the assessment of differences in treatment effects. Currently, both RCTs and observational studies are increasingly used in orthopedic trauma meta-analyses for the evaluation of treatment effects. Provided that observational studies are of high quality, adding information from observational studies to meta-analyses could increases sample size, which could enable the evaluation of small treatment effects and infrequent outcome measures. Furthermore, observational studies might provide insight into a variety of populations, subgroups, and long term effects, therefore having a role in improving the value and best available evidence in orthopedic trauma care.[4,6,11]

Patient-reported outcome measures

Over the last decades, focus has shifted from a volume to a value-based health care system. In the value-based system, achieving high value for patients must become the goal of health care, with value defined as patient relevant health outcomes, relative to the costs of medical care. The relevant health outcomes, which define the value in the equation, are inherently condition-specific and multidimensional.[12] With the shifting focus towards a value-based health care system there is also a shift from physician-reported to patient-reported outcome measures (PROMs). Moreover, while health care evolves from volume to value, there is increasing interest by payers to use patient-reported outcomes to determine value and more specifically, quality from the

patient's perspective.[13] Standardized outcome measures and routine collection of PROMs are needed to monitor and assess present and new treatment approaches and to support clinical trials and evidence-based care.[1]

In the past, research primarily focused on clinical and radiological outcomes and by doing so overlooked the quality of life of surviving patients.[1] Modern trauma care systems consider the recovery of patients from injury through prehospital care, acute care, and rehabilitation. However, the understanding of the degree of recovery, the time needed, and the proportion of the injured population who will experience lifelong disability is limited. Outcome studies are mainly single center, include small samples, and/or are limited to a single moment of short-term follow-up. Multicenter studies that evaluate multiple time points after injury, important for establishing the long-term burden of injury to provide information about prognosis, and guiding treatment decisions are lacking.[3]

Reports suggest that the use of PROMs has rapidly increased over the last decades and it has been noted that this trend will continue. However, despite the advances and use in routine care, there are still substantial challenges regarding implementation and standardization of PROMs. These challenges include the reliability and precision of the instruments used to capture the outcome of interest. Previous orthopedic studies, evaluating similar conditions, have used a variety of different "traditional" legacy PROMs, making it difficult to compare results. Moreover, the completion of the previous legacy measures can be burdensome and time consuming. Hence, the challenges are how to compare outcomes score between groups and studies, and how to increase the effectiveness of measuring different health outcomes, while reducing administration time and lowering responder burden for PROMs.[14-16]

Standardized quality measurements have been difficult to implement for the orthopedic trauma population since there is an almost innumerable combination of injuries secondary to different mechanisms and contexts.[17] The patients in orthopedic trauma care are heterogeneous and their care is complex, thus established methods for measuring outcomes for elective orthopedic procedures are unlikely to translate easily. A common framework for judging post-intervention patient-reported outcome for elective procedures is comparison to pre-operative function. However, because pre-injury status for trauma patients is often unclear and subject to recall bias, this framework is difficult to apply without additional assumptions that vary by patient and injury

context. As a result, these factors make implementation of value-based care models for trauma patients particularly challenging. In this context, less is known about outcome measures and quality when evaluating trauma.[17] Future studies are required to determine which specific quality measures to institutionalize, how to modify them for different injury contexts, and then how to incorporate them into new payment models. In addition, benchmark studies will only produce valid results if adjustments for injury severity and other injury specific covariates are performed.[17]

Aims and outline this thesis

Evidence-based medicine in the field of orthopedic trauma surgery is constrained by two important factors, a lack of high-quality studies and insufficient attention to patient-reported outcomes. To improve the prognosis of orthopedic trauma patients, additional research is needed into the value of different study designs which evaluate the effects of new and existing medical interventions for trauma patients in everyday clinical practice. Also, assessment of the use of PROMs as an integrated part of research practice. The studies presented in this thesis aim to contribute to these two ambitions.

The potential added value of observational studies in orthopedic trauma will be evaluated in **Part 1**. In **Chapter 2**, the potential value of routinely collected data on elective surgical interventions is assessed. **Part 1** also describes four meta-analyses, which included both RCTs and observational studies, of surgical and non-surgical interventions in patients with proximal humeral fractures (**Chapters 3**), humeral shaft fractures (**Chapter 4**), distal radius fractures (**Chapter 5**), and Achilles tendon ruptures (**Chapter 6**), respectively.

Part 2 of this thesis focuses on patient-reported outcome measures. Specifically, the value is assessed of a tool to measure patient-reported outcomes, the PROMIS tool, in comparison to established measurement tools including the shortened version of the Disabilities of the Arm, Shoulder and Hand questionnaire (QuickDASH) in trauma patients. **Chapters 7** and **Chapter 8** describe two benchmark studies that assess the effects of new and existing medical interventions with the use of the QuickDASH scores for two injuries, acromioclavicular joint dislocations and lateral clavicular fractures. Studies presented in **Chapters 9** and **Chapter 10** assess the use of the PROMIS and the QuickDASH scores and explores the association between these PROMs and clinical outcomes in proximal humeral fractures and distal humeral fractures, respectively. **Chapter 11** establishes normative data on the long-term patient-reported functional outcome and

health-related quality of life after surgical treatment of bicondylar tibial plateau fractures. The study described in **Chapter 12**, looks into the relation between the PROMIS tool and the Achilles Tendon Total Rupture Score (ATRS) in patients undergoing treatment for acute Achilles tendon ruptures. This thesis ends with a general discussion including recommendations and directions for future research (**Chapter 13**).

References:

1. Balogh ZJ, Reumann MK, Gruen RL, et al. Advances and future directions for management of trauma patients with musculoskeletal injuries. *Lancet (London, England).* 2012;380(9847):1109-1119.
2. Mock C, Cherian MN. The global burden of musculoskeletal injuries: challenges and solutions. *Clin Orthop Relat Res.* 2008;466(10):2306-2316.
3. Gabbe BJ, Simpson PM, Harrison JE, et al. Return to Work and Functional Outcomes After Major Trauma: Who Recovers, When, and How Well? *Annals of surgery.* 2016;263(4):623-632.
4. Houwert RM, Beks RB, Dijkgraaf MGW, et al. Study methodology in trauma care: towards question-based study designs. *Eur J Trauma Emerg Surg.* 2019.
5. Chapman SJ, Shelton B, Mahmood H, Fitzgerald JE, Harrison EM, Bhangu A. Discontinuation and non-publication of surgical randomised controlled trials: observational study. *Bmj.* 2014;349:g6870.
6. Beks RB, Bhashyam AR, Houwert RM, et al. When observational studies are as helpful as randomized trials: Examples from orthopedic trauma. *The journal of trauma and acute care surgery.* 2019;87(3):730-732.
7. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *Jama.* 2007;297(11):1233-1240.
8. Frieden TR. Evidence for Health Decision Making - Beyond Randomized, Controlled Trials. *The New England journal of medicine.* 2017;377:465-475.
9. Arditi C, Burnand B, Peytremann-Bridevaux I. Adding non-randomised studies to a Cochrane review brings complementary information for healthcare stakeholders: an augmented systematic review and meta-analysis. *BMC health services research.* 2016;16:598.
10. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *Bmj.* 2002;324(7351):1448-1451.
11. Rouleau DM. Designing clinical studies in orthopedic traumatology. *Orthopaedics & traumatology, surgery & research : OTSR.* 2018;104(1s):S1-s7.
12. Porter ME. What is value in health care? *N Engl J Med.* 2010;363(26):2477-2481.
13. Lowry KJ, Brox WT, Naas PL, Tubb CC, Muschler GF, Dunn W. Musculoskeletal-based Patient-reported Outcome Performance Measures, Where Have We Been-Where Are We Going. *J Am Acad Orthop Surg.* 2019;27(13):e589-e595.
14. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *J Am Acad Orthop Surg.* 2016;24(11):744-749.
15. Cheung EC, Moore LK, Flores SE, Lansdown DA, Feeley BT, Zhang AL. Correlation of PROMIS with Orthopaedic Patient-Reported Outcome Measures. *JBJS reviews.* 2019;7(8):e9.
16. Gagnier JJ. Patient reported outcomes in orthopaedics. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2017;35(10):2098-2108.
17. Bhashyam AR, van der Vliet QMJ, Ochen Y, et al. Injury-related variation in patient-reported outcome after musculoskeletal trauma: a systematic review. *Eur J Trauma Emerg Surg.* 2019.

PART 1

STUDY DESIGNS

CHAPTER 2

Potential value of observational studies of surgical orthopedic interventions: two studies of total hip arthroplasty

Yassine Ochen, Maaike G.J. Gademan, Rob G.H.H. Nelissen, Luke P.H. Leenen, R. Marijn Houwert, Rolf H.H. Groenwold

(Submitted)

Abstract

Background

There is a growing debate on the complementary value of well-designed observational studies of surgical interventions compared to randomized controlled trials (RCTs). The aim of this study was to assess the potential value of routinely collected data on elective surgical interventions. Two studies on total hip arthroplasty (THA) were performed to evaluate comparability of treatment groups and what might be the driving force behind (in)comparability.

Methods

Routinely collected data from the nationwide Dutch Arthroplasty Register (LROI) were used. Two studies of THA were conducted comparing (1) surgical approach (posterolateral approach (PLA) versus straight lateral approach (SLA) versus anterior approach (AA), where PLA versus SLA was the primary comparison and (2) fixation method (cemented versus uncemented). Treatment groups were compared regarding preoperative patient characteristics and postoperative patient-reported outcome measures (PROMs). The EuroQol-5 Dimensions (EQ-5D) index score measured 12 months after surgery was considered the outcome of primary interest. Differences between groups were quantified per variable using the standardized mean difference (SMD). Regression analysis was performed with and without adjustment of baseline information and presented as mean difference with 95% confidence interval (CI). The magnitude and direction of the difference between the crude and adjusted mean difference was quantified by means of a Z-score, which provides a standardized measure of the change in effect estimate when adjustment for confounding is made. Sensitivity analyses were conducted to assess the potential for unmeasured confounding.

Results

The comparison of surgical approaches for total hip arthroplasty (PLA vs SLA) showed no meaningful differences in patient characteristics between treatment groups (SMD<0.1) and also no relevant impact of confounding adjustment (Z-scores <1). For the other surgical approach comparisons (PLA versus AA, SLA versus AA), Z-scores >2 were observed. In the study on fixation method (cemented versus uncemented) several meaningful imbalances were observed in patient characteristic between the two treatment groups (SMD>0.1), as well as a relevant impact of confounding adjustment (Z-scores >2).

Conclusion

This study based on the nationwide Dutch Arthroplasty Register (LROI) of patients with THA provides insight in when observational data can be used to compare surgical treatments, provide valuable clinical information, and thus when routinely collected data on elective surgical interventions can be of value in comparing treatment options. Particularly studies of surgical treatments might be less sensitive to confounding if treatment preference is not subject to patient characteristics and 'allocation to' treatment is a close to a random process. The comparison between surgical approaches (PLA vs SLA) for THA is an example of this.

Introduction

Randomized controlled trials (RCTs) are generally considered to provide the highest level of evidence of treatment effects.[1,2] Randomization prevents confounding due to selective prescription of treatment to patients who would potentially benefit most. Blinding, of patients and treating physicians, prevents changes in health care behavior, and efforts can be made to ensure that assessors of the outcome are blinded for the received treatment. Furthermore, in RCTs efforts can be made to enhance the adherence to the received pharmaceutical treatment. Nevertheless, RCTs might not always be ethical, feasible, or necessary to address a specific research question. This is especially apparent in the surgical field, where variation in surgical practice can lead to practical challenges in terms of patient recruitment and randomization.[3,4] Moreover, the patient populations encountered in daily clinical practice can differ from the often highly selected patient populations enrolled in RCTs.[5] Also, RCTs might not always have sufficient follow-up or sample size to assess rare outcomes or long-term treatment effects.[6,7] Consequently, the results of RCTs often don't find their way into surgical practice.[8,9] This has led to a growing debate about the question whether well-designed observational studies of surgical interventions might complement and add valuable information to the results from RCTs.[3,9,10] This debate centers around the question whether the treatment groups that are being compared are inherently different, or whether there might be situations in which comparability can be achieved. Although it is clear that randomization, concealment of allocation, and blinding are not possible in observational studies, the extent to which their absence impacts the validity of an observational study may differ based on the clinical context and research questions.[10]

In daily practice, the allocation of surgical interventions can sometimes be close to a random process, possibly improving the validity of observational study designs in research of surgical interventions. Particularly studies of acute operative treatments might be less sensitive to confounding when the treatment option depends on a surgeon's preference but not on individual patient characteristics.[3] In such cases, one can speculate that groups of patients who underwent different surgical treatments might be rather similar (except for the treatment option).[3] This has been observed in different meta-analyses of various surgical treatments in orthopedic trauma surgery, in which the treatment arms appeared comparable in terms of patient characteristics. Also, the pooled results of observational studies indeed matched those of RCTs on the same comparison.[3,11-13] However, whether this also holds for elective surgery has not been investigated.

The aim of this study was to assess the comparability and potential value of routinely collected data on elective surgical interventions, investigating preoperative comparability of treatment groups in terms of patient characteristics and postoperative differences in terms of patient reported outcomes. Two studies of total hip arthroplasty (THA) were performed: (1) surgical approach (posterolateral versus straight lateral versus anterior) and (2) fixation method (cemented versus uncemented). The aim of this study was not to provide evidence on the relative benefits of the different discussed surgical techniques.

Methods

Study 1: Surgical approach of THA

THA is considered to be one of the most successful orthopedic procedures for patients with osteoarthritis, resulting in relief of pain, improved hip function, and substantial improvement in quality of life. However, there is no consensus regarding the optimal surgical approach.[14-18] Currently, the posterolateral approach (PLA) and the straight lateral approach (SLA) are the most frequently used techniques worldwide. Another approach, which has become more popular in recent years, is the anterior approach (AA).[19] The difference in outcomes seems small and each of the approaches have their own set of complications, benefits, and learning curves.[14-18] Therefore, the decision for the surgical approach is predominantly determined by surgeon preference and experience, as well as local hospital standards.[20] We hypothesize that groups of patients who are operated using either of the three approaches are similar in terms of prognostic relevant characteristics (Figure 1A).

We compared the three groups of patients who were treated with primary THA using the PLA, SLA, and AA. The primary comparison was made between the two traditional approaches, PLA versus SLA. Secondary comparisons were made between the more recent AA approach and each of the two traditional approaches; PLA versus AA, and SLA versus AA. Inclusion criteria were: (1) age 18 years or older, and (2) primary diagnosis osteoarthritis. Exclusion criteria were: (1) revision arthroplasties and (2) metal on metal arthroplasties. The three groups were compared in terms of preoperative patient characteristic, surgical variables, and patient-reported outcome measures (PROMs).

Study 2: Fixation of THA

The success of THA and the worldwide acceptance is largely due to the development of the durable cemented low-friction arthroplasty with high survival rates. Although the initial components were cemented, the use of uncemented components has gained popularity over the years.[21-24] Both the cemented and uncemented techniques result in satisfactory fixation, however, there has been much debate regarding complications of each method.[21-24] The cemented and uncemented fixation methods are used for heterogeneous groups, with different factors that can affect revision and survival rates such as geometry, material shape, surface finish, and bearings, with the choice of the fixation technique based on surgeon preference.[21,22] In the last decade, THA has changed from mainly cementation to mainly uncemented fixation and this trend is still continuing, particularly in younger patients.[25] The choice for the cemented or uncemented method for THA is – to a large extent – based on patient characteristics. Therefore, we hypothesize that groups of patients who are operated using either the cemented or uncemented fixation method differ in terms of their characteristics (Figure 1B).

We compared the two groups of patients who were treated with primary THA using the cemented versus the uncemented fixation method. Inclusion criteria were: (1) age 18 years or older, and (2) primary diagnosis osteoarthritis. Exclusion criteria were: (1) revision arthroplasties, (2) metal on metal arthroplasties, or (3) arthroplasties with a hybrid fixation. The two groups were compared in terms of preoperative patient characteristic, surgical variables, and PROMs.

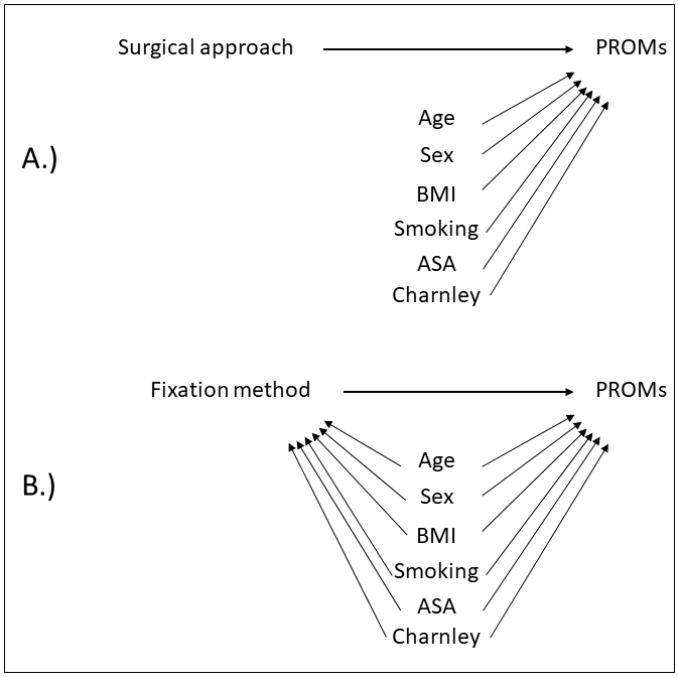

Figure 1. Graphical representation of potential for confounding in observational studies of total hip arthroplasty.
Panel A is causal diagram of possible relations between variables in an observational study of the effect of surgical approach (posterolateral approach vs. straight lateral approach) on patient-reported outcome measures (PROMs). Factors that influence PROMs, including age, sex, body mass index (BMI), smoking status (smoking), American Society of Anaesthesiologists (ASA) classification, and Charnley classification (Charnley), are not expected to influence the choice of surgical approach. Consequently, there are no arrows from these factors to approach.
Panel B is causal diagram of possible relations between variables in an observational study of the effect of fixation method (cemented vs. uncemented) on PROMs. Factors that influence PROMs, including age, sex, BMI, smoking, ASA, and Charnley. These factors are expected to also influence the choice of fixation method. Therefore, arrows from these factors to cementation are included.

Data source

Routinely collected data from the nationwide Dutch Arthroplasty Register (LROI) were extracted for this study.[26] The LROI is a prospective longitudinal cohort containing data on arthroplasties. Data collection started in 2007. The collection of PROMs of patients who underwent THA started in 2014. In 2016, data on primary THAs were provided by up to 99 hospitals and clinics (100% coverage of Dutch hospitals). The completeness of the data is checked against the hospital information systems and currently exceeds 99% for primary THAs. Data on PROMs is provided by up to 80 centers.[26-28]

Data collection

Data were obtained from all adults patients who were treated with primary THA between 2014 and 2018. Information about the following preoperative patient characteristics was collected from the LROI database; age, sex, body mass index (BMI), smoking status, American Society of Anesthesiologists (ASA) classification (I, II, III-IV), Charnley classification (A, B1, B2, C), and previous surgical procedures on the involved hip. In addition, information was collected about surgical approach (PLA, SLA, AA) and fixation method (cemented, hybrid, uncemented). PROMs were collected preoperative, postoperative at 3 months, and 12 months, and consisted of the three-level version EuroQol-5 Dimensions (EQ-5D), Numeric Rating Scale (NRS) for pain (during activity and at rest), Hip disability and Osteoarthritis Outcome Score (HOOS-PS), and Oxford Hip Score (OHS). [29-32] The EQ-5D index score measured 12 months after surgery was considered the outcome of primary interest.

Statistical analysis

In both studies, the same analyses were performed. First, a comparison was made between the intervention groups regarding baseline information about preoperative patient characteristics, surgical variables, and preoperative PROMs. Differences between groups were quantified per variable by means of the standardized mean difference (SMD), where a standardized mean difference of >0.1 was considered a meaningful imbalance in baseline covariates between intervention groups.[33] The relation between the interventions and post-treatment (3 months and 12 months) PROMs were assessed using linear regression analysis. Regression analysis was performed with and without adjustment for baseline information and presented as crude, or adjusted, mean difference with 95% confidence interval (CI). Adjustment was performed for baseline information regarding the preoperative patient characteristic, surgical variables, and

preoperative PROMs. The magnitude and direction of the difference between the crude and adjusted mean difference was quantified by means of a Z-score, which in this case provides a standardized measure of the change in effect estimate when adjustment for potential confounders is made. Z-score values >2 indicate a relevant change.[34] The comparisons have a descriptive nature, focusing on comparability of treatment groups.

Sensitivity analyses were conducted to assess the potential impact of unmeasured confounding. For the primary outcomes, we determined the minimum association that an unmeasured confounder would need to have with both the treatment and outcome, conditional on the measured covariates, to fully explain away a specific treatment–outcome association.[35] All analyses were performed in R version 3.6.1 (R Development Core Team, Released 2013, Vienna, Austria: R Foundation for Statistical Computing).[36]

Results

Study 1: Surgical approach of THA

Patient characteristics

In total, 120,902 patients met the inclusion criteria for study 1. The baseline characteristics are shown in Table 1. The PLA group included 73,750 patients (61%), the SLA group 16,557 patients (14%), and the AA group 30,595 patients (25%). There were no meaningful differences in preoperative patient characteristics between the PLA and the SLA groups (all SMD <0.1). However, the PLA and AA groups differed regarding various preoperative patient characteristics, for example age (SMD 0.109), ASA classification (SMD 0.172), and BMI (SMD 0.178). Also, the SLA and AA groups differed regarding various preoperative patient characteristics, for example age (SMD 0.141), ASA classification (SMD 0.131), and BMI (SMD 0.188).

Outcomes

The results of the crude and adjusted regression analyses are shown in Table 2. The mean EQ-5D index score at 12 months was 0.859 (SD 0.188) in the PLA group, compared to 0.826 (SD 0.200) in the SLA group; crude mean difference -0.033 (95% CI -0.040 to -0.026). The adjusted mean difference in EQ-5D index score at 12 months was -0.036 (95% CI -0.044 to -0.029). The corresponding Z-score for the EQ-5D index score at 12 months between the crude and adjusted differences was 0.613, indicating no relevant change in treatment effect estimate after adjustment for observed potential confounders. Also, for the other outcomes, the change in effect estimate was relatively small, with Z-scores <1. Sensitivity analyses showed that the observed adjusted

Table 1. Baseline characteristics of patients undergoing total hip arthroplasty, stratified by surgical approach

	Posterolateral	**Straight Lateral**	**Anterior**	**SMD**	**SMD**	**SMD**
N	**73750**	**16557**	**30595**	**PLA vs. SLA**	**PLA vs. AA**	**SLA vs. AA**
Age	69.83 (9.75)	70.14 (9.76)	68.78 (9.51)	0.032	0.109	0.141
Sex (%)						
Male	25554 (34.7)	5600 (33.8)	10246 (33.5)	0.018	0.025	0.007
Female	48101 (65.3)	10950 (66.2)	20341 (66.5)			
ASA classification (%)						
ASA I	12250 (16.6)	2857 (17.3)	6537 (21.4)	0.047	0.172	0.131
ASA II	48203 (65.4)	11001 (66.5)	20137 (65.9)			
ASA III-IV	13215 (17.9)	2678 (16.2)	3897 (12.7)			
Previous operation (%)						
Yes	1557 (2.1)	424 (2.6)	269 (0.9)	0.032	0.103	0.133
No	70958 (97.9)	15670 (97.4)	29844 (99.1)			
BMI (%)						
Underweight (<18,5)	448 (0.6)	97 (0.6)	198 (0.6)	0.011	0.178	0.188
Normal weight (18,5-25)	21793 (30.2)	4761 (29.9)	11049 (36.3)			
Overweight (25-30)	31051 (43.0)	6819 (42.8)	13147 (43.2)			
Obese (30-40)	17944 (24.8)	4034 (25.3)	5880 (19.3)			
Class 3 Obese (>40)	1022 (1.4)	227 (1.4)	189 (0.6)			
Charnley classification (%)						
A	31432 (43.6)	7274 (44.7)	13920 (45.8)	0.038	0.097	0.074
B1	21749 (30.2)	4979 (30.6)	9617 (31.7)			
B2	16987 (23.6)	3577 (22.0)	6285 (20.7)			
C	1953 (2.7)	447 (2.7)	544 (1.8)			
Smoking (%)						
Yes	7425 (10.7)	1713 (11.5)	2963 (9.8)	0.025	0.029	0.054
No	61815 (89.3)	13166 (88.5)	27135 (90.2)			
Fixation method (%)						
Cemented	22078 (30.0)	4280 (25.9)	3221 (10.5)	0.094	0.517	0.439`
Hybrid	7384 (10.0)	1881 (11.4)	2607 (8.5)			
Uncemented	44229 (60.0)	10375 (62.7)	24756 (80.9)			
EQ-5D index score	0.54 (0.28)	0.55 (0.28)	0.59 (0.26)	0.009	0.188	0.179
NRS pain score during activity	7.27 (2.07)	7.28 (2.02)	6.88 (2.18)	0.003	0.188	0.192
NRS pain score at rest	5.34 (2.55)	5.14 (2.63)	4.98 (2.60)	0.075	0.139	0.063
HOOS-PS score	50.06 (18.02)	48.92 (18.22)	46.81 (17.66)	0.063	0.182	0.118
OHS score	22.45 (8.67)	22.33 (8.69)	23.78 (8.48)	0.015	0.155	0.170

Continuous variables presented as mean (SD); SMD Standardized Mean Difference; EQ-5D EuroQol-5 Dimensions, NRS Numeric Rating Scale for pain; HOOS-PS Hip disability and Osteoarthritis Outcome Score; OHS Oxford Hip Score.

Table 2. Patient-reported outcome measures of patients receiving total hip arthroplasty, stratified by surgical approach

	Posterolateral		Straight Lateral		Crude			Adjusted			
	Mean	SD	Mean	SD	Mean difference	95% CI		Mean difference	95% CI		Z-score
EQ-5D score											
3 months	0.821	0.188	0.794	0.189	-0.027	-0.034	-0.02	-0.027	-0.034	-0.019	-0.075
12 months	0.859	0.188	0.826	0.200	-0.033	-0.04	-0.026	-0.036	-0.044	-0.029	0.613
NRS activity											
3 months	2.103	2.277	2.458	2.389	0.355	0.271	0.438	0.352	0.262	0.443	0.034
12 months	1.470	2.184	1.961	2.508	0.491	0.403	0.579	0.507	0.410	0.603	-0.252
NRS at rest											
3 months	1.190	1.864	1.315	1.979	0.125	0.056	0.194	0.144	0.07	0.218	-0.379
12 months	0.873	1.745	1.177	2.070	0.305	0.234	0.375	0.342	0.265	0.42	-0.744
HOOS-PS score											
3 months	18.276	14.458	21.707	14.68	3.43	2.871	3.99	3.519	2.923	4.115	-0.219
12 months	13.800	14.824	17.784	16.753	3.984	3.376	4.593	4.142	3.485	4.799	-0.360
OHS score											
3 months	39.003	7.791	37.819	7.339	-1.184	-1.481	-0.887	-1.100	-1.412	-0.788	-0.395
12 months	41.873	7.418	40.173	7.909	-1.699	-1.998	-1.401	-1.630	-1.944	-1.315	-0.324

Z-score magnitude and direction of the change between crude and adjusted mean difference; EQ-5D EuroQol-5 Dimensions; NRS Numeric Rating Scale for pain; HOOS-PS Hip disability and Osteoarthritis Outcome Score; OHS Oxford Hip Score.

difference in mean EQ-5D index score at 12 months (i.e., -0.036) could be explained away by a binary unmeasured confounder that increases the mean EQ-5D by, e.g., 0.36 and has a difference in its prevalence between PLA and SLA of 0.10. Other scenarios that could explain away the observed difference in mean EQ-5D are presented in Figure 2.

For the other comparisons (PLA versus AA, SLA versus AA), larger Z-scores were observed, owing to the observed baseline incomparability (supplementary Table S1 and Table S2). For example, the comparison PLA versus AA, the corresponding Z-score for the EQ-5D index score at 12 months between the crude and adjusted differences was 5.984.

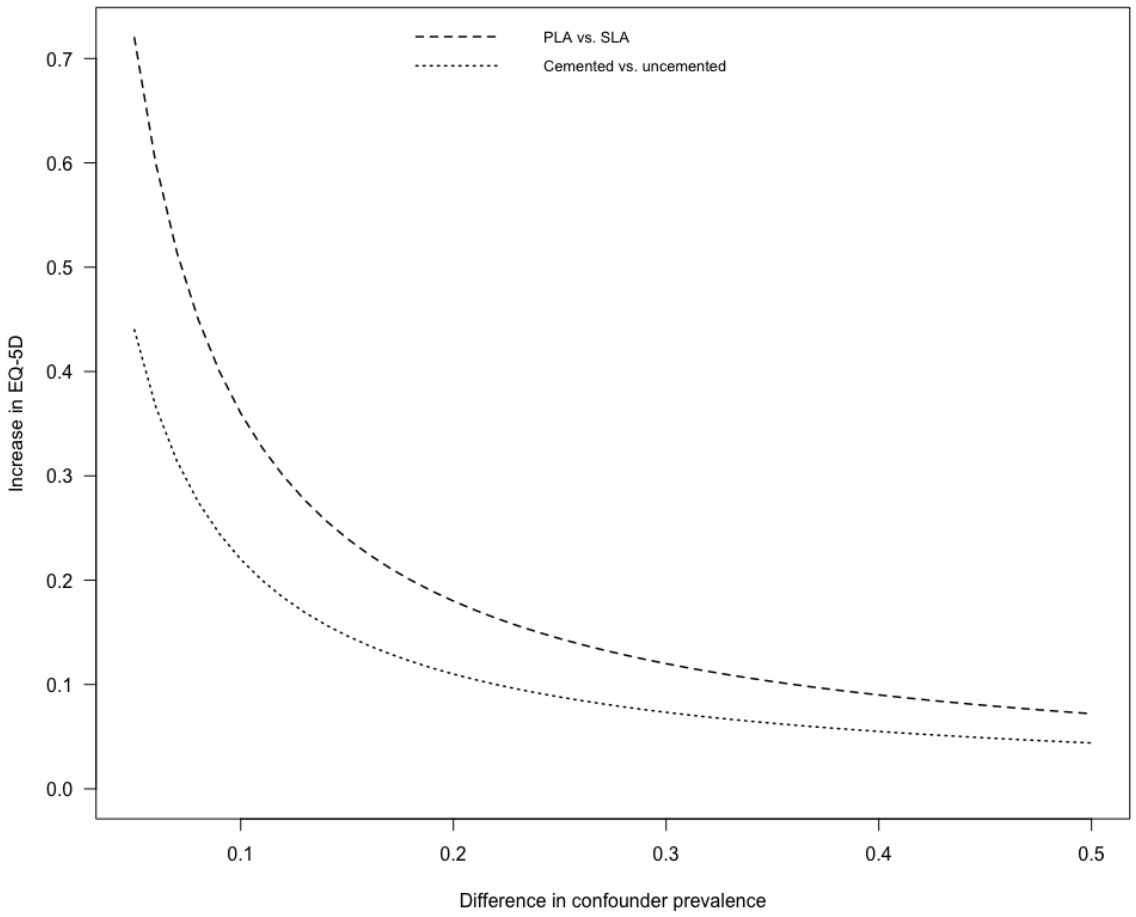

Figure 2. Scenarios of unmeasured confounding that could explain away the observed treatment-outcome relations. The figure shows the minimum association that an unmeasured confounder would need to have with both the treatment (horizontal axis) and outcome (vertical axis), conditional on the measured covariates, to fully explain away the observed treatment–outcome association.

Table 3. Baseline characteristics of patients undergoing total hip arthroplasty, stratified by fixation method

	Cemented	Uncemented	SMD
N	**29579**	**79360**	
Age	75.37 (8.14)	67.39 (9.32)	0.913
Sex (%)			
Male	7685 (26.0)	30071 (37.9)	0.258
Female	21867 (74.0)	49212 (62.1)	
ASA classification (%)			
ASA I	2924 (9.9)	16612 (21.0)	0.384
ASA II	19483 (65.9)	52104 (65.7)	
ASA III-IV	7142 (24.2)	10559 (13.3)	
Previous operation (%)			
Yes	637 (2.2)	1314 (1.7)	0.035
No	28695 (97.8)	76271 (98.3)	
BMI (%)			
Underweight (<18,5)	267 (0.9)	412 (0.5)	0.072
Normal weight (18,5-25)	9632 (33.2)	24251 (31.1)	
Overweight (25-30)	12211 (42.1)	33914 (43.5)	
Obese (30-40)	6504 (22.4)	18511 (23.7)	
Obese (30-40)	6504 (22.4)	18511 (23.7)	
Class 3 Obese (>40)	395 (1.4)	871 (1.1)	
Charnley classification (%)			
A	11794 (40.4)	35275 (45.3)	0.126
B1	9042 (31.0)	24072 (30.9)	
B2	7383 (25.3)	16762 (21.5)	
C	943 (3.2)	1683 (2.2)	
Smoking (%)			
Yes	2367 (8.4)	8511 (11.4)	0.102
No	25905 (91.6)	66097 (88.6)	
Surgical approach (%)			
Posterolateral	22078 (74.6)	44229 (55.7)	0.519
Straight Lateral	4280 (14.5)	10375 (13.1)	
Anterior	3221 (10.9)	24756 (31.2)	
EQ-5D index score	0.51 (0.29)	0.58 (0.27)	0.229
NRS pain score during activity	7.29 (2.11)	7.12 (2.09)	0.080
NRS pain score at rest	5.31 (2.65)	5.14 (2.56)	0.067
HOOS-PS score	50.97 (18.65)	48.08 (17.69)	0.159
OHS score	21.34 (9.01)	23.42 (8.44)	0.238

Continuous variables presented as mean (SD); SMD Standardized Mean Difference; EQ-5D EuroQol-5 Dimensions, NRS Numeric Rating Scale for pain; HOOS-PS Hip disability and Osteoarthritis Outcome Score; OHS Oxford Hip Score.

Study 2: Fixation of THA

Patient characteristics

In total, 108,939 patients were included in study 2. The characteristics are shown in Table 3. The cemented group included 29,579 patients (27%) and the uncemented group 79,360 patients (73%). There were meaningful imbalances of preoperative patient characteristic in the comparison of the cemented versus uncemented regarding age (SMD 0.913), sex (SMD 0.258), ASA classification (SMD 0.384), Charnley classification (SMD 0.126), and smoking (SMD 0.102).

Outcomes

The results of the crude and adjusted regression analyses are shown in Table 4. The mean EQ-5D index score at 12 months was 0.824 (SD 0.206) in the cemented group, compared to 0.877 (SD 0.176) in the uncemented group; crude mean difference -0.053 (95% CI -0.058 to -0.048).

The adjusted mean difference in EQ-5D index score at 12 months was -0.022 (95% CI -0.028 to -0.016). The corresponding Z-score for the EQ-5D index score at 12 months between the crude and adjusted differences was -8.646, indicating a relevant change in treatment effect estimate after adjustment for observed potential confounders. Sensitivity analyses showed that the observed adjusted difference in mean EQ-5D index score at 12 months (i.e., -0.022) could be explained away by a binary unmeasured confounder that increases the mean EQ-5D by, e.g., 0.22 and has a difference in its prevalence between cemented and uncemented of 0.10. Other scenarios that could explain away the observed difference in mean EQ-5D are presented in Figure 2.

Table 4. Patient-reported outcome measures of patients receiving total hip arthroplasty, stratified by fixation method

	Cemented		Uncemented		Crude			Adjusted			
	Mean	SD	Mean	SD	Mean difference	95% CI		Mean difference	95% CI		Z-score
EQ-5D score											
3 months	0.797	0.199	0.841	0.179	-0.044	-0.048	-0.039	-0.022	-0.027	-0.016	-6.686
12 months	0.824	0.206	0.877	0.176	-0.053	-0.058	-0.048	-0.022	-0.028	-0.016	-8.646
NRS activity											
3 months	2.244	2.387	1.965	2.228	0.279	0.224	0.335	0.200	0.134	0.266	1.975
12 months	1.654	2.315	1.364	2.134	0.230	0.168	0.292	0.180	0.110	0.251	2.637
NRS at rest											
3 months	1.356	2.031	1.100	1.799	0.256	0.210	0.301	0.153	0.100	0.207	3.117
12 months	0.991	1.876	0.809	1.695	0.182	0.136	0.227	0.055	-0.002	0.111	3.837
HOOS-PS score											
3 months	20.301	15.223	16.709	14.178	3.593	3.215	3.971	1.602	1.156	2.047	7.298
12 months	16.434	16.091	12.257	14.214	4.177	3.767	4.587	1.330	0.839	1.821	9.631
OHS score											
3 months	37.625	8.359	40.029	7.277	-2.405	-2.601	-2.208	-1.549	-1.779	-1.318	-6.033
12 months	40.228	8.015	42.702	7.065	-2.474	-2.678	-2.271	-1.120	-1.361	-0.879	-9.219

Z-score magnitude and direction of the change between the crude and adjusted mean difference; EQ-5D EuroQol-5 Dimensions, NRS Numeric Rating Scale for pain; HOOS-PS Hip disability and Osteoarthritis Outcome Score; OHS Oxford Hip Score.

Discussion

The aim of this study was to assess the potential value of routinely collected data on surgical interventions by evaluating the comparability of treatment groups and the potential for confounding. In the first study on surgical approach (PLA, SLA, AA) of THA, the primary comparison between the two traditional approaches, PLA versus SLA, showed no meaningful differences in patient characteristics (SMD<0.1) and also no relevant impact of adjustment for baseline characteristics (Z-scores <1). For the other surgical approach comparisons (PLA versus AA, SLA versus AA), Z-scores >2 were observed. In the second study on fixation method (cemented versus uncemented) of total hip arthroplasty, there were several meaningful imbalances in patient characteristic between treatment groups (SMD>0.1), and a relevant impact of adjustment for baseline characteristics (Z-scores >2).

Our aim was to evaluate comparability of patients receiving different elective surgical orthopedic interventions. We did not look into the differences in effect estimates between observational studies and RCTs, which has been investigated in previous research. Ioannidis et al.[37] evaluated the results of randomized and nonrandomized studies for a variety of topics and found that observational studies overestimate treatment effects compared to RCTs. This was confirmed by Hemkens et al.[38]. In contrast, Benson et al.[39] and Concato et al.[40] found little evidence for systematic differences between results of observational studies and RCTs. Focusing on surgical interventions, Abraham et al.[41] found that results of high-quality observational studies were similar to those of RCTs. Clearly, based on these studies, one cannot conclude that results of observational studies are always different from those of RCTs, nor that they always concur. It probably largely depends on the type of interventions being compared, the context in which the comparison is made, and the quality of the observational study including the data being used.[10] In daily practice, the allocation of surgical interventions can sometimes be close to a random process, which might increase the validity of observational studies designs based on the specific research question. Furthermore, both patients and surgeons can have a strong preference for a certain treatment, which forms an obstacle for randomization in surgical trials.[42] A study that evaluated the efficacy of surgical stabilization of rib fracture encountered these challenges, and decided to offer randomization as well as observational follow-up to participants. Nearly 80% of subjects declined randomization, yet no differences were observed between subjects who chose for the different options.[43] Our findings support the viewpoint that, in specific cases, one can speculate that groups of patients who undergo different orthopedic surgical interventions will be

comparable with respect to patient characteristics, and therefore results of such observational studies would be valuable to use in comparative studies, complementary to RCTs.[3] Presence of large observational cohorts as present in regional and national registries underscore the accessibility of these readily available data sources in evaluating treatment modalities.

The potential for confounding in an observational study likely depends on the context, i.e., the type of intervention that is studied and the comparison that is being made.[3,10] Confounding may be more prominent in studies in which pharmacological treatments for surgical patients are compared, than studies comparing for instance different surgical interventions in acute trauma care. To assess the potential impact of unmeasured confounding, a sensitivity analysis of unmeasured confounding was performed, which showed that a binary unmeasured confounder that would result in a relative large difference in mean difference, would decrease depending on the prevalence of the confounder in the comparison groups.

In addition to the potential for confounding, other sources of bias in observational studied should be considered. Electronic health record data may be affected by for instance errors in data linkage, misclassification, and missing values, all of which could also impact the quality of observational research using these data.[44] The data used in this study were extracted from the LROI, a prospective longitudinal cohort containing high-quality data. The completeness of the LROI data (100% coverage of Dutch hospitals) is checked against the hospital information systems and currently exceeds 99% for primary THAs. [26-28] Hence, the phenomena observed in this study are not necessarily to be expected in other observational studies. It is clear that blinding of participants is not possible in observational studies. However, this is also the case in many surgical RCTs and hence often not an argument to overrule evidence from observational studies in favor of that from RCTs. Blinding of the outcome assessor also is typically not implemented in observational studies. However, in case of patient reported outcome measures, like the ones used in this study, such blinding is also not feasibly in RCTs.

In orthopedic trauma, well-designed observational studies might complement and add valuable information to results from RCTs.[10] The (complementary) value of observational studies in addition to randomized trials, has been discussed before.[3,10] Particularly studies of acute operative treatments might be less sensitive to confounding if treatment preference is not subject to patient characteristics, and 'allocation to surgeon' is a random process.[3] This has been observed for

various interventions in acute orthopedic trauma surgery.[3,11-13] The current study extends this to the field of elective surgery.

In this study the comparability and potential value of routinely collected data on elective orthopedic surgical interventions was assessed with the use of two studies. The comparability seemed the most apparent between the two traditional approaches, PLA versus SLA. However, the PLA and AA groups and the SLA and AA groups differed slightly regarding various patient characteristics. The AA is a relative new approach and is thought to include a steep learning curve, which might explain these differences.[45]

Conclusion

The aim of this study was not to provide evidence on the relative benefits of the different discussed surgical techniques, nor do these studies provide evidence for all studies of elective surgical treatment options. It does, however, provide support that there are cases in which observational studies of surgical treatment options are viable and provide valuable information. Particularly studies of surgical treatments might be less sensitive to confounding if treatment preference is not subject to patient characteristics and 'allocation to' treatment is close to a random process. It is up to the researchers of such studies to provide the arguments to substantiate the claim that treatment groups are expected to be comparable and why a particular research question could be answered using an observational study design.

References:

1. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS medicine.* 2008;5(3):e67.
2. Black N. Why we need observational studies to evaluate the effectiveness of health care. *Bmj.* 1996;312(7040):1215-1218.
3. Houwert RM, Beks RB, Dijkgraaf MGW, et al. Study methodology in trauma care: towards question-based study designs. *Eur J Trauma Emerg Surg.* 2019.
4. McCulloch P, Altman DG, Campbell WB, et al. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet (London, England).* 2009;374(9695):1105-1112.
5. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *Jama.* 2007;297(11):1233-1240.
6. Frieden TR. Evidence for Health Decision Making - Beyond Randomized, Controlled Trials. *The New England journal of medicine.* 2017;377:465-475.
7. Arditi C, Burnand B, Peytremann-Bridevaux I. Adding non-randomised studies to a Cochrane review brings complementary information for healthcare stakeholders: an augmented systematic review and meta-analysis. *BMC health services research.* 2016;16:598.
8. Oberkofler CE, Hamming JF, Staiger RD, et al. Procedural Surgical RCTs in Daily Practice: Do Surgeons Adopt Or Is It Just a Waste of Time? *Annals of surgery.* 2019;270(5):727-734.
9. Chapman SJ, Shelton B, Mahmood H, Fitzgerald JE, Harrison EM, Bhangu A. Discontinuation and non-publication of surgical randomised controlled trials: observational study. *Bmj.* 2014;349:g6870.
10. Beks RB, Bhashyam AR, Houwert RM, et al. When observational studies are as helpful as randomized trials: Examples from orthopedic trauma. *The journal of trauma and acute care surgery.* 2019;87(3):730-732.
11. Smeeing DPJ, van der Ven DJC, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older: A Systematic Review, Meta-analysis, and Comparison of Randomized Controlled Trials and Observational Studies. *Am J Sports Med.* 2017;45(8):1937-1945.
12. Beks RB, Ochen Y, Frima H, et al. Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. *Journal of shoulder and elbow surgery.* 2018;27:1526-1534.
13. Ochen Y, Beks RB, van Heijl M, et al. Operative treatment versus nonoperative treatment of Achilles tendon ruptures: systematic review and meta-analysis. *BMJ (Clinical research ed).* 2019;364:k5120.
14. Higgins BT, Barlow DR, Heagerty NE, Lin TJ. Anterior vs. posterior approach for total hip arthroplasty, a systematic review and meta-analysis. *The Journal of arthroplasty.* 2015;30(3):419-434.
15. Putananon C, Tuchinda H, Arirachakaran A, Wongsak S, Narinsorasak T, Kongtharvonskul J. Comparison of direct anterior, lateral, posterior and posterior-2 approaches in total hip arthroplasty: network meta-analysis. *European journal of orthopaedic surgery & traumatology : orthopedie traumatologie.* 2018;28(2):255-267.
16. Wang Z, Hou JZ, Wu CH, et al. A systematic review and meta-analysis of direct anterior approach versus posterior approach in total hip arthroplasty. *Journal of orthopaedic surgery and research.* 2018;13(1):229.
17. Wang Z, Bao HW, Hou JZ. Direct anterior versus lateral approaches for clinical outcomes after total hip arthroplasty: a meta-analysis. *Journal of orthopaedic surgery and research.* 2019;14(1):63.
18. Kucukdurmaz F, Sukeik M, Parvizi J. A meta-analysis comparing the direct anterior with other approaches in primary total hip arthroplasty. *The surgeon : journal of the Royal Colleges of Surgeons of Edinburgh and Ireland.* 2019;17(5):291-299.
19. den Daas A, Reitsma EA, Knobben BAS, Ten Have B, Somford MP. Patient satisfaction in different approaches for total hip arthroplasty. *Orthopaedics & traumatology, surgery & research : OTSR.* 2019;105(7):1277-1282.
20. Peters RM, van Beers L, van Steenbergen LN, et al. Similar Superior Patient-Reported Outcome Measures for Anterior and Posterolateral Approaches After Total Hip Arthroplasty: Postoperative Patient-Reported Outcome Measure Improvement After 3 months in 12,774 Primary Total Hip Arthroplasties Using the Anterior, Anterolateral, Straight Lateral, or Posterolateral Approach. *The Journal of arthroplasty.* 2018;33(6):1786-1793.
21. Morshed S, Bozic KJ, Ries MD, Malchau H, Colford JM, Jr. Comparison of cemented and uncemented fixation in total hip replacement: a meta-analysis. *Acta Orthop.* 2007;78(3):315-326.
22. Toossi N, Adeli B, Timperley AJ, Haddad FS, Maltenfort M, Parvizi J. Acetabular components in total hip arthroplasty: is there evidence that cementless fixation is better? *J Bone Joint Surg Am.* 2013;95(2):168-174.
23. Abdulkarim A, Ellanti P, Motterlini N, Fahey T, O'Byrne JM. Cemented versus uncemented fixation in total hip replacement: a systematic review and meta-analysis of randomized controlled trials. *Orthopedic reviews.* 2013;5(1):e8.
24. Yoon BH, Ha YC, Lee YK, Koo KH. Postoperative Deep Infection After Cemented Versus Cementless Total Hip Arthroplasty: A Meta-Analysis. *The Journal of arthroplasty.* 2015;30(10):1823-1827.

25. Kuijpers MFL, Hannink G, van Steenbergen LN, Schreurs BW. Total Hip Arthroplasty in Young Patients in The Netherlands: Trend Analysis of >19,000 Primary Hip Replacements in the Dutch Arthroplasty Register. *The Journal of arthroplasty.* 2018;33(12):3704-3711.
26. van Steenbergen LN, Denissen GA, Spooren A, et al. More than 95% completeness of reported procedures in the population-based Dutch Arthroplasty Register. *Acta Orthop.* 2015;86(4):498-505.
27. LROI. Completeness of registering hospitals and completeness of registered arthroplasties in the LROI based on the hospital information system in 2016. Accessed 2020 April 3. http://www.lroi-rapportage.nl/data-quality-coverage-andcompleteness.
28. Hesseling B, Mathijssen NMC, van Steenbergen LN, Melles M, Vehmeijer SBW, Porsius JT. Fast Starters, Slow Starters, and Late Dippers: Trajectories of Patient-Reported Outcomes After Total Hip Arthroplasty: Results from a Dutch Nationwide Database. *J Bone Joint Surg Am.* 2019;101(24):2175-2186.
29. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med.* 2001;33(5):337-343.
30. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain.* 2011;152:2399-2404.
31. Davis AM, Perruccio AV, Canizares M, et al. The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative. *Osteoarthritis and cartilage.* 2008;16(5):551-559.
32. Wylde V, Learmonth ID, Cavendish VJ. The Oxford hip score: the patient's perspective. *Health Qual Life Outcomes.* 2005;3:66.
33. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in medicine.* 2009;28(25):3083-3107.
34. Franklin JM, Pawar A, Martin D, et al. Nonrandomized Real-World Evidence to Support Regulatory Decision Making: Process for a Randomized Trial Replication Project. *Clinical pharmacology and therapeutics.* 2020;107(4):817-826.
35. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology.* 2011;22(1):42-52.
36. Team R. R Development Core Team. *R: A Language and Environment for Statistical Computing.* 2013.
37. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *Jama.* 2001;286(7):821-830.
38. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *BMJ (Clinical research ed).* 2016;352:i493-i493.
39. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *The New England journal of medicine.* 2000;342:1878-1886.
40. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England journal of medicine.* 2000;342:1887-1892.
41. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *Journal of clinical epidemiology.* 2010;63:238-245.
42. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *Bmj.* 2002;324(7351):1448-1451.
43. Pieracci FM, Leasia K, Bauman Z, et al. A multicenter, prospective, controlled clinical trial of surgical stabilization of rib fractures in patients with severe, nonflail fracture patterns (Chest Wall Injury Society NONFLAIL). *The journal of trauma and acute care surgery.* 2020;88(2):249-257.
44. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Routinely collected data and comparative effectiveness evidence: promises and limitations. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* 2016;188(8):E158-e164.
45. Meermans G, Konan S, Das R, Volpin A, Haddad FS. The direct anterior approach in total hip arthroplasty: a systematic review of the literature. *The bone & joint journal.* 2017;99-b(6):732-740.

Supplementary materials to Chapter 2

Table S1. Patient-reported outcome measures of patients receiving total hip arthroplasty, stratified by surgical approach

	Posterolateral		Anterior		Crude			Adjusted			
	Mean	**SD**	**Mean**	**SD**	**Mean difference**	**95% CI**		**Mean difference**	**95% CI**		**Z-score**
EQ-5D score											
3 months	0.821	0.188	0.852	0.176	0.031	0.027	0.035	0.014	0.010	0.018	6.006
12 months	0.859	0.188	0.884	0.170	0.025	0.021	0.030	0.007	0.002	0.011	5.984
NRS activity											
3 months	2.103	2.277	1.817	2.213	-0.287	-0.335	-0.238	-0.182	-0.237	-0.128	-2.995
12 months	1.470	2.184	1.249	2.062	-0.221	-0.271	-0.171	-0.107	-0.165	-0.050	-3.137
NRS at rest											
3 months	1.190	1.864	1.118	1.853	-0.072	-0.112	-0.032	0.016	-0.028	0.061	-3.085
12 months	0.873	1.745	0.743	1.638	-0.129	-0.169	-0.089	-0.042	-0.087	0.004	-3.038
HOOS-PS score											
3 months	18.276	14.458	15.409	14.062	-2.867	-3.189	-2.545	-1.641	-2.001	-1.282	-5.28
12 months	13.800	14.824	11.138	13.782	-2.662	-3.013	-2.311	-1.247	-1.641	-0.853	-5.584
OHS score											
3 months	39.003	7.791	40.608	7.180	1.604	1.427	1.782	0.849	0.656	1.042	5.898
12 months	41.873	7.418	42.956	7.246	1.083	0.894	1.272	0.241	0.035	0.446	6.185

Z-score magnitude and direction of the change between the crude and adjusted mean difference; EQ-5D EuroQol-5 Dimensions, NRS Numeric Rating Scale for pain; HOOS-PS Hip disability and Osteoarthritis Outcome Score; OHS Oxford Hip Score.

Table S2. Patient-reported outcome measures of patients receiving total hip arthroplasty, stratified by surgical approach

	Straight Lateral		Anterior		Crude			Adjusted			
	Mean	**SD**	**Mean**	**SD**	**Mean difference**	**95% CI**		**Mean difference**	**95% CI**		**Z-score**
EQ-5D score											
3 months	0.794	0.189	0.852	0.176	0.058	0.051	0.065	0.039	0.031	0.046	3.930
12 months	0.826	0.200	0.884	0.170	0.058	0.051	0.066	0.038	0.029	0.046	4.002
NRS activity											
3 months	2.458	2.389	1.817	2.213	-0.641	-0.727	-0.555	-0.504	-0.602	-0.406	-2.215
12 months	1.961	2.508	1.249	2.062	-0.712	-0.801	-0.622	-0.538	-0.643	-0.432	-2.699
NRS at rest											
3 months	1.315	1.979	1.118	1.853	-0.197	-0.269	-0.125	-0.075	-0.155	0.006	-2.364
12 months	1.177	2.070	0.743	1.638	-0.434	-0.506	-0.362	-0.342	-0.427	-0.258	-1.764
HOOS-PS score											
3 months	21.707	14.680	15.409	14.062	-6.297	-6.869	-5.726	-4.764	-5.406	-4.122	-3.717
12 months	17.784	16.753	11.138	13.782	-6.646	-7.259	-6.033	-4.561	-5.273	-3.849	-4.712
OHS score											
3 months	37.819	7.339	40.608	7.180	2.789	2.496	3.082	1.916	1.588	2.243	4.126
12 months	40.173	7.909	42.956	7.246	2.782	2.465	3.100	1.633	1.261	2.004	5.020

Z-score magnitude and direction of the change between the crude and adjusted mean difference; EQ-5D EuroQol-5 Dimensions, NRS Numeric Rating Scale for pain; HOOS-PS Hip disability and Osteoarthritis Outcome Score; OHS Oxford Hip Score.

CHAPTER 3

Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials

Reinier B. Beks, Yassine Ochen, Herman Frima, Diederik P.J. Smeeing,
Olivier van der Meijden, Tim K. Timmers, Detlef van der Velde,
Mark van Heijl, Luke P.H. Leenen, Rolf H.H. Groenwold, R. Marijn Houwert

(Journal of Shoulder and Elbow Surgery)

Abstract

Background

There is no consensus on the choice of treatment for displaced proximal humeral fractures in older (>65 years) patients. The aim of this systematic review and meta-analysis was (1) to compare operative with nonoperative management of displaced proximal humeral fractures and (2) to compare effect estimates obtained from randomized controlled trials (RCTs) and observational studies.

Methods

The databases of MEDLINE, Embase, CENTRAL, and CINAHL were searched on September 5, 2017 for studies comparing operative versus nonoperative treatment of proximal humeral fractures; both RCTs and observational studies were included. The MINORS criteria, a validated instrument for methodological quality assessment, were used to assess study quality. The primary outcome measure was physical function as measured by the absolute Constant-Murley score after operative or nonoperative treatment. Secondary outcome measures were major reinterventions, nonunion, and avascular necrosis.

Results

We included 22 studies comprising 7 RCTs and 15 observational studies, resulting in 1743 patients total: 910 treated operatively and 833 nonoperatively. The average age was 68.3 years, and 75% were female. There was no difference in functional outcome between operative and nonoperative treatment with a mean difference of -0.87 (CI, -5.13 – 3.38; P=0.69; I^2=69%). Major reinterventions occurred more often in the operative group. Pooled effects of RCTs were similar to pooled effects of observational studies for all outcome measures.

Conclusion

We recommend nonoperative treatment for the average elderly (aged >65 years) patient with a displaced proximal humeral fracture. Pooled effects of observational studies were similar to those of RCTs, and including observational studies led to more generalizable conclusions.

Introduction

The proximal humeral fracture is the third most common fracture seen in elderly persons, with an incidence of 82 per 100,000 person-years, with an annual increase in the rate by 13.7% over the past 33 years[1–3] The typical patient is a female aged 65 or older.[4] Nearly 75% of patients are treated nonoperatively, and one out of five will undergo surgery depending on fracture type and displacement.[5]

Depending on related factors such as patient age, activity, and fracture pattern, operative treatment options include minimally invasive reduction and intramedullary fixation, open reduction and internal plate fixation, or arthroplasty of the glenohumeral joint. Nonoperative treatment usually starts with immobilization followed by passive and active rehabilitation.[5] Despite the fact that the available literature is inconclusive regarding the superiority of either treatment option, it is common practice to attempt joint-saving operative procedures in younger patients.[5,6] In addition, there is no consensus on whether surgery is beneficial for the older patient with a displaced proximal humeral fracture.

Increasing scientific evidence has demonstrated that meta-analyses of both high-quality observational studies and randomized controlled trials (RCTs) can be similar in value to meta-analyses of RCTs alone in the field of orthopedic trauma surgery.[7–10] Observational studies may give better insight into infrequent outcome measures, rare complications, and small effects of operative treatment while also increasing the generalizability of the results owing to an increase in patient numbers available for analysis or meta-analysis.

The aims of this systematic review and meta-analysis were (1) to compare operative versus nonoperative treatment of displaced proximal humeral fractures and (2) to compare effect estimates obtained from RCTs and observational studies. We hypothesized that (1) operative treatment of proximal humeral fractures does not improve functional outcomes as compared with nonoperative treatment and (2) including observational studies in this meta-analysis will lead to more robust conclusions without decreasing the quality of the results.

Methods

This systematic review and meta-analysis followed guidelines published by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and MOOSE (Meta-Analysis of

Observational Studies in Epidemiology).[11,12] These checklists aim to improve the reporting of systematic reviews and meta-analyses for RCTs and observational studies, respectively.

Search strategy and eligibility criteria

Two reviewers (R.B.B. and Y.O.) independently searched the MEDLINE, Embase, CENTRAL, and CINAHL databases on September 5, 2017, for studies comparing operative and nonoperative treatment of proximal humeral fractures. The search syntax is provided in supplementary Table S1. Both RCTs and observational studies were included. After screening of the titles and abstracts of identified records, studies were independently assessed based on full text. The eligibility criteria were proximal humeral fracture; operative versus nonoperative treatment; and reporting of functional outcomes, as well as complications. The exclusion criteria were language other than English, Dutch, or German; no availability of full text; inclusion of patients younger than 18 years; letters, meeting proceedings, and case reports; and external osteosynthesis as operative treatment. Disagreement over eligibility was resolved by discussion with a third reviewer (R.M.H.). The references of the included studies were screened for eligibility, and citation tracking was performed by using Web of Science to identify articles not found in the original search. Authors were approached via ResearchGate when no full-text version of the article was available.

Data extraction

Data extraction was done independently by two reviewers (R.B.B. and Y.O.) with a data extraction file. The following data were extracted: first author, journal, year of publication, study period, study design, country or countries in which the study was performed, fracture displacement, fracture classification system (Neer classification), follow-up, treatment groups, operative treatment, nonoperative treatment, number of patients, loss to follow-up, implant removal, and outcome measures. Definitions of fracture characteristics, such as displacement, were applied according to the description in the original study. A major reintervention was defined as an additional, initially unplanned surgical procedure for implant failure, deep infection, symptomatic nonunion, subacromial impingement, or avascular necrosis. Planned implant removal was not considered a major reintervention. Fjalestad et al.[13,14] reported additional follow up of previously published data that were merged with the original article for this meta-analysis.

Quality assessment

Two reviewers (R.B.B. and H.F.) independently assessed the methodological quality of all included studies with the Methodological Index for Non-Randomized Studies (MINORS).[15] The MINORS is a validated instrument for methodological quality assessment and clear reporting of observational studies of surgical interventions.[15] Other quality assessment tools focus on a specific study design, while the MINORS is externally validated on RCTs by comparison with the CONSORT statement, making it a suitable instrument for meta-analyses of different study designs. The MINORS score ranges from 0 – 24; a higher score represents better methodological quality. Further details on the MINORS criteria and scoring system are provided in supplementary Table S2. Disagreements were resolved by involving a third reviewer (R.M.H.).

Outcome measures

The primary outcome measure was physical function as measured by the absolute Constant-Murley score[16] at least one year after initialization of either treatment. Normalized (sex- and age-adjusted) Constant-Murley scores were converted to absolute Constant-Murley scores using normal population-based values.[17] Secondary outcome measures were major reinterventions, nonunion, and avascular necrosis. If available, other functional outcome measures, such as the American Shoulder and Elbow Surgeons Shoulder Score[18] or the Neer score[19], were extracted as well.

Statistical analysis

Statistical analyses were performed using Review Manager (RevMan, Version 5.3.5. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). All continuous variables were converted to means and standard deviations (SD) when sufficient information was available using methods described in the Cochrane Handbook for Systematic Reviews of Interventions.[20]
All analyses were performed stratified by study design (i.e. RCTs and observational studies separately) as well as including both designs. Outcomes reported by two or more studies were pooled in a meta-analysis. Pooled effects of operative versus nonoperative treatment of dichotomous outcome measures were presented as risk ratios with confidence intervals (CI) using the Mantel-Haenszel method.[20] Pooled effects of continuous outcome measures were presented as mean differences with CI using the inverse variance weighting method.[20] Heterogeneity between studies was assessed by visual inspection of the forest plots and by estimating statistical measures for heterogeneity, i.e. the I^2 statistic and the Chi-square test. The main quantitative

assessment of heterogeneity was the I^2 statistic where the following interpretation was used: 0% to 40% might not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; and 75% to 100% considerable heterogeneity.[20] When heterogeneity was present a random-effects models was used instead of a fixed-effects model. Inspection of a funnel plot of the primary outcome measure against its standard error was done to detect potential publication bias.

Sensitivity analyses

Several sensitivity analyses were performed for study quality, year of publication, osteosynthesis by (locking) plate fixation and arthroplasty, and Neer classification. For the analysis of study quality only studies with an arbitrarily chosen MINORS score of 16 or higher were included, similar to previously published meta-analyses in orthopedic trauma surgery studying both study designs.[8,21] To assess the influence of the period in time in which the study was performed (and, consequently, development of different operative techniques), only studies published after 2005 were included in a separate analysis. Since the locking plate is the most commonly used type of osteosynthesis, another sensitivity analysis was conducted with studies where at least 80% of patients were treated with a locking plate. Furthermore, a sensitivity analysis was performed for all studies in which arthroplasty was the operative intervention. Finally, to explore the impact of fracture type on the functional outcome, a sensitivity analysis was performed including only Neer 3-part and 4-part fractures.

Different methods of meta-analysis may be differentially sensitive to studies with zero events on one or both study arms. Therefore, a sensitivity analysis to the choice of method of analysis was performed by means of the DerSimonian Laird method with correction and the inverse variance with and without correction for zero event data.[22]

Results

Figure 1 shows a flowchart of the literature search. In the end, 22 studies were included.[4,13,14,23–42] There were seven RCTs and 15 observational studies, of which nine were retrospective, four prospective, and two a combination of retrospective and prospective design.

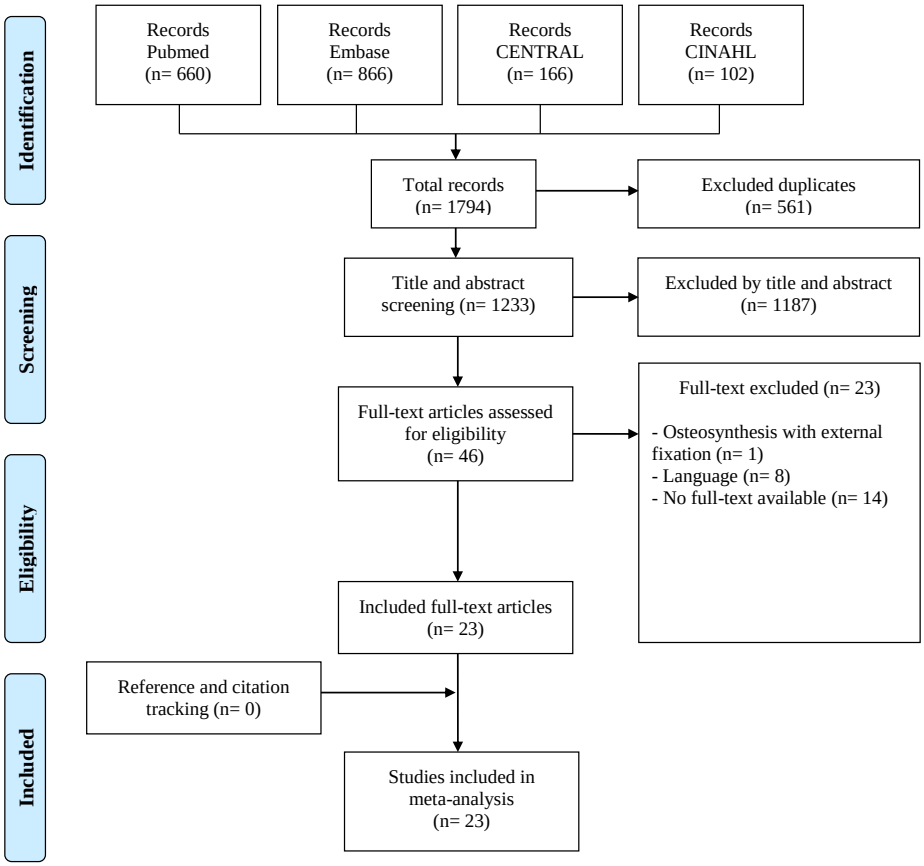

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram representing search and screen process of studies comparing operative versus nonoperative treatment of proximal humeral fractures. Central, Cochrane Central Register of Controlled Trials; CINAHL, Cumulative Index to Nursing and Allied Health Literature.

Quality assessment

The MINORS score for all included studies ranged from 12 to 22 with a median of 17.5 (IQR 14-21). The MINORS score ranged from 16 to 22 with a median of 21 (IQR 21-22) for RCTs and from 12 to 21 with a median of 16 (IQR 14-18) for observational studies. Study-specific MINORS scores are provided in supplementary Table S3. The MINORS criteria for unbiased assessment of study endpoints and prospective calculation of study size were rarely met.

Baseline characteristics of study participants

Details of the included studies and patients are provided in Table 1. The 22 studies included a total of 1743 patients for meta-analysis: 910 treated operatively and 833 nonoperatively. The weighted average age was 68.3 years, and 75% were female. Follow-up ranged from 12 to 86 months.

Table 1. Baseline characteristics of studies included in a meta-analysis of proximal humerus fractures comparing operative with nonoperative treatment

Study	Study design	Country	Fracture classification	Treatment groups	Number patients	Follow-up (months)	Age (years, range/SD)	Female/ Male
Boons 2012	RCT	Netherlands	Neer 4	Operative: arthroplasty	25	12	76,4 (5,6)	24/1
				Nonoperative: Sling	25	12	79,9 (7,7)	23/2
Fjalestad 2012-14*	RCT	Norway	AO type B2-C2	Operative: LP	25	24	72,2 (60-86)	20/5
				Nonoperative: Sling + closed reduction	25	24	73,1 (60-88)	24/1
Olerud 2011a	RCT	Sweden	Neer 4	Operative: hemiarthroplasty	27	24	75,8 (58-90)	23/4
				Nonoperative: Sling	28	24	77,5 (60-92)	24/4
Olerud 2011b	RCT	Sweden	Neer 3	Operative: LP	30	24	72,9 (56-92)	24/6
				Nonoperative: Sling	29	24	74,9 (58-88)	24/5
Rangan 2015	RCT	England	Neer 2,3,4	Operative: PHN, LP, TB, arthroplasty, screw	125	24	66,6 (11,8)	97/28
				Nonoperative: Sling or hanging cast	125	24	65,43 (12,09)	95/30
Stableforth 1984	RCT	England	Neer 4	Operative: arthroplasty	16	All: 18-144	65,6 (52-88)	12/4
				Nonoperative: Sling	16		70,1 (60-85)	13/3
Zyto 1997	RCT	Sweden	Neer 3,4	Operative: Tension band	20	50	73 (7,5)	18/2
				Nonoperative: Sling	20	50	75 (6,7)	17/3
Court-Brown 2001	PC	Scotland	Neer 2	Operative: PHN + tension band fixation	18	12	73	NR
				Nonoperative: Sling	31	12	78	
Hauschild 2013	PC	Germany	AO type A2, A3	Operative: PHN, LP	133	12	62,9 (17,2)	97/36
				Nonoperative: Sling	31	12	65,6 (13,3)	22/9
Innocenti 2013	PC	Italy	Neer 2,3,4	Operative: K-wire	23	All: 86	73,92 (6,01)	All: 38/13
				Nonoperative: Sling	19		77,47 (6,95)	
Noureai 2014	PC	Iran	Neer 2,3,4	Operative: LP, Tension band, K-wire	57	12	All: 52,9 (15,0)	All: 70/44
				Nonoperative: Sling	57	12		

*Fjalestad 2012 and 2014 were analyzed as one study as both described the same patient cohort. RC retrospective cohort study; RCT randomized controlled trial; PC prospective cohort study; NR not reported; TB tension band; PHN proximal humerus nail; LP locking plate

Table 1. Continued

Study	Study design	Country	Fracture classification	Treatment groups	Number patients	Follow-up (months)	Age (years, range, SD)	Female/ Male
Fjalestad 2005	RC+PC	Norway	AO type A,B,C	Operative: K-wire(n=4), LP(n=5), Screws(n=4), Screws + cerclage(n=2)	15	12	All: 70 (25-95)	All: 50/20
				Nonoperative: Sling	55	12		
Ilchman 1998	RC+PC	Sweden / Swiss	Neer 3,4	Operative: Tension band	18	63	61 (23-80)	13/5
				Nonoperative: Sling (n=10), Closed reduction(n=4), Open reduction(n=2)	16	70	70 (23-91)	13/3
Blonna 2009	RC	Italy	AO type A2,2	Operative: K-wire	42	32	73 (7,83)	20/12
				Nonoperative: Sling	37	35	75,1 (8,0)	26/9
vd Broek 2007	RC	Netherlands	Neer 4,5,6	Operative: PHN	27	16	64,6 (27-87)	NR
				Nonoperative: Sling	17	69	69,4 (35-84)	
Hageman 2016	RC	Netherlands	Neer 2,3,4	Operative: PHN (n=3); LP (n=23), K-wire (n=2), Screws (n=5)	33	37	59,0 (12,5)	22/11
				Nonoperative: Sling	33	70	60,1 (15,3)	24/9
Kollig 2003	RC	Germany	Neer 4,5,6	Operative: LP(n=2), Screw + cerclage(n=7), K-wire(n=4)	13	82	52,5 (14,7)	NR
				Nonoperative: Sling	9	76	52,7 (11,5)	
Lange 2016	RC	Germany	Neer 2,3,4	Operative: PHN	41	All: 55	69,1 (37-88)	35/6
				Nonoperative: hanging cast or dessault dressing	41		68,9 (42-93)	35/6
Okike 2015	RC	US	Neer 2,3,4	Operative: LP	61	All:40	All: 76,9	109/37
				Nonoperative: Sling	146			46/15
Roberson 2017	RC	US	Neer 3,4	Operative: reversed arthroplasty	20	53	All: 71	19/1
				Nonoperative: Sling	19	29	(52-88)	15/4
Sanders 2011	RC	Australia	Neer 2,3,4	Operative: LP	17	37	58 (14)	9/8
				Nonoperative: Sling	18	42	64 (15)	12/6
Tamimi 2015	RC	Canada	Neer 2,3,4	Operative: PHN(n=19), LP(n=44), K-wire(n=25)	88	26	All: 65,3 (15,2)	All: 57/31
				Nonoperative: Sling	25	28		

*Fjalestad 2012 and 2014 were analyzed as one study as both described the same patient cohort. RC retrospective cohort study; RCT randomized controlled trial; PC prospective cohort study; NR not reported; TB tension band; PHN proximal humerus nail; LP locking plate

All studies but one included displaced proximal humeral fractures in their study. The majority of the included studies excluded patients with pathological fractures, open fractures, fractures of the skeletally immature, and other sustained injury to the affected side. Most studies (n=18, 82%) used the Neer classification and included patients with a Neer 2,3 or 4-part proximal humeral fracture. In seven studies at least 80% of patients were treated with a locking plate.[13,14,25,27,33,35,41,42] Four studies investigated arthroplasty; three hemiarthroplasty and one reverse shoulder arthroplasty[29,30,32,34]; three studies assessed proximal humeral nails[4,28,39], and eight studied fixation by means of Kirschner wires, screws, tension band, or a combination of techniques.

Functional outcome

Fourteen studies (64%, n=817) reported the Constant-Murley score after at least one year of follow-up (supplementary Table S4).[13,14,37,39,40,42,23,25–28,32–34] In patients with a proximal humeral fracture, the functional outcome as measured by the Constant-Murley score showed no difference in operative versus nonoperative treatment with a mean difference of -0.87 (CI, -5.13 – 3.38; P=0.69; I^2=69%) (Figure 2). Pooled effects of RCTs were similar to those of observational studies for all outcome measures (Table 2). Figure 3 shows a funnel plot of the mean difference and standard error of the included studies using the Constant-Murley score; there was no important asymmetry observed.

For studies that did not use the Constant-Murley score, we performed additional analysis with the standardized mean difference of different functional outcome measures which yielded the same result as the primary analysis (SMD -0.06; CI, -0.25 – 0.12; P=0.52; I^2=53%) (supplementary Figure S1). Seven studies (n=327) reported functional outcome of patients treated with a Neer 3-part or 4-part fracture.[14,28,29,31–34] Forty-three percent of patients with Neer 4-part fractures were initially treated with arthroplasty (Table 1). A subgroup analysis of these studies showed no difference in standardized mean difference of functional outcome measures between operative and nonoperative treatment with a mean difference of 0.02 (CI, -0.20 – 0.24; P=0.86; I^2=0%) (supplementary Figure S2).

Major reinterventions

Fifteen studies (68%, n=938) reported on major reinterventions (supplementary Table S4).[13,14,34–36,39–41,23,25,27–30,32,33] Two studies had no major reintervention in either treatment arm

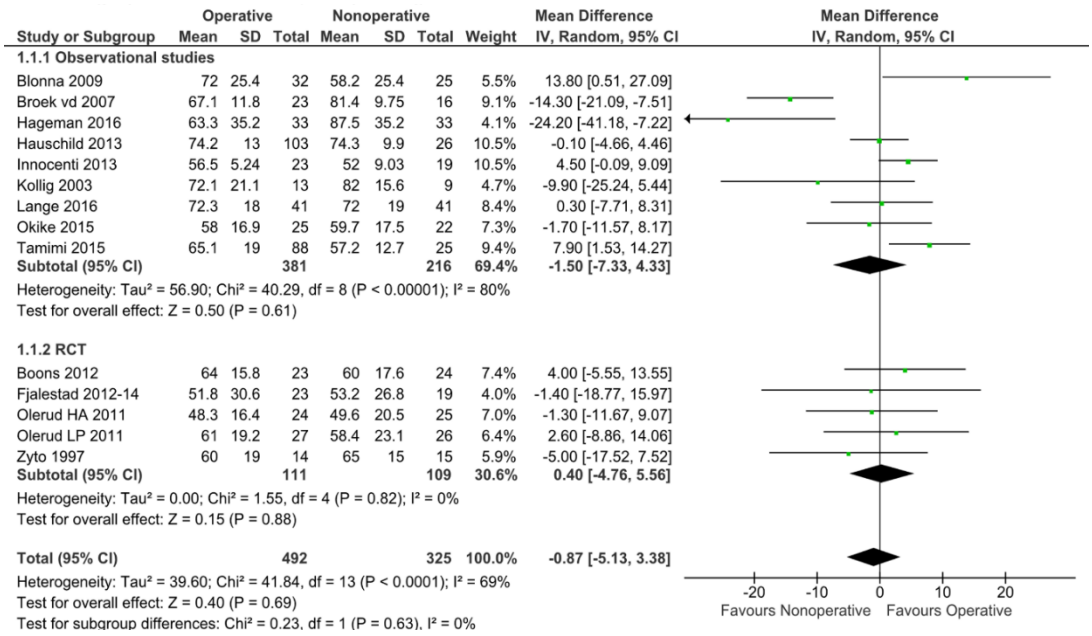

Figure 2. Functional outcome as measured with Constant-Murley score in systematic review of proximal humeral fractures comparing operative versus nonoperative treatment. *SD*, standard deviation; *IV*, inverse variance; *CI*, confidence interval; *RCT*, randomized controlled trial.

at follow-up. Major reinterventions occurred more often in the operative group than the nonoperative group with a risk ratio (RR) of 2.72 (CI, 1.71 – 4.34; P< 0.0001; I^2=0%) (supplementary Figure S3). Using different methods of incorporating studies in the meta-analysis with zero event data in one or both arms yielded similar results (supplementary Table S5). Implant removal was reported in 10 studies (45%). The mean percentage of implant removal across studies was 21% (range 0–100%). When stratified by study design, observational studies showed a greater risk for major reinterventions in the operative treatment group compared with the nonoperative group (RR 5.43; CI 2.51–11.74; P< 0.0001; I^2=0%) (Table 2). Five studies specified their reinterventions for nonoperatively treated patients: four patients received arthroplasty for displacement and malunion, two patients received ORIF for displacement, and two patients received acromioplasty for impingement complaints.

Nonunion

A total of thirteen studies (59%) reported on nonunion (supplementary Table S4). Operative treatment of proximal humeral fractures resulted in fewer nonunion than nonoperative treatment with a RR of 0.45 (CI, 0.23–0.89; P=0.02; I^2=0%) (supplementary Figure S4). When stratified by study design, both subgroups showed a similar, non-significant, pooled effect (Table 2).

Table 2. Subgroup & sensitivity analyses of studies included in a meta-analysis of proximal humerus fractures comparing operative to nonoperative treatment

Analysis description	CS				MR				Nonunion				AVN			
	n	MD	(95% CI)	P-value	n	RR	(95% CI)	P-value	n	RR	(95% CI)	P value	n	RR	(95% CI)	P-value
All studies	14	-0.87	(-5.13; 3.38)	0.69	15	2.72	(1.71; 4.34)	< 0.0001	13	0.45	(0.23; 0.89)	0.02	13	1.24	(0.87; 1.77)	0.24
Subgroup analysis																
RCTs	5	0.40	(-4.76; 5.56)	0.88	6	1.45	(0.78; 2.70)	0.25	6	0.48	(0.19; 1.20)	0.12	6	0.88	(0.55; 1.41)	0.59
Observational studies	9	-1.50	(-7.33; 4.33)	0.61	7	5.43	(2.51; 11.74)	< 0.0001	7	0.41	(0.15; 1.16)	0.09	7	1.93	(1.11; 3.37)	0.02
Sensitivity analysis																
High-quality studies	11	0.55	(-2.93; 4.03)	0.76	11	2.52	(1.55; 4.11)	0.0002	11	0.44	(0.21; 0.93)	0.03	10	1.14	(0.74; 1.74)	0.55
Studies after 2005	12	-0.14	(-4.65; 4.38)	0.95	14	2.58	(1.59; 4.20)	0.0001	12	0.41	(0.18; 0.89)	0.03	10	1.10	(0.72; 1.69)	0.65
Locking plate	5	-0.15	(-0.43; 0.13)	0.30	7	1.81	(1.04; 3.16)	0.04	6	0.37	(0.12; 1.17)	0.09	6	1.35	(0.86; 2.11)	0.19
Arthroplasty	2	1.50	(-5.24; 8.23)	0.66	4	2.66	(0.72; 9.77)	0.14	3	0.52	(0.13; 1.99)	0.34	2	0.17	(0.02; 1.37)	0.10

N Number; CS Constant score MR Major reintervention; AVN avascular necrosis; RCT randomized controlled trial; RR risk ratio; MD mean difference; CI confidence interval; Sensitivity analysis of locking plate includes studies comparing locking plate to nonoperative treatment; Sensitivity analysis of arthroplasty includes studies comparing hemiarthroplasty and reversed arthroplasty to nonoperative treatment

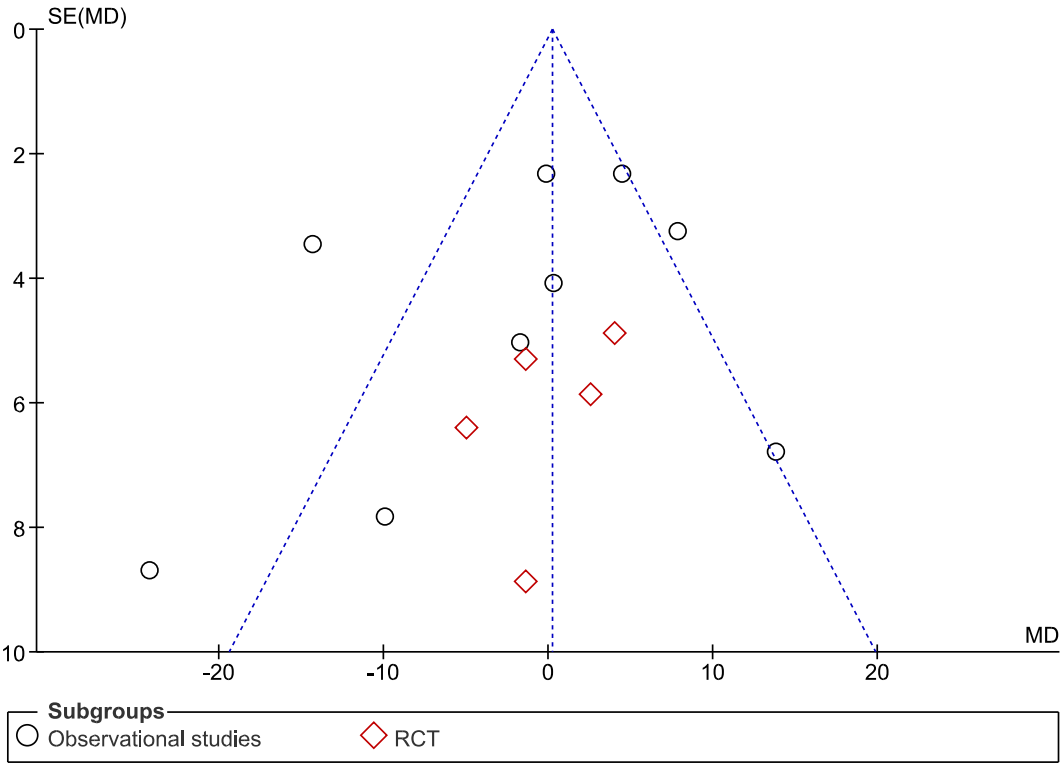

Figure 3. Funnel plot of studies included in meta-analysis reporting Constant-Murley scores after operative or nonoperative treatment of proximal humeral fractures. *SE*, standard error; *MD*, mean difference; *RCT*, randomized controlled trial.

Avascular necrosis

A total of thirteen studies (59%) reported on avascular necrosis (supplementary Table S4). There was no difference in the rate of avascular necrosis between operative and nonoperative treatment for proximal humeral fractures with a RR of 1.24 (CI, 0.87–1.77; P=0.24; I^2=24%) (supplementary Figure S5). When stratified by study design, observational studies showed a higher risk of avascular necrosis for the operative group compared with the nonoperative group (RR 1.93; CI 1.11–3.37; P=0.02; I^2=9%) (Table 2).

Sensitivity analysis

Sensitivity analysis did not significantly alter the primary and secondary outcome measures (Table 2).

Discussion

In this systematic review and meta-analysis of patients with displaced proximal humeral fractures, there was no difference in physical function as measured with the Constant-Murley score after operative or nonoperative treatment. Subgroup analysis for Neer 3-part or 4-part fractures

neither showed differences in functional outcome. Results of the primary and secondary outcome measures were similar from the pooled effects of RCTs and observational studies. There was a higher risk for major reinterventions and a lower risk of nonunion after operative treatment compared with nonoperative treatment. This the largest meta-analysis in the current literature by including both RCTs and observational studies.

Compared with nonoperative treatment, there is no improved in functional outcome after operative treatment for displaced proximal humeral fractures, which confirms findings of previous meta-analyses.[6,43] A recent systematic review of displaced proximal humeral fractures is based on only 7 RCTs with just over 500 patients.[6] With a total of 250 patients, the PROFHER trial represents the most substantial evidence currently available.[35] The patient demographic characteristics of the PROFHER trial are comparable with those of the included studies in this meta-analysis (Table 1). However, only 4.4% of patients in the PROFHER trial suffered a Neer 4-part fracture compared with 21% of patients in this meta-analysis. Therefore, compared with previous, smaller magnitude meta-analyses, this review contributes substantially to the current evidence and enables recommendations for a broader patient population. Furthermore, this is the first meta-analysis in which subgroup analysis for Neer 3-part and 4-part fractures was possible, and the results showed no differences in operative versus nonoperative treatment.

This review showed similar pooled effects of observational studies and RCTs for the primary and secondary outcome measures. This finding is similar to previous meta-analyses in orthopedic trauma surgery including both study designs.[7–10,44] As such, this review speaks to the growing potential of observational studies in orthopedic trauma surgery and contributes to the expanding discussion about the value of different study designs.[45]

In this review, the major reintervention rate included every additional surgery except for implant removal because of patient preference, implant-related irritation, or a stiff shoulder. Therefore, the major reintervention rate in this review is a surrogate marker for severe complications (e.g. implant failure, deep infection, nonunion, impingement, or avascular necrosis) after operative and nonoperative treatment of displaced proximal humeral fractures. This is the first review to show significantly more severe complications requiring surgical re-intervention after operative treatment of displaced proximal humeral fractures. These procedures add up to the additional surgery for implant removal for 21% of the patients for a less serious indication.

Another new finding is the higher risk of nonunion for nonoperatively treated patients. RCTs and observational studies alone were not able to detect a significant difference in this outcome. This demonstrates the added value of increasing study power by including observational studies in order to detect rare outcomes. It is important to note that this difference is supported by the sensitivity analysis including only high-quality RCTs and observational studies (Table 2).

This review found no difference in the rate of avascular necrosis between the nonoperative and operative management. However, it should be noted that three of the 15 studies reporting on avascular necrosis had a follow-up of 12 months while avascular necrosis can be detected up to two years of follow-up. For this outcome measure, the pooled effect of observational studies was significantly different than the pooled effect of RCTs. However, in the sensitivity analysis with high quality studies, this contrasting result did not yield, and pooled effects of both study designs were similar again. This demonstrates the importance of evaluating the quality of the included studies (Table 2). Therefore, including a study in a meta-analysis should be based on the quality of the study regardless of the study design.[44] Generally, RCTs will be of higher quality and thus included for analysis, however, a high-quality observational study should be chosen over a low-quality RCT.

The results of this systematic review and meta-analysis should be interpreted in the light of several limitations. First, the results of the meta-analysis may be influenced by missed studies in the database search or by publication bias. However, an extensive search was performed using multiple databases, and the citations and references of included studies were also screened. Furthermore, a funnel plot of the primary outcome measure did not suggest possible bias due to selective publication. Second, results of observational studies are more heterogeneous than those of RCTs in the meta-analysis of the Constant-Murley score. Still, it should be noted despite heterogeneity in mean differences of the observational studies, the observed effects all are within a range of the Constant-Murley score which is clinically nonimportant.[46] Third, in the analysis of functional outcome, we did not distinguish between 12 or more than 12 months of follow-up since prior studies have shown the greatest increase in functional outcome takes place in the first six months and no significant improvement is to be expected after 12 months.[4,14,32,33,35] This is further supported by an additional sensitivity analysis that showed no differences in functional outcome at 12 months and at 24 or more months. Fourth, the Neer classification for proximal humeral fractures is the most frequently used

classification system in the literature even though it has been considered to have important limitations. However, no other system for evaluating these fractures is consistently more reliable than the Neer classification.[47] Fifth, The majority of the included studies were European, and only three studies described patients from Northern America, let alone other continents. However, subgroup analyses revealed no differences for the primary and secondary outcome measures between these continents (data not shown). Finally, it should be noted that the majority of studies in this review excluded patients with pathological fractures, patients with open fractures, fractures of skeletal immature patients, and patients with other sustained injuries to the affected arm. As a result, recommendations from this review are not applicable to these patients.

Although we acknowledge the vast amount of existing systematic reviews on this topic[6,43,48,49], we believe that the several unique qualities of this meta-analysis contribute to the existing knowledge. Strengths of this study include the consistent results of the different sensitivity analyses for time of publication, type of osteosynthesis, and arthroplasty. Furthermore, by including observational studies in addition to the highly selective patient population of RCTs, the analyzed patients may be more representative of patients encountered in daily clinical practice and improve the generalizability of our results. We also demonstrated that the findings were consistent across study designs with respect to different outcome measures. Although no subgroup analysis of elderly patients (aged > 65 years) could be performed, the mean age of all patients in this review was 68 years, with a relatively small standard deviation for the majority of the included studies; therefore, we feel confident that recommendations from this review apply to the average elderly patient. Finally, this is the largest meta-analysis in the literature with the highest number of patients available for analysis of proximal humeral fractures.

Conclusion

We recommend nonoperative treatment for the average elderly patient (aged > 65 years) with a displaced proximal humeral fracture. Pooled effects of observational studies were similar to those of RCTs, and the inclusion of observational studies improved the generalizability of findings.

References:

1. Palvanen M, Kannus P, Niemi S, Parkkari J. Update in the epidemiology of proximal humeral fractures. *Clin Orthop Relat Res*. 2006;442:87-92.
2. Roux A, Decroocq L, El Batti S, et al. Epidemiology of proximal humerus fractures managed in a trauma center. *Orthop Traumatol Surg Res*. 2012;98(6):715-719.
3. Launonen AP, Lepola V, Saranko A, Flinkkilä T, Laitinen M, Mattila VM. Epidemiology of proximal humerus fractures. *Archives of Osteoporosis*. 2015;10(1):2.
4. Court-Brown CM, Garg A, McQueen MM. The translated two-part fracture of the proximal humerus. Epidemiology and outcome in the older patient. *Journal of Bone and Joint Surgery - Series B*. 2001;83(6):799-804.
5. Jawa A, Burnikel D. Treatment of Proximal Humeral Fractures. *JBJS Reviews*. 2016;4(June):1-9.
6. Handoll HH, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev*. 2015;(11):CD000434.
7. Smeeing DPJ, van der Ven DJC, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older: A Systematic Review, Meta-analysis, and Comparison of Randomized Controlled Trials and Observational Studies. *Am J Sports Med*. 2016.
8. Houwert RM, Smeeing DP, Ahmed Ali U, Hietbrink F, Kruyt MC, van der Meijden OA. Plate fixation or intramedullary fixation for midshaft clavicle fractures: a systematic review and meta-analysis of randomized controlled trials and observational studies. *J Shoulder Elbow Surg*. 2016;25(7):1195-1203.
9. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *J Clin Epidemiol*. 2010;63(3):238-245.
10. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database Syst Rev*. 2014;(4):MR000034.
11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6(7):e1000100.
12. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of Observational Studies in Epidemiology<SUBTITLE>A Proposal for Reporting</SUBTITLE> *JAMA*. 2000;283(15):2008.
13. Fjalestad T, Hole MO, Hovden IA, Blucher J, Stromsoe K. Surgical treatment with an angular stable plate for complex displaced proximal humeral fractures in elderly patients: a randomized controlled trial. *J Orthop Trauma*. 2012;26(2):98-106.
14. Fjalestad T, Hole MO. Displaced proximal humeral fractures: operative versus non-operative treatment--a 2-year extension of a randomized controlled trial. *Eur J Orthop Surg Traumatol*. 2014;24(7):1067-1073.
15. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg*. 2003;73(9):712-716.
16. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res*. 1987;(214):160-164.
17. Constant CR, Gerber C, Emery RJ, Sojbjerg JO, Gohlke F, Boileau P. A review of the Constant score: modifications and guidelines for its use. *J Shoulder Elbow Surg*. 2008;17(2):355-361.
18. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: Reliability, validity, and responsiveness. *Journal of Shoulder and Elbow Surgery*. 2002;11(6):587-594.
19. Neer CS. Displaced Proximal Humeral Fractures. *J Bone Joint Surg Am*. 1970;52(6):1077-1089. 20. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley; 2011.
21. Smeeing DP, van der Ven DJ, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older. *Am J Sports Med*. 2016:363546516673615.
22. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med*. 2007;26(1):53-77.
23. Innocenti M, Carulli C, Civinini R, Matassi F, Tani M, Muncibi F. Displaced fragility fractures of proximal humerus in elderly patients affected by severe comorbidities: percutaneous fixation and conservative treatment. *Aging Clin Exp Res*. 2013;25(4):447-452.
24. Nouraei MH, Majd DA, Zamani F. Comparing the treatment results of proximal humerus fracture based on surgical or nonsurgical methods. *Adv Biomed Res*. 2014;3:253.
25. Okike K, Lee OC, Makanji H, Morgan JH, Harris MB, Vrahas MS. Comparison of locked plate fixation and nonoperative management for displaced proximal humerus fractures in elderly patients. *Am J Orthop (Belle Mead NJ)*. 2015;44(4):E106-12.
26. Tamimi I, Montesa G, Collado F, et al. Displaced proximal humeral fractures: when is surgery necessary? *Injury*. 2015;46(10):1921-1929.
27. Hageman MGJS, Meijer D, Stufkens SA, Ring D, Doornberg JN, Steller EP. Proximal humeral fractures: Nonoperative versus operative treatment. *Arch Trauma Res*. 2017;6(1).

28. Lange M, Brandt D, Mittlmeier T, Gradl G. Proximal humeral fractures: non-operative treatment versus intramedullary nailing in 2-, 3- and 4-part fractures. *Injury*. 2016;47 Suppl 7:S14-s19.
29. Roberson TA, Granade CM, Hunt Q, et al. Nonoperative management versus reverse shoulder arthroplasty for treatment of 3- and 4-part proximal humeral fractures in older adults. *Journal of Shoulder and Elbow Surgery*. 2017;26(6):1017-1022.
30. Stableforth PG. Four-part fractures of the neck of the humerus. *J Bone Joint Surg Br*. 1984;66(1):104-108.
31. Zyto K, Törnkvist CH, Ahrengart L, et al. Treatment of displaced proximal humeral fractures in elderly patients. *J Bone Joint Surg Br*. 1997;79(3):412-417.
32. Olerud P, Ahrengart L, Ponzer S, Saving J, Tidermark J. Hemiarthroplasty versus nonoperative treatment of displaced 4-part proximal humeral fractures in elderly patients: a randomized controlled trial. *J Shoulder Elbow Surg*. 2011;20(7):1025-1033.
33. Olerud P, Ahrengart L, Ponzer S, Saving J, Tidermark J. Internal fixation versus nonoperative treatment of displaced 3-part proximal humeral fractures in elderly patients: a randomized controlled trial. *J Shoulder Elbow Surg*. 2011;20(5):747-755.
34. Boons HW, Goosen JH, van Grinsven S, van Susante JL, Van Loon CJ. Hemiarthroplasty for humeral four-part fractures for patients 65 years and older: a randomized controlled trial. *Clin Orthop Relat Res*. 2012;470(12):3483-3491.
35. Rangan A, Handoll H, Brealey S, et al. Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. *JAMA*. 2015;313(10):1037.
36. Ilchmann T, Ochsner PE, Wingstrand H, Jonsson K. Non-operative treatment versus tension-band osteosynthesis in three- and four-part proximal humeral fractures. A retrospective study of 34 fractures from two different trauma centers. *Int Orthop*. 1998;22(5):316-320.
37. Kollig E, Kutscha-Lissberg F, Roetman B, Dielenschneider D, Muhr G. [Complex fractures of the humeral head: which long-term results can be expected?]. *Zentralbl Chir*. 2003;128(2):111-118.
38. Fjalestad T, Strømsøe K, Blücher J, et al. Fractures in the proximal humerus: functional outcome and evaluation of 70 patients treated in hospital. *Archives of orthopaedic and trauma surgery*. 2005;125(5):310-316.
39. Van-den-Broek CM, van den Besselaar M, Coenen JMF, Vegt PA. Displaced proximal humeral fractures: Intramedullary nailing versus conservative treatment. *Arch Orthop Trauma Surg*. 2007;127(6):459-463.
40. Blonna D, Rossi R, Fantino G, Maiello A, Assom M, Castoldi F. The impacted varus (A2.2) proximal humeral fracture in elderly patients: is minimal fixation justified? A case control study. *J Shoulder Elbow Surg*. 2009;18(4):545-552.
41. Sanders RJ, Thissen LG, Teepen JC, van Kampen A, Jaarsma RL. Locking plate versus nonsurgical treatment for proximal humeral fractures: better midterm outcome with nonsurgical treatment. *J Shoulder Elbow Surg*. 2011;20(7):1118-1124.
42. Hauschild O, Konrad G, Audige L, et al. Operative versus non-operative treatment for two-part surgical neck fractures of the proximal humerus. *Arch Orthop Trauma Surg*. 2013;133(10):1385-1393.
43. Xie L, Ding F, Zhao Z, Chen Y, Xing D. Operative versus non-operative treatment in complex proximal humeral fractures: a meta-analysis of randomized controlled trials. *Springerplus*. 2015;4:728.
44. Smeeing DPJ, Houwert RM, Kruyt MC, van der Meijden OAJ, Hietbrink F. Clinical research on postoperative trauma care: has the position of observational studies changed? *European journal of trauma and emergency surgery : official publication of the European Trauma Society*. 2017;43(1):43-51.
45. Frieden TR. Evidence for Health Decision Making — Beyond Randomized, Controlled Trials. Drazen JM, Harrington DP, McMurray JJV, Ware JH, Woodcock J, eds. *New England Journal of Medicine*. 2017;377(5):465-475.
46. van de Water AT, Shields N, Davidson M, Evans M, Taylor NF. Reliability and validity of shoulder function outcome measures in people with a proximal humeral fracture. *Disabil Rehabil*. 2014;36(13):1072-1079.
47. Carofino BC, Leopold SS. Classifications in brief: The neer classification for proximal humerus fractures. *Clinical Orthopaedics and Related Research*. 2013;471(1):39-43.
48. Fu T, Xia C, Li Z, Wu H. Surgical versus conservative treatment for displaced proximal humeral fractures in elderly patients: a meta-analysis. *International journal of clinical and experimental medicine*. 2014;7(12):4607-4615.
49. Rabi S, Evaniew N, Sprague SA, Bhandari M, Slobogean GP. Operative *vs* non-operative management of displaced proximal humeral fractures in the elderly: A systematic review and meta-analysis of randomized controlled trials. *World Journal of Orthopedics*. 2015;6(10):838.

Supplementary materials to Chapter 3

Table S1. Search syntax performed last on March 30, 2017

Database	Syntax
PubMed/MEDLINE (n= 660)	(Humeral Fractures[MeSH Terms] OR Shoulder Fractures[MeSH Terms] OR ((humeral[Title/Abstract] OR humerus[Title/Abstract] OR humeri[Title/Abstract] OR humor[Title/Abstract] OR (upper[Title/Abstract] AND arm[Title/Abstract] AND bone[Title/Abstract]) OR (upperarm[Title/Abstract] AND bone[Title/Abstract])) AND fractur*[Title/Abstract])) AND (proximal[Title/Abstract] OR sub-capital[Title/Abstract] OR subcapital[Title/Abstract] OR neck[Title/Abstract]) AND (surgery[subheading] OR Fracture Healing[MeSH Terms] OR Fracture Fixation[MeSH Terms] OR Surgical Procedures, Operative[MeSH Terms] OR orthopedics[MeSH Terms] OR orthopedics[Title/Abstract] OR orthopaedics[Title/Abstract] OR orthopedic[Title/Abstract] OR orthopaedic[Title/Abstract] OR surgery[Title/Abstract] OR surgical[Title/Abstract] OR operative[Title/Abstract] OR operate[Title/Abstract] OR operating[Title/Abstract] OR operated[Title/Abstract] OR operation[Title/Abstract]) AND (conservative[Title/Abstract] OR conventional[Title/Abstract] OR non-operative[Title/Abstract] OR non-surgical[Title/Abstract] OR non surgical[Title/Abstract] OR nonoperative[Title/Abstract] OR Physical Therapy Modalities[MeSH Terms] OR sling[Title/Abstract] OR collar[Title/Abstract] OR cuff[Title/Abstract] OR bandages[Title/Abstract] OR bandage[Title/Abstract])
Embase (n= 866)	('humerus'/exp OR humerus:ti,ab OR humeri:ti,ab OR humer:ti,ab OR humor:ti,ab OR 'corpus humeri':ti,ab OR 'upper arm bone':ti,ab OR 'upperarm bone':ti,ab OR humeral:ti,ab) AND ('fracture'/exp OR fracture:ti,ab OR fractured:ti,ab OR fractures:ti,ab) AND (proximal:ti,ab OR 'sub capital':ti,ab OR 'subcapital':ti,ab OR neck:ti,ab) AND ('surgery'/exp OR surgery:ti,ab OR surgical:ti,ab OR operative:ti,ab OR operation:ti,ab OR 'Fracture Healing':ti,ab OR 'Fracture fixation':ti,ab OR 'Surgical Procedures':ti,ab OR orthopedics:ti,ab OR orthopedic:ti,ab OR orthopaedics:ti,ab OR orthopaedic:ti,ab OR operate:ti,ab OR operating:ti,ab OR operated:ti,ab) AND ('conservative treatment'/exp OR 'conservative treatment':ti,ab OR conservative:ti,ab OR conventional:ti,ab OR 'non-operative':ti,ab OR nonoperative:ti,ab OR non-surgical:ti,ab OR 'non surgical':ti,ab OR sling:ti,ab OR collar:ti,ab OR cuff:ti,ab OR bandages:ti,ab OR bandage:ti,ab)
CENTRAL (n= 166)	humerus AND fracture AND (proximal OR neck OR sub capital OR subcapital)
CINAHL (n= 102)	(humerus OR humeri OR humer OR humor OR corpus humeri OR upper arm bone OR upperarm bone OR humeral) AND (fracture OR fractured OR fractures) AND (proximal OR sub capital OR neck OR subcapital) AND (surgery OR surgical OR operative OR operation OR Fracture Healing OR Fracture fixation OR Surgical Procedures OR orthopedics OR orthopedic OR orthopaedics OR orthopaedic OR operate OR operating OR operated) AND (conservative treatment OR conservative OR conventional OR non-operative OR nonoperative OR non-surgical OR non surgical OR sling OR collar OR cuff OR bandages OR bandage)

Table S2. Quality assessment according to the MINORS criteria in a meta-analysis of proximal humeral fractures

Criteria	Reported and adequate (2)	Reported but inadequate (1)	Not reported (0)
Clearly stated aim	Aim including outcomes reported	Aim reported without outcomes	Not reported
Inclusion consecutive patients	Inclusion/exclusion criteria reported	Unclear description inclusion/exclusion criteria	Not reported
Prospective collection data	Prospective	Retrospective	Not reported
Appropriate endpoints	Appropriate endpoints to aim study	Endpoints not appropriate to aim study	Not reported
Unbiased assessment	Blinded evaluation of outcomes	Reason not blinding stated	Not reported
Appropriate follow-up	≥ 1 year	< 1 year	Not reported
Loss to follow-up < 5%	≤ 5%	> 5% and ≤ 20%	Not reported/>20%
Prospective calculation study size	Power-analysis performed	Explanation number without power-analysis	Not reported
Adequate control group	Operative versus nonoperative treatment	Not applicable	Not reported
Contemporary groups	Study/control group managed during same period	Study/control not managed during same period	Not reported
Baseline equivalence groups	Baseline characteristics described and comparable	Baseline characteristics not comparable	Not reported
Adequate statistical analyses	Statistical analysis described including type of analyses	Inadequate description statistical analysis	Not reported

Items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S3. Quality assessment of all included studies in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment

Criteria	Blonna 2009	Boons 2012	vd Broek 2007	Court-Brown 2001	Fjalestad 2005	Fjalestad 2012-14	Hageman 2016	Hauschild 2013	Ilchman 1998	Innocenti 2013	Kollig 2003	Lange 2016	Noureai 2014	Okike 2015	Olerud 2011a	Olerud 2011b	Rangan 2015	Roberson 2017	Sanders 2011	Stableforth 1984	Tamimi 2015	Zyto 1997
Clearly stated aim	2	2	2	1	1	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
Inclusion of consecutive patients	2	2	2	2	1	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2
Prospective collection of data	0	2	0	2	1	2	0	2	0	2	0	0	2	0	2	2	2	0	0	2	0	2
Appropriate endpoints	2	2	2	2	1	2	2	2	1	2	2	2	1	2	2	2	2	2	2	1	2	2
Unbiased assessment endpoints	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	0	1	0	0	1
Appropriate follow-up	2	2	1	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2
Loss to follow-up < 5%	1	1	1	0	0	2	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	2
Prospective calculation study size	0	2	0	0	0	1	0	1	0	0	0	0	0	0	1	1	2	0	0	0	0	0
Adequate control group	2	2	2	2	2	2	2	2	1	2	1	2	2	2	2	2	2	2	2	2	1	2
Contemporary groups	2	2	1	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2
Baseline equivalence of groups	2	2	1	2	1	2	2	2	2	2	0	2	0	1	2	2	2	2	2	2	1	2
Adequate statistical analysis	2	2	0	2	2	2	2	2	1	2	0	2	1	2	2	2	2	2	2	0	2	2
Total MINORS score	**18**	**22**	**13**	**18**	**14**	**22**	**17**	**21**	**14**	**19**	**12**	**16**	**13**	**16**	**21**	**21**	**22**	**16**	**18**	**16**	**14**	**21**

Items are scored 0 (not reported/ not applicable), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S4. Outcome measures in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment

Study		Constant score (± SD)	Revision surgery	Non-union	AVN	DASH score (± SD)	Implant removal
Blonna 2009	Operative	72 (25.4)	0	0	NR	15.0 (3.0)	32
	Nonoperative	58.2 (25.4)	0	0		30.5 (5.1)	
Boons 2012	Operative	64 (15.8)	1	2	0	NR	0
	Nonoperative	60 (17.6)	1	3	2		
vd Broek 2007	Operative	67.1 (11.8)	3	0	0	NR	5
	Nonoperative	81.4 (9.8)	0	1	0		
Court-Brown 2001	Operative	NR	NR	1	NR	NR	NR
	Nonoperative			4			
Fjalestad 2005	Operative	NR	NR	1	3	NR	NR
	Nonoperative			5	2		
Fjalestad 2012-14*	Operative	51.8 (30.6)	1	1	12	NR	7
	Nonoperative	53.2 (26.8)	1	2	15		
Hageman 2016	Operative	63.3 (35.2)	5	NR	1	22 (13.9)	2
	Nonoperative	87.5 (35.2)	2		0	10.3 (13.9)	
Hauschild 2013	Operative	74.2 (13)	NR	1	1	NR	NR
	Nonoperative	74.3 (9.9)		0	0		
Ilchman 1998	Operative	NR	4	NR	9	NR	NR
	Nonoperative		1		7		
Innocenti 2013	Operative	56.5 (5.2)	0	NR	0	NR	23
	Nonoperative	52 (9.0)	0		0		
Kollig 2003	Operative	72.1 (21.1)	NR	NR	NR	NR	NR
	Nonoperative	82 (15.6)					
Lange 2016	Operative	72.3 (18)	13	NR	NR	NR	NR
	Nonoperative	72 (19)	0				

*In this analysis Fjalestad 2012 and 2014 were seen as one study as both studies describe the same patient cohort AVN avascular necrosis; NR not reported; SD standard deviation

Table S4. Continued

Study		Constant score (± SD)	Revision surgery	Non-union	AVN	DASH score (± SD)	Implant removal
Noureai 2014	Operative	NR	NR	NR	NR	NR	NR
	Nonoperative						
Okike 2015	Operative	58 (16.9)	8	0	10	26.5 (17.8)	NR
	Nonoperative	59.7 (17.5)	2	2	3	25.1 (18.2)	
Olerud 2011a	Operative	48.3 (16.4)	2	0	0	30.2 (18.3)	1
	Nonoperative	49.6 (20.5)	1	1	3	36.9 (21.3)	
Olerud 2011b	Operative	61 (19.2)	4	1	3	26.4 (25.2)	5
	Nonoperative	58.4 (23.1)	1	1	2	35 (26.8)	
Rangan 2015	Operative	NR	11	0	4	NR	NR
	Nonoperative		11	5	1		
Roberson 2017	Operative	NR	3	NR	NR	NR	0
	Nonoperative		0				
Sanders 2011	Operative	NR	3	0	8	NR	7
	Nonoperative		0	1	5		
Stableforth 1984	Operative	NR	1	NR	NR	NR	1
	Nonoperative		0				
Tamimi 2015	Operative	65.1 (19)	NR	NR	NR	33 (21.8)	NR
	Nonoperative	57.2 (12.7)				38.4 (19.2)	
Zyto 1997	Operative	60 (19)	NR	1	1	NR	1
	Nonoperative	65 (15)		0	0		

*In this analysis Fjalestad 2012 and 2014 were seen as one study as both studies describe the same patient cohort AVN avascular necrosis; NR not reported; SD standard deviation

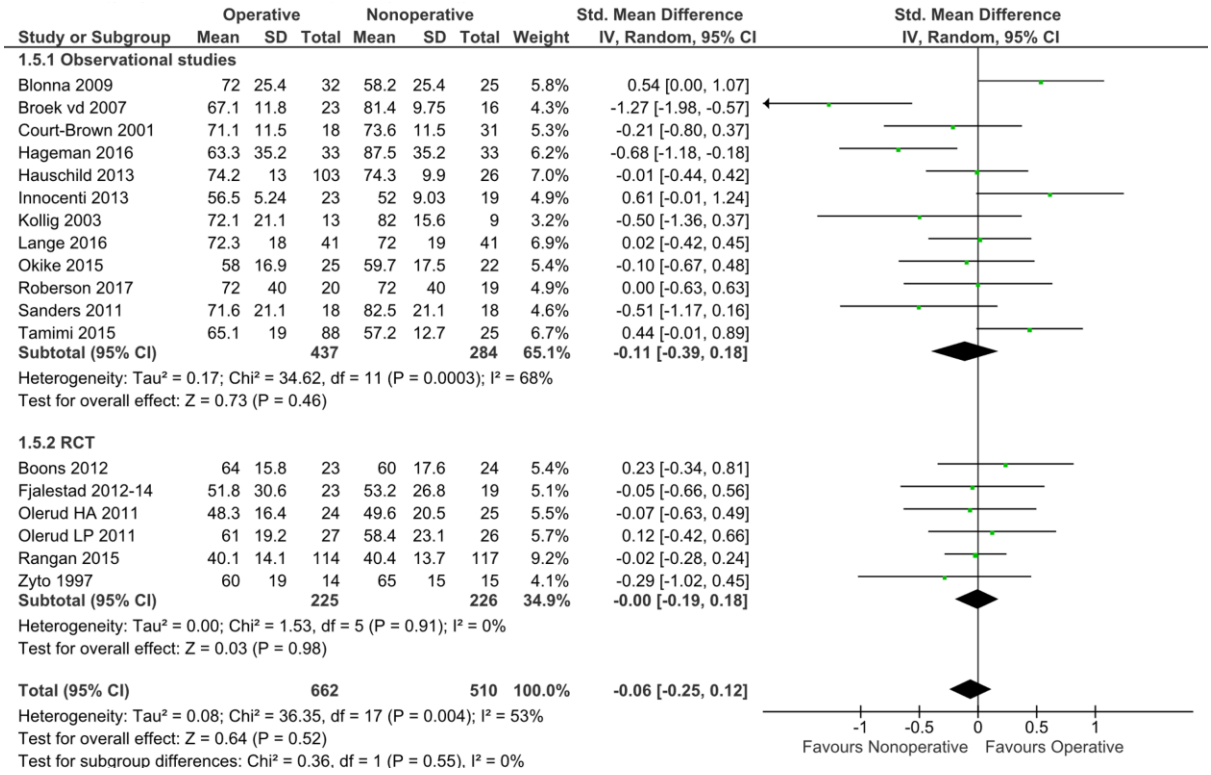

Figure S1. Standardized mean difference of functional outcome scores in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

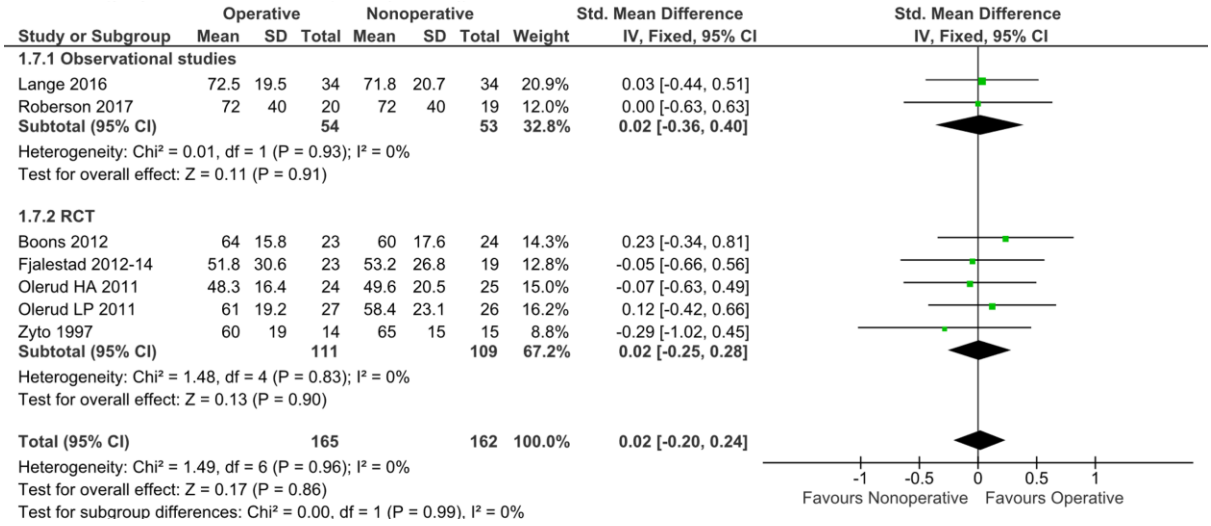

Figure S2. Subgroup analyses looking at standardized mean difference for functional outcome measures including only studies reporting on Neer 3-part or 4-part fractures in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

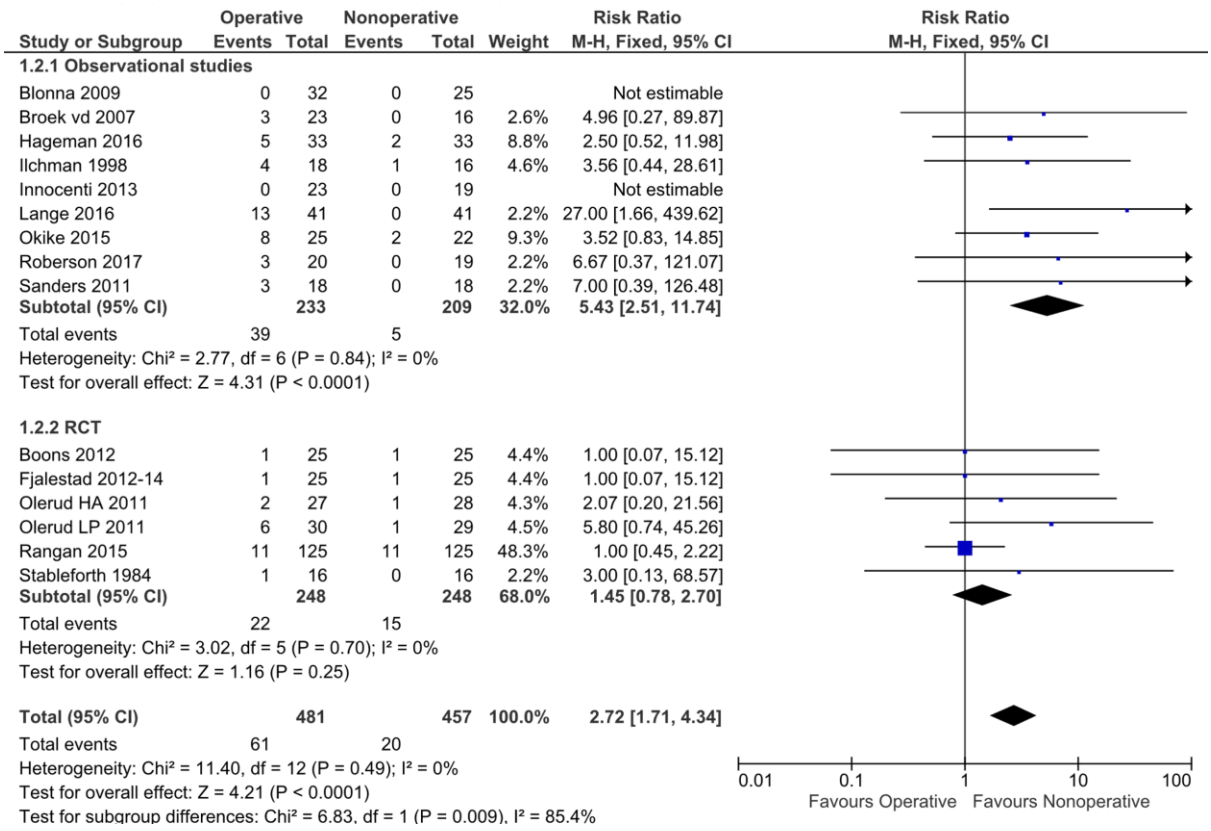

Figure S3. Revision surgery in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

Table S5. Impact of different methods to handle zero-event data

Method	Observational studies	RCT	Total
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Mantel-Haenzel*	5.46 (2.29, 13.01)	1.37 (0.85, 2.77)	2.32 (1.34, 4.02)
Inverse variance - no correction	3.76 (1.30, 10.91)	1.32 (0.64, 2.71)	1.83 (1.01, 3.33)
Inverse variance - with correction	4.64 (2.03, 10.62)	1.37 (0.68, 2.77)	2.29 (1.30, 7.28)
DerSimonian Laird with correction	4.75 (1.43, 15.73)	1.71 (0.57, 5.13)	2.96 (1.26, 7.00)

* Method used in meta-analysis; OR odds-ratio; CI confidence interval. In a model with correction 0.5 is added to every table of the 2x2 table

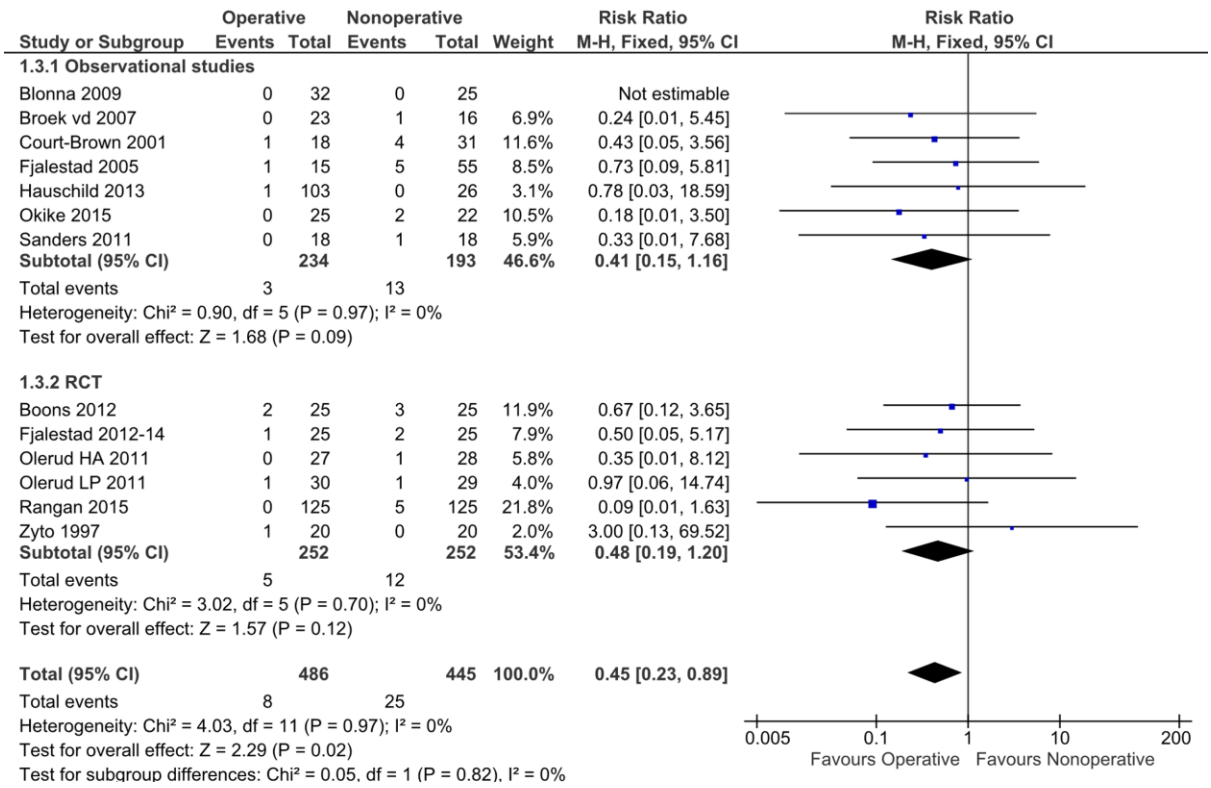

Figure S4. Nonunion in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

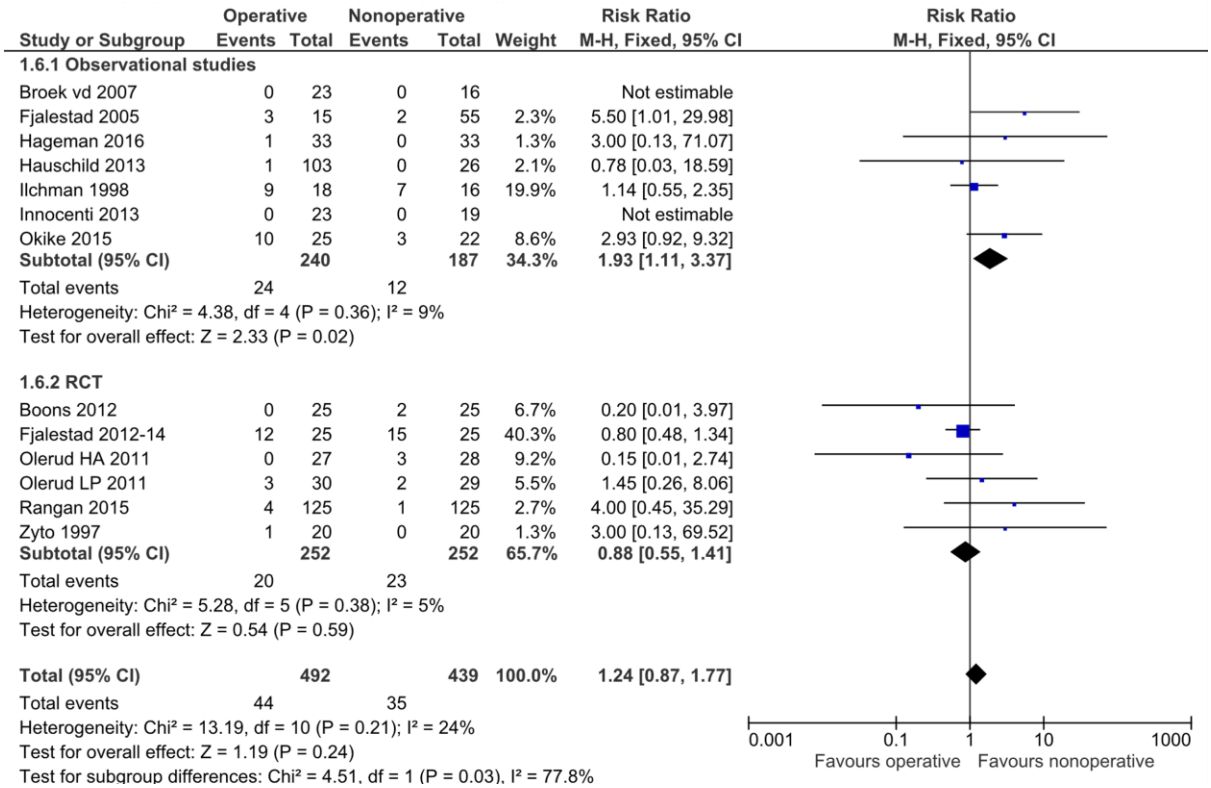

Figure S5. Avascular necrosis in a systematic review of proximal humerus fractures comparing operative with nonoperative treatment.

CHAPTER 4

Conservative versus operative treatment for humeral shaft fractures: a meta-analysis and systematic review of randomized clinical trials and observational studies

Bryan J.M. van de Wall, Yassine Ochen, Frank J.P. Beeres, Reto Babst, Björn C. Link, Marilyn Heng, Detlef van der Velde, Matthias Knobe, Rolf H.H. Groenwold, R. Marijn Houwert

(Journal of Shoulder and Elbow Surgery)

Abstract

Background

This meta-analysis aimed to compare conservative vs. operative treatment for humeral shaft fractures in terms of the nonunion rate, reintervention rate, permanent radial nerve palsy rate, and functional outcomes. Secondarily, effect estimates from observational studies were compared with estimates of randomized clinical trials (RCTs).

Methods

The PubMed/Medline, Embase, CENTRAL (Cochrane Central Register of Controlled Trials), and CINAHL (Cumulative Index to Nursing and Allied Health Literature) databases were searched for both RCTs and observational studies comparing conservative with operative treatment for humeral shaft fractures.

Results

A total of 2 RCTs (150 patients) and 10 observational studies (1262 patients) were included. The pooled nonunion rate of all studies was higher in patients treated conservatively (15.3%) vs. operatively (6.4%) (risk difference, 8%; odds ratio [OR], 2.9; 95% confidence interval [CI], 1.8-4.5; $I^2 = 0\%$). The reintervention rate was also higher for conservative treatment (14.3%) than for operative treatment (8.9%) (risk difference, 6%; OR, 1.9; 95% CI, 1.1-3.5; $I^2 = 30\%$). The higher reintervention rate was predominantly attributable to the higher nonunion rate in patients treated conservatively. The permanent radial nerve palsy rate was equal in both groups (OR, 0.6; 95% CI, 0.2-1.9; $I^2 = 18\%$). There appeared to be no difference in mean time to union and mean Disabilities of the Arm, Shoulder and Hand scores between the treatment groups. No difference was found between effect estimates form observational studies and RCTs.

Conclusion

This systematic review shows that satisfactory results can be achieved with both conservative and operative management; however, operative treatment reduces the risk of nonunion compared with conservative treatment, with comparable reintervention rates (for indications other than nonunion). Furthermore, operative treatment results in a similar permanent radial nerve palsy rate, despite its inherent additional surgery-related risks. No difference in mean time-to-union and short-term functional results was detected.

Introduction

Humeral shaft fractures represent 1%-3% of all fractures.[1] Traditionally, patients with humeral shaft fractures have been treated conservatively.[2] In the past few decades, however, operative treatment has become more popular, with more than half of patients undergoing either plate fixation or nailing.[3]

The optimal treatment of humeral shaft fractures remains a topic of debate. Two meta-analyses have previously been published.[4,5] Because of the lack of randomized clinical trials and the existence of only observational studies at the time, both concluded that the superiority of one treatment over the other could not be determined.

Meta-analyses of randomized clinical trials are considered the highest level of evidence for evaluation of treatment effects. Multiple studies have shown that the estimates of the effects of certain surgical treatments estimated from randomized clinical trials and observational studies tend to be similar.[6-8] The addition of observational studies to meta-analyses increases the sample size and could increase the power for detecting small differences in treatment effects. As randomized clinical trials usually include a highly selective study population, including observational studies in meta-analyses might improve the generalizability of results. Notably, randomized clinical trials and observational studies are increasingly being combined in orthopedic trauma meta-analyses for evaluation of treatment effects.[9-12]

The primary aim of this meta-analysis was to compare the nonunion rate, reintervention rate, permanent radial nerve palsy rate, and functional outcomes after conservative and operative treatment for humeral shaft fractures by considering evidence from randomized clinical trials as well as observational studies. The secondary aim was to determine whether there is a difference in effect estimates obtained from observational studies and from randomized clinical trials in this field of research.

Methods

This systematic review with meta-analysis was performed and reported according to the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines and Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) checklist.[13,14] A published protocol for this review does not exist.

Search strategy and selection criteria

The PubMed/Medline, Embase, CENTRAL (Cochrane Central Register of Controlled Trials), and CINAHL (Cumulative Index to Nursing and Allied Health Literature) databases were searched on March 23, 2019, for studies comparing conservative with operative treatment for humeral shaft fractures. The search syntax is described in Supplementary Table S1. Duplicate articles were removed. Two reviewers (B.J.M.v.d.W. and Y.O.) independently screened titles and abstracts for eligibility. All published studies consisting of observational and randomized clinical trials and comparing conservative with operative treatment for humeral shaft fractures were included.

The same two reviewers independently performed the full-text screening. The inclusion criteria were humeral shaft fracture, conservative treatment (cast immobilization and/or functional bracing), operative treatment (minimally invasive or open plating, nail fixation, and external fixator), age 16 years or older, and reporting of outcomes of interest (nonunion, reintervention, time to union, radial nerve palsy, and functional outcomes). The exclusion criteria were pathologic fractures; treatment for delayed union or nonunion; studies with an average follow-up period of less than 6 months; languages other than English, French, German, or Dutch; no availability of full text; and letters, meeting proceedings, and case reports. Disagreements on the eligibility of full-text articles were resolved by consensus or by discussion with a third reviewer (M.R.H.). References of all included studies were screened to identify studies not found in the original literature search.

Data extraction

Two reviewers (B.J.M.v.d.W. and Y.O.) independently performed data extraction using a predefined data extraction sheet. The following baseline characteristics were extracted from the included studies: first author, year of publication, study period, country in which study was performed, study design, number of included patients, conservative method, operative method, sex, age, open or closed fracture, Arbeitsgemeinschaft für Osteosynthesefragen/Orthopaedic Trauma Association (AO/OTA) Fracture and Dislocation Classification, low- or high-energy trauma, and follow-up duration.[15,16]

Quality assessment

Two reviewers (B.J.M.v.d.W. and Y.O.) independently assessed the methodologic quality of included studies using the Methodological Index for Non-Randomized Studies (MINORS).[17] The MINORS is a validated instrument for assessing the methodologic quality of cohort studies, resulting in a score between 0 and 24. Randomized studies were appraised using the same tool to measure quality on the same scale as observational studies. Disagreements were resolved by consensus. Details on methodologic quality assessment are provided in Supplementary Table S2.

Primary and secondary outcomes

The primary outcome was the nonunion rate after conservative or operative treatment. Nonunion was defined as the absence of fracture consolidation 6 months after treatment with the absence of radiologic bridging callus at 3 of 4 cortices.[18,19] Secondary outcome measures included reintervention, radial nerve palsy, infection, and functional outcome scores. Functional outcome scores included the Disabilities of the Arm, Shoulder and Hand (DASH) score.[20] Measurements of the DASH score were subdivided according to follow-up, into short term (≤1 year) and long term (>1 year). Reintervention included all surgical procedures performed during follow-up. Radial nerve palsy was categorized into palsy at presentation (primary radial nerve palsy), palsy after surgery (secondary radial nerve palsy), or persistent radial nerve palsy at the end of the follow-up period (persistent radial nerve palsy). In other words, permanent radial nerve palsy encompassed all patients in whom nerve function was not restored following either primary or secondary nerve palsy. Infection was classified as either superficial or deep according the definition of the Centers for Disease Control and Prevention.

Statistical analysis

Data for continuous variables were presented as means with standard deviations (SDs) or ranges. The mean and SD were calculated for studies that presented descriptive statistics other than the mean, SD, or range using the methods described in the Cochrane Handbook for Systematic Reviews of Interventions.[21] Dichotomous variables were presented as counts and percentages. Effects of treatment options on binary outcomes were pooled using the (random-effects) Mantel-Haenszel method and presented as odds ratios (ORs) with 95% confidence intervals (CIs). In case of zero-cell counts in 1 of the 2 treatment groups, 0.5 was added to all cells of the contingency table of treatment and outcome of those studies in which this occurred. Effects of treatment options on continuous outcomes were pooled using the (random-effects) inverse-

variance weighting method and presented as mean differences with 95% CIs. None of the observational studies were corrected for confounding. Therefore, the estimated relations between treatment and outcome presented for these studies are unadjusted for possible confounding.

Heterogeneity between studies was assessed for all ORs by visual inspection of forest plots and by the I^2 statistic for heterogeneity. All analyses were stratified according to study design, that is, randomized clinical trials or observational studies. The difference in effect estimates between the 2 subgroups were assessed using the $\chi 2$ test as described in the Cochrane Handbook for Systematic Reviews of Interventions.[21] $P < 0.05$ was considered statistically significant. Publication bias was assessed by visual inspection of funnel plots.12 Review Manager (RevMan, version 5.3.5; The Cochrane Collaboration, London, UK) was used for all statistical analyses.

Sensitivity analyses

Sensitivity analysis for the primary outcome was performed on different types of operative fixation methods. The effect estimates of the primary meta-analysis were compared with the effect estimates of studies using only plate fixation as operative treatment. We performed additional sensitivity analyses using information from studies in which the mean age of included subjects was older than 50 years, as well as from high-quality studies. The cutoff point for age was based on the upper quartile of studies with the highest mean age of participants. High-quality studies were defined as those with a MINORS score (range, 0-24) of 16 or higher. Additional sensitivity analysis was performed on the secondary outcome of reintervention. The effect estimates of the primary meta-analysis on reintervention for all indications (including nonunion) were compared with the risk estimates of reintervention excluding nonunion.

Results

Search

Figure 1 presents the flowchart of the literature search and study selection. The full text could not be obtained for 1 observational study.[22] A total of 12 articles could be included for analyses in this study: 2 randomized clinical trials and 10 observational studies.[23-34]

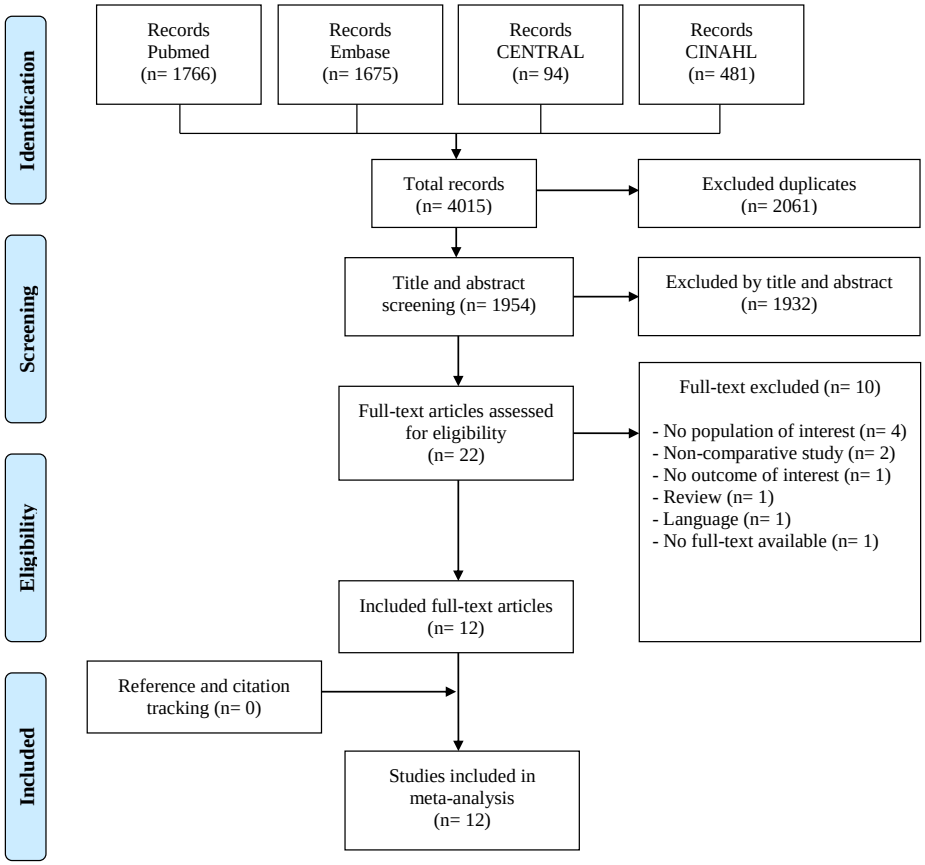

Figure 1. Flow diagram of search and selection of studies comparing operative vs. conservative treatment for humeral shaft fractures. CENTRAL, Cochrane Central Register of Controlled Trials; CINAHL, Cumulative Index to Nursing and Allied Health Literature.

Baseline study characteristics

The 12 studies included 1412 patients: 628 treated conservatively and 784 treated operatively. The overall weighted mean age was 42 years (range, 16-103 years), with 43 years in the conservative group and 42 years in the operative group. The studies included 380 female patients (26.9%). The overall mean follow-up period ranged from 6 to 72 months. Table 1 shows the baseline characteristics of all studies including AO/OTA Fracture and Dislocation Classification, fractures with a concomitant open wound (open fractures), energy of trauma, and treatment type.

The 2 randomized clinical trials included 150 patients, of whom 78 were treated operatively.[23,24] The weighted mean age, as well as age per treatment group, was 37 years (range, 18-83 years). The operative fixation method in both studies was plate fixation. As conservative management, bracing was used in one study and splinting in the other.

The 10 observational studies—1 prospective study and 9 retrospective studies—included 1262 patients, of whom 706 were treated operatively. [25-34] The weighted mean age was 44 years (range, 16-103 years), with 45 years in the conservative group and 43 years in the operative group. Conservative management consisted of bracing in 7 studies and a combination of bracing and splinting in 2, whereas 1 study did not further specify the type of conservative treatment. Operative treatment consisted of a combination of plating, nailing, and external fixation in 7 studies, of which 1 study also included intramedullary flexible nails. In the other 3 studies, either solely plating or nailing was used.

Quality assessment

The details and distribution of the MINORS scores are described in Supplementary Table S3. The overall mean MINORS score was 15.6 (SD, 2.6; range, 13-23), where the 2 randomized clinical trials had scores of 17 and 23.

Primary outcome measure

Nonunion rate

The nonunion rate was reported in 11 studies—2 randomized clinical trials and 9 observational studies.[23-33] The overall pooled effect showed that conservative treatment was associated with a higher nonunion rate compared with operative treatment (OR, 2.9; 95% CI, 1.8-4.5; $I^2 = 0\%$) (Figure 2). The pooled effect for randomized clinical trials showed an OR of 5.7 (95% CI, 0.6-53.6; $I^2 = 29\%$). The pooled effect estimate of observational studies demonstrated an OR of 2.8 (95% CI, 1.7-4.4; $I^2 = 0\%$). Nonunion occurred in 15.3% of patients treated conservatively and 6.4% treated operatively (risk difference [RD], 8%; 95% CI, 4%-12%). No difference in pooled effect estimates was found between randomized clinical trials and observational studies ($P = .43$, test for subgroup difference; $I^2 = 0\%$). The funnel plot is described in Supplementary Figure S1.

Secondary outcome measures

Intervention or reintervention rate

Reintervention was reported in 11 studies—2 randomized clinical trials and 9 observational studies.[23-33] The overall pooled effect showed that the reintervention rate was higher among patients treated conservatively than those treated operatively (OR, 1.9; 95% CI, 1.1-3.5; $I^2 = 30\%$) (Figure 3). The pooled effect for randomized clinical trials was 2.7 (95% CI, 0-156.6; $I^2 =$

Table 1. Baseline characteristics of studies included in systematic review of conservative vs. operative treatment for humeral shaft fracture

Authors	Year	Design	Country	Study period	Total n:	Type of treatment	
					Cons/Op	Cons	Op
RCTs							
Kumar et al	2017	RCT	India	2012-2014	20/20	Splint	Plate
Matsunaga et al	2017	RCT	Brazil	2012-2015	52/58	Brace	Plate
Observational studies							
Harkin and Large et al	2017	RCS	Australia	2008-2015	96/30	Brace	Plate/nail
Westrick et al	2017	RCS	United States	2000-2012	69/227	Brace	Plate/nail/FX
Dielwart et al	2017	RCS	United States	2006-2011	31/40	Brace	Nail/ORIF
Mahabier et al	2013	RCS	The Netherlands	2002-2008	91/95	Brace	Nail/ORIF/FX
Broadbent et al	2010	PCS	United Kingdom	2006-2009	89/21	Brace/cast	Plate/nail/FX
Denard et al	2010	RCS	United States	2001-2006	63/150	Brace	Plate
Ekholm et al	2008	RCS	Sweden	1998-1999	20/7	NR	Plate/nail
Jawa et al	2006	RCS	United States	2000-2004	21/19	Brace	Plate
Osman et al	1998	RCS	France	1994-1997	32/72	Splint/brace	Plate/wires/nail
Wallny et al	1997	RCS	Germany	1990-1994	44/45	Brace	Nail

SD, standard deviation; Cons, conservative treatment; Op, operative treatment; NR, not reported; AO, arbeitsgemeinschaft osteosynthese; RCT, randomised clinical trial; PCS, prospective cohort study; RCS, retrospective cohort study; NA, not applicable; FX, fixator external; ORIF, open reduction internal fixation;

Table 1. Continued

Authors	Sex: female/male		Mean age (SD), yr		Open fracture		AO type: A/B/C		High-energy trauma		Mean follow-up, mo
	Cons	Op	Cons	Op	Cons	Op	Cons	Op	Cons	Op	
RCTs											
Kumar et al	16/6	5/15	33 (11)	38 (16)	0	0	20/0/0	20/0/0	NR	NR	6
Matsunaga et al	14/38	23/35	40 (17)	37 (15)	0	0	28/17/6	38/15/3	NR	NR	12
Observational studies											
Harkin and Large et al	64/33	21/9	NA	NA	0	4	49/14/17	16/8/3	NR	NR	>6
Westrick et al	35/34	75/152	41 (29)	37 (29)	7	92	140/112/40		46	180	>12
Dielwart et al	8/23	22/18	39 (18)	38 (18)	0	0	16/7/8	23/8/9	31	40	10
Mahabier et al	55/36	51/44	61 (24)	59 (26)	0	0	43/40/8	46/32/7	10	22	>6
Broadbent et al	68/42		59 (19)		0	3	52/46/12		NR	NR	12
Denard et al	29/34	68/82	36 (17)	35 (15)	NR	NR	NR	NR	NR	NR	7.9
Ekholm et al	15/5	3/4	53 (29)	48 (27)	0	0	NR	NR	8	2	72
Jawa et al	12/9	8/11	41 (17)	50 (19)	NR	NR	NR	NR	NR	NR	>6
Osman et al	44/60		48 (22)		0		60/38/6			39	18
Wallny et al	20/24	19/26	59 (20)	56 (17)	0	0	NR	NR	NR	NR	27

SD, standard deviation; Cons, conservative treatment; Op, operative treatment; NR, not reported; AO, arbeitsgemeinschaft osteosynthese; RCT, randomised clinical trial; PCS, prospective cohort study; RCS, retrospective cohort study; NA, not applicable; FX, fixator external; ORIF, open reduction internal fixation;

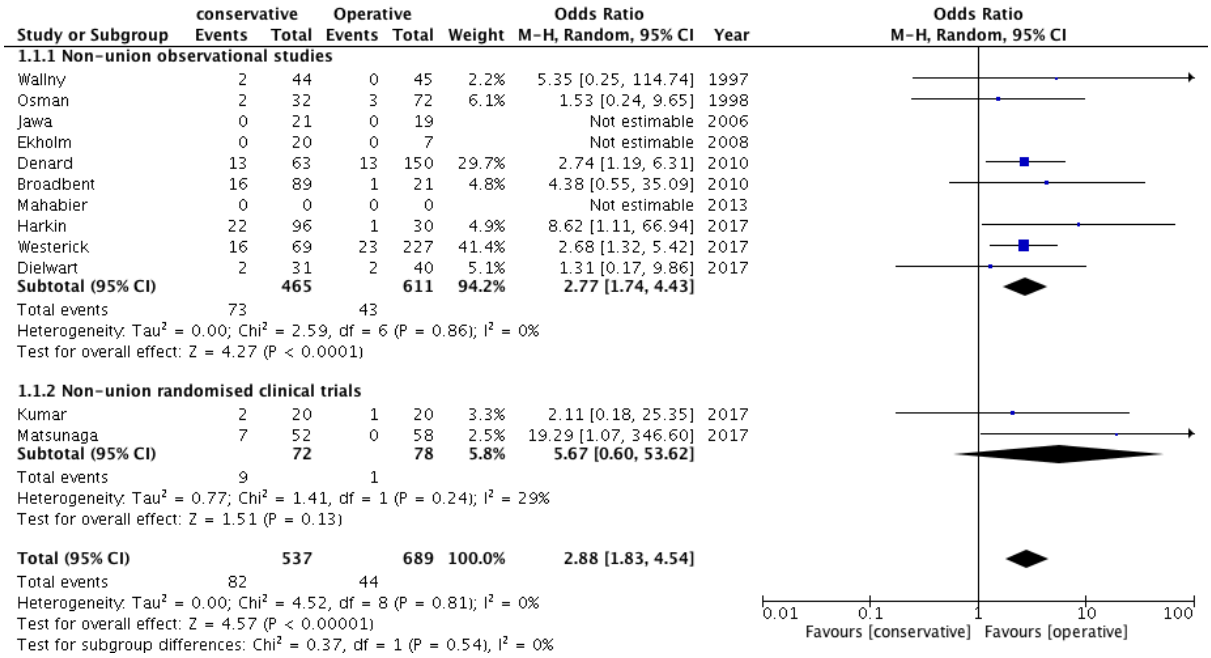

Figure 2. Forest plot of nonunion rate after conservative vs. operative treatment for humeral shaft fractures. CI, confidence interval; M-H, Mantel Haenszel.

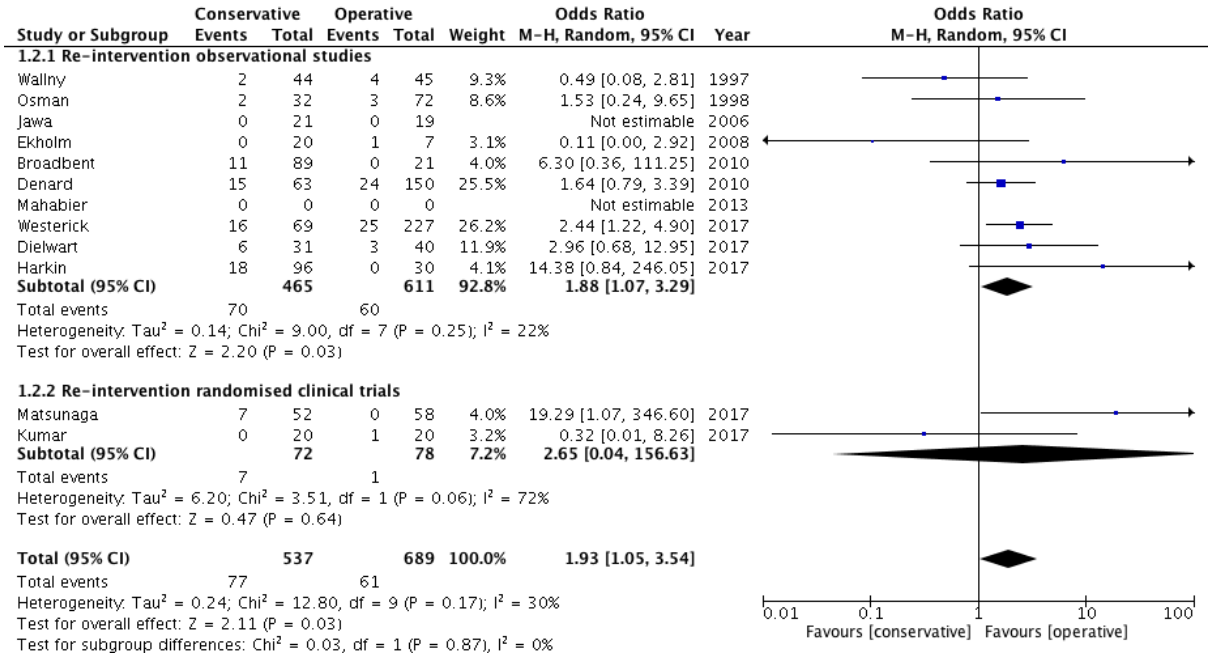

Figure 3. Forest plot of intervention (or reintervention) rate after conservative vs. operative treatment for humeral shaft fractures. CI, confidence interval; M-H, Mantel Haenszel.

72%). The pooled effect estimate of observational studies demonstrated an OR of 1.9 (95% CI, 1.1-3.3; I^2 = 22%). Reintervention occurred in 14.3% of patients treated conservatively and 8.9% treated operatively (absolute RD, 6%; 95% CI, 1%-12%). The most frequent indication for surgical intervention among patients treated conservatively was nonunion. Other indications included malalignment and intolerance of bracing (Supplementary Table S4). The most frequent indication for reintervention among patients treated surgically was nonunion as well. Other indications included infection, implant migration (only for nails), and implant irritation (Supplementary Table S5). No difference in pooled effect estimates was found between randomized clinical trials and observational studies (P = .83, test for subgroup difference; I^2 = 0%). The funnel plot is described in Supplementary Figure S2.

Mean time to union

Five studies reported on mean time to union—1 randomized clinical trial and 4 observational studies.[23,26,27,33,34] The overall pooled time to union did not differ between the treatment groups (mean difference, –1.2 weeks; 95% CI, –4.3 to 2.0 weeks; I^2 = 84%) (Figure 4). The weighted mean time to union was 16 weeks in the conservative group and 17 weeks in the operative group. Subgroup analysis was not possible as only 1 randomized clinical trial reported on time to union. The funnel plot is described in Supplementary Figure S3.

DASH score

Only the 2 randomized clinical trials reported on short-term DASH scores, both at 6 months.[23,24] The overall pooled DASH score did not differ between conservative and operative treatment (mean difference, 10.7; 95% CI, –0.7 to 22.2; I^2 = 68%) (Figure 5). The weighted mean DASH score was 27 among patients treated conservatively and 15 among those treated operatively. The funnel plot is described in Supplementary Figure S4. Long-term functional outcomes using the DASH score were not reported in the included studies.

Radial nerve palsy

Eleven studies reported on radial nerve palsy—2 randomized clinical trials and 9 observational studies.[23,24,26-34] Radial nerve palsy at presentation (primary radial nerve palsy) was found among 9.6% of patients treated conservatively (n = 52). Only 7 of these patients (1.5%) had permanent radial nerve palsy at the end of the study period. Among patients treated operatively, 16.1% (n = 123) had primary radial nerve palsy; of these, 19 (2.5%) had permanent palsy (Table 2). Radial

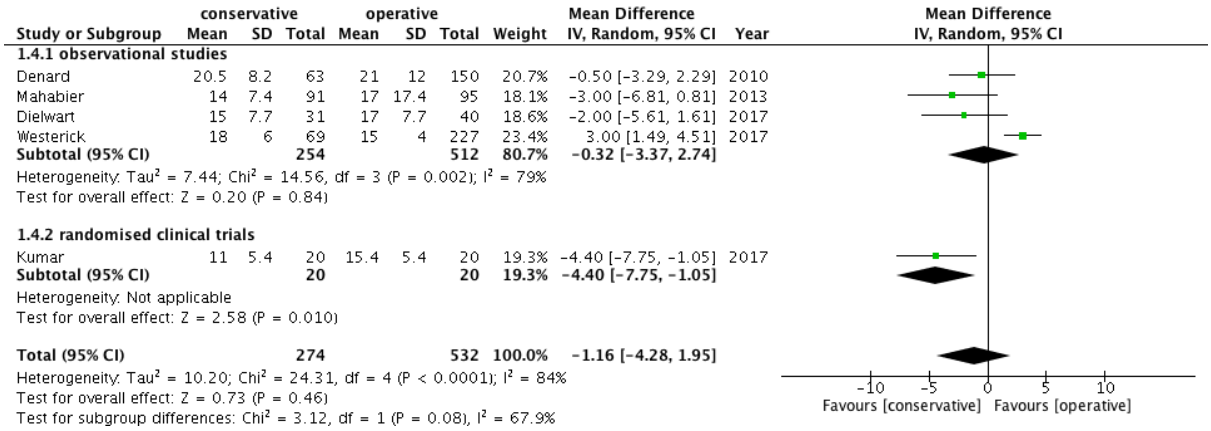

Figure 4. Forest plot of mean time to union after conservative vs. operative treatment for humeral shaft fractures. CI, confidence interval; IV, weighted mean difference

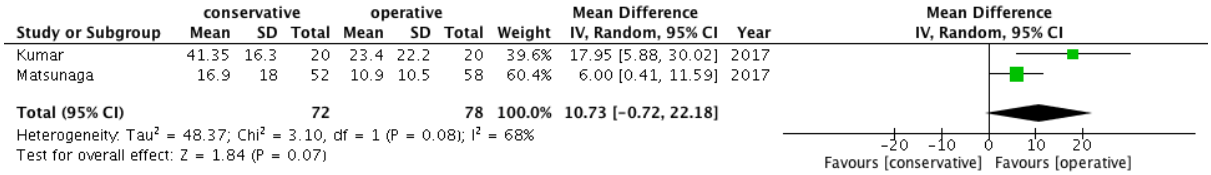

Figure 5. Forest plot of Disabilities of the Arm, Shoulder and Hand (DASH) score at 6 months after conservative vs. operative treatment for humeral shaft fractures. CI, confidence interval; IV, weighted mean difference.

nerve palsy due to the operation was found in 3.5% of patients in the operative group (n = 27). Only 1 patient had permanent damage. The other patients had full recovery of nerve function. The overall pooled permanent radial nerve palsy rate at the end of follow-up was equal in both groups (OR, 0.6; 95% CI, 0.2-1.9; I^2 = 18%) (Figure 6). Subgroup analysis could not be performed because of insufficient numbers of events between the randomized clinical trials. The funnel plot is described in Supplementary Figure S5.

Infection

Seven studies reported on postoperative infections in the operative group.[24,26,27,30-33] No distinction could be made between deep or superficial infection as none of the studies clearly defined infection or applied the definition of the Centers for Disease Control and Prevention. Infection was reported in 0.6% of patients treated conservatively (n = 2). In both, infection developed following a humeral shaft fracture caused by a gunshot injury. Symptoms resolved after antibiotic treatment in both patients. Infection occurred in 3.1% of patients treated operatively (n = 19). Twelve of these patients underwent subsequent wound débridement. The other 7 patients were treated conservatively with antibiotics.

Table 2. Primary, secondary, and persistent radial nerve palsy in studies of conservative vs. operative treatment for humeral shaft fractures

			Total n:	Primary radial nerve palsy at presentation, n				Secondary radial nerve palsy after surgery, n	
Authors	**Type**		**Cons/Op**	**Temporary**	**Persistent***	**Temporary**	**Persistent***	**Temporary**	**Persistent***
	Cons	**Op**		**Cons**		**Op**			
Randomized clinical trials									
Kumar et al	Splint	Plate	20/20	0	0	0	0	1	0
Matsunaga et al	Bracing	Plate	52/58	0	0	0	0	2	0
Observational studies									
Harkin and Large	Bracing	Plate/nail	96/30	7	1	5	2	4	0
Westrick et al	Bracing	Plate/nail/FX	69/227	14	1	82	14	2	0
Dielwart et al	Bracing	Nail/ORIF	31/40	5	1	10	1	2	0
Mahabier et al	Bracing	Nail/ORIF/FX	91/95	8	2	5	1	4	NR
Broadbent et al	Bracing/splint	Plate/nail/FX	89/21	NR	NR	NR	NR	NR	NR
Denard et al	Bracing	Plate	63/150	6	0	1	0	4	0
Ekholm et al	NR	Plate/nail	20/7	2	0	4	0	0	0
Jawa et al	Bracing	Plate	21/19	2	0	4	1	3	1
Osman et al	Splint/brace	Plate/wires/nail	32/72	2	0	6	0	4	0
Wallny et al	Bracing	Nail	44/45	6	0	6	0	1	0
Total				**52**	**7**	**123**	**19**	**27**	**1**

Cons, conservative treatment; Op, operative treatment; FX, fixator external; ORIF, open reduction internal fixation.* Persistent indicates the number of patients with primary or secondary radial nerve palsy in whom radial nerve palsy did not recover during follow-up.

Table 3. Sensitivity analysis on primary outcome (nonunion) after conservative vs. operative treatment for humeral shaft fractures

	n	RD	OR (95% CI)	P value	I^2, %
All studies	11	8%	2.9 (1.8-4.5)	<.001	0
Studies with plate fixation	4	8%	3.1 (1.4-6.6)	.004	0
Studies with age > 50 years	3	6%	4.7 (0.8-26.1)	.08	0
High-quality studies	5	8*	2.8 (1.4-5.6)	.005	0

RD, risk difference; OR, odds ratio; CI, confidence interval.

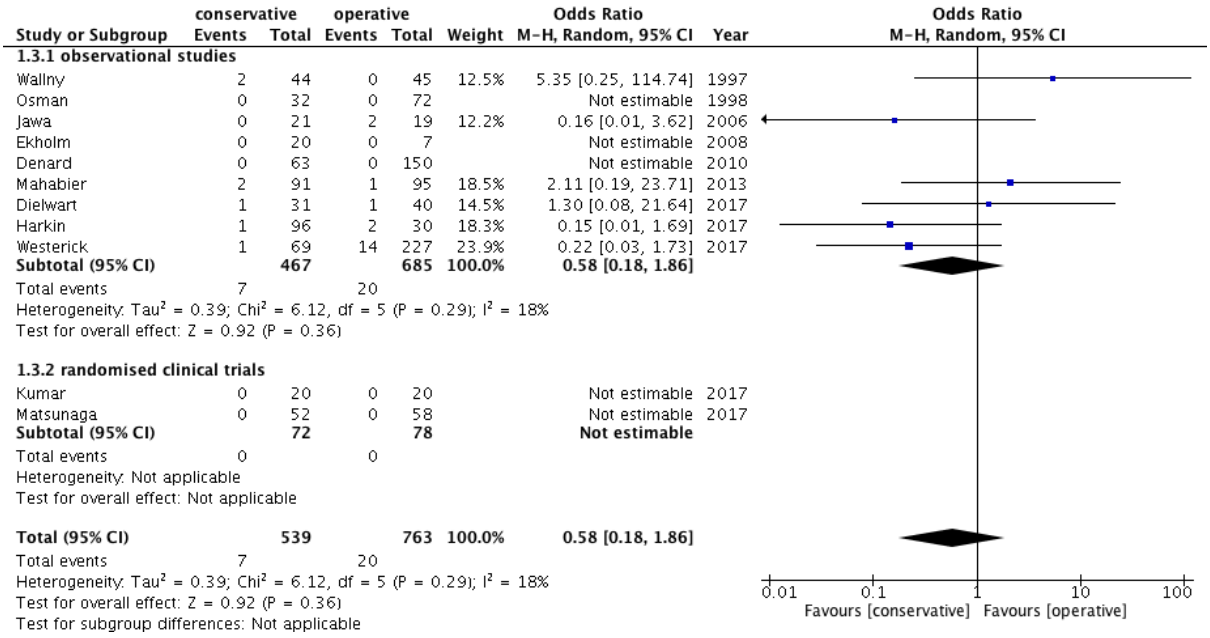

Figure 6. Forest plot of permanent radial nerve palsy rate after conservative vs. operative treatment for humeral shaft fractures. CI, confidence interval; M-H, Mantel Haenszel.

Other complications

All other reported complications are listed in Supplementary Table S6.

Sensitivity analysis

Table 3 shows the results of the sensitivity analysis on the primary outcome (nonunion). A total of 4 studies compared plate fixation with conservative treatment—2 randomized clinical trials and 2 observational studies.[23,24,26,30] The pooled estimate showed that the nonunion rate was higher among patients treated conservatively than among those treated by plate fixation (RD, 8%; OR, 3.1; 95% CI, 1.4-6.6; I^2 = 0%; Supplementary Figure S6). Only 3 studies—all observational studies—had a study population with a mean age older than 50 years.[28,32,34] The pooled analysis did not demonstrate a difference in nonunion rates between conservative and operative treatment (OR, 4.7; 95% CI, 0.8-26.1; I^2 = 0%; Supplementary Figure S7). There were 5 high-quality studies—2 randomized clinical trials and 3 observational studies.[23,24,26-28] The nonunion rate was higher among patients treated conservatively than those treated operatively (OR, 2.8; 95% CI, 1.4-5.6; I^2 = 0%) (Supplementary Figure S8). Reintervention for indications other than nonunion (Supplementary Tables S4 and S5) was reported in 11 studies—2 randomized clinical trials and 9 observational studies.[23,24,26-34] The pooled analysis showed no difference between groups (OR, 1.0; 95% CI, 0.4-2.8; I^2 = 53%) (Supplementary Figure S9).

Discussion

This systematic review and meta-analysis, including both randomized clinical trials and observational studies, compared conservative with operative treatment for humeral shaft fractures. The pooled effect estimates demonstrated that conservative treatment was associated with higher nonunion and reintervention rates compared with operative treatment. There appeared to be no difference in mean time to union and DASH scores. The pooled analysis also found no difference in the rate of persistent radial nerve palsy between the two treatment groups. Sensitivity analysis on the secondary outcome of reintervention showed that the higher reintervention rate in the conservative group was mainly caused by a high rate of intervention for nonunion. There appeared to be no difference in effect estimates from randomized clinical trials and observational studies for either the nonunion or reintervention rate.

To date, only 2 systematic reviews have been published comparing operative with conservative treatment for humeral shaft fractures.[4,5] Gosler et al.[4] performed a systematic review in 2012 but could not identify any randomized clinical trials. They therefore did not perform any formal analysis and concluded that there was insufficient evidence to support either of the 2 treatment modalities. Clement et al.[5] published a systematic review in 2015 and reached the same conclusion as Gosler et al. Clement, however, identified 1 ongoing randomized clinical trial, the results of which were unavailable at that time.[24] In contrast to the present meta-analysis, both previous meta-analyses did not include observational studies.

Our findings of a higher nonunion rate among patients treated conservatively compared with those treated operatively are in line with the general consensus in the literature. Nonunion rates among patients treated conservatively are usually found to be between 0% and 22.6% in noncomparative studies.[35] These rates range from 0% to 9% for operative management.[36] Given the large number of patients included in our meta-analysis, we were able to more reliably determine these incidences. We found an incidence of 15.3% in the conservative group vs. 6.4% in the operative group.

The reintervention rate appeared to be higher in patients treated conservatively. This was mainly caused by a higher reintervention rate for nonunion. The reintervention rate was equal for indications other than nonunion as described in the sensitivity analysis. It is interesting to note that operative treatment exposes patients to surgery-related complications that do not occur in

patients treated conservatively (e.g. infections requiring débridement, implant removal, or migration). Despite the additional risk, the overall reintervention rate for indications other than nonunion was equal. This means that a great number of patients initially treated conservatively ultimately require surgery, with malalignment being the most frequent indication. In addition, it should be acknowledged that performing surgery in patients initially managed conservatively is generally less complex than that in patients initially treated operatively. In the conservative group, surgery is performed for treatment failure, and in the operative group, reintervention is performed for the treatment of complications. The lower complexity of performing reintervention in patients initially treated by conservative means might also explain the relatively high reintervention rate.

Surgical fixation of humeral shaft fractures carries a risk of 3.5% for radial nerve palsy following surgery, as found in our meta-analysis. Despite the added risk, the rate of persistent radial nerve palsy is equally rare in both patients treated conservatively and those treated operatively. Radial nerve palsy following surgery therefore appears to be a mostly temporary issue and rarely leads to permanent damage. In addition, this study emphasizes that the presence of radial nerve palsy in patients with humeral shaft fractures does not necessarily mandate exploration. As seen in our study and described in the literature, primary radial nerve palsy usually resolves spontaneously.[37,38]

Only the 2 randomized clinical trials reported on validated functional outcome scores (DASH score).[23,24] The other studies either did not report functional results or reported results of non-validated instruments. The pooled analysis showed a trend toward better functional results in patients treated operatively. This difference, however, did not reach statistical significance. As both randomized clinical trials found comparable results in favor of operative treatment, it is likely that the failure to detect a difference is mainly a result of underpowering rather than due to the fact that there is no actual difference.

The present meta-analysis found no difference in pooled effect estimates between randomized clinical trials and observational studies. Observational studies may provide valuable information about treatment effects.[39-41] Including this information in a meta-analysis increases the sample size and thus allows for evaluation of effects in subgroups of patients or effects on rare clinical endpoints. The benefit of including observational data has been previously demonstrated in meta-analyses on surgical interventions.[6,9,10,12,42] Similarly to our study, these meta-analyses found

no difference in pooled treatment effects between observational studies and randomized clinical trials, although effect estimates of observational studies were more heterogeneous.

An important aspect in incorporating observational data in meta-analyses is that the chances of confounding should be deemed small. In this meta-analysis, the observed baseline patient characteristics were comparable between treatment groups, from which we inferred that this may also be the case for unobserved patient characteristics. On the basis of this observation, we consider the potential for confounding acceptably low to allow for the inclusion of observational data in the meta-analysis.

Several potential limitations in this review should be considered. First, the results might have been influenced by missing articles. There appeared to be some visual asymmetry in the funnel plot for the outcome of nonunion. This, however, might also have been caused by the relatively low number of studies. Second, a limited number of randomized clinical trials were available for comparison of risk estimates of observational studies and randomized clinical trials. Although less robust, our findings, suggesting comparable risk estimates between the 2 study designs, are in line with those of previous studies. Third, this meta-analysis investigated the difference between conservative and operative treatment, irrespective of type of operative management (nail, plate, minimally invasive techniques). Finally, to increase the power of the pooled analysis, we used a compound endpoint for reintervention. In other words, we did not take the severity of the indication or reintervention itself into account.

A trend is observed toward the increased use of operative fixation.[3] Possible reasons for this include a perceived quicker return to work, earlier initiation of shoulder and elbow rehabilitation, and avoidance of potential troublesome brace wear during the recovery period.[3] However, evidence supporting this is scarce. Investigating whether these patient-related outcomes truly exist would require prospective studies measuring these outcomes on a daily basis (e.g. patient diary) and not at a fixed point in time (e.g. during outpatient clinic visits), as frequently used in the studies in our meta-analysis. This would complement the already existing data indicating more favorable outcomes for surgical treatment.

The next step in determining optimal management for humeral shaft fractures would be to determine which type of surgical treatment is superior. Multiple meta-analyses have been

performed comparing plate fixation with minimally invasive plating and nailing.[43-45] Although these meta-analyses found differences in procedure-related complications (e.g. shoulder complaints with nailing or radial nerve palsy with plate fixation), they failed to detect differences in other important outcomes including nonunion, infection, reintervention, and functional scores.

Conclusion

This systematic review shows that satisfactory results can be achieved with both conservative and operative management. However, operative treatment reduces the risk of nonunion compared with conservative treatment, with comparable reintervention rates (for indications other than nonunion). Intervention (or reintervention) is mostly performed because of treatment failure in the conservative group and for the treatment of complications in the operative group, which logically differ in complexity. Furthermore, operative treatment results in a similar permanent radial nerve palsy rate, despite its inherent additional surgery-related risks. There is also a trend toward better functional results for operative treatment.

References:

1. Ekholm R, Adami J, Tidermark J, et al. Fractures of the shaft of the humerus: an epidemiological study of 401 fractures. 2006;88(11):1469-1473.
2. Sarmiento AJOTD. Functional bracing of fractures of the shaft of the humerus. 2008;6(01):33-37.
3. Schoch BS, Padegimas EM, Maltenfort M, Krieg J, Namdari SJJoO, Traumatology. Humeral shaft fractures: national trends in management. 2017;18(3):259.
4. Gosler MW, Testroote M, Morrenhof J, Janzing HMJCDoSR. Surgical versus non-surgical interventions for treating humeral shaft fractures in adults. 2012(1).
5. Clement NDJAotr. Management of humeral shaft fractures; non-operative versus operative. 2015;4(2).
6. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *Journal of clinical epidemiology.* 2010;63:238-245.
7. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *The New England journal of medicine.* 2000;342:1878-1886.
8. Boyko EJJJoD, Complications i. Observational research—opportunities and limitations. 2013;27(6):642-648.
9. Frieden TRJNEJoM. Evidence for health decision making—beyond randomized, controlled trials. 2017;377(5):465-475.
10. Smeeing DPJ, van der Ven DJC, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older: A Systematic Review, Meta-analysis, and Comparison of Randomized Controlled Trials and Observational Studies. *The American journal of sports medicine.* 2017;45:1937-1945.
11. Beks RB, Peek J, de Jong MB, et al. Fixation of flail chest or multiple rib fractures: current evidence and how to proceed. A systematic review and meta-analysis. *European journal of trauma and emergency surgery : official publication of the European Trauma Society.* 2019;45:631-644.
12. Ochen Y, Beks RB, van Heijl M, et al. Operative treatment versus nonoperative treatment of Achilles tendon ruptures: systematic review and meta-analysis. *BMJ (Clinical research ed).* 2019;364:k5120.
13. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000;283:2008-2012.
14. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of clinical epidemiology.* 2009;62:1006-1012.
15. ATLS S, trauma IAWGJTjo, surgery ac. Advanced trauma life support (ATLS®): the ninth edition. 2013;74(5):1363.
16. Meinberg EG, Agel J, Roberts CS, Karam MD, Kellam JFJJoot. Fracture and dislocation classification compendium—2018. 2018;32:S1-S10.
17. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ journal of surgery.* 2003;73:712-716.
18. Hammer R, Hammerby S, Lindholm BJCo, research r. Accuracy of radiologic assessment of tibial shaft fracture union in humans. 1985(199):233-238.
19. Whelan D, Bhandari M, McKee M, et al. Interobserver and intraobserver variation in the assessment of the healing of tibial fractures after intramedullary fixation. 2002;84(1):15-18.
20. Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC musculoskeletal disorders.* 2003;4:11.
21. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. *The Cochrane Collaboration.* 2011;0:10-11.
22. Nast-Kolb D, Knoefel W, Schweiberer LJDU. The treatment of humeral shaft fractures. Results of a prospective AO multicenter study. 1991;94(9):447-454.
23. Kumar S, Shanmugam N, Kumar S, Ramanusan KJIJRO. Comparison between operative and non-operative treatment of fracture shaft of humerus: An outcome analysis. 2017;3(3):445-450.
24. Matsunaga FT, Tamaoki MJS, Matsumoto MH, Netto NA, Faloppa F, Belloti JCJJ. Minimally invasive osteosynthesis with a bridge plate versus a functional brace for humeral shaft fractures: a randomized controlled trial. 2017;99(7):583-592.
25. Broadbent M, Will E, McQueen MJI. Prediction of outcome after humeral diaphyseal fracture. 2010;41(6):572-577.
26. Denard A, Richards JE, Obremskey WT, Tucker MC, Floyd M, Herzog GAJO. Outcome of nonoperative vs operative treatment of humeral shaft fractures: a retrospective study of 213 patients. 2010;33(8).
27. Dielwart C, Harmer L, Thompson J, Seymour RB, Karunakar MAJJoOT. Management of closed diaphyseal humerus fractures in patients with injury severity score≥ 17. 2017;31(4):220-224.
28. Ekholm R, Ponzer S, Törnkvist H, Adami J, Tidermark JJJoot. The Holstein-Lewis humeral shaft fracture: aspects of radial nerve injury, primary treatment, and outcome. 2008;22(10):693-697.
29. Harkin FE, Large RJJJoS, Surgery E. Humeral shaft fractures: union outcomes in a large cohort. 2017;26(11):1881-1888.

30. Jawa A, McCarty P, Doornberg J, Harris M, Ring DJJ. Extra-articular distal-third diaphyseal fractures of the humerus: a comparison of functional bracing and plate fixation. 2006;88(11):2343-2347.
31. Osman N, Touam C, Masmejean E, Asfazadourian H, Alnot J. Results of non-operative and operative treatment of humeral shaft fractures A series of 104 cases. Paper presented at: Annales de Chirurgie de la Main et du Membre Supérieur1998.
32. Wallny T, Sagebiel C, Westerman K, Wagner U, Reimer MJIo. Comparative results of bracing and interlocking nailing in the treatment of humeral shaft fractures. 1998;21(6):374-379.
33. Westrick E, Hamilton B, Toogood P, Henley B, Firoozabadi RJIo. Humeral shaft fractures: results of operative and non-operative treatment. 2017;41(2):385-395.
34. Mahabier KC, Vogels LM, Punt BJ, Roukema GR, Patka P, Van Lieshout EMJI. Humeral shaft fractures: retrospective results of non-operative and operative treatment of 186 patients. 2013;44(4):427-430.
35. Papasoulis E, Drosos GI, Ververidis AN, Verettas D-AJI. Functional bracing of humeral shaft fractures. A review of clinical studies. 2010;41(7):e21-e27.
36. Hohmann E, Glatt V, Tetsworth KJJos, surgery e. Minimally invasive plating versus either open reduction and plate fixation or intramedullary nailing of humeral shaft fractures: a systematic review and meta-analysis of randomized controlled trials. 2016;25(10):1634-1642.
37. Belayneh R, Lott A, Haglin J, et al. Final outcomes of radial nerve palsy associated with humeral shaft fracture and nonunion. 2019;20(1):1-6.
38. Schwab T, Stillhard P, Schibli S, Furrer M, Sommer CJEJoT, Surgery E. Radial nerve palsy in humeral shaft fractures with internal fixation: analysis of management and outcome. 2018;44(2):235-243.
39. Ioannidis JP, Haidich A-B, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. 2001;286(7):821-830.
40. Khan AY, Preskorn SH, Baker BJJocp. Effect of study criteria on recruitment and generalizability of the results. 2005;25(3):271-275.
41. Van Spall HG, Toren A, Kiss A, Fowler RAJJ. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. 2007;297(11):1233-1240.
42. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England journal of medicine*. 2000;342:1887-1892.
43. Hu X, Xu S, Lu H, et al. Minimally invasive plate osteosynthesis vs conventional fixation techniques for surgically treated humeral shaft fractures: a meta-analysis. 2016;11(1):59.
44. Zhang Q, Sun N, Huang Q, Zhu S, Wu XJJoIS. Minimally invasive plating osteosynthesis in the treatment of humeral shaft fractures: a meta-analysis. 2017;30(2):133-142.
45. Zhao J-G, Wang J, Meng X-H, Zeng X-T, Kan S-LJPO. Surgical interventions to treat humerus shaft fractures: a network meta-analysis of randomized controlled trials. 2017;12(3):e0173634.

.

Supplementary materials to Chapter 4

Table S1. Search syntax performed last on March 23, 2019

Database	Syntax
PubMed/MEDLINE (n= 1766)	((((("shaft"[Title/Abstract]) OR "Diaphysis"[Title/Abstract]) OR "diaphyseal"[Title/Abstract]) OR "mid shaft"[Title/Abstract])) AND (("fracture"[Title/Abstract]) OR "fractures"[Title/Abstract])) AND (("humeral"[Title/Abstract]) OR "humerus"[Title/Abstract]))
Embase (n= 1675)	(humeral:ti,ab OR humerus:ab,ti) AND (fracture:ti,ab OR fractures:ab,ti) AND (shaft:ti,ab OR diaphysis:ab,ti OR diaphyseal:ab,ti OR mid) AND shaft:ab,ti
CENTRAL (n= 94)	(AB (humerus OR humeral) AND AB (fracture or fractures) AND AB (shaft OR diaphysis OR diaphyseal OR mid shaft)) OR TI ((humerus OR humeral) AND (fracture or fractures) AND (shaft OR diaphysis OR diaphyseal OR mid shaft))
CINAHL (n= 481)	(AB (humerus OR humeral) AND AB (fracture or fractures) AND AB (shaft OR diaphysis OR diaphyseal OR mid shaft)) OR TI ((humerus OR humeral) AND (fracture or fractures) AND (shaft OR diaphysis OR diaphyseal OR mid shaft))

Table S4. Indications for (re)interventions other than non-union for patients treated conservatively

	Conservative (N)	
	Mal-alignment	**Non-tolerance brace**
Kumar et al		
Westerick et al		
Dielwart et al	4	
Osmann et al	8	
Jawa et al	2	
Denard et al		
Matsunaga et al	1	1
Wallny et al		
Ekholm et al		

Table S5. Indications for re-interventions other than non-union in patients treated operatively

	Implant irritation	**Infection**	**Implant migration**	**Mal-reduction**	**Elbow stiffness****	**Hematoma**	**Secondary dislocation**
Kumar et al	1						
Westerick et al		2					
Dielwart et al		1					
Osmann et al			6*				
Jawa et al		1	1				
Denard et al		7		3	1		
Matsunaga et al							
Wallny et al		1				1	2*
Ekholm et al				1			

* intra-medullary nails, ** Elbow function impairment due to scar tissue. Under general anesthesia stretching scar tissue.

Table S2. Quality assessment according to the MINORS criteria

Criteria	Reported and adequate (2)	Reported but inadequate (1)	Not reported (0)
Clearly stated aim	Aim including outcomes reported	Aim reported without outcomes	Not reported
Inclusion consecutive patients	Inclusion/exclusion criteria reported	Unclear description inclusion/exclusion criteria	Not reported
Prospective collection data	Prospective	Not applicable	Not applicable
Appropriate endpoints	Appropriate endpoints to aim study	Endpoints not appropriate to aim study	Not reported
Unbiased assessment	Blinded evaluation of outcomes	Reason not blinding stated	Not reported
Appropriate follow-up	≥ 1 year	< 1 year	Not reported
Loss to follow-up < 5%	≤ 5%	> 5%	Not applicable
Prospective calculation study size	Prospective power-analysis performed	Prospective calculation without power-analysis	Not applicable
Adequate control group	Operative versus nonoperative treatment	Not applicable	Not applicable
Contemporary groups	Study/control group managed during same period	Study/control not managed during same period	Not reported
Baseline equivalence groups	Baseline characteristics described and comparable	Baseline characteristics not comparable	Not reported
Adequate statistical analyses	Statistical analysis described including type of analyses	Inadequate description statistical analysis	Not reported

Items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S3. Quality assessment of included studies

MINORS criteria	**RCTs**	Kumar et al	Matsunaga et al	**Observational studies**	Harkin et al	Westerick et al	Dielwart et al	Mahabier et al	Broadbent et al	Denard et al	Ekholm et al	Jawa et al	Osman et al	Wallny et al
Clearly stated aim		2	2		2	2	2	2	2	2	2	2	2	2
Inclusion of consecutive patients		2	2		2	2	2	2	2	2	2	2	2	2
Prospective collection of data		2	2		0	0	0	0	2	0	0	0	0	0
Appropriate endpoints		2	2		2	2	2	1	2	2	2	2	1	1
Unbiased assessment endpoints		0	2		0	0	0	0	0	0	0	0	0	0
Appropriate follow-up		1	2		0	2	1	0	2	1	2	1	2	2
Loss to follow-up < 5%		0	1		1	0	1	0	0	1	1	1	1	0
Prospective calculation study size		0	2		0	0	0	0	0	0	0	0	0	0
Adequate control group		2	2		2	2	2	2	0	2	2	2	1	2
Contemporary groups		2	2		2	2	2	2	2	2	2	2	2	2
Baseline equivalence of groups		2	2		2	1	2	2	0	2	1	1	0	1
Adequate statistical analysis		2	2		2	2	2	2	2	2	2	2	2	2
Total MINORS score		**17**	**23**		**15**	**15**	**16**	**13**	**14**	**16**	**16**	**15**	**13**	**14**

Items are scored 0 (not reported/ not applicable), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S6. Other complications reported in studies

Author	Year	Cons | Op (total)	Mal-alignment*		DVT		Contact dermatitis		Hypertrophic scare	
			Cons	Op	Cons	Op	Cons	Op	Cons	Op
Kumar et al	2017	20 | 20								
Harkin et al	2017	96 | 30								
Westerick et al	2017	69 | 227								
Matsunaga et al	2017	52 | 58	1				5			4
Dielwart et al	2017	31 | 40	4		1	1				
Mahabier et al	2013	91 | 95								
Broadbent et al	2010	89 | 21								
Denard et al	2010	63 | 150	8	12						
Ekholm et al	2008	20 | 7								
Jawa et al	2006	21 | 19	2							
Osman et al	1998	32 | 72								
Wallny et al	1997	44 | 45								

* Supplementary table 4 and 5 describe the number of patients with mal-alignment who were subsequently treated operatively. All other complications described in this table were treated conservatively. DVT Deep venous thrombosis

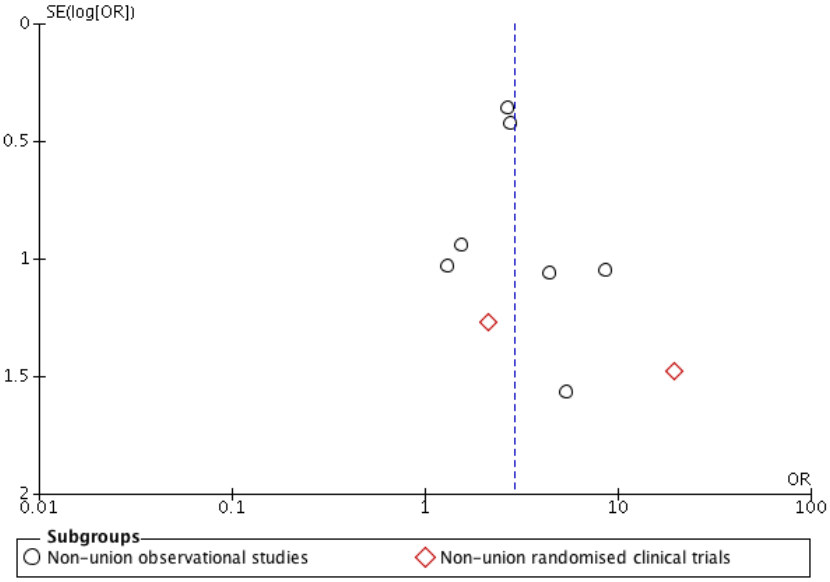

Figure S1. Funnel-plot of non-union rate (OR odds ratio; SE standard error).

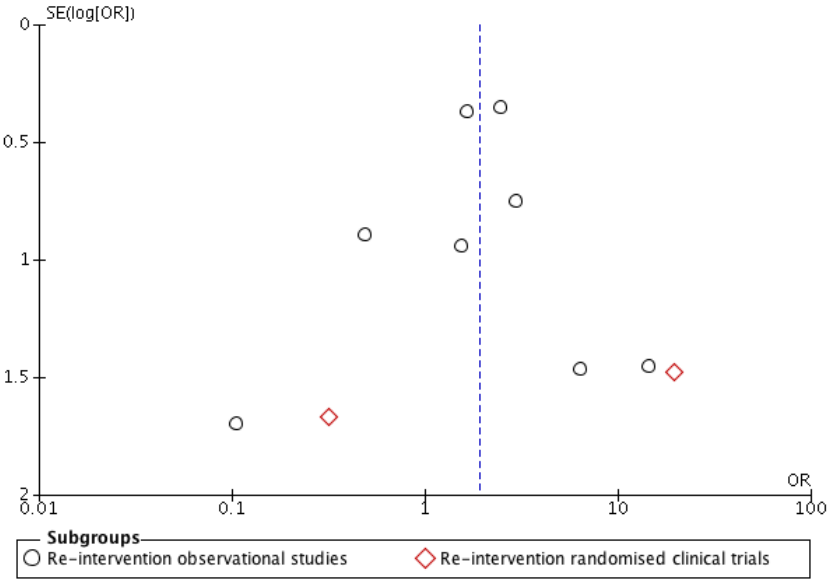

Figure S2. Funnel-plot of (re)intervention rate (OR odds ratio; SE standard error).

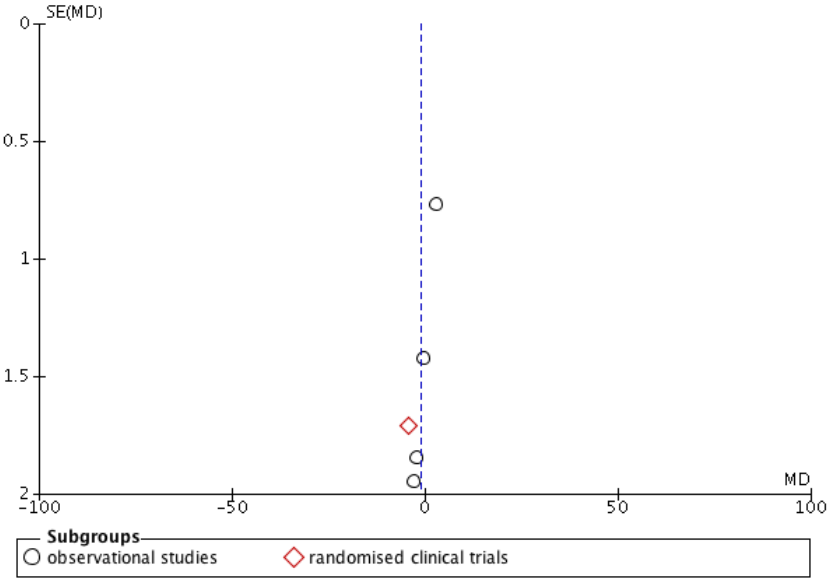

Figure S3. Funnel-plot mean time to union (MD mean difference; SE standard error).

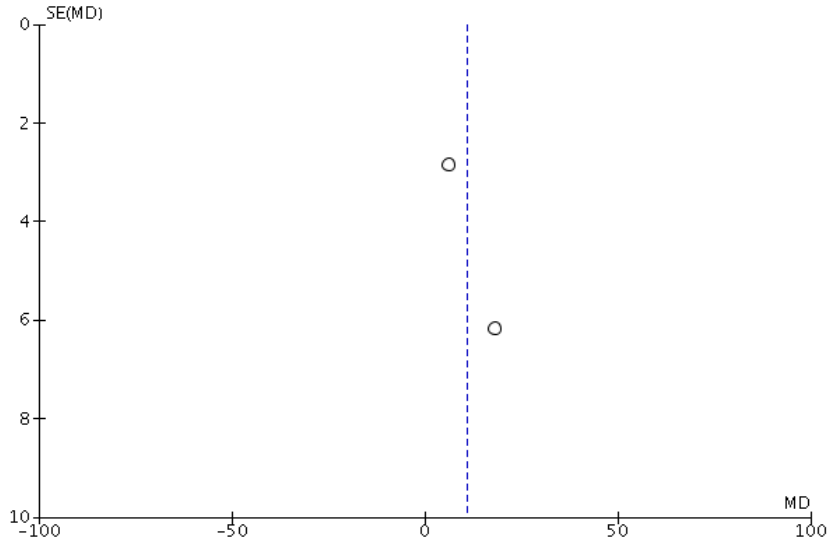

Figure S4. Funnel-plot DASH score at 6 months (MD mean difference; SE standard error).

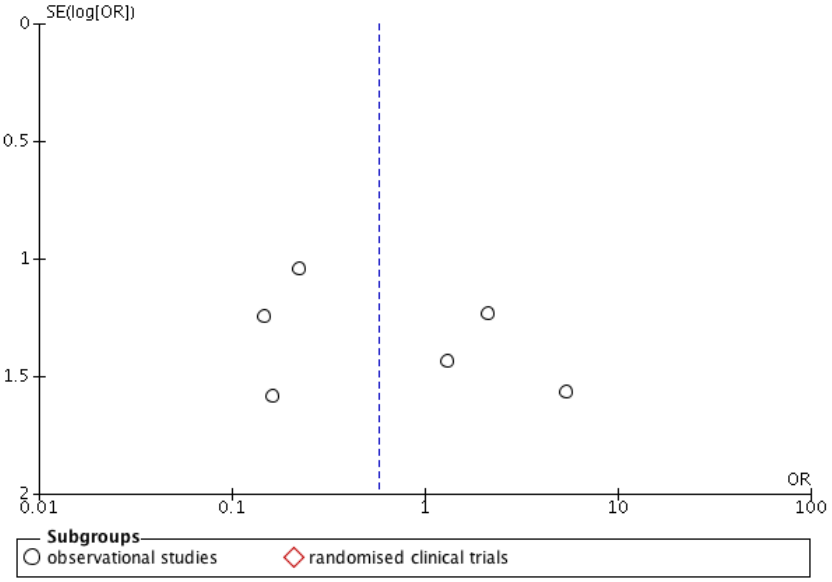

Figure S5. Funnel-plot permanent radial nerve palsy rate (OR odds ratio; SE standard error).

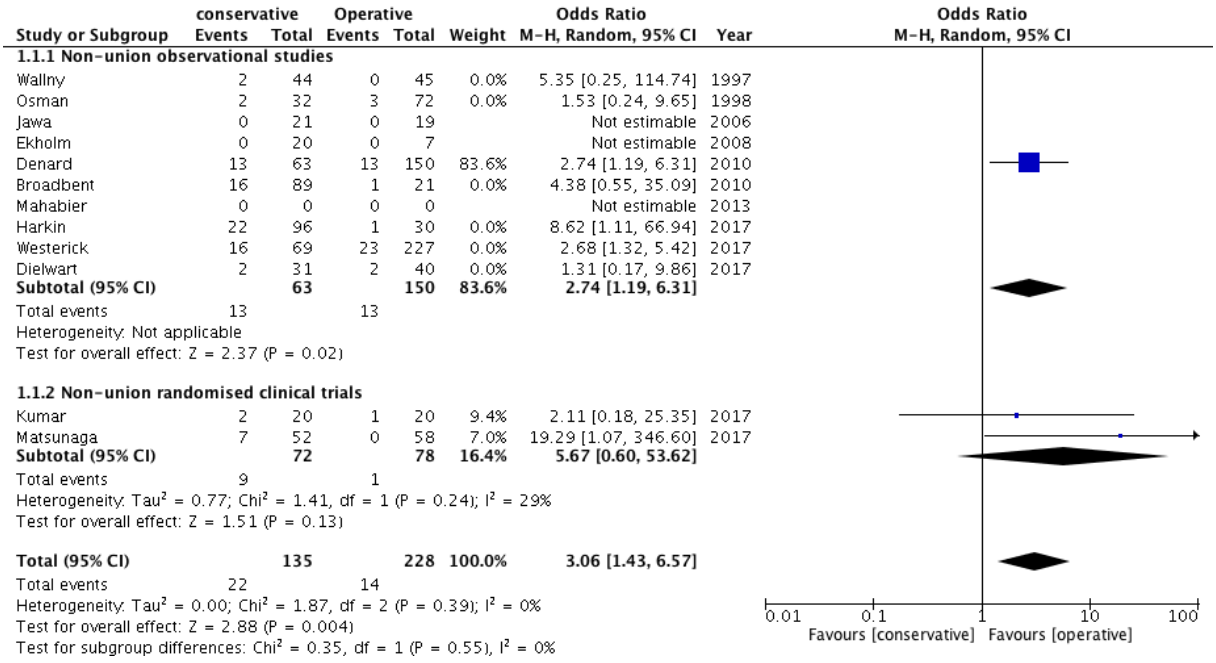

Figure S6. Forest-plot of non-union rate for studies comparing conservative treatment to plate fixation.

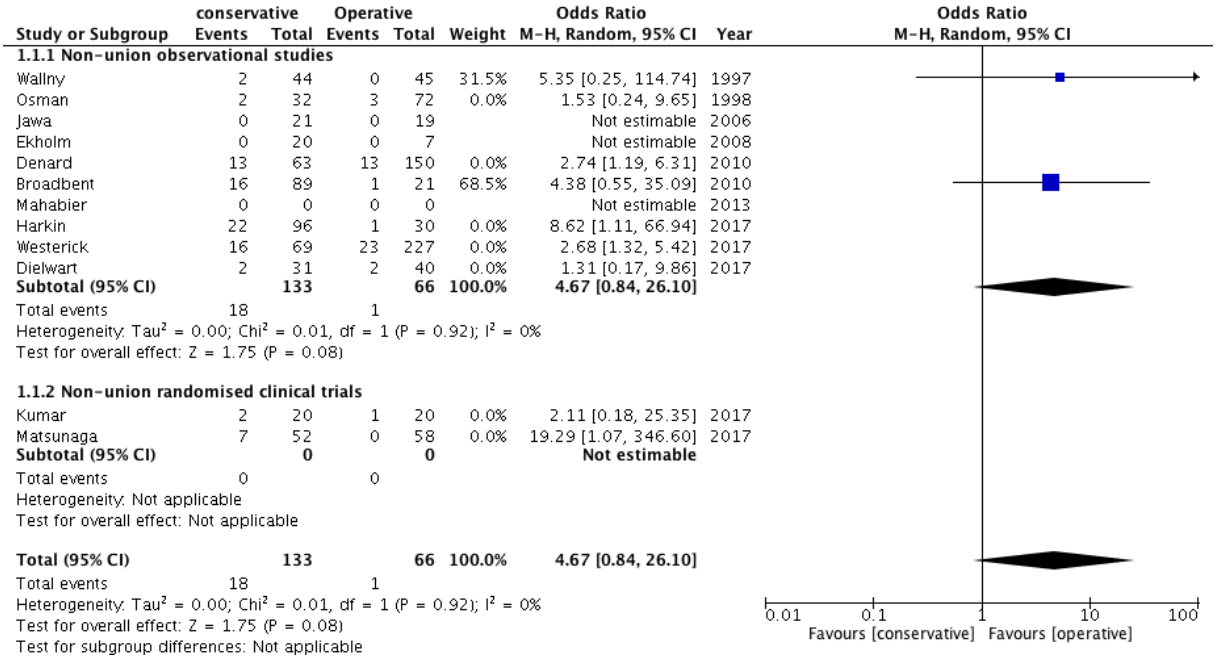

Figure S7. Forest-plot of non-union rate for studies comparing conservative treatment to plate fixation in patients older than 50 years.

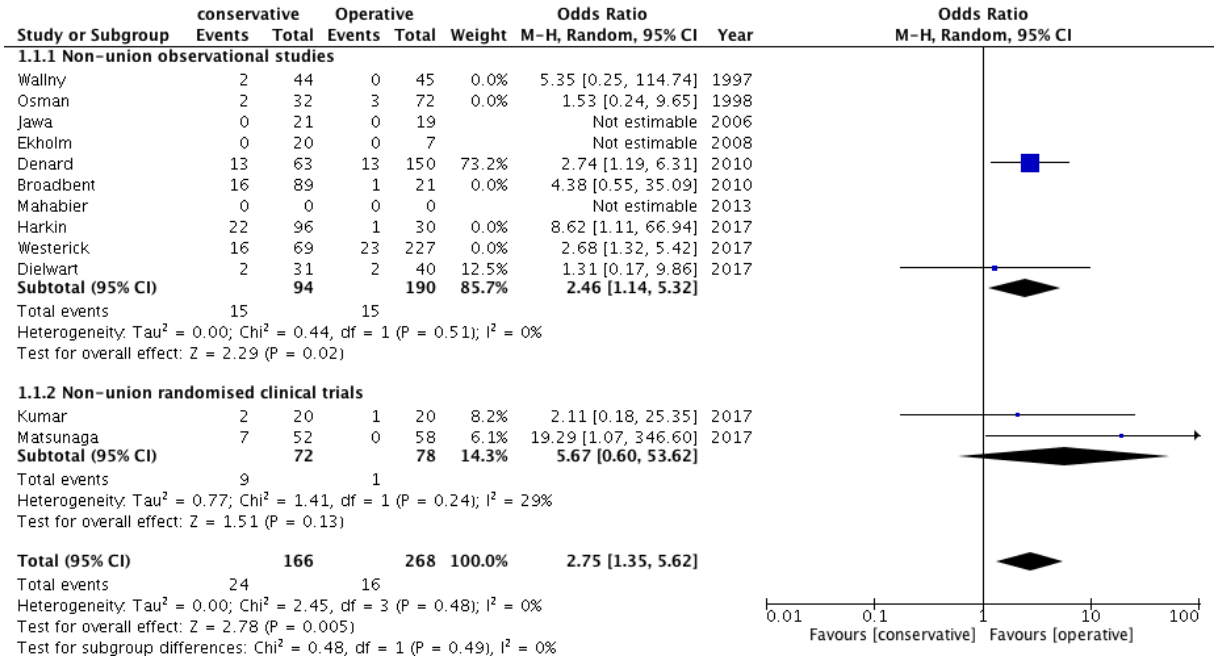

Figure S8. Forest-plot of non-union rate for high quality studies.

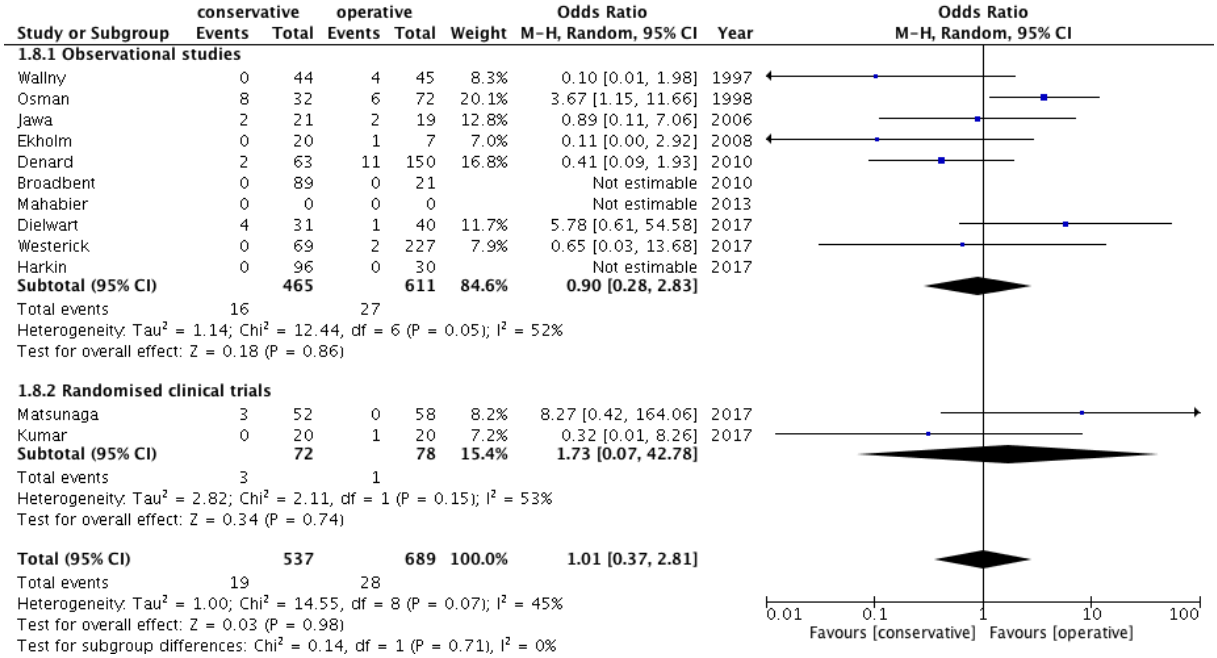

Figure S9. Forest-plot of (re)intervention rate for indications other than non-union.

CHAPTER 5

Operative versus nonoperative treatment of distal radius fractures in adults: a systematic review and meta-analysis

Yassine Ochen, Jesse Peek, Detlef van der Velde, Frank J.P. Beeres, Mark van Heijl, Rolf H.H. Groenwold, R. Marijn Houwert, Marilyn Heng

(JAMA Network Open)

Abstract

Background

No consensus has been reached to date regarding the optimal treatment for distal radius fractures. The international rate of operative treatment has been increasing, despite higher costs and limited functional outcome evidence to support this shift. The aim of this study was to compare functional, clinical, and radiologic outcomes after operative vs nonoperative treatment of distal radius fractures in adults.

Methods

The PubMed/MEDLINE, Embase, CENTRAL, and CINAHL databases were searched from inception to June 15, 2019, for studies comparing operative vs nonoperative treatment of distal radius fractures. Randomized clinical trials (RCTs) and observational studies reporting on the following: acute distal radius fracture with operative treatment (internal or external fixation) vs nonoperative treatment (cast immobilization, splinting, or bracing); patients 18 years or older; and functional outcome were included. Studies in a language other than English or reporting treatment for refracture were excluded. Data extraction was performed independently by 2 reviewers. Effect estimates were pooled using random-effects models and presented as risk ratios (RRs) or mean differences (MDs) with 95% CIs. Data were analyzed in September 2019. The primary outcome measures included medium-term functional outcome measured with the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH) and the overall complication rate after operative and nonoperative treatment.

Results

A total of 23 unique studies were included, consisting of 8 RCTs and 15 observational studies, that described 2254 unique patients. Among the studies that presented sex data, 1769 patients were women [80.6%]. Overall weighted mean age was 67 [range, 22-90] years). The RCTs included 656 patients (29.1%); observational studies, 1598 patients (70.9%). The overall pooled effect estimates the showed a significant improvement in medium-term (≤1 year) DASH score after operative treatment compared with nonoperative treatment (MD, −5.22 [95% CI, −8.87 to −1.57]; P = .005; I^2 = 84%). No difference in complication rate was observed (RR, 1.03 [95% CI, 0.69-1.55]; P = .87; I^2 = 62%). A significant improvement in grip strength was noted after operative treatment, measured in kilograms (MD, 2.73 [95% CI, 0.15-5.32]; P = .04; I^2 = 79%) and as a percentage of the unaffected side (MD, 8.21 [95% CI, 2.26-14.15]; P = .007; I^2 = 76%).

No improvement in medium-term DASH score was found in the subgroup of studies that only included patients 60 years or older (MD, −0.98 [95% CI, −3.52 to 1.57]; P = .45; I^2 = 34%]), compared with a larger improvement in medium-term DASH score after operative treatment in the other studies that included patients 18 years or older (MD, −7.50 [95% CI, −12.40 to −2.60]; P = .003; I^2 = 77%); the difference between these subgroups was statically significant (test for subgroup differences, P = .02).

Conclusion

This meta-analysis suggests that operative treatment of distal radius fractures improves the medium-term DASH score and grip strength compared with nonoperative treatment in adults, with no difference in overall complication rate. The findings suggest that operative treatment might be more effective and have a greater effect on the health and well-being of younger, nonelderly patients.

Introduction

The fracture of the distal radius is the most common injury in adults, accounting for approximately 17.5% of fractures.[1] Distal radius fractures have a bimodal age distribution in the population, with a peak incidence seen in patients younger than 18 years and a second peak in patients 50 years or older. Recent studies indicate the worldwide incidence of distal radius fractures is increasing each year owing to the overall potential to live longer with comorbidities such as osteoporosis.[2] Although the elderly population is at greatest risk, distal radius fractures still have a significant effect on the health and well-being of nonelderly adults. Reports have shown a significant increase of distal radius fractures in patients aged 17 to 64 years.[2]

The management of distal radius fractures consists of operative or nonoperative treatment. However, no consensus has been reached regarding the optimal treatment method. Several meta-analyses have been published on the comparison between operative and nonoperative treatment.[3-5] Recent meta-analyses have focused specifically on patient populations 60 years or older.[4,5] These meta-analyses found no difference in functional outcome between operative and nonoperative treatment in elderly patients. However, the international rate of operative treatment of distal radius fractures has been increasing, despite higher cost and limited functional outcome evidence to support this shift.[6]

At present, no meta-analysis, to our knowledge, has evaluated functional outcome in patients younger than 60 years by including all patients 18 years or older. Moreover, the high incidence of distal radius fractures and the inconsistencies in treatment practices indicate further investigation is warranted to understand current treatment methods and outcomes.[7]

Randomized clinical trials (RCTs) and observational studies are both increasingly used in orthopedic trauma meta-analyses for the evaluation of treatment effects.[8-12] Growing evidence shows that meta-analyses of RCTs and observational studies can be of value compared with meta-analyses of RCTs alone. Provided that observational studies are of high quality, the addition of observational studies in meta-analyses increases sample size and might provide a better insight into small treatment effects and infrequent outcome measures. Furthermore, observational studies might provide insight into treatment effects in a more heterogeneous patient population compared with the usually highly selected patient populations in RCTs.[13-18] The addition of observational studies in this meta-analysis could increase sample size and heterogeneity in patient

characteristics, which could lead to the evaluation of different age groups, compared with the previous highly selected meta-analyses focusing on the elderly.

The primary aim of this systematic review and meta-analysis was to compare functional, clinical, and radiologic outcomes after operative vs nonoperative treatment of distal radius fractures in adults. As a secondary aim, we sought to compare outcomes in studies that only included patients 60 years or older and other studies that included patients 18 years or older. Finally, we compared effect estimates from RCTs and observational studies.

Methods

This systematic review and meta-analysis was performed and reported according to the Meta-analysis of Observational Studies in Epidemiology (MOOSE) and the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guidelines.[19-21] This review of the literature did not require approval from the independent ethics committee or institutional review board of the participating institutions.

Search strategy and selection criteria

The PubMed/MEDLINE, Embase, CENTRAL (Cochrane Central Register of Controlled Trials), and CINAHL (Cumulative Index to Nursing and Allied Health Literature) databases were searched from inception to June 15, 2019, for studies comparing operative vs nonoperative treatment of distal radius fractures by 2 reviewers (Y.O. and J.P.). The search syntax is provided in supplementary Table S1. Duplicate articles were removed, and 2 reviewers (Y.O. and J.P.) independently performed title and abstract screening for eligibility of identified studies. All published comparative studies, including RCTs and observational studies, reporting on the comparison of operative vs nonoperative treatment of distal radius fractures were eligible for inclusion.

After title and abstract screening, full-text articles were reviewed independently by the same 2 reviewers (Y.O. and J.P.). Inclusion criteria consisted of (1) acute distal radius fracture, (2) operative treatment (internal or external fixation) vs nonoperative treatment (cast immobilization, splinting, or bracing), (3) patients 18 years or older, and (4) reporting of functional outcome. Exclusion criteria consisted of (1) treatment for refracture, (2) language other than English, (3) no availability of full text, and (4) letters, meeting proceedings, and case reports. Disagreements

on eligibility of full-text articles were resolved by consensus or by discussion with a third reviewer (M.H.). References of included studies were screened, and backward citation tracking was performed using Web of Science to identify articles not found in the original literature search.

Data extraction

Data extraction was performed independently by 2 reviewers (Y.O. and J.P.) with the use of a predefined data extraction form. The following characteristics were extracted from the included studies: first author, year of publication, study design, country in which the study was performed, study period, number of included patients, follow-up period, included age groups, AO fracture classification, operative method, and nonoperative method. Studies reporting on patient cohorts described in previously published articles were excluded or merged.

Quality assessment

The methodological quality of included studies was independently assessed by 2 reviewers (Y.O. and J.P.) using the Methodological Index for Non-randomized Studies (MINORS).[22] The MINORS is a validated instrument for the assessment of methodological quality and clear reporting of nonrandomized surgical studies, resulting in a score ranging from 0 to 24 (higher scores indicate better quality) for comparative studies.[22] Details on the methodological quality assessment are provided in supplementary Table S2. Disagreements were resolved by consensus.

Primary outcome measures

The primary outcome measures included medium-term functional outcome measured with the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH) and the overall complication rate after operative and nonoperative treatment. The DASH is a patient-reported outcome instrument developed to measure upper extremity disability and symptoms, resulting in a score ranging from no disability (0) to most severe disability (100).[23] Functional outcome scores were subdivided according to follow-up as medium term (≤1 year) and long term (>1 year). Complication rate was defined as the overall rate of complications and included reports of infection, nerve injury, chronic pain, complex regional pain syndrome, implant failure, and fracture healing disorders.

Secondary outcomes

Secondary functional outcome measures included the Patient-Rated Wrist Evaluation score[24] and the visual analogue scale score.[25] Secondary clinical outcome measures included grip strength, range of wrist extension (in degrees), range of wrist flexion (in degrees), range of wrist pronation (in degrees), range of wrist supination (in degrees), radial deviation (in degrees), and ulnar deviation (in degrees). Secondary radiologic outcome measures included volar tilt (in degrees), radial inclination (in degrees), radial height (in millimeters), articular step-off (in millimeters), and ulnar variance (in millimeters).

Statistical analysis

Data were analyzed in September 2019. Continuous variables are presented as means with SDs or ranges. Continuous variables were converted to mean (SD) if sufficient information was available, using the methods described in the Cochrane Handbook for Systematic Reviews of Interventions.[26] Dichotomous variables were extracted as absolute number and percentage. Dichotomous outcomes were pooled using the Mantel-Haenszel method and presented as risk ratios (RRs) with 95% CIs. Continuous outcomes were pooled using the inverse variance weighting method and presented as mean differences (MDs) with 95% CIs.[26] All analyses were performed using random-effects models. Statistical heterogeneity between studies was assessed by visual inspection of forest plots and by the I^2 and χ^2 statistics for heterogeneity. The significance level for treatment effects was determined by the overall-effect z test. All analyses were performed stratified by study design (RCT or observational study). Differences in effect estimates between the 2 subgroups were assessed, as described in the Cochrane Handbook for Systematic Reviews of Interventions.[26] The significance level for difference in effect estimates across the subgroups was determined by the test for subgroup differences. The significance level for treatment effects and differences across the subgroups was defined as 2-sided $P < .05$. Potential publication bias was assessed by visual inspection of funnel plots with MD or RR and standard error and Egger statistical tests.[27,28] Statistical meta-analyses were performed using Review Manager (RevMan, version 5.3.5).[29] Additional random-effects meta-regression analyses and Egger statistical tests for publication bias were performed in R, version 3.6.1 (R Project for Statistical Computing).[30]

Subgroup analyses

Subgroup analyses were performed for the primary outcome measures, the medium-term DASH score and complication rate, by stratifying by studies that only included patients 60 years or older and the other studies that included patients 18 years or older. In addition, random-effects meta-regression was performed, in which the reported mean difference in medium-term DASH score was regressed according to the mean age of the different study populations. Secondary subgroup analyses were performed including only high-quality studies and according to year of the study period. High-quality studies were defined as having a MINORS score of 16 or higher. The subgroup analyses for study period were performed with studies that included patients after 2008 to account for the development of new operative techniques and nonoperative treatment modalities during the past decade.

Results

Search

A flowchart of the literature search and study selection is shown in supplementary Figure S1. In total, 23 unique studies were included in this systematic review and meta-analysis, including 8 RCTs and 15 observational studies.[31-53]

Study characteristics

The 23 studies included 2254 unique patients, of whom 1040 were treated operatively and 1214 nonoperatively. The overall weighted mean age was 67 (range, 22-90) years (66 years in the operative group and 67 years in the nonoperative group). Overall, the studies that presented sex data included 425 men (19.4%) and 1769 women (80.6%). The overall follow-up ranged from 6 to 156 months. The baseline characteristics for RCTs and observational studies are presented in Table 1. In addition, supplementary Table S3 presents the treatment and fracture characteristics of all included studies. The studies included 851 patients (37.8%) who sustained an AO fracture type A; 164 (7.3%), type B; 689 (30.6%), type C; and 550 (24.4%), unknown type.

The 8 RCTs[31,35,36,38,46-49] included 656 patients (29.1%), of whom 322 were treated operatively and 334 nonoperatively. The weighted mean age was 67 years (67 years in the operative group and 68 years in the nonoperative group). The studies included 130 men (19.8%). The operative method was open reduction and internal fixation with a volar plate in 6 studies,[35,38,46-49] external fixation in

1 study,[31] and percutaneous pinning in 1 study.[36] The conservative method was cast immobilization in all studies.

The 15 observational studies (3 prospective[39,41,44] and 12 retrospective[32-34,37,40,42,43,45,50-53] cohort studies) included 1598 patients (70.9%). Operative treatment was performed in 718 patients (44.9%), and 880 (55.1%) were treated nonoperatively. The weighted mean age in the studies was 67 years (66 years in the operative group and 67 years in the nonoperative group). The studies that presented sex data included 295 men (19.2%). The operative method was open reduction and internal fixation with a volar plate in 6 studies,[34,39,41,42,51,53] external fixation in 1 study,[32] percutaneous pinning in 1 study,[33] intramedullary nail fixation in 1 study,[50] k-wire fixation in 1 study,[43] and unclear or a combination of methods in 5 studies.[37,40,44,45,52] The conservative method was cast immobilization in 13 studies[32,34,37,39-45,50,51,53] and unclear in 2 studies.[33,52]

Quality assessment

The overall mean MINORS score was 17.2 (SD, 3.6; range, 11-23). The mean MINORS score for the RCTs was 20.9 (SD, 2.0; range, 17-23). The mean MINORS score for the observational studies was 15.2 (SD, 2.5; range, 11-20). The details and distribution of MINORS scores are provided in supplementary Table S4.

Primary outcome measures

Medium-term (≤1 year) functional outcome assessed according to the DASH score was reported in 10 studies, including 4 RCTs[35,38,47,48] and 6 observational studies,[39-41,44,50,51] with 845 patients. The AO fracture type was known for 716 patients. Of these, 402 patients (56.1%) sustained an AO fracture type A; 55 (7.7%), type B; and 259 (36.2%), type C. The overall pooled effect revealed that operative treatment was associated with a significant improvement in the medium-term DASH score compared with nonoperative treatment (MD, −5.22 [95% CI, −8.87 to −1.57]; $P = .005$; $I^2 = 84\%$) (Figure 1). There was no difference in effect estimates from RCTs compared with observational studies (test for subgroup differences, $\chi^2_1 = 0.08$; $P = .78$). There was no visual asymmetry in the funnel plot (supplementary Figure S2). The Egger linear regression test (slope, 1.51; $t = 1.61$; $P = .15$) indicated no evidence of publication bias.

Table 1. Baseline characteristics of included studies in a meta-analysis of distal radius fractures

Study	Year	Study period	Design	Country	Overall number	Number		Age group included	Age (years) #		Male (%)		FU (months) #	
RCTs						**OP**	**NON**		**OP**	**NON**	**OP**	**NON**	**OP**	**NON**
Abbaszadegan et al.	1990	N/A	RCT	Sweden	47	23	24	>18	63 (22-75)		11 (23)		12	
Arora et al.	2011	2005-2008	RCT	Austria	73	36	37	>65	75.9 (65-88)	77.4 (65-89)	8 (22)	10 (27)	12	
Azzopardi et al.	2005	1997-2000	RCT	Scotland	54	27	27	>60	72 (8)	71 (9)	4 (15)	2 (7)	12	
Bartl et al.	2014	2008-2012	RCT	Germany	149	68	81	>65	75.3 (6.7)	74.4 (7.1)	9 (13)	12 (15)	12	
Martinez-Mendez et al.	2018	2012-2015	RCT	Spain	97	50	47	>60	67 (8)	70 (7)	11(22)	10 (21)	29 (24-48)	
Mulders et al.	2019	2013-2016	RCT	Netherlands	92	48	44	18-75	59 (42-66)*	60 (52-65)*	16 (33)	7(16)	12	
Sharma et al.	2014	2009-2010	RCT	India	64	32	32	22-55	52.4 (9.1)	48.1 (10.3)	12 (38)	14 (44)	24	
Sirniö et al.	2019	2008-2014	RCT	Finland	80	38	42	>50	62 (50-79)	64 (50-82)	1 (3)	3 (7)	24	
Observational studies														
Aktekin et al.	2010	N/A	RCS	Turkey	46	22	24	>65	69.8 (4.5)	71.2 (5.2)	9 (41)	5 (21)	27 (10.9)	23 (11.0)
Alm-Paulsen et al.	2012	1997-2006	RCS	Norway	60	30	30	30-85	61 (37-80)	60 (34-78)	N/A		72 (36-84)	84 (36-156)
Arora et al.	2009	2000-2005	RCS	Austria	114	53	61	>70	75.9 (4.8)	80.9 (5.7)	17 (32)	19 (31)	52 (12-64)	62 (12-81)
Barai et al.	2018	2014-2015	RCS	New Zealand	116	29	87	>18	58 (47-70)*	56 (29-68)*	10 (34)	25 (29)	18	
Chan et al.	2014	2009-2010	PCS	Singapore	75	40	35	>65	71.5 (5.2)	75.8 (9.3)	6 (15)	5 (14)	12	
Egol et al.	2010	2004-2008	RCS	United States	90	44	46	>65	73 (6.2)	76 (7.0)	8 (18)	6 (13)	12	
Gong et al.	2011	2008-2009	PCS	South Korea	50	26	24	>18	53 (13)	58 (13)	6 (23)	3 (13)	6	
Hung et al.	2015	2010-2013	RCS	China	57	26	31	61-80	65 (61-80)	64 (61-80)	5 (19)	7 (23)	12 (6-24)	
Jordan et al.	2016	2011-2013	RCS	United Kingdom	159	74	85	>50	66.3 (10.7)	68.7 (11.8)	12 (16)	6 (7)	24 (17-36)	
Larouche et al.	2016	N/A	PCS	Canada	129	70	59	>55	64.6 (7.6)		12 (9)		12	
Leerdam et al.	2019	2012	RCS	Netherlands	272	87	185	>18	62 (16)		69 (25)		46 (4)	
Lutz et al.	2014	1995-2011	RCS	United Kingdom	258	129	129	>65	74 (5, 65-90)		21 (8)		11.3 (9.3)	14.9 (8.9)
Tan et al.	2012	2006-2009	RCS	United States	63	31	32	>18	65 (15)	63 (18)	2 (6)	3 (9)	13 (12-17)	14 (11-23)
Toon et al.	2017	2011-2012	RCS	Singapore	60	32	28	>21	52.1 (23-77)	57.4 (26-79)	14 (44)	11 (39)	12	
Zengin et al.	2019	2014-2016	RCS	Turkey	49	25	24	>60	66.6 (7.4)	68.9 (8.7)	7 (28)	7 (29)	16.5 (3.1)	15.6 (4.4)

OP|NON operative/nonoperative; # SD or range; FU follow-up; N/A not available; RCT randomized controlled trial; PCS prospective cohort study; RCS retrospective cohort study; * median (IQR)

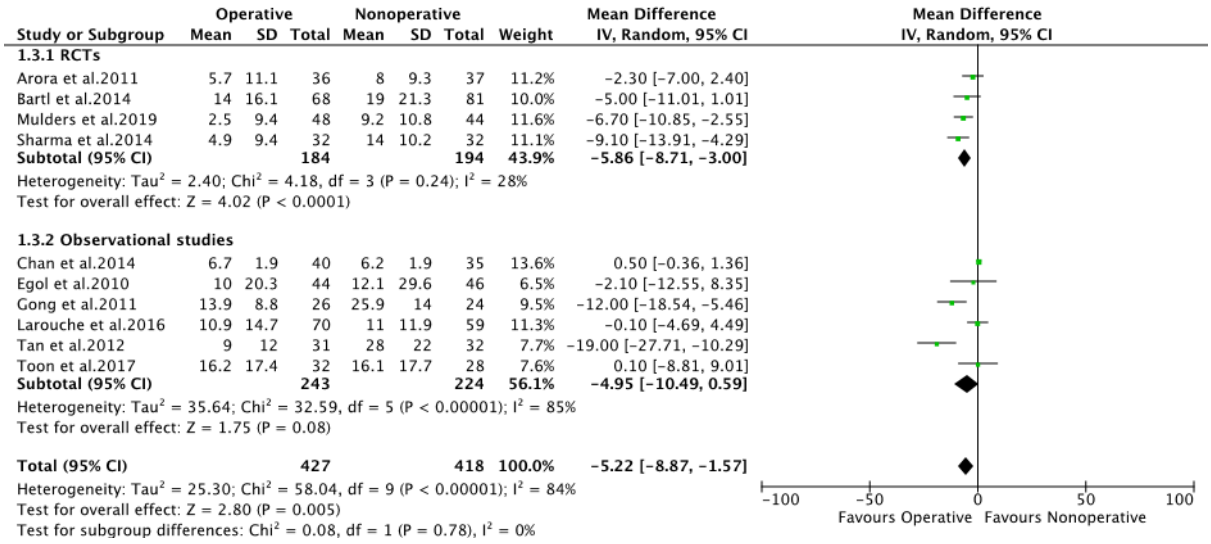

Figure 1. Forest Plot of Medium-Term Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH) Score. Medium term indicates 1 year or less. Results are reported using inverse-variance weighted random-effects methods. MD indicates mean difference; RCT, randomized clinical trial. Size of diamond markers indicates weight.

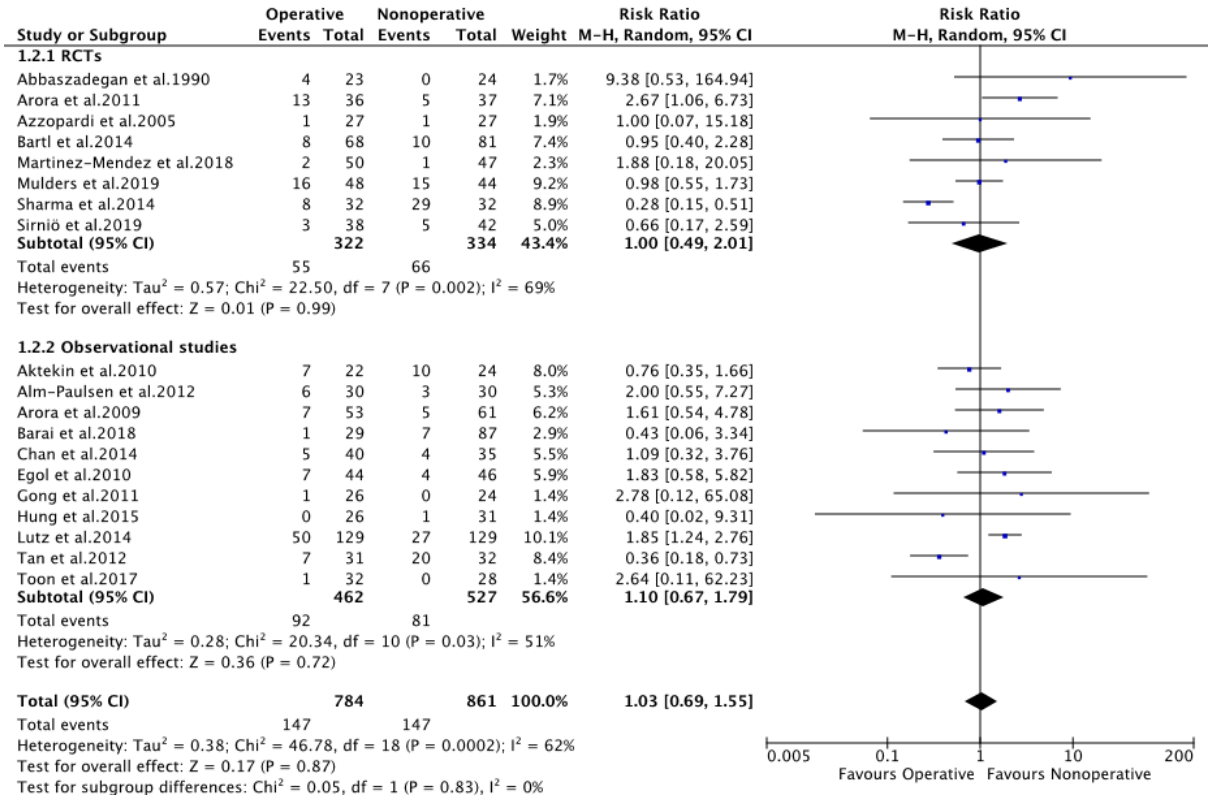

Figure 2. Forest Plot of Complication Rate of Distal Radius Fractures. Results are reported using inverse-variance weighted random-effects methods. RCT indicates randomized clinical trial; RR, risk ratio. Size of diamond markers indicates weight.

Complication rate was reported in 19 studies, including 8 RCTs[31,35,36,38,46-49] and 11 observational studies.[32-34,37,39-42,45,50,51] The overall pooled effect showed no difference in complication rate between operative and nonoperative treatment with an RR of 1.03 (95% CI, 0.69-1.55; $P = .87$; $I^2 = 62\%$) (Figure 2). No difference was found in effect estimates from RCTs compared with observational studies (test for subgroup differences, $\chi^2_1 = 0.05$; $P = .83$). There was no visual asymmetry in the funnel plot (supplementary Figure S3). The Egger linear regression test (slope, 1.11; $t = 0.02$; $P = .99$) indicated no evidence of publication bias. The incidence of complications was 18.8% (147 of 784) after operative treatment compared with 17.1% (147 of 861) after nonoperative treatment. Complication classification and incidence are presented in Table 2. The main complications after operative treatment were nerve injury or symptoms (26 of 784 [3.3%]) and infection (25 of 784 [3.2%]). The main complications after nonoperative treatment were nerve injury or symptoms (57 of 861 [6.6%]) and chronic pain or complex regional pain syndrome (33 of 861 [3.8%]).

Secondary functional outcome measures

No difference was found regarding the secondary functional outcome measures (supplementary Figures S4-S8). Descriptive details on functional outcome measures are provided in supplementary Table S5.

Secondary clinical outcome measures

Grip strength was reported in 13 studies, including 6 RCTs[35,36,46-49] and 7 observational studies,[33,34,39,40,50,51,53] and was assessed in kilograms (509 patients) and percentage of the unaffected side (462 patients). Both methods revealed an improvement of the grip strength in favor of operative treatment in grip strength measured in kilograms (MD, 2.73 [95% CI, 0.15-5.32]; $P = .04$; $I^2 = 79\%$) and grip strength as a percentage of the unaffected side (MD, 8.21 [95% CI, 2.26-14.15]; $P = .007$; $I^2 = 76\%$) (supplementary Figure S9 and Figure S10).

There was no difference regarding range of wrist extension, range of wrist flexion, range of wrist pronation, range of wrist supination, radial deviation, and ulnar deviation (supplementary Figures S11-S16). Descriptive details on clinical outcome measures are provided in in supplementary Table S6.

Table 2. Complications of included studies in a meta-analysis of distal radius fractures

Complication classification	Operative (n)	Incidence (%)	Nonoperative (n)	Incidence (%)
Infection	25	3.18	0	0
Nerve injury/symptoms	26	3.31	57	6.62
Carpal tunnel syndrome	8	1.02	12	1.39
Chronic pain/CRPS	21	2.67	33	3.83
Tendon injury	16	2.04	4	0.46
Implant failure	2	0.25	0	0
Wound dehiscence	1	0.12	0	0
Tenosynovitis	23	2.93	4	0.46
NP/other	22	2.80	14	1.62
Malunion/ nonunion/ malposition	3	0.38	23	2.67
Total	**147**	**18.75**	**147**	**17.07**

NP not specified; n number; CRPS complex regional pain syndrome

Secondary radiologic outcome measures

There was a significant improvement in favor of operative treatment regarding volar tilt (MD, 5.49° [95% CI, 2.94°-8.03°]; $P < .001$; $I^2 = 90\%$), radial inclination (MD, 3.46° [95% CI, 2.73°-4.18°]; $P = .001$; $I^2 = 54\%$), radial height (MD, 2.36 [95% CI, 1.87-2.85] mm; $P < .001$; $I^2 = 54\%$), and articular step-off (MD, −0.27 [95% CI, −0.51 to –0.03] mm; $P = .03$; $I^2 = 83\%$) (supplementary Figures S17-S20). There was no difference between treatment groups regarding the ulnar variance (MD, −0.29 [95% CI, −0.97 to 0.40] mm; $P = .41$; $I^2 = 92\%$) (supplementary Figure S21). Descriptive details on radiologic outcome measures are provided in supplementary Table S7

Subgroup analyses

The results of the subgroup analyses are presented in Table 3. The medium-term DASH score for studies that only included patients 60 years or older was reported in 4 studies (2 RCTs[35,38] and 2 observational studies[39,40]), with 387 patients and an overall mean age of 75 years. These studies included 247 patients (63.8%) who sustained an AO fracture type A; 9 (2.3%), type B; and 131 (33.9%), type C. The overall pooled effect showed no difference in the medium-term DASH score (MD, −0.98 [95% CI, −3.52 to 1.57]; $P = .45$; $I^2 = 34\%$) (supplementary Figure S22). The medium-term DASH score for other studies that included patients 18 years or older was reported in 6 studies (2 RCTs[47,48] and 4 observational studies[41,44,50,51]), with 458 patients and an overall mean age of 59 years. The AO fracture type was known for 329 patients, including 155 (47.1%) who sustained an AO fracture type A; 46 (14.0%), type B; and 128 (38.9%), type C. The overall pooled effect revealed operative treatment was associated with a significant improvement of the medium-term DASH score compared with nonoperative treatment (MD, −7.50 [95% CI, −12.40 to −2.60]; $P = .003$; $I^2 = 77\%$) (supplementary Figure S22). There was a significant difference in effect estimates from studies that only included patients 60 years or older compared with the

Table 3. Subgroup analyses of included studies in a meta-analysis of distal radius fractures.

	Short-term DASH score					Complication rate				
	n	**MD**	**95% CI**	**p-value**	**I²**	**n**	**RR**	**95% CI**	**p-value**	**I²**
All studies	10	-5.22	-8.87 to -1.57	0.005	84%	19	1.03	0.69 to 1.55	0.87	62%
Studies only age >60 y	4	-0.98	-3.52 to 1.57	0.45	34%	10	1.51	1.15 to 2.00	0.003	0%
Other studies age >18 y	6	-7.50	-12.40 to -2.60	0.003	77%	9	0.73	0.39 to 1.38	0.34	60%
High-quality studies	7	-6.98	-11.80 to -2.17	0.004	90%	11	0.88	0.50 to 1.55	0.66	64%
Study period (≥2008)	6	-5.31	-10.20 to -0.43	0.03	87%	10	0.72	0.44 to 1.17	0.18	34%

n number of studies; y years; MD mean difference; RR risk ratio; 95% CI confidence interval; I² heterogeneity

other studies that included patients 18 years or older (test for subgroup differences, $\chi^2_1 = 5.37$; $P = .02$) (supplementary Figure S22).

Results of the random-effects meta-regression analysis are shown in Figure 3; the trend of the MD in medium-term DASH score appears to decrease by 0.28 per year increase in the mean age of the study population (estimated regression coefficient, 0.28 [95% CI, −0.03 to 0.59]; $P = .07$). In the studies that only included patients 60 years or older, there was a significant difference in complication rate in favor of nonoperative treatment (RR, 1.51 [95% CI, 1.15-2.00]; $P = .003$; $I^2 = 0\%$), compared with other studies that included patients 18 years or older (RR, 0.73 [95% CI, 0.39-1.38]; $P = .34$; $I^2 = 60\%$) (test for subgroup differences: $P = .04$) (supplementary Figure S23). The results of all the secondary subgroup analyses are presented in Table 3 and supplementary Figures S24- S27).

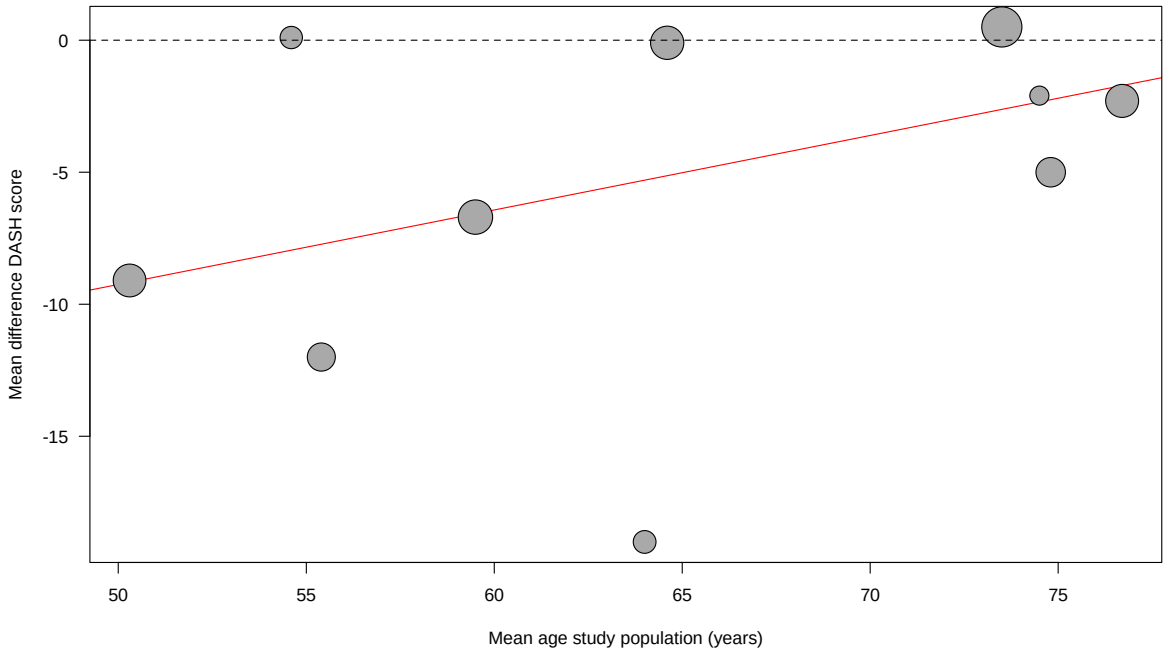

Figure 3. Random-Effects Meta-regression Plot. Data are expressed as medium-term (≤1 year) Disabilities of the Arm, Shoulder and Hand questionnaire (DASH) score (operative vs nonoperative groups) according to mean age of the study population in a meta-analysis of distal radius fractures. Circles represent the different studies, with circle size corresponding to the study weight. The black line represents the null value. MD indicates mean difference.

Discussion

Operative treatment of distal radius fractures was associated with an improvement in medium-term DASH score compared with nonoperative treatment in adults. No difference was observed in complication rate between treatment groups. There was also an improvement of grip strength in favor of operative treatment. However, no difference was found in medium-term DASH score in the subgroup of studies that only included patients 60 years or older. Furthermore, in the studies that only included these patients, a significant difference in complication rate favored nonoperative treatment. Subgroup analyses with high-quality studies and studies with a study period after 2008 showed similar results, compared with the primary analyses. No difference was found between effect estimates from RCTs and observational studies regarding the primary outcome measures (medium-term DASH score and complication rate).

The pooled effect estimates showed that operative treatment was associated with an improvement in medium-term DASH score compared with nonoperative treatment, which is in contrast to findings of previous meta-analyses.[3-5] Song et al[3] pooled functional outcome according to the medium-term DASH score at 12 months from 2 studies with 133 patients and found no difference between treatment groups. Ju et al[4] pooled the DASH score from 6 studies with 577 patients and reported no difference. Chen et al[5] found no difference in DASH score between treatment groups after they evaluated 7 studies with 600 patients. The present review included 10 studies with 845 patients in the medium-term DASH analysis, which resulted in an increased number of patients available for analyses, thus exceeding the samples of previous meta-analyses. Furthermore, only the meta-analysis by Song et al[3] evaluated the DASH score at 12 months. The meta-analyses by Ju et al[4] and Chen et al[5] did not distinguish between medium-term and long-term DASH scores, including the studies by Arora et al[34] and Aktekin et al[32] in their analyses. In the present review, the DASH scores reported by Arora et al[34] and Aktekin et al[32] were used for the evaluation of the long-term DASH score owing to their long-term follow-up periods to 81 months. In general, medium-term functional outcome can be assumed to reflect the effect of treatment, with long-term follow-up being influenced by other conditions, events, or patient factors that in turn could influence functional outcome scores. Reports have shown that the DASH score after distal radius fracture treatment tends to plateau after 12 months.[54,55]

The previous meta-analyses have mainly focused on elderly patients. Ju et al[4] and Chen et al[5] specifically focused on patient populations 60 years and older. Song et al[3] included only studies

with patients 45 years or older, with most of the patients in their DASH analyses 60 years or older. These findings are in accordance with our subgroup analyses of the studies that only included patients 60 years or older, showing no difference in medium-term DASH score. However, we found a significant improvement in medium-term DASH score in the subgroup of other studies that included patients 18 years or older. To our knowledge, with the analyses of 6 studies with 458 patients, this study is the first meta-analysis to evaluate functional outcome focusing on patient populations 18 years or older. The random-effects meta-regression plot confirmed this trend; however, with only 10 studies and based on the mean age of the complete population, the regression is underpowered. Meta-regression is an extension to subgroup analyses that allows the effect of characteristics to be investigated. However, this is rarely possible owing to inadequate numbers of studies, and meta regression should generally not be considered when there are fewer than 10 studies, as described in the Cochrane Handbook for Systematic Reviews of Interventions.[26] This trend shows that, to improve personalized care, further evaluation of individual patient data meta-analyses is needed.

We found no difference in the overall complication rate between operative and nonoperative treatment, in accordance with the studies by Song et al[3] and Yu et al.[56] However, in our analyses with studies that only included patients 60 years or older, a significant difference favored nonoperative treatment. These findings could indicate that operative treatment results in a higher risk of complications in the elderly population. The study by Chen et al[5] subdivided complications into minor and major, classifying minor as not requiring surgical treatment. They found no significant difference in minor complications; however, there was a significant difference in major complications, with the most common major complications being nerve and tendon injuries. In the present review, we did not subdivide major and minor complications; however, we did present complication classifications with incidence, showing that nerve injury or symptoms were the main complications in both groups. In the present review, we were not able to accurately compare major and minor complications or specify nerve injuries and symptoms. Unfortunately, this remains difficult owing to limited or missing information regarding the presentation and treatment of complications in studies.

We found a significant improvement of grip strength in favor of operative treatment, which is in contrast with 2 previous meta-analyses. Ju et al[4] found no significant difference in grip strength in their analysis of 4 studies with 337 patients. Song et al[3] evaluated grip strength at 12 months with

the results of 2 studies with 133 patients and found no difference. However, both the meta-analyses by Ju et al[4] and Song et al[3] could be limited by the number of included patients in their grip strength analyses. On the contrary, Chen et al[5] reported grip strength was significantly greater in the operative group in their analyses of 5 studies with 398 patients. In the present review, grip strength was reported in 13 studies and assessed in kilograms and percentage of the unaffected side with 509 and 462 patients, respectively.

We found no significant difference between treatment groups regarding range of wrist motions. These findings are also in accordance with those of Chen et al,[5] who reported wrist range of motion did not differ significantly at final follow-up between the 2 treatment groups.

Subgroup analyses including only high-quality studies or studies performed after 2008 showed similar results regarding the primary outcome measures, medium-term DASH score and complication rate, compared with the primary analyses. Furthermore, no difference was observed in effect estimates from RCTs and observational studies regarding the primary outcome measures. These results are in line with previous orthopedic trauma meta-analyses,[9-12] including RCTs and observational studies, showing high-quality observational studies to result in similar treatment effects compared with RCTs. Reports[9,11-15,18] have shown that differences in effect estimates between RCTs and observational studies tend to be small. Randomized clinical trials require strict conditions such as participant selection, inclusion and exclusion criteria, randomization method, and outcome measurements. Patient population in daily clinical practice might differ from the often highly selected patient populations in RCTs.[57-59] The results of observational studies, representing daily clinical practice with various levels of surgical experience and differences in operative techniques, could complement those of RCTs, provided that confounding has been adequately addressed.[17,18] Including observational studies in meta-analyses that evaluate surgical interventions increases sample size and may facilitate subgroup analysis. These results could help to understand the generalizability of previous results and improve existing guidelines.

Operative treatment of distal radius fractures results in a significant improvement of the medium-term DASH score and grip strength in adults, with no significant difference in overall complication rate. These results might support the international increase of operative treatment of distal radius fractures.[6] Operative treatment might be the preferred treatment for distal radius

fractures in younger patients. However, patient- and fracture-specific factors (patient preference, handedness, occupation, comorbidities, fracture displacement, etc) should always be taken into consideration, and patients should be counseled regarding incidence of complications. Studies have shown an increase of distal radius fractures in patients aged 17 to 64 years.[2] Hence, future studies should also focus on the nonelderly population, because traditionally most studies on this topic solely include patient populations 60 years or older. Further investigation is warranted to understand the optimal treatment methods and outcomes in this nonelderly, generally healthy, and still working age group. Furthermore, for the evaluation of the effect on the health and well-being of nonelderly adults, future studies could also focus on return to sporting activity and return to work, aside from traditional outcomes. Unfortunately, comparison of literature remains difficult owing to a wide variety of AO fracture types, different age groups, operative treatments, the use of different functional outcome measures, and duration of follow-up. Further research is needed for the development of patient- and fracture-specific guidelines.

Potential limitations in this review need to be acknowledged. First, analyses could be influenced by missing results; however, an extensive electronic database search was performed, and funnel plots did not indicate evidence of publication bias. Second, the subgroup analyses regarding age were stratified based on the inclusion criteria of studies, which resulted in overlap of the age distributions between the subgroup analyses. Nevertheless, there still was a substantial difference in the overall mean age in both subgroups (59 years vs 75 years). Furthermore, it should be noted that the cutoff of 60 years or older is arbitrarily chosen to compare our findings with the previous meta-analyses that mainly focused on patient populations 60 years and older. We acknowledge that better evidence is lacking, and further evaluation using individual patient data meta-analysis is needed. Third, we were not able to accurately classify all complications. Unfortunately, this remains difficult owing to insufficient or missing information. In addition, this review included a variety of fracture types. The AO fracture types A, B, and C seemed equally distributed throughout the different functional outcome analyses, with most studies including AO types A and C fractures. However, reports have shown patient-reported outcomes to vary in the setting of multiple-trauma or high-energy injury mechanisms. In addition to demographic and fracture characteristics, factors related to injury context (multiple-trauma, high-energy mechanism) could also account for differences in patient-reported wrist function after distal radius fractures.[60,61]

Conclusion

This meta-analysis found that operative treatment of distal radius fractures improved the medium-term DASH score and grip strength compared with nonoperative treatment in adults. There was no difference in complication rate between treatment groups. However, there was no difference in medium-term DASH score in the subgroup of studies that only included patients 60 years or older. Furthermore, in this subgroup, operative treatment resulted in a significantly higher complication rate. Our findings suggest that operative treatment might be more effective and have a greater effect on the health and well-being of younger, nonelderly patients. However, to improve personalized care, this trend needs to be confirmed with patient-level data. Further evaluation of individual patient data meta-analyses is needed.

References:

1. Court-Brown CM, Caesar B. Epidemiology of adult fractures: A review. *Injury.* 2006;37:691-697.
2. Corsino Cb Fau - Sieg RN, Sieg RN. Distal Radius Fractures *StatPearls.* 2019.
3. Song J, Yu A-X, Li Z-H. Comparison of conservative and operative treatment for distal radius fracture: a meta-analysis of randomized controlled trials. *International journal of clinical and experimental medicine.* 2015;8:17023-17035.
4. Ju J-H, Jin G-Z, Li G-X, Hu H-Y, Hou R-X. Comparison of treatment outcomes between nonsurgical and surgical treatment of distal radius fracture in elderly: a systematic review and meta-analysis. *Langenbeck's archives of surgery.* 2015;400:767-779.
5. Chen Y, Chen X, Li Z, Yan H, Zhou F, Gao W. Safety and Efficacy of Operative Versus Nonsurgical Management of Distal Radius Fractures in Elderly Patients: A Systematic Review and Meta-analysis. *The Journal of hand surgery.* 2016;41:404-413.
6. Armstrong KA, von Schroeder HP, Baxter NN, Zhong T, Huang A, McCabe SJ. Stable rates of operative treatment of distal radius fractures in Ontario, Canada: a population-based retrospective cohort study (2004-2013). *Canadian Journal of Surgery.* 2019(1488-2310 (Electronic)).
7. Mauck BM, Swigler CW. Evidence-Based Review of Distal Radius Fractures. *The Orthopedic clinics of North America.* 2018;49:211-222.
8. Houwert RM, Smeeing DPJ, Ahmed Ali U, Hietbrink F, Kruyt MC, van der Meijden OA. Plate fixation or intramedullary fixation for midshaft clavicle fractures: a systematic review and meta-analysis of randomized controlled trials and observational studies. *Journal of shoulder and elbow surgery.* 2016;25:1195-1203.
9. Smeeing DPJ, van der Ven DJC, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older: A Systematic Review, Meta-analysis, and Comparison of Randomized Controlled Trials and Observational Studies. *The American journal of sports medicine.* 2017;45:1937-1945.
10. Beks RB, Ochen Y, Frima H, et al. Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. *Journal of shoulder and elbow surgery.* 2018;27:1526-1534.
11. Beks RB, Peek J, de Jong MB, et al. Fixation of flail chest or multiple rib fractures: current evidence and how to proceed. A systematic review and meta-analysis. *European journal of trauma and emergency surgery : official publication of the European Trauma Society.* 2019;45:631-644.
12. Ochen Y, Beks RB, van Heijl M, et al. Operative treatment versus nonoperative treatment of Achilles tendon ruptures: systematic review and meta-analysis. *BMJ (Clinical research ed).* 2019;364:k5120.
13. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *The New England journal of medicine.* 2000;342:1878-1886.
14. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England journal of medicine.* 2000;342:1887-1892.
15. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *Journal of clinical epidemiology.* 2010;63:238-245.
16. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *The Cochrane database of systematic reviews.* 2014:MR000034.
17. Arditi C, Burnand B, Peytremann-Bridevaux I. Adding non-randomised studies to a Cochrane review brings complementary information for healthcare stakeholders: an augmented systematic review and meta-analysis. *BMC health services research.* 2016;16:598.
18. Frieden TR. Evidence for Health Decision Making - Beyond Randomized, Controlled Trials. *The New England journal of medicine.* 2017;377:465-475.
19. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000;283:2008-2012.
20. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Journal of clinical epidemiology.* 2009;62:1006-1012.
21. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of clinical epidemiology.* 2009;62:e1-34.
22. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ journal of surgery.* 2003;73:712-716.
23. Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC musculoskeletal disorders.* 2003;4:11.
24. MacDermid JC, Turgeon T, Richards RS, Beadle M, Roth JH. Patient rating of wrist pain and disability: a reliable and valid measurement tool. *Journal of orthopaedic trauma.* 1998;12:577-586.
25. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain.* 2011;152:2399-2404.

26. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. *The Cochrane Collaboration.* 2011;0:10-11.
27. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ (Clinical research ed).* 1997;315:629-634.
28. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *Jama.* 2006;295(6):676-680.
29. Review Manager, Cochrane Collaboration(2014).
30. Team R. R Development Core Team. *R: A Language and Environment for Statistical Computing.* 2013.
31. Abbaszadegan H, Jonsson U. External fixation or plaster cast for severely displaced Colles' fractures? Prospective 1-year study of 46 patients. *Acta orthopaedica Scandinavica.* 1990;61:528-530.
32. Aktekin CN, Altay M, Gursoy Z, Aktekin LA, Ozturk AM, Tabak AY. Comparison between external fixation and cast treatment in the management of distal radius fractures in patients aged 65 years and older. *The Journal of hand surgery.* 2010;35:736-742.
33. Alm-Paulsen PS, Rod O, Rod K, Rajabi B, Russwurm H, Finsen V. Percutaneous pinning of fractures of the distal radius. *Journal of plastic surgery and hand surgery.* 2012;46:195-199.
34. Arora R, Gabl M, Gschwentner M, Deml C, Krappinger D, Lutz M. A comparative study of clinical and radiologic outcomes of unstable colles type distal radius fractures in patients older than 70 years: nonoperative treatment versus volar locking plating. *Journal of orthopaedic trauma.* 2009;23:237-242.
35. Arora R, Lutz M, Deml C, Krappinger D, Haug L, Gabl M. A prospective randomized trial comparing nonoperative treatment with volar locking plate fixation for displaced and unstable distal radial fractures in patients sixty-five years of age and older. *The Journal of bone and joint surgery American volume.* 2011;93:2146-2153.
36. Azzopardi T, Ehrendorfer S, Coulton T, Abela M. Unstable extra-articular fractures of the distal radius: a prospective, randomised study of immobilisation in a cast versus supplementary percutaneous pinning. *The Journal of bone and joint surgery British volume.* 2005;87:837-840.
37. Barai A, Lambie B, Cosgrave C, Baxter J. Management of distal radius fractures in the emergency department: A long-term functional outcome measure study with the Disabilities of Arm, Shoulder and Hand (DASH) scores. *Emergency medicine Australasia : EMA.* 2018;30:530-537.
38. Bartl C, Stengel D, Bruckner T, Gebhard F. The treatment of displaced intra-articular distal radius fractures in elderly patients. *Deutsches Arzteblatt international.* 2014;111:779-787.
39. Chan Y-H, Foo T-L, Yeo C-J, Chew WY-C. Comparison between cast immobilization versus volar locking plate fixation of distal radius fractures in active elderly patients, the Asian perspective. *Hand surgery : an international journal devoted to hand and upper limb surgery and related research : journal of the Asia-Pacific Federation of Societies for Surgery of the Hand.* 2014;19:19-23.
40. Egol KA, Walsh M, Romo-Cardoso S, Dorsky S, Paksima N. Distal radial fractures in the elderly: operative compared with nonoperative treatment. *The Journal of bone and joint surgery American volume.* 2010;92:1851-1857.
41. Gong HS, Lee JO, Huh JK, Oh JH, Kim SH, Baek GH. Comparison of depressive symptoms during the early recovery period in patients with a distal radius fracture treated by volar plating and cast immobilisation. *Injury.* 2011;42:1266-1270.
42. Hung LP, Leung YF, Ip WY, Lee YL. Is locking plate fixation a better option than casting for distal radius fracture in elderly people? *Hong Kong medical journal = Xianggang yi xue za zhi.* 2015;21:407-410.
43. Jordan RW, Naeem R, Jadoon S, Parsons H, Shyamalan G. Cast Immobilisation versus Wire Fixation in the Management of Middle-aged and Elderly Patients with Distal Radius Fractures. *The journal of hand surgery Asian-Pacific volume.* 2016;21:18-23.
44. Larouche J, Pike J, Slobogean GP, et al. Determinants of Functional Outcome in Distal Radius Fractures in High-Functioning Patients Older Than 55 Years. *Journal of orthopaedic trauma.* 2016;30:445-449.
45. Lutz K, Yeoh KM, MacDermid JC, Symonette C, Grewal R. Complications associated with operative versus nonsurgical treatment of distal radius fractures in patients aged 65 years and older. *The Journal of hand surgery.* 2014;39:1280-1286.
46. Martinez-Mendez D, Lizaur-Utrilla A, de-Juan-Herrero J. Intra-articular distal radius fractures in elderly patients: a randomized prospective study of casting versus volar plating. *The Journal of hand surgery, European volume.* 2018;43:142-147.
47. Mulders MAM, Walenkamp MMJ, van Dieren S, Goslings JC, Schep NWL. Volar Plate Fixation Versus Plaster Immobilization in Acceptably Reduced Extra-Articular Distal Radial Fractures: A Multicenter Randomized Controlled Trial. *The Journal of bone and joint surgery American volume.* 2019;101:787-796.
48. Sharma H, Khare GN, Singh S, Ramaswamy AG, Kumaraswamy V, Singh AK. Outcomes and complications of fractures of distal radius (AO type B and C): volar plating versus nonoperative treatment. *Journal of orthopaedic science : official journal of the Japanese Orthopaedic Association.* 2014;19:537-544.
49. Sirnio K, Leppilahti J, Ohtonen P, Flinkkila T. Early palmar plate fixation of distal radius fractures may benefit patients aged 50 years or older: a randomized trial comparing 2 different treatment protocols. *Acta orthopaedica.* 2019;90:123-128.

50. Tan V, Bratchenko W, Nourbakhsh A, Capo J. Comparative analysis of intramedullary nail fixation versus casting for treatment of distal radius fractures. *The Journal of hand surgery.* 2012;37:460-468.e461.
51. Toon DH, Premchand RAX, Sim J, Vaikunthan R. Outcomes and financial implications of intra-articular distal radius fractures: a comparative study of open reduction internal fixation (ORIF) with volar locking plates versus nonoperative management. *Journal of orthopaedics and traumatology : official journal of the Italian Society of Orthopaedics and Traumatology.* 2017;18:229-234.
52. van Leerdam RH, Huizing F, Termaat F, et al. Patient-reported outcomes after a distal radius fracture in adults: a 3-4 years follow-up. *Acta orthopaedica.* 2019;90:129-134.
53. Zengin EC, Ozcan C, Aslan C, Bulut T, Sener M. Cast immobilization versus volar locking plate fixation of AO type C distal radial fractures in patients aged 60 years and older. *Acta orthopaedica et traumatologica turcica.* 2019;53:15-18.
54. Abramo A, Kopylov P, Tagil M. Evaluation of a treatment protocol in distal radius fractures: a prospective study in 581 patients using DASH as outcome. *Acta Orthop.* 2008;79(3):376-385.
55. Landgren M, Abramo A, Geijer M, Kopylov P, Tagil M. Similar 1-year subjective outcome after a distal radius fracture during the 10-year-period 2003-2012. *Acta Orthop.* 2017;88(4):451-456.
56. Yu GS, Lin YB, Le LS, Zhan MF, Jiang XX. Internal fixation vs conservative treatment for displaced distal radius fractures: a meta-analysis of randomized controlled trials. *Ulusal travma ve acil cerrahi dergisi = Turkish journal of trauma & emergency surgery : TJTES.* 2016;22:233-241.
57. Boyko EJ. Observational research--opportunities and limitations. *Journal of diabetes and its complications.* 2013;27(6):642-648.
58. Khan AY, Preskorn SH, Baker B. Effect of study criteria on recruitment and generalizability of the results. *Journal of clinical psychopharmacology.* 2005;25(3):271-275.
59. Van Spall HGC, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA.* 2007;297(11):1233-1240.
60. Ferree S, van der Vliet QMJ, Nawijn F, et al. Epidemiology of distal radius fractures in polytrauma patients and the influence of high traumatic energy transfer. *Injury.* 2018;49:630-635.
61. van der Vliet QMJ, Sweet AAR, Bhashyam AR, et al. Polytrauma and High-energy Injury Mechanisms Are Associated with Worse Patient-reported Outcomes After Distal Radius Fractures. *Clinical orthopaedics and related research.* 2019.

Supplementary materials to Chapter 5

Table S1. Search syntax performed last on June 15, 2019

Database	Syntax
PubMed/MEDLINE (n= 1838)	(((((((((radius fractures[MeSH Terms]) AND distal[Title/Abstract])) OR colles' fracture[MeSH Terms]) OR wrist injuries[MeSH Terms])) OR (((((((radius[Title/Abstract]) OR radial[Title/Abstract])) AND distal[Title/Abstract])) AND fractur*[Title/Abstract])) OR ((((colles[Title/Abstract]) OR smith[Title/Abstract]) OR barton[Title/Abstract]) OR wrist[Title/Abstract]))) AND fractur*[Title/Abstract])))))) AND (((((((surgical procedure, operative[MeSH Terms]) OR fracture fixation[MeSH Terms]) OR orthopedic procedure[MeSH Terms]) OR orthopedics[MeSH Terms])) OR ((((((((((surg*[Title/Abstract]) OR operat*[Title/Abstract]) OR orthop*[Title/Abstract]) OR pin*[Title/Abstract]) OR nail*[Title/Abstract]) OR screw*[Title/Abstract]) OR plat*[Title/Abstract]) OR rod*[Title/Abstract]) OR wire*[Title/Abstract]) OR fix*[Title/Abstract]) OR ORIF[Title/Abstract]) OR ExFix[Title/Abstract]))) AND ((((conservative treatment[MeSH Terms]) OR physical therapy modalities[MeSH Terms])) OR ((((((((((((((conserv*[Title/Abstract]) OR conven*[Title/Abstract]) OR non-operat*[Title/Abstract]) OR "non operative"[Title/Abstract]) OR nonoperat*[Title/Abstract]) OR non-surg*[Title/Abstract]) OR "non surgical"[Title/Abstract]) OR nonsurg*[Title/Abstract]) OR cast*[Title/Abstract]) OR splint*[Title/Abstract]) OR brace*[Title/Abstract]) OR bracing[Title/Abstract]) OR plaster[Title/Abstract]) OR bandage*[Title/Abstract]) OR tape*[Title/Abstract]) OR taping[Title/Abstract]))))
Embase (n= 1713)	('distal radius fracture'/exp OR 'colles fracture'/exp OR (('radius':ab,ti OR 'radial':ab,ti) AND 'distal':ab,ti AND 'fractur*':ab,ti) OR (('colles':ab,ti OR 'smith':ab,ti OR 'barton':ab,ti OR 'wrist':ab,ti) AND 'fractur*':ab,ti)) AND ('surgery'/de OR 'orthopedic surgery'/de OR 'surg*':ab,ti OR 'operat*':ab,ti OR 'orthop*':ab,ti OR 'pin*':ab,ti OR 'nail*':ab,ti OR 'screw*':ab,ti OR 'plate*':ab,ti OR 'rod*':ab,ti OR 'wire*':ab,ti OR 'fix*':ab,ti OR 'orif':ab,ti OR 'exfix':ab,ti) AND ('conservative treatment'/de OR 'conservative':ab,ti OR 'conventional':ab,ti OR 'non-operative':ab,ti OR 'non operative':ab,ti OR 'nonoperative':ab,ti OR 'non-surgical':ab,ti OR 'non surgical':ab,ti OR 'nonsurgical':ab,ti OR 'cast*':ab,ti OR 'splint*':ab,ti OR 'brace*':ab,ti OR 'bracing':ab,ti OR 'plaster*':ab,ti OR 'bandage':ab,ti OR 'tape*':ab,ti OR 'taping*':ab,ti)
CENTRAL (n= 837)	Radius AND distal AND fracture
CINAHL (n= 272)	(((MH distal radius OR TI distal radius OR AB distal radius OR TI radius OR AB radius OR TI radial OR AB radial OR TI colles OR AB colles OR TI smith OR AB smith OR TI barton OR AB barton OR TI wrist OR AB wrist) AND (MH fracture OR MH fractures OR TI fractur* OR AB fractur*)) AND ((MH surgical procedures, operative OR MH orthopedics OR TI surg* OR AB surg* OR TI operat* OR AB operat* OR TI orthop* OR AB orthop* OR TI pin* OR AB pin* OR TI nail* OR AB nail* OR TI screw* OR AB screw* OR TI plate* OR AB plate* OR TI rod* OR AB rod* OR TI wire* OR AB wire* OR TI fix* OR AB fix* OR TI ORIF OR AB ORIF OR TI ExFix OR AB ExFix) AND (MH Conservative Treatment OR MH physical therapy modalities OR TI conservative OR AB conservative OR TI conventional OR AB conventional OR TI non-operative OR AB non-operative OR TI non operative OR AB non operative OR TI nonoperative OR AB nonoperative OR TI non-surgical OR AB non-surgical OR TI non surgical OR AB non surgical OR TI nonsurgical OR AB nonsurgical OR TI cast* OR AB cast* OR TI brace* OR AB brace* OR TI splint* OR AB splint* OR TI bracing OR AB bracing OR TI bandage* OR AB bandage* OR TI tape* OR AB tape* OR TI taping OR AB taping OR TI plaster* OR AB plaster*))

Table S2. Quality assessment according to the MINORS criteria in a meta-analysis of distal radius fractures

Criteria	Reported and adequate (2)	Reported but inadequate (1)	Not reported (0)
Clearly stated aim	Aim including outcomes reported	Aim reported without outcomes	Not reported
Inclusion consecutive patients	Inclusion/exclusion criteria reported	Unclear description inclusion/exclusion criteria	Not reported
Prospective collection data	Prospective	Not applicable	Not applicable
Appropriate endpoints	Appropriate endpoints to aim study	Endpoints not appropriate to aim study	Not reported
Unbiased assessment	Blinded evaluation of outcomes	Reason not blinding stated	Not reported
Appropriate follow-up	≥ 1 year	< 1 year	Not reported
Loss to follow-up < 5%	≤ 5%	> 5%	Not applicable
Prospective calculation study size	Prospective power-analysis performed	Prospective calculation without power-analysis	Not applicable
Adequate control group	Operative versus nonoperative treatment	Not applicable	Not applicable
Contemporary groups	Study/control group managed during same period	Study/control not managed during same period	Not reported
Baseline equivalence groups	Baseline characteristics described and comparable	Baseline characteristics not comparable	Not reported
Adequate statistical analyses	Statistical analysis described including type of analyses	Inadequate description statistical analysis	Not reported

Items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S3. Treatment characteristics of included studies in a meta-analysis of distal radius fractures

Study	Year	Overall number	Number		Included age group (years)	Fracture type (AO), n (%)			Operative treatment	Nonoperative treatment
RCTs			**OP**	**NON**		**A**	**B**	**C**		
Abbaszadegan et al.	1990	47	23	24	>18	N/A	N/A	N/A	External fixation	Cast
Arora et al.	2011	73	36	37	>65	22 (30)	0	51 (70)	ORIF volar locking plate	Cast
Azzopardi et al.	2005	54	27	27	>60	54 (100)	0	0	Percutaneous pinning	Cast
Bartl et al.	2014	149	68	81	>65	149 (100)	0	0	ORIF volar locking plate	Cast
Martinez-Mendez et al.	2018	97	50	47	>60	0	0	97 (100)	ORIF volar locking plate	Cast
Mulders et al.	2019	92	48	44	18-75	92 (100)	0	0	ORIF volar locking plate	Cast
Sharma et al.	2014	64	32	32	22-55	0	28 (44)	36 (56)	ORIF volar locking plate	Cast
Sirniö et al.	2019	80	38	42	>50	48 (60)	0	32 (40)	ORIF volar locking plate	Cast
Observational studies										
Aktekin et al.	2010	46	22	24	>65	22 (48)	0	24 (52)	External fixation	Cast
Alm-Paulsen et al.	2012	60	30	30	30-85	34 (57)	0	26 (43)	Percutaneous pinning	N/A
Arora et al.	2009	114	53	61	>70	59 (52)	0	55 (48)	ORIF volar locking plate	Cast
Barai et al.	2018	116	29	87	>18	N/A	N/A	N/A	ORIF	Cast
Chan et al.	2014	75	40	35	>65	33 (44)	0	42 (56)	ORIF volar locking plate	Cast
Egol et al.	2010	90	44	46	>65	43 (48)	9 (10)	38 (42)	Volar plate/external fixation	Cast
Gong et al.	2011	50	26	24	>18	26 (52)	0	24 (48)	ORIF volar locking plate	Cast
Hung et al.	2015	57	26	31	61-80	23 (40)	20 (35)	14 (25)	ORIF volar locking plate	Cast
Jordan et al.	2016	159	74	85	>50	102 (64)	3 (2)	54 (34)	K-wire fixation + cast	Cast
Larouche et al.	2016	129	70	59	>55	N/A	N/A	N/A	ORIF	Cast
Leerdam et al.	2019	272	87	185	>18	107 (39)	86 (32)	79 (29)	N/A	N/A
Lutz et al.	2014	258	129	129	>65	N/A	N/A	N/A	Multiple	Cast
Tan et al.	2012	63	31	32	>18	37 (59)	2 (3)	24 (38)	Intramedullary nail fixation	Cast
Toon et al.	2017	60	32	28	>21	0	16 (27)	44 (73)	ORIF volar locking plate	Cast
Zengin et al.	2019	49	25	24	>60	0	0	49 (100)	ORIF volar locking plate	Cast

OP|NON operative/nonoperative; N/A not available; n number; ORIF Open Reduction Internal Fixation; AO Classification of Fractures

Table S4. Quality assessment of included studies in a meta-analysis of distal radius fractures

MINORS criteria	**RCTs**	Abbaszadegan et al. 1190	Arora et al. 2011	Azzopardi et al. 2005	Bartl et al. 2014	Martinez-Mendez et al. 2018	Mulders et al.2019	Sharma et al. 2014	Sirniö et al. 20119	**Observational studies**	Aktekin et al. 2010	Alm-Paulsen et al. 2012	Arora et al. 2009	Barai et al. 2018	Chan et al. 2014	Egol et al. 2010	Gong et al. 2011	Hung et al. 2015	Jordan et al. 2016	Larouche et al. 2016	Leerdam et al. 2019	Lutz et al. 2014	Tan et al. 2012	Toon et al. 2017	Zengin et al. 2019
Clearly stated aim		1	2	2	2	2	2	2	2		2	2	2	2	2	2	1	2	2	2	2	2	2	2	2
Inclusion of consecutive patients		2	2	2	2	2	2	2	2		2	1	2	2	2	1	2	2	2	1	1	1	2	2	2
Prospective collection of data		2	2	2	2	2	2	2	2		1	1	1	1	2	1	2	0	1	2	0	1	1	1	0
Appropriate endpoints		2	2	2	2	2	2	2	2		2	2	2	2	2	2	2	1	2	2	2	2	2	2	2
Unbiased assessment endpoints		0	1	1	2	1	1	0	0		1	0	1	0	0	0	1	0	2	0	0	0	1	0	0
Appropriate follow-up		2	2	2	2	2	2	2	2		2	2	2	1	2	2	1	1	2	2	2	1	2	2	2
Loss to follow-up < 5%		2	1	1	1	2	2	2	1		1	2	1	1	2	1	1	0	1	0	0	2	2	1	0
Prospective calculation study size		0	2	2	0	2	2	0	2		0	0	0	2	0	0	2	0	0	2	0	0	0	2	0
Adequate control group		2	2	2	2	2	2	2	2		2	2	2	2	2	2	2	2	2	2	0	2	2	2	2
Contemporary groups		2	2	2	2	2	2	2	2		0	1	0	0	2	2	2	2	2	0	2	2	2	0	2
Baseline equivalence of groups		0	2	1	2	2	2	2	1		2	1	2	2	1	2	2	2	2	0	0	1	2	2	2
Adequate statistical analysis		2	2	2	2	2	2	2	2		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Total MINORS score		**17**	**22**	**21**	**21**	**23**	**23**	**20**	**20**		**17**	**16**	**17**	**17**	**19**	**17**	**20**	**14**	**20**	**15**	**11**	**16**	**20**	**18**	**16**

Items are scored 0 (not reported/ not applicable), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S5. Functional outcome measures of included studies in a meta-analysis of distal radius fractures

Study	DASH ≤ 1 year#		PRWE ≤ 1 year#		VAS ≤ 1 year#		DASH > 1 year#		PRWE > 1 year#		VAS > 1 year#	
RCTs	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**
Abbaszadegan et al. 1990	N/A		N/A		0	1	N/A		N/A		N/A	
Arora et al. 2011	5,7 (11,1)	8,0 (9,3)	12,8 (23,2)	14,6 (22,8)	0,1 (0,3)	0,1 (0,5)	N/A		N/A		N/A	
Azzopardi et al. 2005	N/A		N/A		0,7 (1,3)	1,2 (1,6)	N/A		N/A		N/A	
Bartl et al. 2014	14 (16,1)	19 (21,3)	N/A		N/A		N/A		N/A		N/A	
Martinez-Mendez et al. 2018	N/A		N/A		N/A		16 (14)	28 (21)	17 (13)	30 (25)	2 (2)	3 (2)
Mulders et al. 2019	2,5 (9,4)*	9,2 (10,8)*	4,0 (9,3)*	8,0 (11,9)*	N/A		N/A		N/A		N/A	
Sharma et al. 2014	4,9 (9,4)	14,0 (10,2)	N/A		N/A		N/A		N/A		N/A	
Sirniö et al. 2019	7,2	14,4	N/A		N/A		N/A		N/A		N/A	
Observational studies												
Aktekin et al. 2010	N/A		N/A		N/A		21,9 (21,7)	20,3 (13)	N/A		N/A	
Alm-Paulsen et al. 2012	N/A		N/A		N/A		13,5 (15)	10 (13)	11 (15)	8 (12)	5 (11)	2 (7)
Arora et al. 2009	N/A		N/A		N/A		11,1 (12,9)*	11,6 (13,4)*	9,3 (9,3)*	16,9 (12,1)*	1,7 (1,4)	0,7 (1,4)
Barai et al. 2018	N/A		N/A		N/A		12,1 (14,4)	6 (10,9)	N/A		N/A	
Chan et al. 2014	6,7 (1,9)	6,2 (1,9)	N/A		N/A		N/A		N/A		N/A	
Egol et al. 2010	10 (20,3)	12,1 (29,6)	N/A		1,2 (1,7)	1,5 (2,1)	N/A		N/A		N/A	
Gong et al. 2011	13,9 (8,8)	25,9 (14,0)	N/A		N/A		N/A		N/A		N/A	
Hung et al. 2015	N/A		N/A		N/A		4,5	13,6	N/A		N/A	
Jordan et al. 2016	N/A		N/A		N/A		26,6 (6,5)	27,1 (7,7)	N/A		N/A	
Larouche et al. 2016	10,9 (14,7)	11,0 (11,9)	12,3 (15,8)	10,9 (14,1)	N/A		N/A		N/A		N/A	
Leerdam et al. 2019	N/A		N/A		N/A		N/A		17 (22)	8 (15)	N/A	
Lutz et al. 2014	N/A		N/A		N/A		N/A		17 (23)	16 (18)	N/A	
Tan et al. 2012	9 (12)	28 (22)	N/A		N/A		7 (9)	25 (24)	N/A		N/A	
Toon et al. 2017	16,2 (17,4)	16,1 (17,7)	N/A		1,8 (1,6)	1,1 (1,1)	N/A		N/A		N/A	
Zengin et al. 2019	N/A		N/A		N/A		11,7 (8)	17,6 (14,2)	N/A		N/A	

OP|NON operative/nonoperative; n number; # mean (SD); N/A not available; * median

Table S6. Clinical outcome measures of included studies in a meta-analysis of distal radius fractures

Study	Complication$		Grip strength (kg)#		Range wrist extension (°)#		Range wrist flexion (°)#	
RCTs	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**
Abbaszadegan et al. 1990	4 (17)	0	N/A		N/A		N/A	
Arora et al. 2011	13 (36)	5 (14)	22,2 (6,3)	18,8 (5,8)	59 (10)	61 (7)	55 (11)	57 (10)
Azzopardi et al. 2005	1 (4)	1 (4)	77 (21)*	72 (17)*	94 (11)*	95 (9)*	87 (12)*	82 (15)*
Bartl et al. 2014	8 (12)	10 (12)	N/A		7,5 (11,7)**	7,5 (10)**	8,2 (11,9)**	11,5 (12,8)**
Martinez-Mendez et al. 2018	2 (4)	1 (2)	73 (27)*	64 (33)*	57 (11)	54 (13)	54 (13)	60 (16)
Mulders et al. 2019	16 (33)	25 (57)	26 (1,9)^	20 (8,4)^	85 (7,4)^	80 (14,8)^	80 (11,9)^	70 (14,8)^
Sharma et al. 2014	8 (25)	29 (91)	89,0 (4,3)*	72,1 (4,4)*	84,3 (2,4)	69,0 (3,91)	83,8 (2,9)	65,9 (7,5)
Sirniö et al. 2019	3 (8)	5 (12)	27 (5)	26 (7)	69 (6)	68 (7)	71 (7)	64 (11)
Observational studies								
Aktekin et al. 2010	7 (32)	10 (42)	19	18	63	44	61	53
Alm-Paulsen et al. 2012	6 (20)	3 (10)	30 (12)	27 (10)	N/A		N/A	
Arora et al. 2009	7 (13)	5 (8)	19,4 (6)	21,1 (7)	57 (11,6)	59,8 (7,0)	44,6 (10,4)	49,6 (9,8)
Barai et al. 2018	1 (3)	7 (8)	N/A		N/A		N/A	
Chan et al. 2014	5 (13)	4 (11)	82 (19)*	77 (17)*	N/A		N/A	
Egol et al. 2010	7 (16)	4 (9)	17,7 (7,3)	12,7 (6,5)	54,8 (18,7)	54,6 (14,9)	47,8 (13,1)	51,8 (11,1)
Gong et al. 2011	1 (4)	0	N/A		N/A		N/A	
Hung et al. 2015	0	1 (3)	N/A		60^	60^	60^	60^
Jordan et al. 2016	N/A		N/A		N/A		N/A	
Larouche et al. 2016	14 (11)		N/A		N/A		N/A	
Leerdam et al. 2019	N/A		N/A		N/A		N/A	
Lutz et al. 2014	50 (39)	27 (21)	N/A		N/A		N/A	
Tan et al. 2012	7 (23)	20 (63)	83 (17)*	78 (24)*	N/A		N/A	
Toon et al. 2017	1 (3)	0	83,3 (14,1)*	81,3 (22,9)*	67,5 (13,7)	72,9 (13,2)	63,1 (10,2)	64,1 (13)
Zengin et al. 2019	N/A		67,7 (11,7)*	57,5 (19,6)*	N/A		N/A	

OP|NON operative/nonoperative; $ number (%); # mean (SD); N/A not available; * percentage of unaffected side; ** difference between fractured side and unaffected side; ^ median

Table S6. Continued

Study	Range wrist pronation (°)#		Range wrist supination (°)#		Radial deviation (°)#		Ulnar deviation (°)#	
RCTs	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**
Abbaszadegan et al. 1990	N/A		N/A		N/A		N/A	
Arora et al. 2011	84 (7)	85 (8)	85(8)	85 (8)	24 (6)	25 (7)	35 (8)	35 (8)
Azzopardi et al. 2005	100 (2)*	97 (6)*	91 (19)*	95 (7)*	89 (15)*	80 (31)*	93 (12)*	76 (26)*
Bartl et al. 2014	2,8 (5,6)**	2,6 (9,4)**	2,5 (5,9)**	3,2 (8,3)**	3,9 (6,3)**	3,0 (5,7)**	4,4 (7,5)**	5,9 (8,0)**
Martinez-Mendez et al. 2018	84 (10)	71 (19)	85 (5)	72 (20)	N/A		N/A	
Mulders et al. 2019	90 (7,4^)	85 (11,1)^	85 (11,1)^	75 (11,1)^	15 (7,4)^	15 (3,7)^	25 (4,4)^	25 (7,4)^
Sharma et al. 2014	34,1 (2,6)	32,0 (2,9)	43,4 (3,5)	41,9 (3,9)	79,1 (3,7)	62,8 (6,6)	79,6 (3,0)	65,9 (5,4)
Sirniö et al. 2019	88 (5)	88 (6)	88 (5)	84 (10)	22 (6)	22 (5)	28 (5)	25 (6)
Observational studies								
Aktekin et al. 2010	62	60	51	53	13	14	18	11
Alm-Paulsen et al. 2012	N/A		N/A		N/A		N/A	
Arora et al. 2009	82,2 (8,9)	81,4 (8,6)	83,0 (9,9)	82,5 (6,8)	20,6 (8,6)	21,2 (8,4)	38,0 (9,4)	36,4 (9,2)
Barai et al. 2018	N/A		N/A		N/A		N/A	
Chan et al. 2014	N/A		N/A		N/A		N/A	
Egol et al. 2010	82,9 (6,8)	84,4 (3,8)	80,6 (8,1)	83,9 (3,0)	18,7 (7,9)	22,9 (13,4)	29,9 (8,8)	30,3 (7,1)
Gong et al. 2011	N/A		N/A		N/A		N/A	
Hung et al. 2015	85^	90^	90^	80^	N/A		N/A	
Jordan et al. 2016	N/A		N/A		N/A		N/A	
Larouche et al. 2016	N/A		N/A		N/A		N/A	
Leerdam et al. 2019	N/A		N/A		N/A		N/A	
Lutz et al. 2014	N/A		N/A		N/A		N/A	
Tan et al. 2012	N/A		N/A		N/A		N/A	
Toon et al. 2017	N/A		N/A		15,6 (7,3)	15,7 (5,2)	22,8 (8,0)	17,9 (6,0)
Zengin et al. 2019	N/A		N/A		N/A		N/A	

OP|NON operative/nonoperative; $ number (%); # mean (SD); N/A not available; * percentage of unaffected side; ** difference between fractured side and unaffected side; ^ median

Table S7. Radiologic outcome measures of included studies in a meta-analysis of distal radius fractures

Study	Volar tilt (°)#		Radial inclination (°)#		Radial height (mm)#		Articular step-off (mm)#		Ulnar variance (mm)#	
RCTs	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**
Abbaszadegan et al. 1990	N/A		N/A		N/A		N/A		N/A	
Arora et al. 2011	3,0 (7,2)	−10,4 (19,1)	21,2 (2,6)	15,9 (9,0)	N/A		0,2 (0,5)	0,6 (1,1)	0,7 (1,8)	3,2 (2,9)
Azzopardi et al. 2005	N/A		22,5 (5)	19 (6)	8 (3)	5 (4)	N/A		3 (2)	3 (2)
Bartl et al. 2014	N/A		20,3 (4,5)	17,7 (6,3)	N/A		N/A		N/A	
Martinez-Mendez et al. 2018	8 (8)	4 (8)	19 (6)	13 (6)	9 (2)	5 (3)	0,4 (1)	1 (3)	4 (3)	1 (2)
Mulders et al. 2019	N/A		N/A		N/A		N/A		N/A	
Sharma et al. 2014	8,3 (1,0)	5,1 (0,5)	17,8 (0,8)	15,2 (0,7)	8,2 (0,6)	6,1 (1,1)	N/A		−0,2 (0,2)	0,2 (0,0)
Sirniö et al. 2019	N/A		24 (4)	21 (8)	N/A		N/A		−0,7 (0,9)	−2,0 (1,7)
Observational studies										
Aktekin et al. 2010	11	8	18	20	10	8	N/A		0,27	0,5
Alm-Paulsen et al. 2012	N/A		21 (6)	19 (7)	7 (4)	6 (5)	N/A		3 (2)	3 (3)
Arora et al. 2009	N/A		23,6 (3,8)	19,2 (6,9)	N/A		N/A		1,5 (1,9)	3,8 (2,6)
Barai et al. 2018	N/A		N/A		N/A		N/A		N/A	
Chan et al. 2014	N/A		N/A		N/A		N/A		N/A	
Egol et al. 2010	6,2 (9,2)	−5,8 (10,4)	22,3 (4,7)	18,0 (4,0)	10,6 (2,5)	8,7 (1,6)	N/A		1,5 (2,2)	2,8 (1,8)
Gong et al. 2011	8,1 (1,5)	5,3 (3,4)	19,2 (2,2)	14,3 (5,6)	9,5 (1,3)	6,3 (3,2)	N/A		N/A	
Hung et al. 2015	9,6 (7,3)	−5,2 (15)	21,0 (4,6)	17,9 (5,0)	9,8 (2,4)	8,6 (2,6)	N/A		N/A	
Jordan et al. 2016	N/A		N/A		10,9 (2,6)	8,2 (2,2)	0,9 (0,5)	0,9 (0,6)	N/A	
Larouche et al. 2016	8,3 (7,9)	1,3 (12,3)	24,2 (5,2)	21,2 (6,0)	11,1 (3,4)	9,7 (3,4)	0,1 (0,4)	0,1 (0,4)	−0,8 (2,4)	−1,6 (2,8)
Leerdam et al. 2019	N/A		N/A		N/A		N/A		N/A	
Lutz et al. 2014	2 (9)	7 (13)	20 (5)	19 (6)	N/A		0,3 (0,7)	0,2 (0,6)	1,3 (2,1)	2,6 (2,2)
Tan et al. 2012	N/A		N/A		N/A		N/A		N/A	
Toon et al. 2017	5,6 (8,9)	0,1 (11,6)	21,6 (6,1)	16,9 (6,3)	9,6 (3,3)	7,2 (3,4)	0,7 (0,6)	1,5 (0,9)	N/A	
Zengin et al. 2019	N/A		21,5 (2,6)	16,6 (5,3)	10,4 (2,8)	7,8 (2,4)	0,6 (0,5)	1,4 (0,9)	1,7 (1,7)	2,1 (2,1)

OP|NON operative/nonoperative; n number; # mean (SD); N/A not available

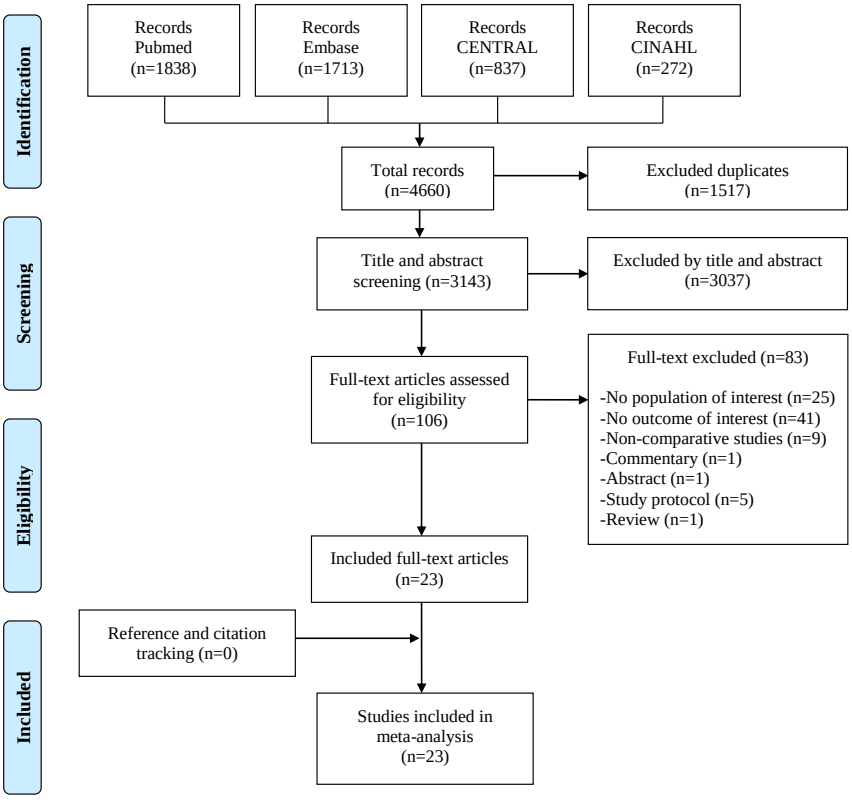

Figure S1. PRISMA flow diagram representing the search and selection of studies comparing operative versus nonoperative treatment of distal radius fractures.

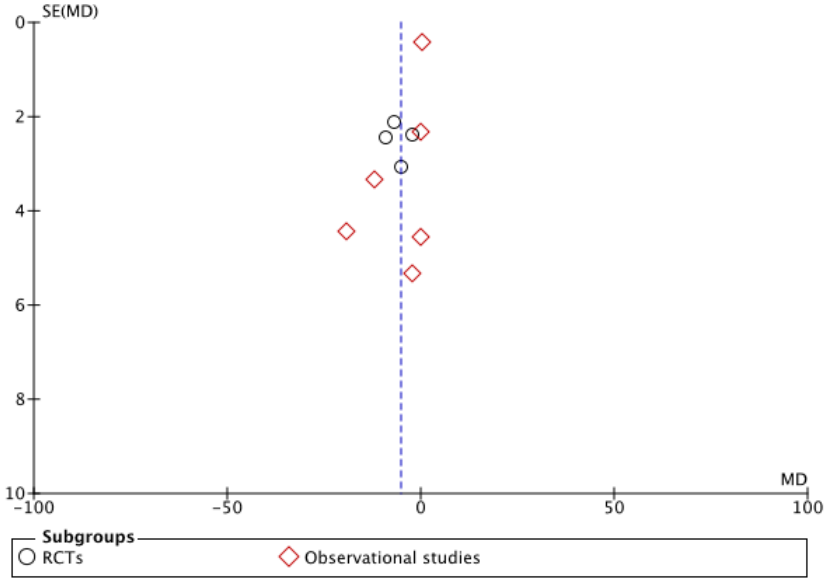

Figure S2. Funnel plot of medium-term (≤ 1 year) DASH score in a meta-analysis of distal radius fractures (MD mean difference; SE standard error).

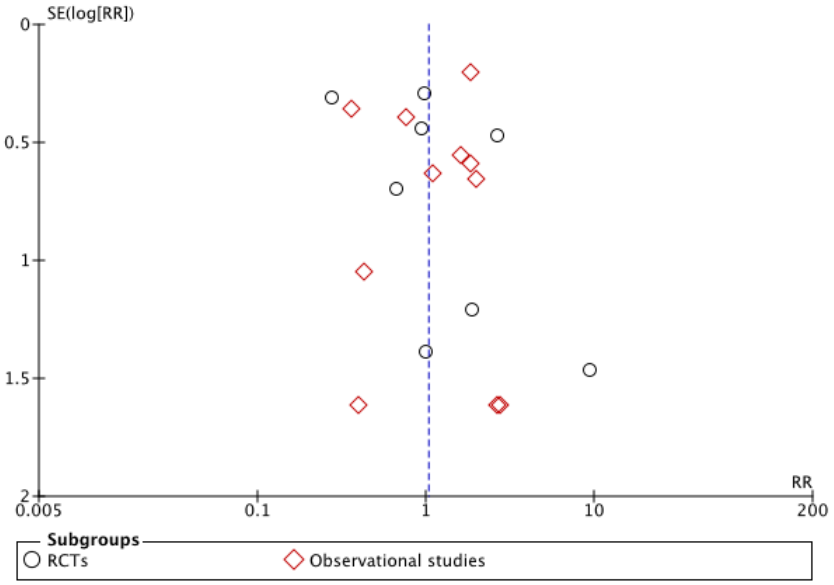

Figure S3. Funnel plot of complication rate in a meta-analysis of distal radius fractures (RR risk ratio; SE standard error).

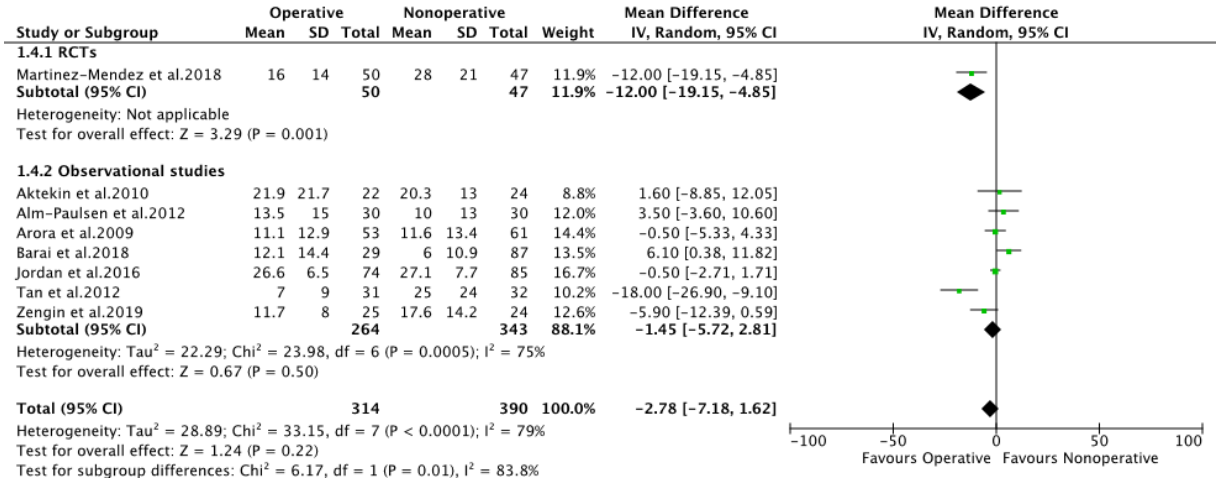

Figure S4. Forest plot of long-term (> 1 year) DASH score in a meta-analysis of distal radius fractures.

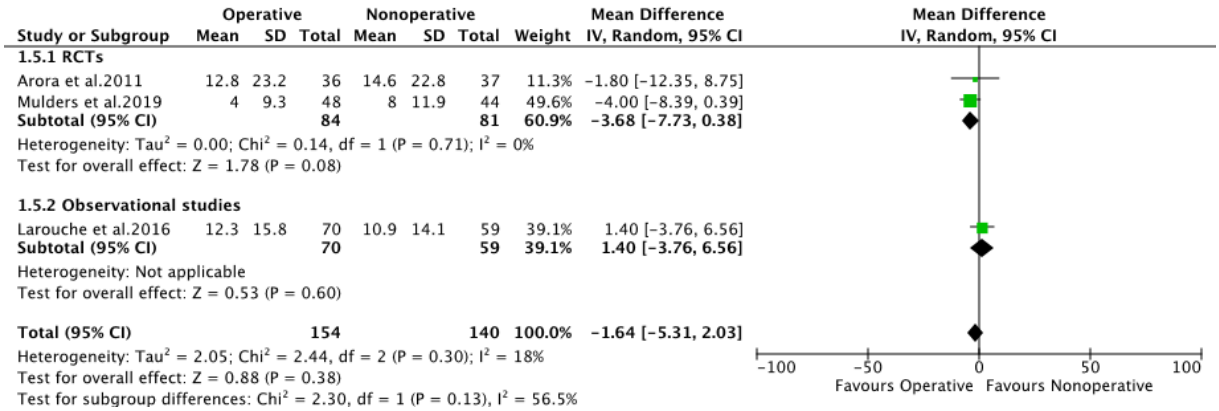

Figure S5. Forest plot of medium-term (≤ 1 year) PRWE score in a meta-analysis of distal radius fractures.

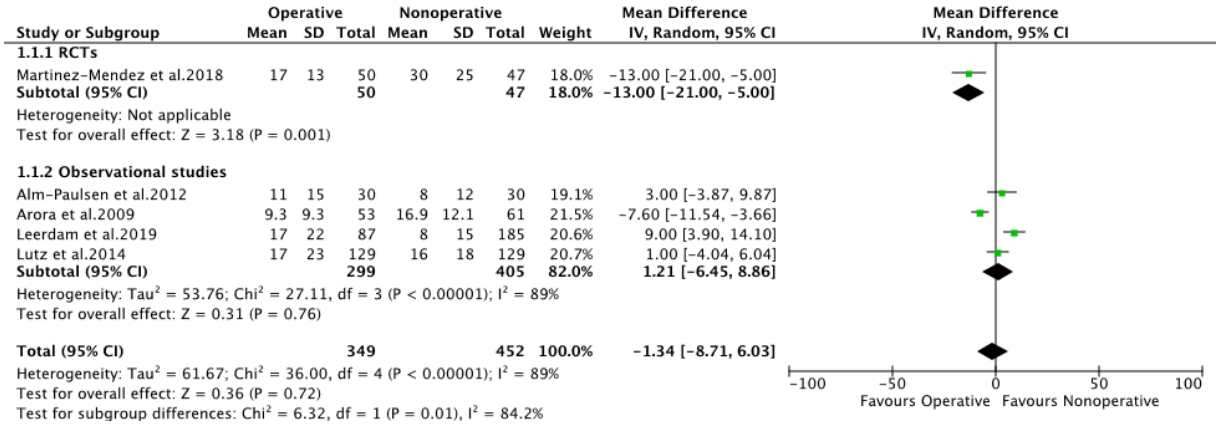

Figure S6. Forest plot of long-term (> 1 year) PRWE score in a meta-analysis of distal radius fractures.

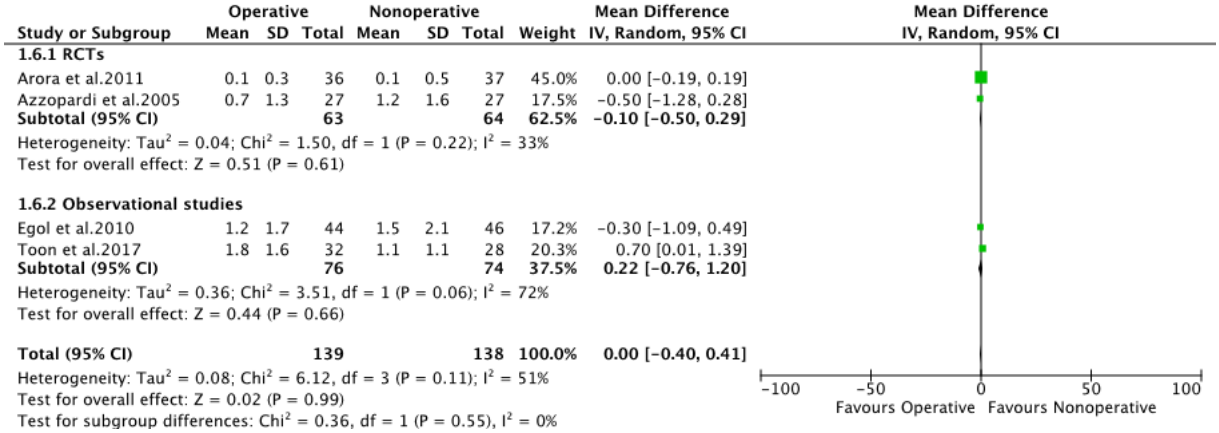

Figure S7. Forest plot of medium-term (≤ 1 year) VAS score in a meta-analysis of distal radius fractures.

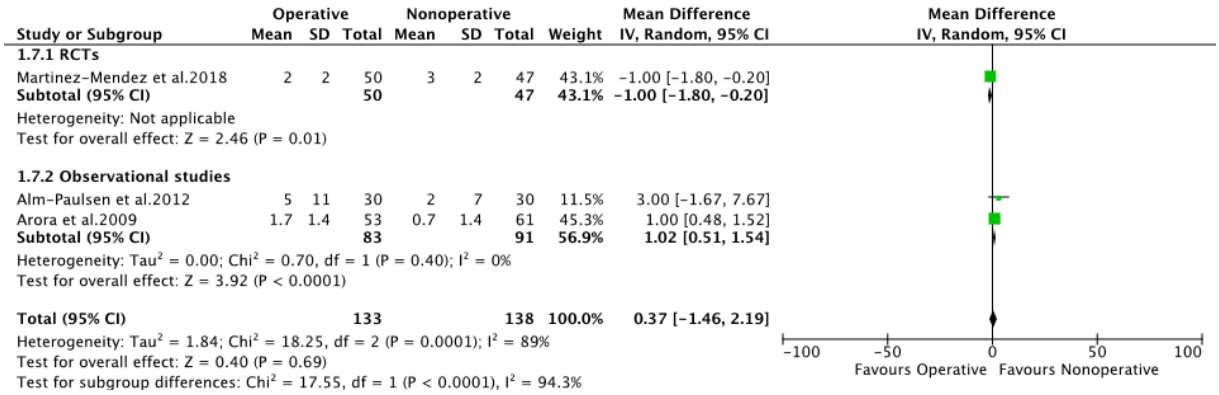

Figure S8. Forest plot of long-term (> 1 year) VAS score in a meta-analysis of distal radius fractures.

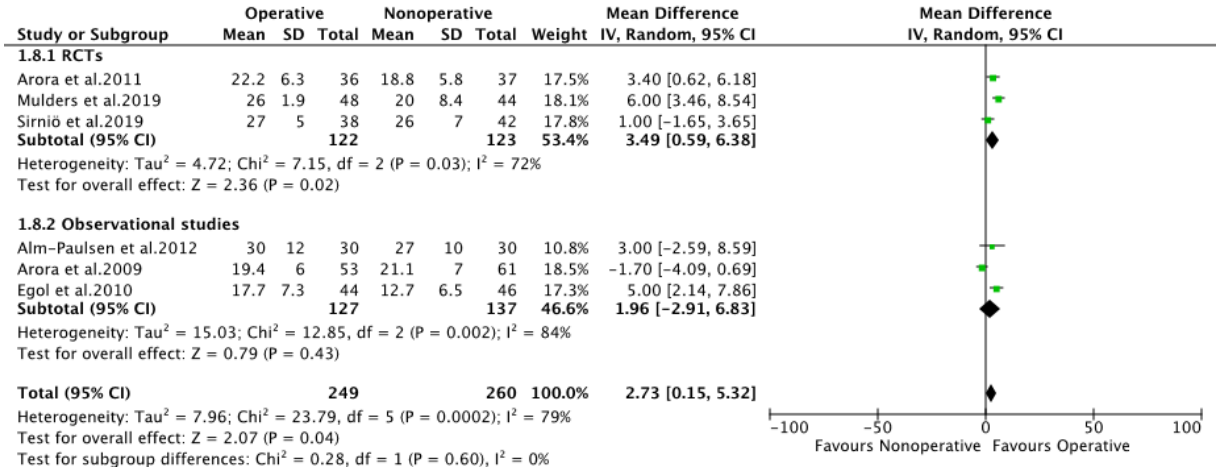

Figure S9. Forest plot of grip strength in kg in a meta-analysis of distal radius fractures.

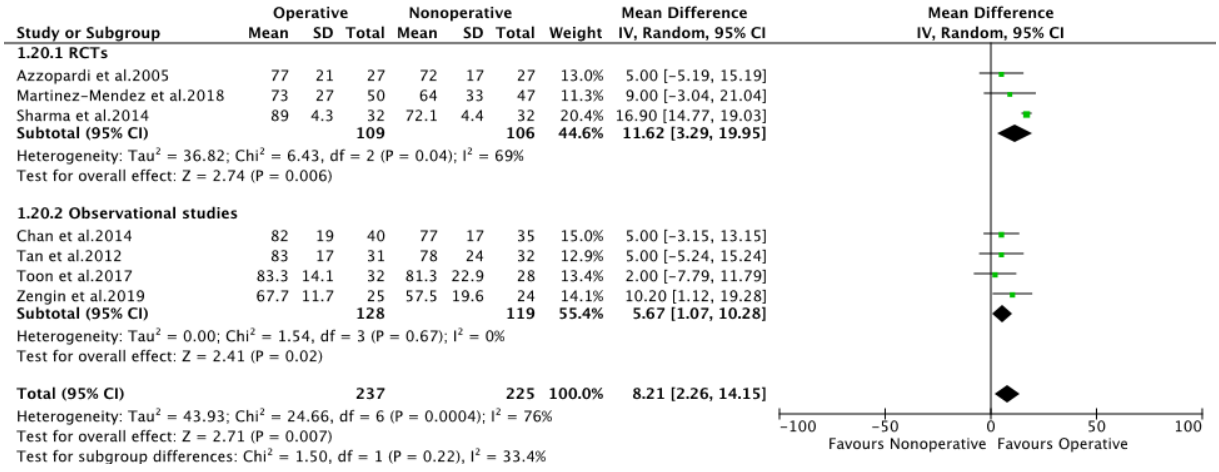

Figure S10. Forest plot of grip strength as percentage of unaffected side in a meta-analysis of distal radius fractures.

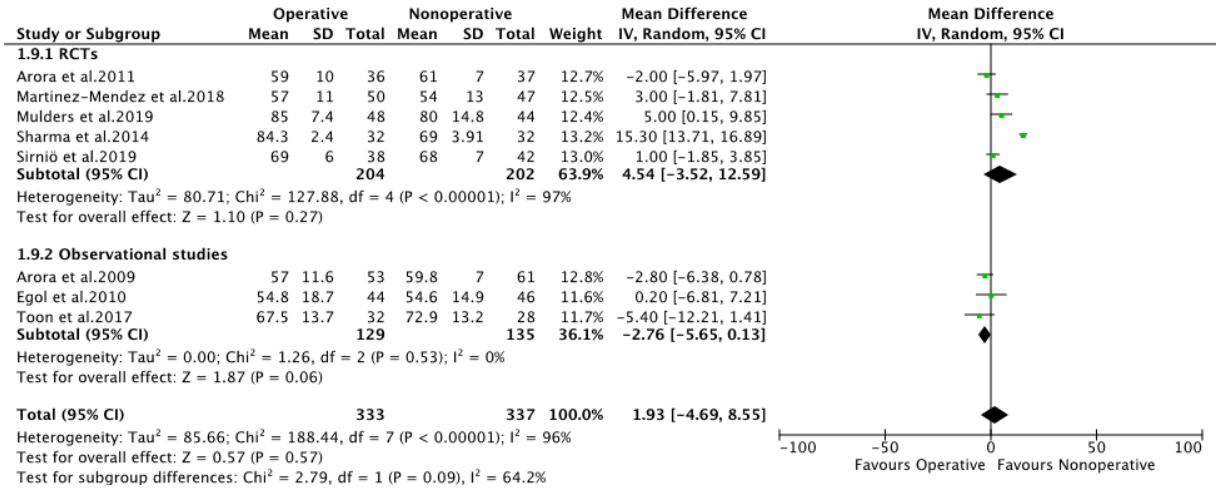

Figure S11. Forest plot of range of wrist extension (°) in a meta-analysis of distal radius fractures.

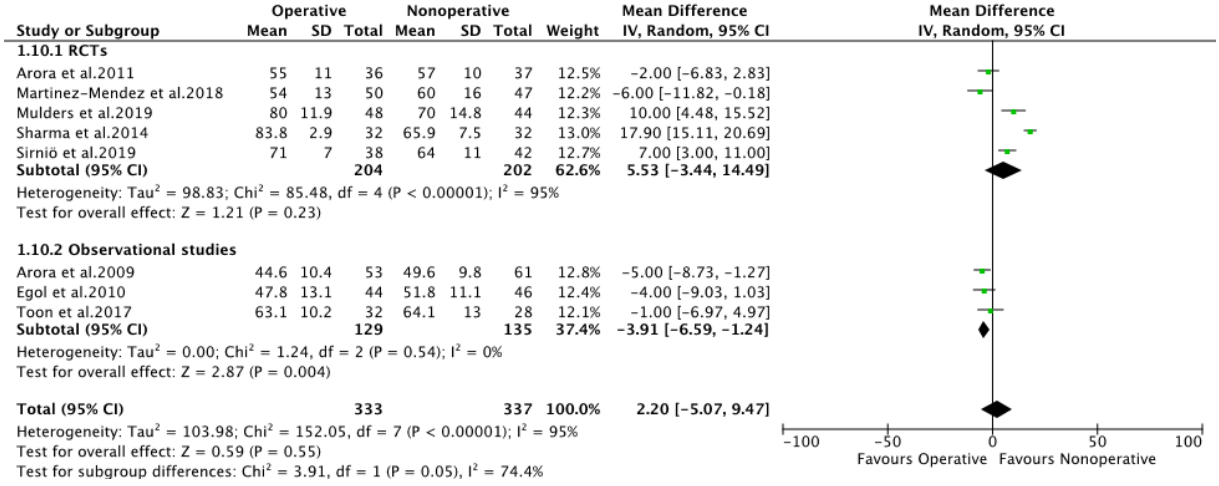

Figure S12. Forest plot of range of wrist flexion (°) in a meta-analysis of distal radius fractures.

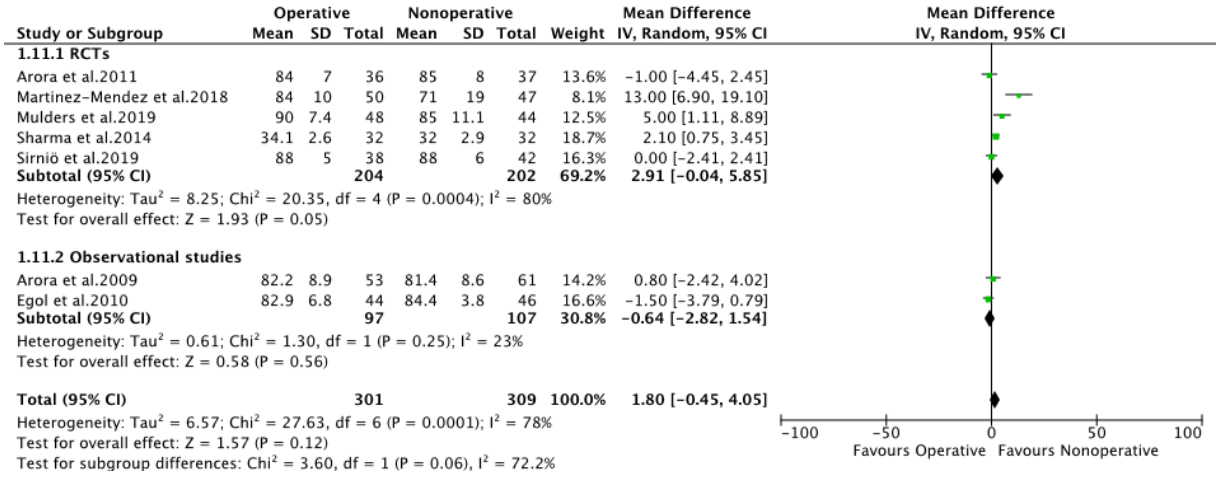

Figure S13. Forest plot of range of wrist pronation (°) in a meta-analysis of distal radius fractures.

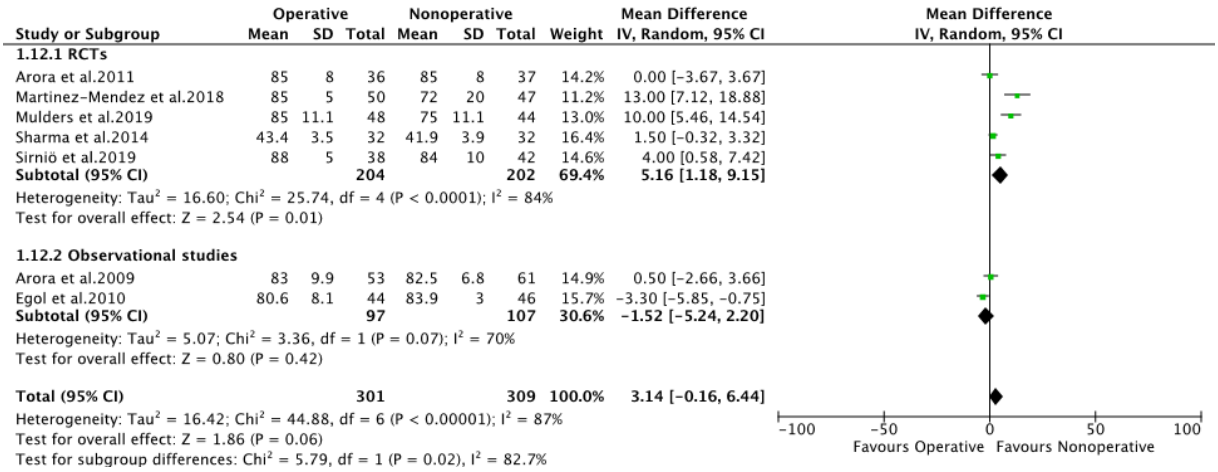

Figure S14. Forest plot of range of wrist supination (°) in a meta-analysis of distal radius fractures.

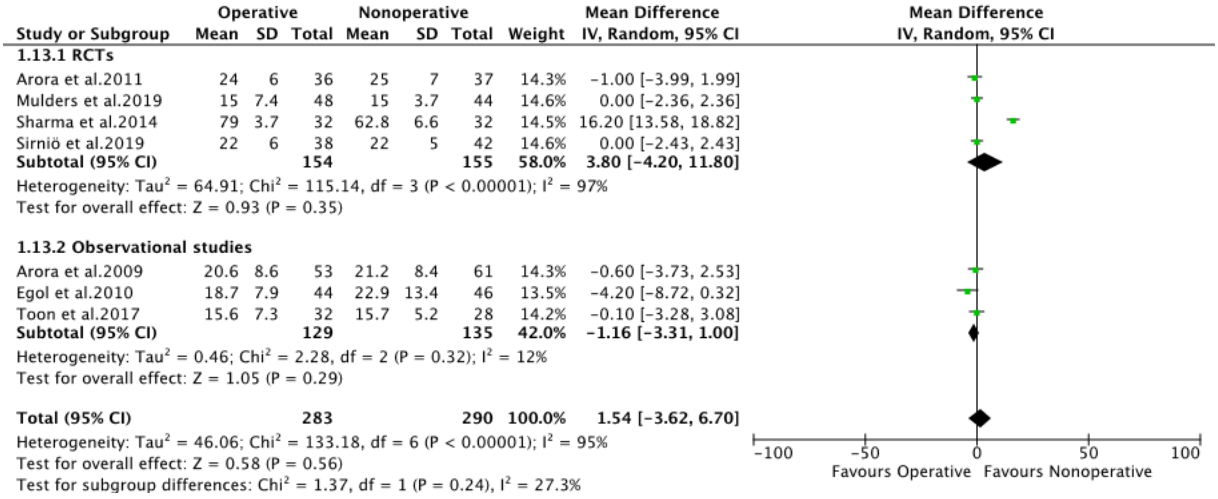

Figure S15. Forest plot of radial deviation (°) in a meta-analysis of distal radius fractures.

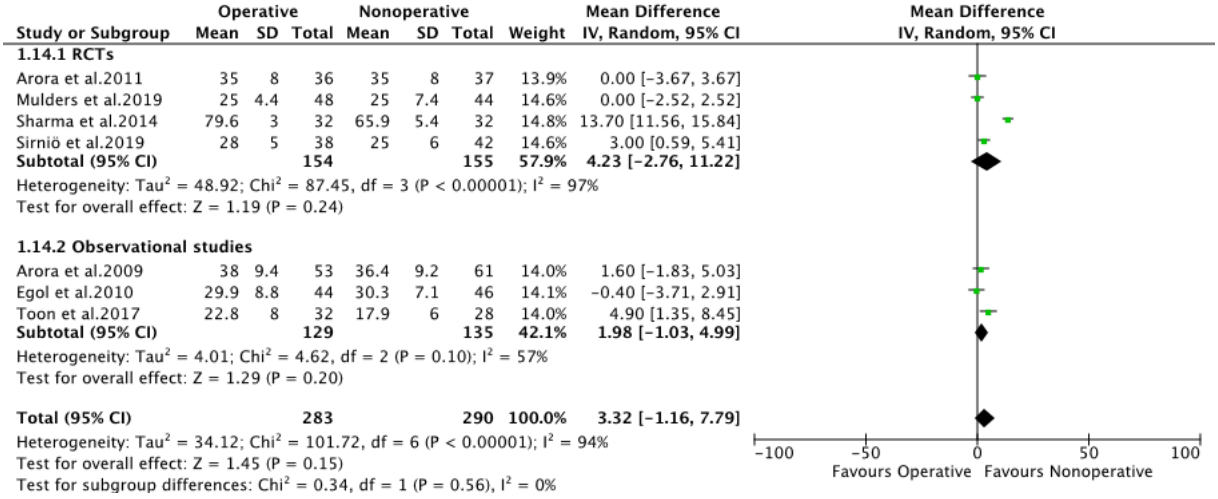

Figure S16. Forest plot of ulnar deviation (°) in a meta-analysis of distal radius fractures.

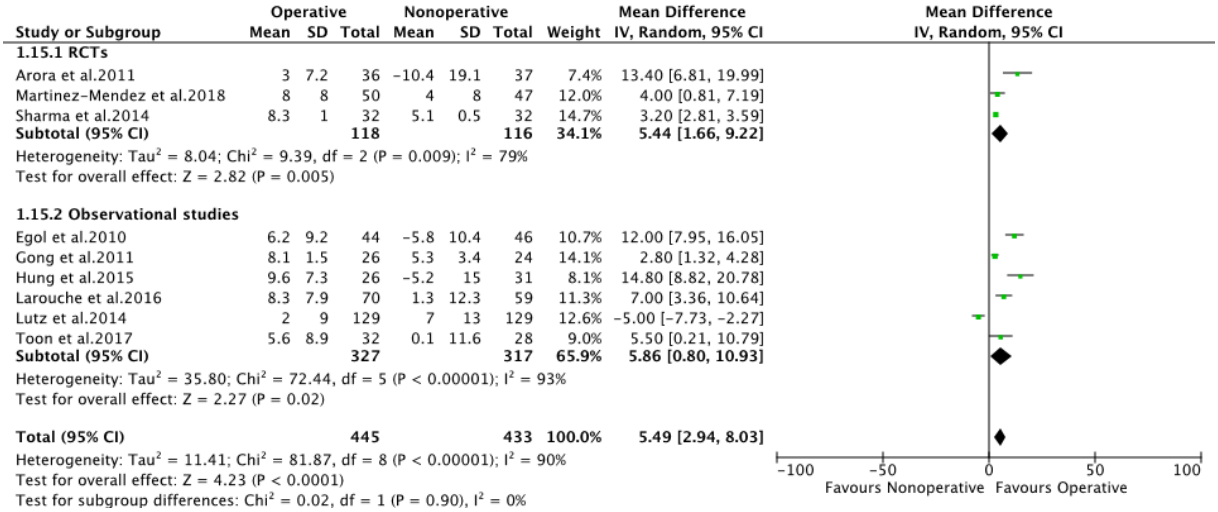

Figure S17. Forest plot of volar tilt (°) in a meta-analysis of distal radius fractures.

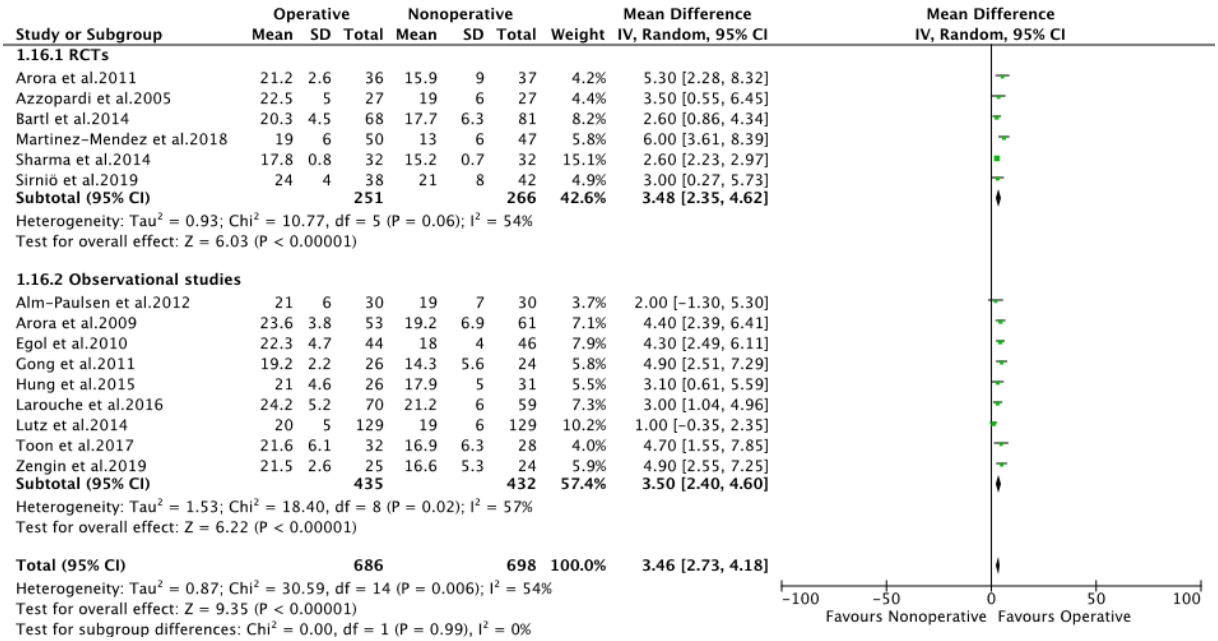

Figure S18. Forest plot of radial inclination (°) in a meta-analysis of distal radius fractures.

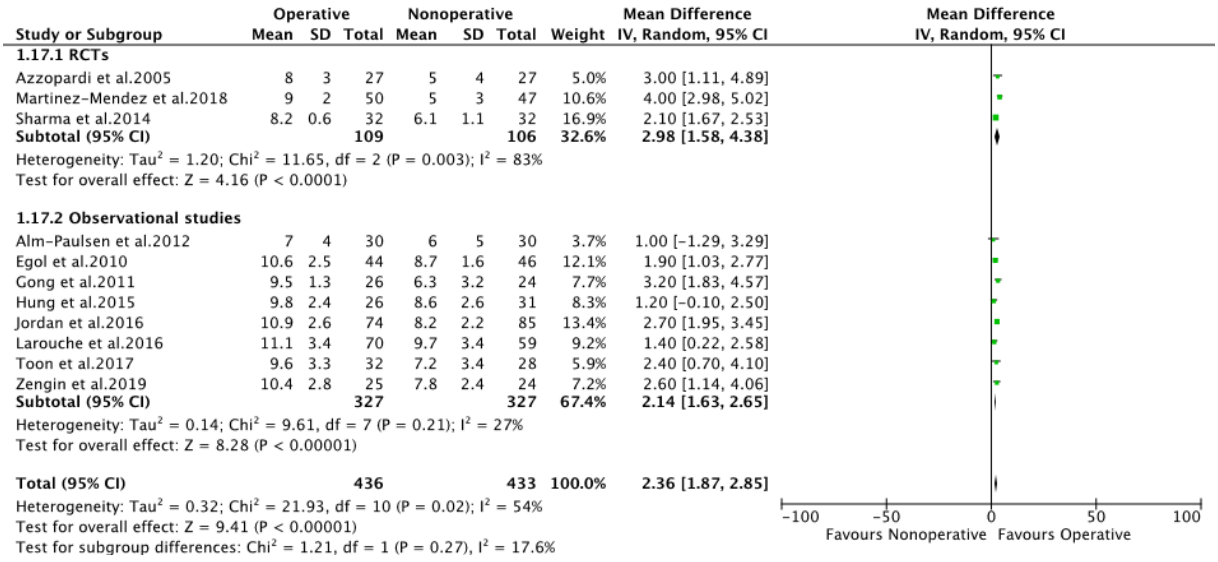

Figure S19. Forest plot of radial height (mm) in a meta-analysis of distal radius fractures.

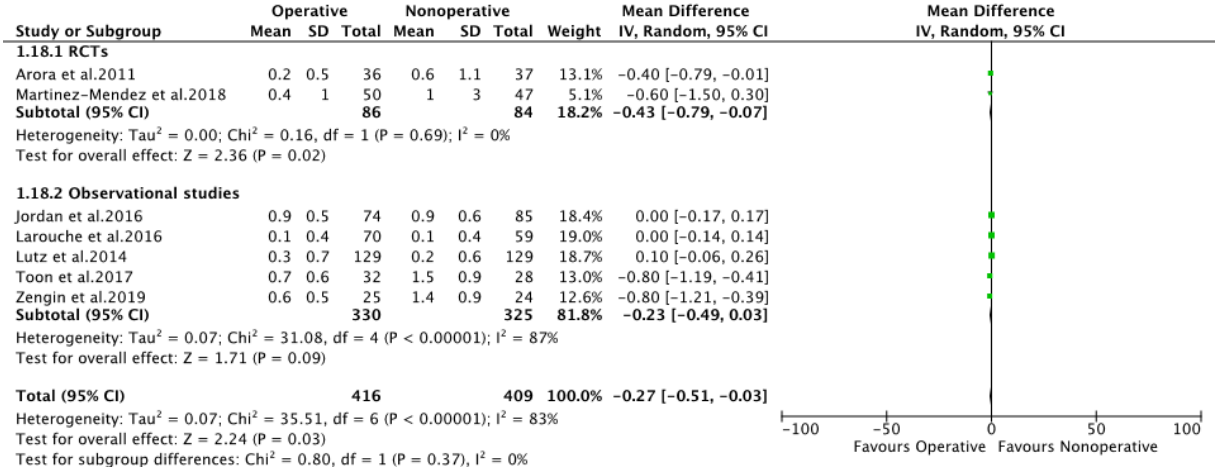

Figure S20. Forest plot of articular step-off (mm) in a meta-analysis of distal radius fractures.

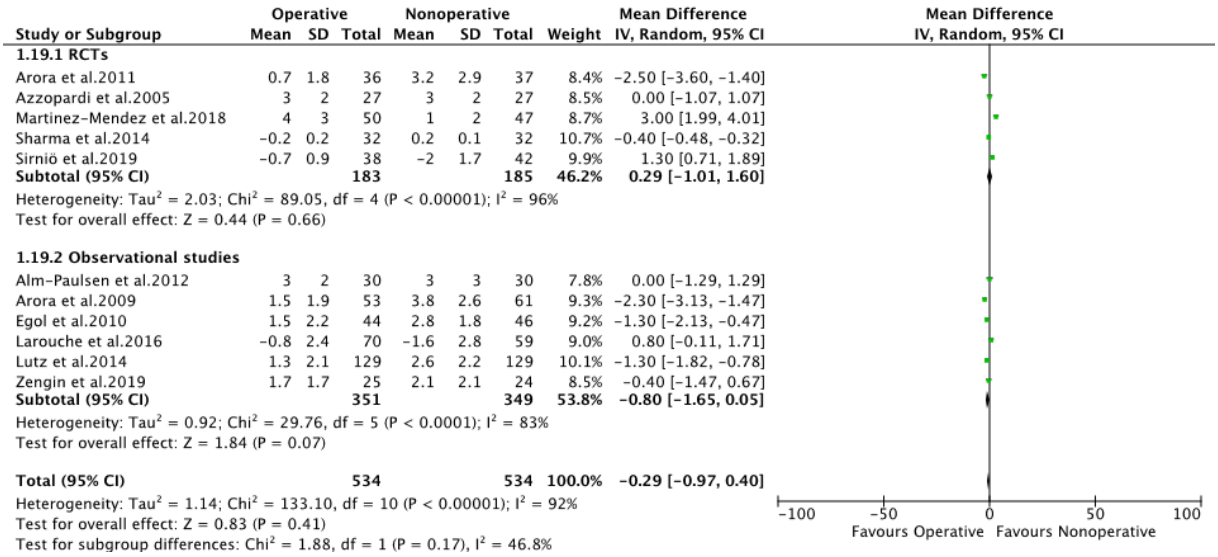

Figure S21. Forest plot of ulnar variance (mm) in a meta-analysis of distal radius fractures.

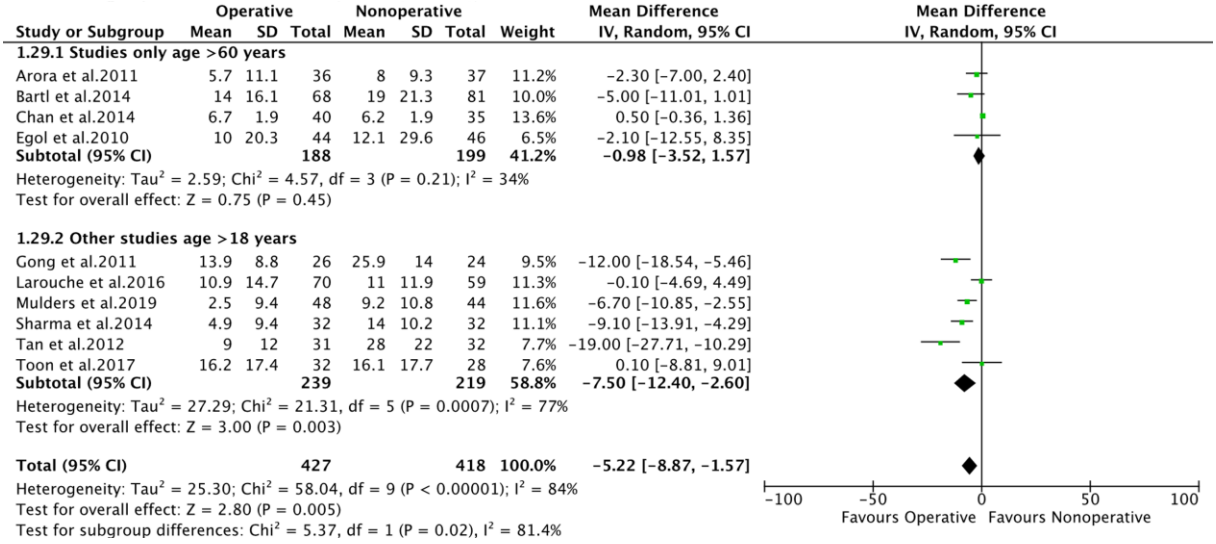

Figure S22. Forest plot of medium-term (≤ 1 year) DASH score for studies that only included patients with age >60 years and other studies that included patients with age >18 years in a meta-analysis of distal radius fractures.

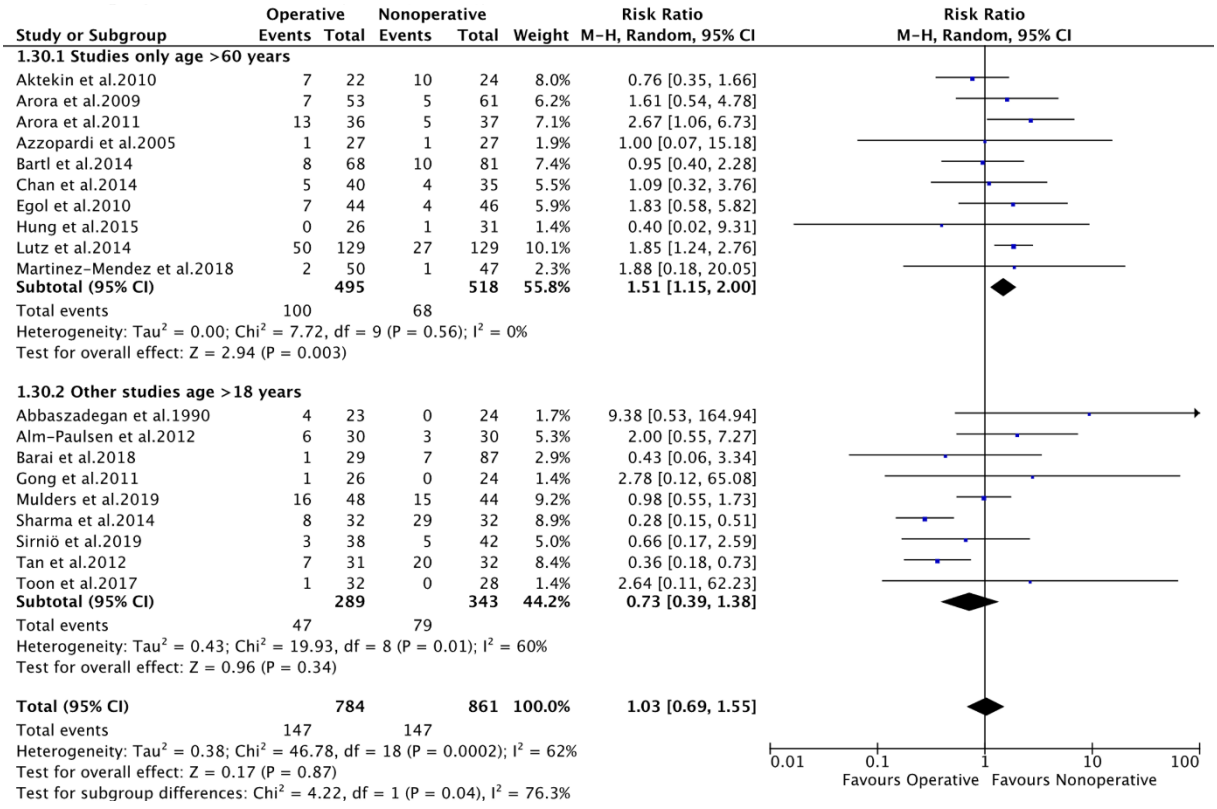

Figure S23. Forest plot of complication rate for studies that only included patients with age >60 years and other studies that included patients with age >18 years in a meta-analysis of distal radius fractures.

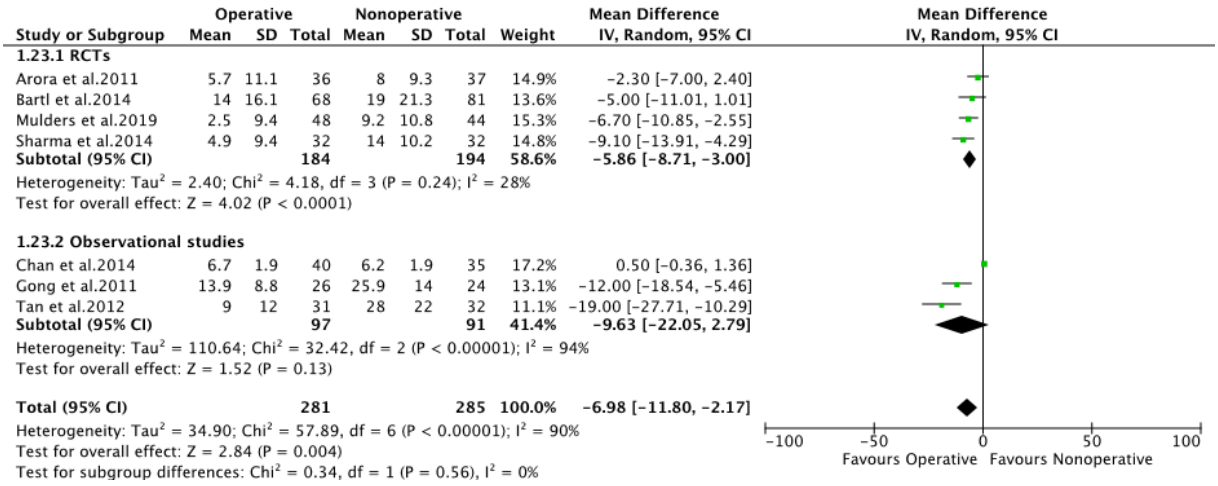

Figure S24. Forest plot of medium-term (≤ 1 year) DASH score in high-quality studies in a meta-analysis of distal radius fractures.

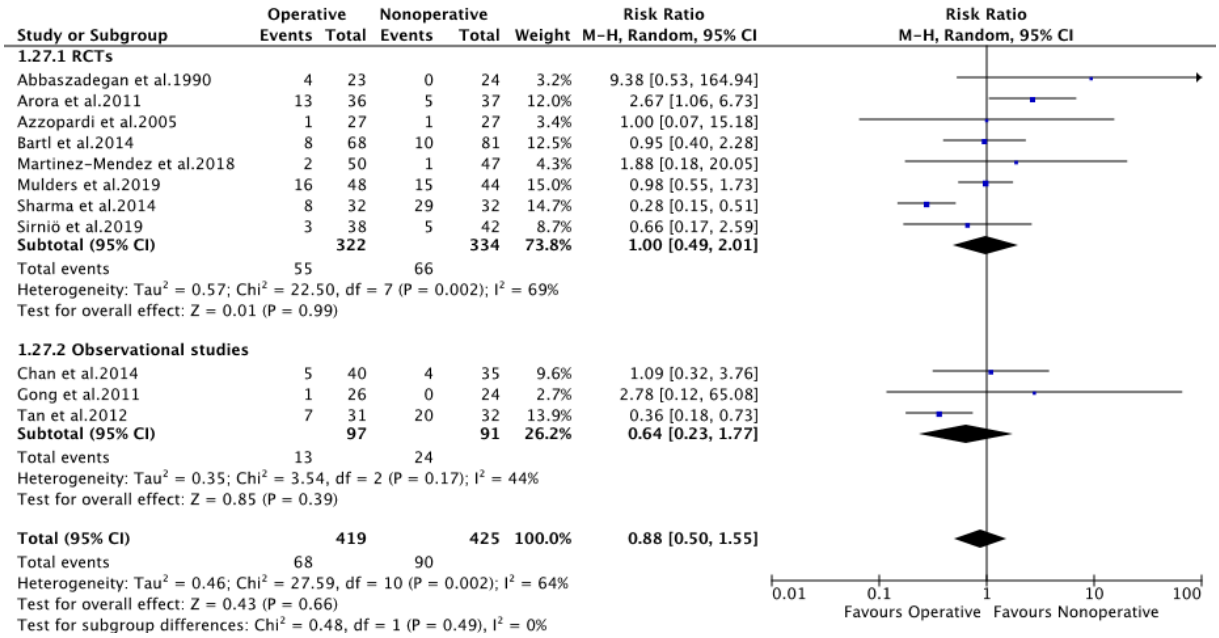

Figure S25. Forest plot of complication rate in high-quality studies in a meta-analysis of distal radius fractures.

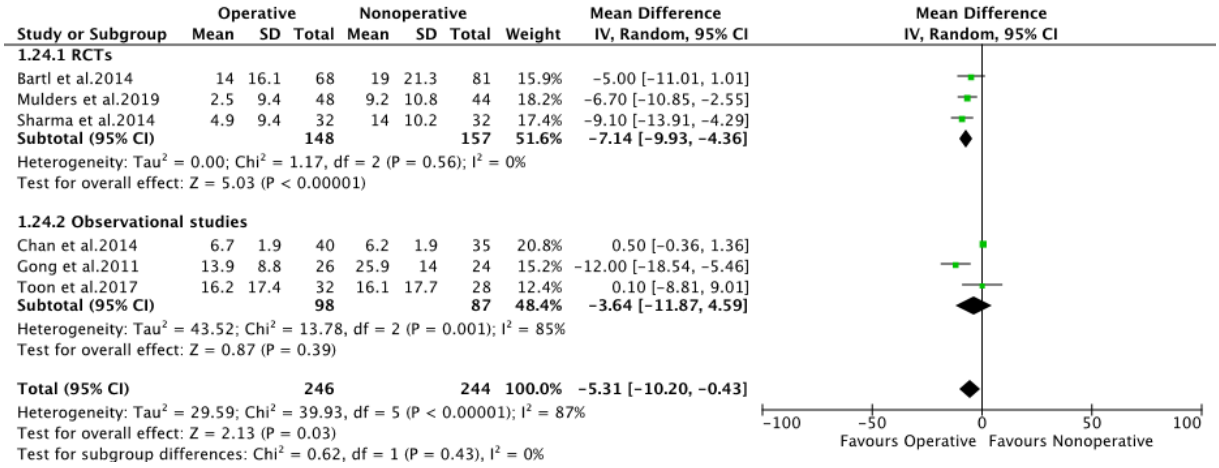

Figure S26. Forest plot of medium-term (≤ 1 year) DASH score in studies with a study period after the year 2008 in a meta-analysis of distal radius fractures.

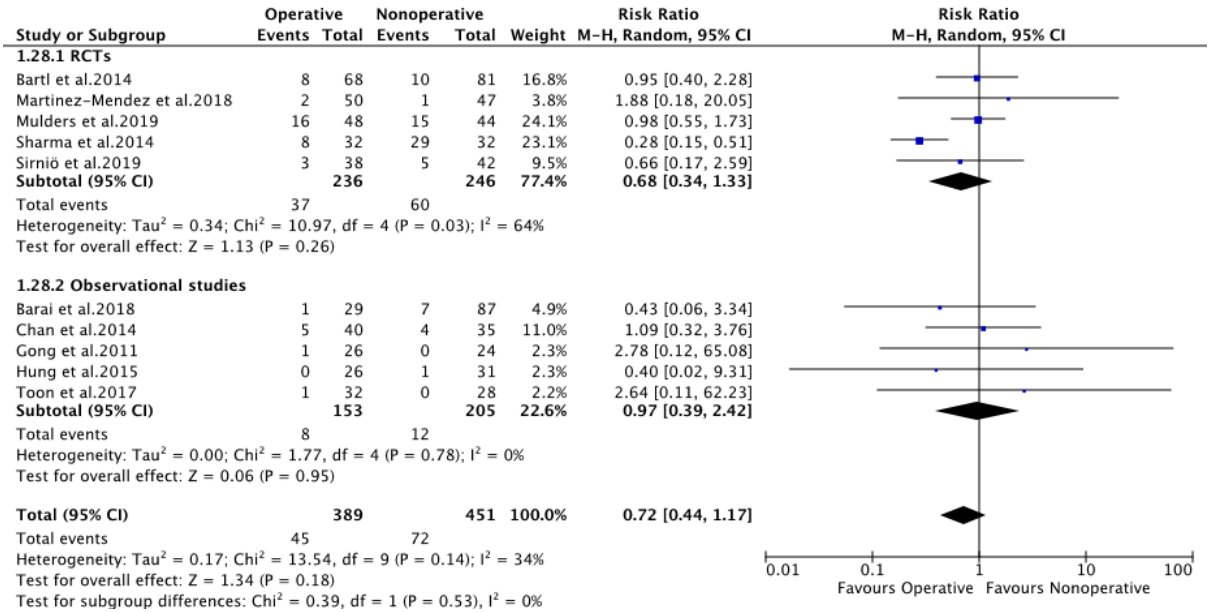

Figure S27. Forest plot of complication rate in studies with a study period after the year 2008 in a meta-analysis of distal radius fractures.

CHAPTER 6

Operative treatment versus nonoperative treatment of Achilles tendon ruptures: systematic review and meta-analysis

Yassine Ochen, Reinier B. Beks, Mark van Heijl, Falco Hietbrink,
Luke P.H. Leenen, Detlef van der Velde, Marilyn Heng,
Olivier van der Meijden, Rolf H.H. Groenwold, R. Marijn Houwert

(The BMJ)

Abstract

Background

The management of acute Achilles tendon ruptures—operative or nonoperative treatment—is much debated. The aim of this study was to compare re-rupture rate, complication rate, and functional outcome after operative versus nonoperative treatment of Achilles tendon ruptures; to compare re-rupture rate after early and late full weight bearing; to evaluate re-rupture rate after functional rehabilitation with early range of motion; and to compare effect estimates from randomized controlled trials and observational studies.

Methods

The PubMed/Medline, Embase, CENTRAL, and CINAHL databases were last searched on 25 April 2018 for studies comparing operative versus nonoperative treatment of Achilles tendon ruptures. Randomized controlled trials and observational studies reporting on comparison of operative versus nonoperative treatment of acute Achilles tendon ruptures were included. Data extraction was performed independently in pairs, by four reviewers, with the use of a predefined data extraction file. Outcomes were pooled using random effects models and presented as risk difference, risk ratio, or mean difference, with 95% confidence interval.

Results

29 studies were included—10 randomized controlled trials and 19 observational studies. The 10 trials included 944 (6%) patients, and the 19 observational studies included 14 918 (94%) patients. A significant reduction in re-ruptures was seen after operative treatment (2.3%) compared with nonoperative treatment (3.9%) (risk difference 1.6%; risk ratio 0.43, 95% confidence interval 0.31 to 0.60; $P<0.001$; $I^2=22\%$). Operative treatment resulted in a significantly higher complication rate than nonoperative treatment (4.9% *v* 1.6%; risk difference 3.3%; risk ratio 2.76, 1.84 to 4.13; $P<0.001$; $I^2=45\%$). The main difference in complication rate was attributable to the incidence of infection (2.8%) in the operative group. A similar reduction in re-rupture rate in favor of operative treatment was seen after both early and late full weight bearing. No significant difference in re-rupture rate was seen between operative and nonoperative treatment in studies that used accelerated functional rehabilitation with early range of motion (risk ratio 0.60, 0.26 to 1.37; $P=0.23$; $I^2=0\%$). No difference in effect estimates was seen between randomized controlled trials and observational studies.

Conclusion

This meta-analysis shows that operative treatment of Achilles tendon ruptures reduces the risk of re-rupture compared with nonoperative treatment. However, re-rupture rates are low and differences between treatment groups are small (risk difference 1.6%). Operative treatment results in a higher risk of other complications (risk difference 3.3%). The final decision on the management of acute Achilles tendon ruptures should be based on patient specific factors and shared decision making. This review emphasizes the potential benefits of adding high quality observational studies in meta-analyses for the evaluation of objective outcome measures after surgical treatment.

Introduction

Rupture of the Achilles tendon is a frequently encountered injury, with an incidence of 31 per 100 000 per year, and is most common in the young to middle aged active population, with a reported mean age ranging from 37 to 44 years.[1,2] Recent studies indicate that the incidence of Achilles tendon rupture is still increasing owing to a more active older population.[2] Injury of the Achilles tendon can be debilitating because of its role in ambulation and activity, affecting both athletes and non-athletes. The management of acute Achilles tendon ruptures—operative or nonoperative treatment—is much debated.[2]

Several meta-analyses of randomized controlled trials (RCTs) have shown that operative treatment significantly reduces the risk of tendon re-rupture compared with nonoperative treatment, with a reported risk difference in re-rupture rate varying from 5% to 7%.[3-6] However, operative treatment leads to a significant increase in other complications such as infection, deep vein thrombosis, and sural nerve injury, with a reported risk difference varying from 16% to 21%.[3,4,6] The incidence of operative treatment has declined over the past decade as a result of multiple RCTs showing comparable results between operative and nonoperative treatment.[1,2]

A recent systematic review of overlapping meta-analyses evaluated nine meta-analyses that compared operative and nonoperative treatment of Achilles tendon ruptures. The discordance found among the nine meta-analyses indicated that further investigation is warranted as rehabilitation protocols, weight bearing restrictions, and treatment modalities have evolved.[7]

Systematic reviews and meta-analyses of RCTs are considered the highest level of evidence for the evaluation of treatment effects. However, several reports have shown that little evidence exists for significant differences in effect estimates between RCTs and observational studies.[8-11] The addition of observational studies in meta-analyses increases sample size, which could enable the evaluation of small treatment effects and infrequent outcome measures. Furthermore, observational studies might provide insight into a variety of populations and long-term effects compared with the usually highly selected patient populations in RCTs.[12,13] Both RCTs and observational studies are increasingly used in orthopedic trauma meta-analyses for the evaluation of treatment effects.[14-17]

The primary aim of this systematic review and meta-analysis was to compare re-rupture rate, complication rate, and functional outcome after operative versus nonoperative treatment of acute Achilles tendon ruptures. Secondly, we sought to evaluate re-rupture rate after early and late full weight bearing and compare re-rupture rate after functional rehabilitation with early range of motion. Finally, we compared effect estimates obtained from RCTs and observational studies.

Methods

This systematic review and meta-analysis was performed and reported according to the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklists.[18-20] A published protocol for this review does not exist.

Search strategy and selection criteria

We last searched the PubMed/Medline, Embase, CENTRAL, and CINAHL databases on 25 April 2018 for studies comparing operative versus nonoperative treatment of Achilles tendon ruptures. The search syntax is provided in supplementary Table S1. Duplicate articles were removed. Two reviewers (Y.O., R.H.H.G.) independently screened titles and abstracts for eligibility of identified studies. All published comparative studies, both RCTs and observational studies, reporting on the comparison of operative versus nonoperative treatment of acute Achilles tendon ruptures were eligible for inclusion.

After title and abstract screening, the same two reviewers (Y.O., R.H.H.G.) independently reviewed full text articles. Inclusion criteria were acute Achilles tendon rupture, operative treatment (open or minimally invasive surgery) versus nonoperative treatment (cast immobilization or functional bracing), treatment within four weeks of rupture, age 16 years or older, and reporting of re-rupture rate, complication rate, or functional outcome. Exclusion criteria were delayed presentation (treatment more than four weeks after rupture), treatment for re-rupture, language other than English, no availability of full text article, and letters, meeting proceedings, and case reports. We had no inclusion restrictions based on weight bearing status or functional rehabilitation protocol. Disagreements on eligibility of full text articles were resolved by consensus or by discussion with a third reviewer (R.M.H.). References of included studies were screened, and backwards citation tracking was performed using Web of Science to identify articles not found in the original literature search.

Data extraction

Four reviewers (Y.O., R.H.H.G., R.M.H., R.B.B.) extracted data independently in pairs, using a predefined data extraction file. The following baseline characteristics were extracted from the included studies: first author, year of publication, study design, country in which the study was performed, study period, number of included patients, operative method, nonoperative method, full weight bearing status, functional rehabilitation protocol, and mean follow-up. Studies reporting on patient cohorts described in previously published articles were excluded or merged.

Quality assessment

The same four reviewers (Y.O., R.H.H.G., R.M.H., R.B.B.), in pairs, independently assessed the methodological quality of included studies by using the Methodological Index for Non-Randomized Studies (MINORS).[21] The MINORS is a validated instrument for the assessment of methodological quality and clear reporting of non-randomized surgical studies, resulting in a score ranging from 0 to 24 for comparative studies.[21] In this study the assessment of methodological quality resulted in a score ranging from 0 to 24 for RCTs and prospective cohort studies. The methodological quality of retrospective cohort studies resulted in a score ranging from 0 to 18. The MINORS criteria for prospective collection of data, loss to follow-up, and prospective calculation of study size were not applicable to the retrospective cohort studies. Details on the methodological quality assessment are provided in supplementary Table S2. Disagreements were resolved by consensus.

Primary and secondary outcomes

The primary outcome measure was re-rupture rate after operative or nonoperative treatment. Secondary outcome measures included complication rate, functional outcome scores, return to sporting activity, and return to work after operative or nonoperative treatment. We defined complication rate as the rate of complications other than re-rupture. Complications included reports of wound infection, sural nerve injury, deep vein thrombosis, and pulmonary embolism. Functional outcome scores included the Achilles Tendon Rupture Score (ATRS).[22] We subdivided functional outcome scores according to follow-up, into short term (one year of less) and long-term (more than one year). We defined return to sporting activity as the duration in months before resumption of sports and return to work as the duration in weeks before resuming work. In studies that reported on both open and minimally invasive surgery, we used the combined outcome measures.

Statistical analysis

We present all continuous variables as mean value with standard deviation or range. We converted continuous variables to mean and standard deviation if sufficient information was available, using the methods described in the Cochrane Handbook for Systematic Reviews of Interventions.[23] We extracted dichotomous variables as absolute number and percentage, pooled them using the Mantel-Haenszel method, and presented them as risk difference and risk ratio with 95% confidence interval. We pooled continuous outcomes by using the inverse variance weighting method and presented them as mean difference with 95% confidence interval. We used random effects models for all analyses. We assessed statistical heterogeneity between studies by visual inspection of forest plots and by the I^2 and χ^2 statistics for heterogeneity. We used the overall effect Z test to determine the significance level for treatment effects. All analyses were stratified according to study design—RCTs or observational studies. We assessed difference in effect estimates between the two subgroups as described in the Cochrane Handbook for Systematic Reviews of Interventions.[23] The significance level for difference in effect estimates across the subgroups was determined by the test for subgroup differences. We defined the significance level for treatment effects and differences across the subgroups as a P value below 0.05. We assessed potential publication bias by visual inspection of funnel plots with risk ratio and standard error.[24] We used Review Manager (RevMan, version 5.3.5) for all statistical analyses.[25] We further assessed publication bias with Begg's and Egger's statistical tests using Stata 13.1.

Primary sensitivity analyses

We did sensitivity analyses for the primary outcome, including studies with an early (four weeks or less) and late (more than four weeks) full weight bearing status after treatment. Studies reporting on both an early and a late full weight bearing cohort were accordingly divided for sensitivity analysis. We did an additional sensitivity analysis for the primary outcome with studies that included an accelerated functional rehabilitation protocol. We defined accelerated functional rehabilitation as the start of early range of motion within three weeks after nonoperative treatment. Rehabilitation with functional bracing systems with successive fixed degrees of plantar flexion, which did not allow for free range of motion, were not considered as accelerated rehabilitation.

Secondary sensitivity analyses

We did secondary sensitivity analyses for high quality studies and year of study period, regarding re-rupture rate and complication rate. We defined high quality studies as RCTs or prospective cohort studies with a MINORS score of 16 or higher (range 0-24) or retrospective cohort studies with a MINORS score of 12 or higher (range 0-18). We did additional sensitivity analyses with studies that included patients after the study period 2000, to account for the development of new rehabilitation protocols, operative techniques, and nonoperative treatment modalities.

Results

Search

Figure 1 shows a flowchart of the literature search and study selection. Full text articles could not be obtained for three studies.[26-28] Four studies reported on patient cohorts described in previously published articles and were excluded or merged with the original studies.[29-32] This resulted in the final inclusion of 29 studies for analyses in this systematic review and meta-analysis—10 RCTs and 19 observational tudies.[33-61]

Baseline study characteristics

The 29 studies included 15 862 patients, of whom 9375 were treated operatively and 6487 nonoperatively. The overall weighted mean age was 41 (range 17-86) years, 41 years in the operative group and 44 years in the nonoperative group. Overall, the studies included 11 779 (74%) males. Overall follow-up ranged from 10 to 95 months. Table 1 shows the baseline characteristics for both RCTs and observational studies. In addition, supplementary Table S3 shows the treatment characteristics of all included studies.

The 10 RCTs included 944 (6%) patients; 469 patients were treated operatively and 475 nonoperatively. The weighted mean age was 40 years in both treatment groups, and 779 (83%) males were included. The operative method was open surgery in nine studies and minimally invasive surgery in one study.

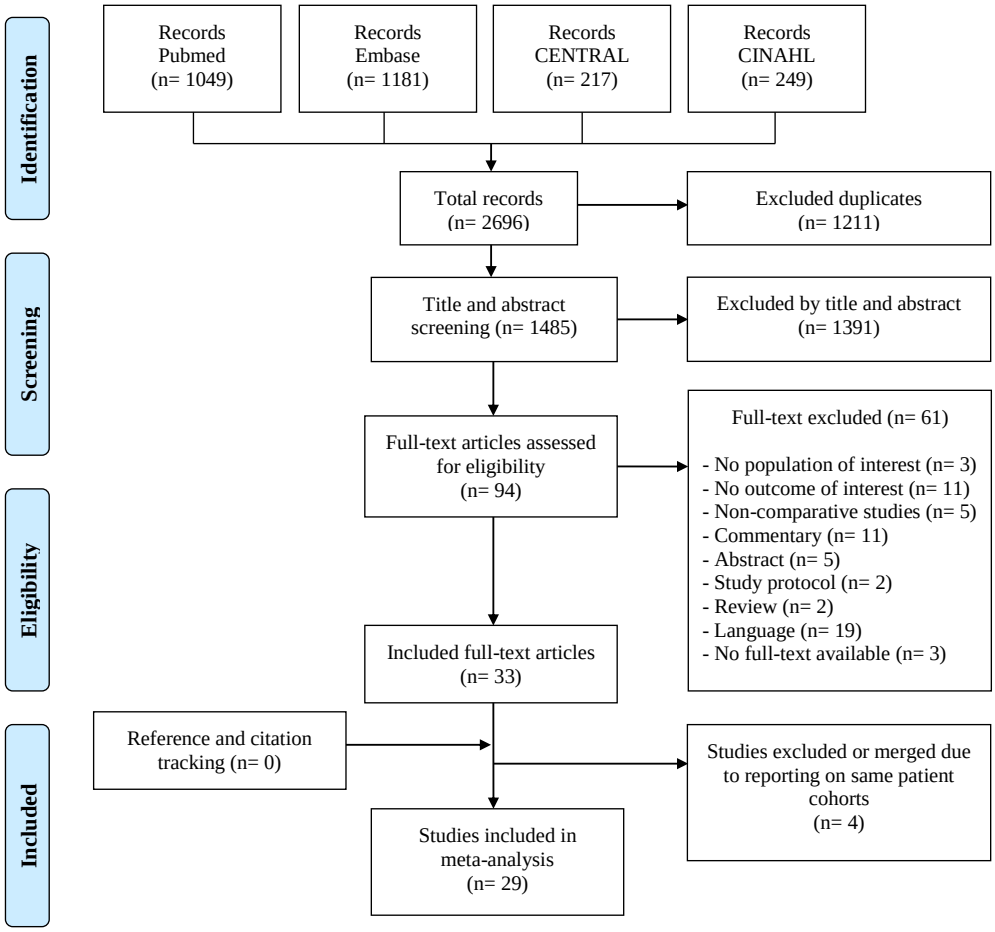

Figure 1. PRISMA flow diagram representing the search and selection of studies comparing operative versus nonoperative treatment of Achilles tendon ruptures.

The 19 observational studies—three prospective and 16 retrospective cohort studies—included 14 918 (94%) patients. Operative treatment was performed in 8906 patients, and 6012 were treated nonoperatively. The weighted mean age in the studies was 42 (range 17-86) years, 40 years in the operative group and 44 years in the nonoperative group, and 11 000 (74%) patients were male. The operative method was open surgery in nine studies, unclear in four studies, and a combination of open and minimally invasive surgery in six studies.

Quality assessment

The overall mean MINORS score was 14.3 (SD 5.2; range 5-23). The mean MINORS score for the RCTs was 20.3 (2.6; 16-23). The mean MINORS score for the observational studies was 11.2 (2.8; 5-16), 14 (2; 12-16) for the prospective cohort studies and 10.6 (2.6; 5-15) for the retrospective cohort studies. The details and distribution of MINORS scores are provided in supplementary Table S4.

Table 1. Baseline characteristics of included studies in a meta-analysis of Achilles tendon ruptures

Study	Year	Study period	Design	Country	Overall number	Number		Age (years, SD or range)		Gender (female/male)		FU (months, SD or range)	
RCTs						**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**
Cetti et al.	1993	1982-1984	RCT	Denmark	111	56	55	37.2 (21-62)	37.8 (21-65)	9/47	10/45	12	12
Keating et al.	2011	2000-2004	RCT	United Kingdom	80	39	41	41.2 (27-59)	39.5 (21-58)	11/28	9/32	12	12
Lantto et al.	2016	2009-2013	RCT	Finland	60	32	28	40 (27-57)	39 (28-60)	2/30	3/25	18	18
Metz et al.	2008	2004-2005	RCT	Netherlands	83	42	41	40 (23-63)	41 (25-62)	11/31	6/35	12	12
Möller et al.	2001	1995-1997	RCT	Sweden	112	59	53	39.6 (21-63)	38.5 (26-59)	8/51	5/48	24	24
Nilsson-Helander et al.	2010	2004-2007	RCT	Sweden	97	49	48	40.9 (8.8)	41.2 (9.5)	9/40	9/39	12	12
Nistor et al.	1981	1973-1977	RCT	Sweden	107	46	61	Overall 41 (21-77)		Overall 11/96		Overall 30 (12-60)	
Olsson et al.	2013	2009-2010	RCT	Sweden	100	49	51	39.8 (8.9)	39.5 (9.7)	10/39	4/47	12	12
Twaddle et al.	2007	1997-2002	RCT	New Zealand	50	25	25	41.8	40.3	6/14*	8/14*	12	12
Willits et al.	2010	2000-2005	RCT	Canada	144	72	72	39.7 (11)	41.1 (8.0)	13/59	13/59	24	24
Observational studies													
Bergkvist et al.	2012	2002-2006	RCS	Sweden	487	220	267	43 (11)	47 (14)	Overall 78/409		Overall 43 (12-97)	
Carden et al.	1987	1969-1981	RCS	United Kingdom	71	35	36	42.7 (26-68)	43 (22-70)	10/26	12/25	48 (12-204)	64 (12-120)
Costa et al.	2006	2001-2002	PCS	United Kingdom	96	48	48	42 (28-69)	53 (21-79)	7/40*	16/32	12	12
Cukelj et al.	2015	1998-2013	RCS	Croatia	90	60	30	34.8 (4.7	35.1 (4.7)	9/51	9/21	Overall 12	
Ebinesan et al.	2008	2001-2003	RCS	United Kingdom	63	51	12	44.8	52.1	14/37	6/6	N/A	
Fahlström et al.	1998	1990-1994	RCS	Sweden	31	22	9	34.6 (23-50)	39.4 (28-51)	Overall 4/27		Overall 39 (16-67)	
Grubor et al.	2012	2003-2010	RCS	Bosnia	42	34	8	N/A		Overall 5/37		Overall 12	
Gwynne-Jones et al.	2011	1999-2008	RCS	New Zealand	363	143	220	37.4	40.9	59/84	107/113	N/A	
Jaakkola et al.	2001	1985-1999	RCS	United States	73	35	38	37.3 (25-64)	38.0 (21-62)	3/32	6/32	43	54
Jackson et al.	2013	2002-2008	PCS	United Kingdom	80	29	51	37 (24-55)	47 (27-80)	3/26	16/35	N/A	
Kotnis et al.	2006	2000-2005	PCS	United Kingdom	125	67	58	41.0 (26-80)	43.9 (26-85)	19/48	18/40	12	12
Lim et al.	2017	N/A	RCS	New Zealand	200	99	101	40.1	42	21/41*	32/38*	Overall 78 (24-156)	
Miller et al.	2005	1990-1996	RCS	United Kingdom	172	140	32	45	49	23/117	11/21	Overall 95 (53)	
Nestorson et al.	2000	1992-1997	RCS	Sweden	24	14	10	72 (65-79)	71 (65-86)	3/11	1/9	Overall 39 (13-65)	
Rajasekar et al.	2005	1997-2001	RCS	United Kingdom	35	21	14	N/A		Overall 10/25		Overall 24 (9-48)	
Renninger et al.	2016	2011-2014	RCS	United States	57	27	30	32.3 (25-40)	29.7 (23-44)	0/27	0/30	Overall 10	
Van der Linden et al.	2004	1990-2001	RCS	Netherlands	292	212	80	37 (9.4)	42 (12)	58/154	21/59	Overall 72 (36)	
Wang et al.	2015	2007-2011	RCS	United States	12570	7625	4945	N/A		1514/6111	1737/3208	N/A	
Weber et al.	2003	1993-1998	RCS	Switzerland	47	24	23	38 (28-51)	39 (17-55)	4/13*	8/15	49 (30-79)	23 (12-42)

OP|NON operative/nonoperative; SD standard deviation; FU follow-up; N/A not available; RCT randomized controlled trial; PCS prospective cohort study; RCS retrospective cohort study; * ratio may not add up to total number of patients due to loss to follow-up

Primary outcome measure

Re-rupture rate

Re-rupture rate was reported in all 29 studies. The overall pooled effect showed that operative treatment was associated with a significant reduction in re-rupture rate compared with nonoperative treatment (risk ratio 0.43, 95% confidence interval 0.31 to 0.60; P<0.001; I^2=22%) (Figure 2). The pooled effect of RCTs showed a risk ratio of 0.40 (0.24 to 0.69; P<0.001; I^2=0%). The pooled effect of observational studies showed a risk ratio of 0.42 (0.28 to 0.64; P<0.001; I^2=31%). Re-rupture occurred in 2.3% of patients after operative treatment compared with 3.9% after nonoperative treatment (risk difference 1.6%). We found no significant difference in effect estimates from RCTs and observational studies (test for subgroup differences: P=0.91; I^2=0%). There was no visual asymmetry in the funnel plot (supplementary Figure S1). The Begg rank correlation test (P=0.66) and Egger linear regression test (P=0.16) indicated no evidence of publication bias.

Secondary outcome measures

Complication rate

Complication rate was reported in 26 (90%) studies—10 RCTs and 16 observational studies. The overall pooled effect showed a risk ratio of 2.76 (1.84 to 4.13; P<0.001; I^2=45%) in favor of nonoperative treatment compared with operative treatment (Figure 3). The pooled effect of RCTs showed a risk ratio of 3.26 (1.26 to 8.41; P=0.01; I^2=74%). The pooled effect of observational studies showed a risk ratio of 2.93 (2.28 to 3.75; P<0.001; I^2=0%). The incidence of complications was 4.9% after operative treatment compared with 1.6% after nonoperative treatment (risk difference 3.3%). Table 2 shows the classification and incidence of complications. The main complication after operative treatment was infection, which occurred in 2.8% of patients. The main complication after nonoperative treatment was deep vein thrombosis, which occurred in 1.2% of patients compared with 1.0% after operative treatment. We found no significant difference between effect estimates from RCTs and observational studies (test for subgroup differences: P=0.83; I^2=0%). There was no visual asymmetry in the funnel plot (supplementary Figure S2). The Begg rank correlation test (P=0.50) and Egger linear regression test (P=0.11) indicated no evidence of publication bias.

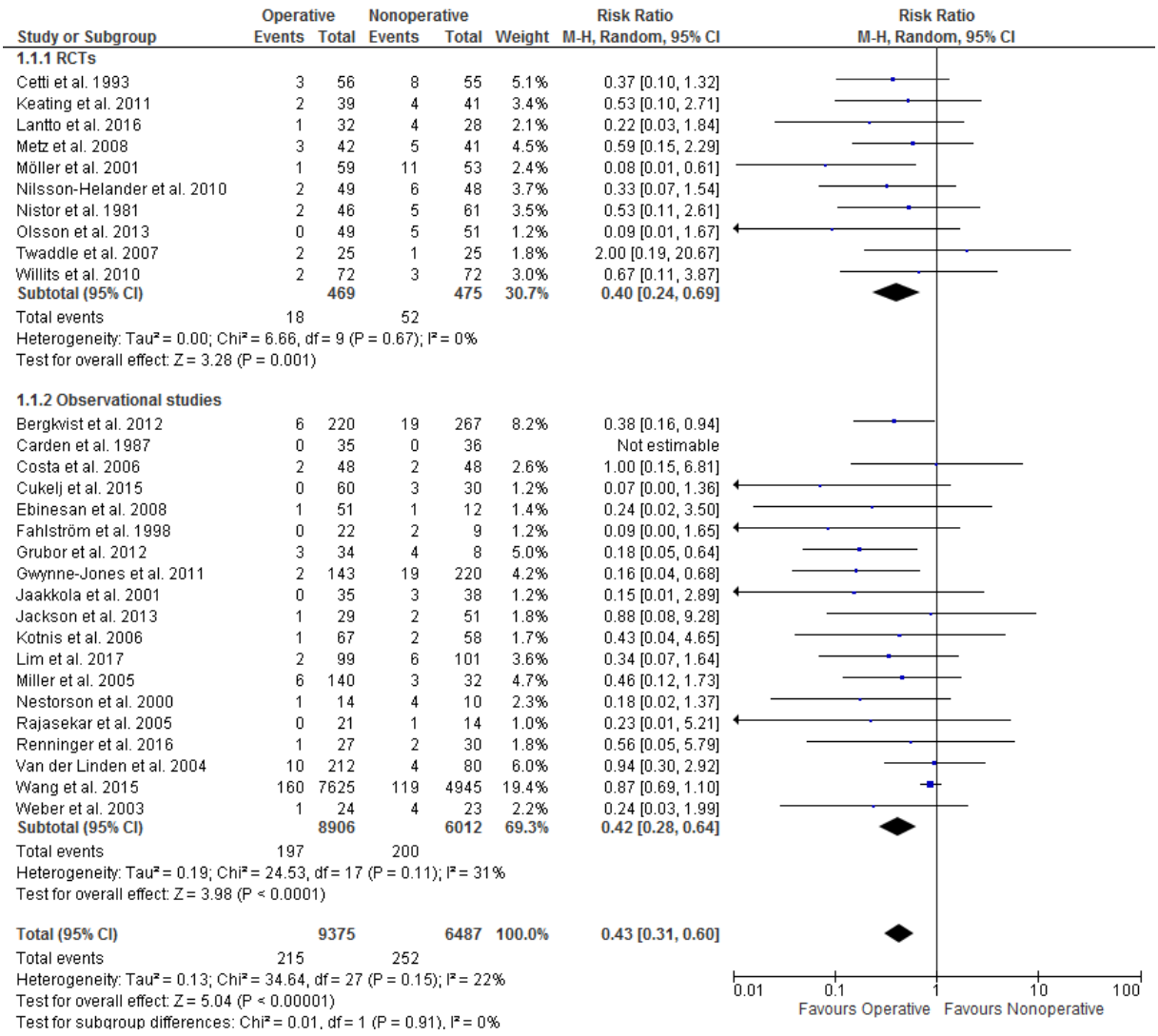

Figure 2. Forest plot of re-rupture rate in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

Functional outcome

Short term functional outcome assessed according to the ATRS score was reported in three (10%) studies. Nilsson-Helander et al reported a median ATRS score of 75 (range 31-100) in the operative group and 90 (31-100) in the nonoperative group.[58] Olsson et al reported a mean ATRS score of 82 (SD 20) in the operative group compared with 80 (23) in the nonoperative group.[59] In both RCTs, the differences found were non-significant. The observational study by Jackson et al reported a statistical significant difference in median ATRS score—94 (range 23-100) in the operative group and 84 (25-100) in the nonoperative group.[34]

Long-term functional outcome using the ATRS score was assessed in two observational studies. Bergkvist et al reported a mean ATRS score of 83 (SD 19) in the operative group and 78 (22) in

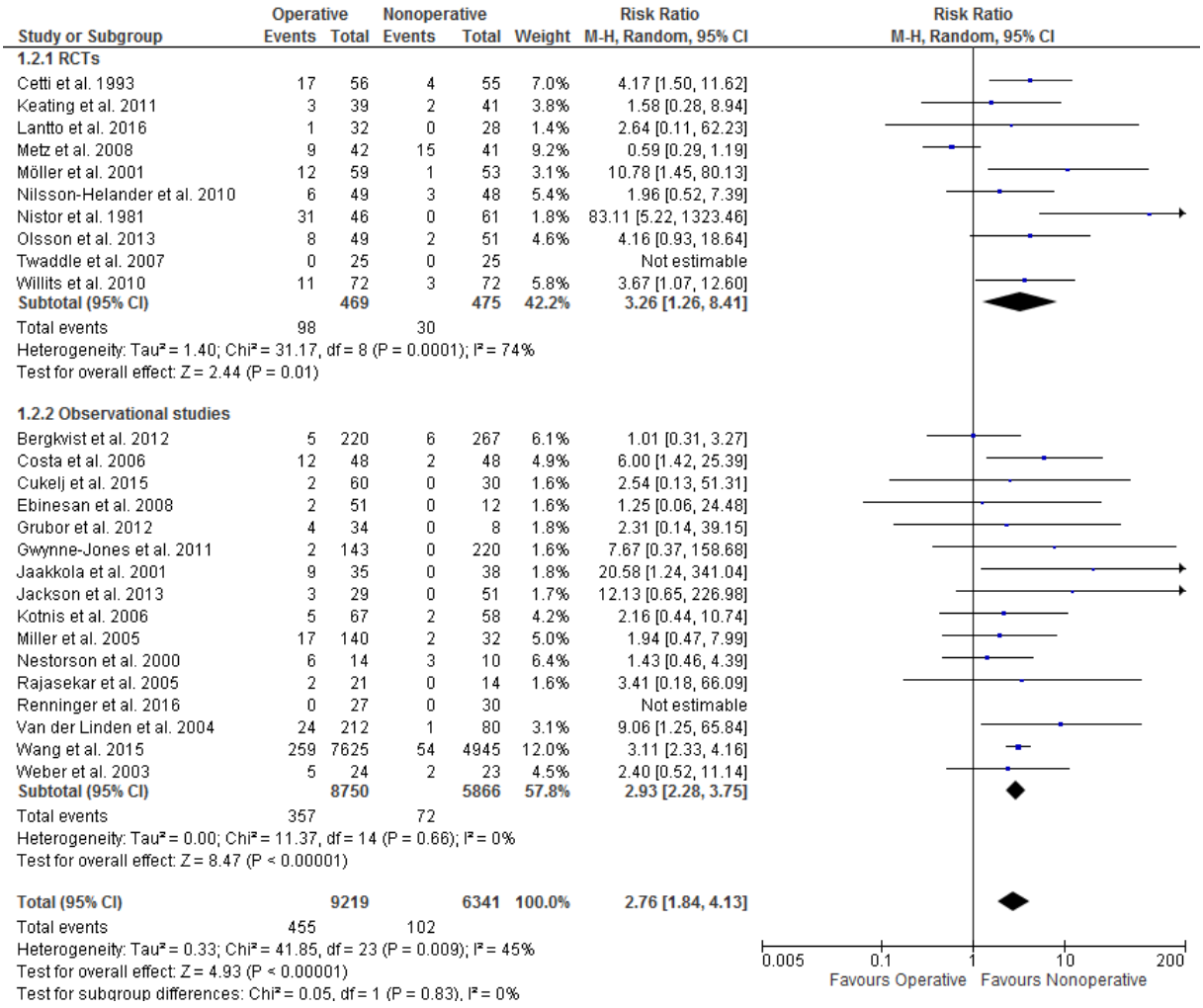

Figure 3. Forest plot of complication rate in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

the nonoperative group.[36] Lim et al reported a mean ATRS score of 85 in both groups.[44] No significant difference was found in either study. We did not pool functional outcome data owing to a wide variety in ATSR score reports and insufficient information to convert data. Descriptive details on functional outcome measures are provided in supplementary Table S5.

Table 2. Complications of included studies in a meta-analysis of Achilles tendon ruptures

Complication classification	Operative (n)	Incidence (%)	Nonoperative (n)	Incidence (%)
Pulmonary embolism	2	0.02	2	0.03
Deep vein thrombosis	89	0.97	74	1.17
Wound/skin infection	258	2.80	1	0.02
Sural nerve injury	39	0.42	5	0.08
Chronic pain	3	0.03	2	0.03
Scar/skin adhesion	35	0.38	15	0.24
Wound dehiscence	8	0.09	0	0
NP/other	21	0.23	3	0.05
Total	**455**	**4.94**	**102**	**1.61**

NP not specified; n number

Return to sports and work

Return to sports was reported by four (14%) studies—one RCT and three observational studies (supplementary Table S5). The mean time varied between six and nine months after operative treatment and between six and eight months after nonoperative treatment. We could not pool data on return to sports in a meta-analysis, as only one study reported a mean and standard deviation.

Return to work was reported in nine (31%) studies—four RCTs and five observational studies (supplementary Table S5). The outcome data of six studies could not be pooled owing to insufficient reporting of information. The pooled effect estimates of three studies—two RCTs and one observational study—showed no significant mean difference between operative and nonoperative treatment groups (supplementary Figure S3).

Primary sensitivity analysis

Weight bearing status

Early (four weeks or less) weight bearing status was reported in nine (31%) studies—five RCTs and four observational studies. The overall pooled effect showed a significant reduction in re-rupture rate after operative treatment compared with nonoperative treatment in the early (four weeks or less) full weight bearing studies (risk ratio 0.49, 0.26 to 0.93; P=0.03; I^2=9%) (supplementary Figure S4). Late (more than four weeks) weight bearing status was reported in 15 (52%) studies—four RCTs and 11 observational studies. The overall pooled effect of the late (more than four weeks) full weight bearing studies also showed a significant reduction in re-rupture rate in favor of operative treatment (risk ratio 0.33, 0.21 to 0.50; P<0.001; I^2=0%) (supplementary Figure S5).

Accelerated functional rehabilitation

Accelerated functional rehabilitation with early range of motion was performed in six (21%) studies—three RCTs and three observational studies. The overall pooled effect showed no significant difference between operative and nonoperative treatment regarding re-rupture rate (risk ratio 0.60, 0.26 to 1.37; P=0.23; I^2=0%) (Figure 4).

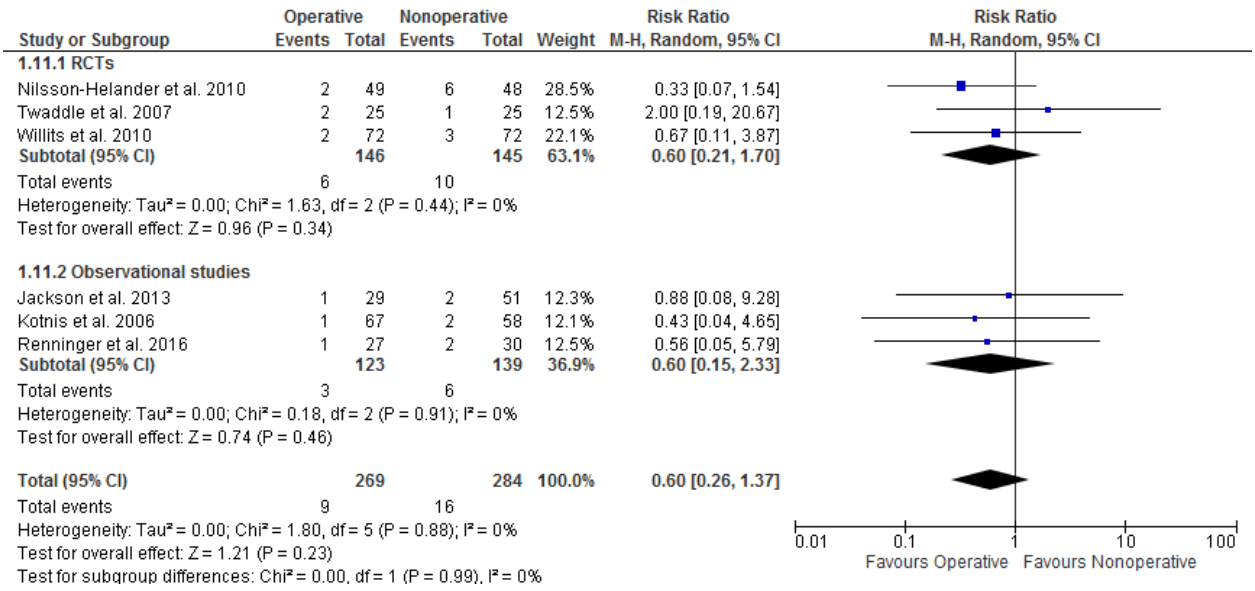

Figure 4. Forest plot of re-rupture rate in studies that included accelerated functional rehabilitation in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

Secondary sensitivity analyses

Table 3 shows the results of the secondary sensitivity analyses. Re-rupture rate was reported in 17 (59%) high quality studies—10 RCTs and seven observational studies. The overall pooled effect showed that operative treatment was associated with a significant reduction in re-rupture rate compared with nonoperative treatment (risk difference 5.1%; risk ratio 0.44, 0.30 to 0.64; P<0.001; I^2=0%) (supplementary Figure S6). Re-rupture rate was reported in 14 studies (48%) with a study period after the year 2000—six RCTs and eight observational studies. The overall pooled effect showed a significant reduction in re-rupture rate after operative treatment compared with nonoperative treatment (risk difference 0.9%; risk ratio 0.59, 0.42 to 0.83; P=0.002; I^2=10%) (supplementary Figure S7).

Table 3. Secondary sensitivity analyses of included studies in a meta-analysis of Achilles tendon ruptures

	Re-rupture					
	n	RD	RR	95% CI	p-value	I^2
All studies	29	1.6%	0.43	0.31 to 0.60	<0.001	22%
High-quality studies	17	5.1%	0.44	0.30 to 0.64	<0.001	0%
Study period (≥ 2000)	14	0.9%	0.59	0.42 to 0.83	0.002	10%
	Complication					
All studies	26	3.3%	2.76	1.84 to 4.13	<0.001	45%
High-quality studies	16	8.8%	2.72	1.44 to 5.12	0.002	62%
Study period (≥ 2000)	14	2.4%	2.15	1.28 to 3.60	0.004	52%

n number of studies; RD risk difference; RR risk ratio; 95% CI confidence interval; I^2 heterogeneity

Complication rate was reported in 16 (55%) high quality studies—10 RCTs and six observational studies. The overall pooled effect showed a risk ratio of 2.72 (1.44 to 5.12; P=0.002; I^2=62%) in favor of nonoperative treatment compared with operative treatment (risk difference 8.8%) (supplementary Figure S8). Complication rate was reported in 14 (48%) studies with a study period after the year 2000—six RCTs and eight observational studies. The overall pooled effect showed a risk ratio of 2.15 (1.28 to 3.60; P=0.004; I^2=52%) in favor of nonoperative treatment compared with operative treatment (risk difference 2.4%) (supplementary Figure S9).

Discussion

This systematic review and meta-analysis, including both RCTs and observational studies, compared outcomes after operative versus nonoperative treatment of acute Achilles tendon ruptures. The pooled effect estimate showed that operative treatment was associated with a significant reduction in re-rupture rate compared with nonoperative treatment. However, operative treatment resulted in a significantly higher rate of other complications. Sensitivity analyses showed a similar reduction in re-rupture rate after both early and late full weight bearing in favor of operative treatment compared with nonoperative treatment. However, we found no significant difference in re-rupture rate if accelerated functional rehabilitation with early range of motion was used. Sensitivity analyses with high quality studies and studies with a study period after the year 2000 also showed operative treatment to be associated with a significant reduction in re-rupture rate but a higher risk of other complications. We found no significant difference in effect estimates from RCTs and observational studies, for either re-rupture rate or complication rate.

Operative treatment reduces the risk of re-rupture compared with nonoperative treatment, but it also results in a higher risk of other complications. These findings are in accordance with those of previous meta-analyses.[3,4,6] Our review included 10 RCTs with a total of 944 patients, which resulted in an increased number of patients available for analyses, thus exceeding previous meta-analyses. Furthermore, the inclusion of observational studies resulted in an additional 14 918 patients for analyses. The previous meta-analyses reported a risk difference in re-rupture rate varying from 5% to 7% and a risk difference of other complications varying from 16% to 21%.[3-6] However, with the addition of observational studies, this review shows that differences between treatment groups are small, with a risk difference in re-rupture rate of 1.6% and a risk difference of 3.3% for other complications.

Functional outcome measures included the ATRS score, return to sports, and return to work. The ATRS score is the most commonly used patient reported instrument to evaluate limitations after treatment for an acute Achilles tendon rupture.[2] ATRS scores were not pooled in this study, but most studies showed no significant difference in ATRS score between the operative and nonoperative treatment groups. Resumption of sports was reported by only four studies; the results indicate no difference between operative treatment (six to nine months) and nonoperative treatment (six to eight months). The pooled effect of return to work showed no significant difference between treatment groups. Wilkins et al pooled return to work data from four studies and also found no statistical significant difference.[5] Soroceanu et al reported a statistically significant difference with the pooled data from four studies; operatively treated patients returned to work 19 days earlier than nonoperatively treated patients (P=0.0014).[6] Wilkins et al included return to work data in their pooled results from the studies by Nistor et al and Cetti et al.[5,52,57] In our study, we did not use the return to work data from these two studies owing to reporting of mean and range and the absence of standard deviations. Soroceanu et al also included the study by Cetti et al, as well as the study by Majewski et al,[6,52,62] which we excluded as it was in a language other than English. However, both our meta-analysis and the studies by Wilkins et al and Soroceanu et al are limited by the number of included patients in the return to work subgroup analyses.[5,6] Unfortunately, accurate comparison of functional outcome measures remains difficult owing to differences in protocols, patient oriented outcome measures, duration of follow-up, and presentation of data.

We found a lower re-rupture rate after both early and late full weight bearing in favor of operative treatment; this is in contrast to a previous meta-analysis by Van der Eng et al,[63] which found no difference in re-rupture rate. The previous meta-analysis could be limited by the number of included patients in the subgroup analyses. In our review, with the addition of observational studies, sensitivity analysis showed a significant difference in re-rupture rate after both early and late full weight bearing in favor of operative treatment. However, regardless of re-rupture rate, timing of weight bearing might influence other outcome measures as shown in different lower extremity injuries. De Boer et al found that early weight bearing regimens did not negatively affect functional outcome after treatment for displaced intra-articular calcaneal fractures.[64] Previously, Smeeing et al showed that early weight bearing tended to accelerate return to work and daily activities compared with late weight bearing, after internal fixation of ankle fractures.[65] Eliasson et al evaluated tendon elongation, mechanical properties, and functional

outcomes during the first 12 months after operative treatment of acute Achilles tendon ruptures.[66] However, they found that different rehabilitation regimens did not affect the outcome measures. Further research could focus on the effect of early weight bearing and long-term functional outcome after treatment of Achilles tendon ruptures.

Soroceanu et al found no significant difference in re-rupture rate in their subgroup analysis if functional rehabilitation with early range of motion was used (risk difference 1.7%; P=0.45).[6] However, they did not define the specific inclusion criteria and definition of early range of motion and functional rehabilitation. Unfortunately, evaluation of the effect of accelerated functional rehabilitation remains difficult owing to use of a wide variety of definitions and protocols. Our review found no significant difference in re-rupture rate if accelerated functional rehabilitation with early range of motion within three weeks was used after nonoperative treatment. These findings indicate that nonoperative management is acceptable for acute Achilles tendon ruptures, if patients are instructed and monitored according to a standardized rehabilitation protocol. However, both our review and the study by Soroceanu et al could be limited by the number of included patients in the subgroup analyses.[6]

The sensitivity analyses including high quality studies resulted in similar risk ratios and significance levels for re-rupture and complication rate. The results showed a risk difference of 5.1% for re-rupture rate, comparable to previous results of meta-analyses of RCTs alone. However, the risk difference of other complications (8.8%) in the high-quality sensitivity analysis was still considerably lower than in previous reports. This difference in other complications could be attributable to the inclusion of studies with both open and minimally invasive surgical techniques. The complication sensitivity analysis with high quality studies included one RCT with minimally invasive surgery and three observational studies that included both open and minimally invasive surgery. A meta-analysis by Yang et al,[67] including five RCTs and four cohort studies, found a significantly lower rate of deep infection with percutaneous treatment (0.6%) than with open treatment (3.6%) (P=0.04). However, the authors reported no significant difference in the rate of re-rupture between percutaneous and open treatment.[67]

The sensitivity analyses including studies with a study period after the year 2000 showed similar risk ratios and significance levels regarding re-rupture and complication rate. However, the risk differences between treatment groups were smaller than in all other analyses. The study period

sensitivity analyses included one RCT with minimally invasive surgery and four observational studies that included both open and minimally invasive surgery. These findings might indicate an overall reduction in complications after treatment of Achilles tendon ruptures due to the development of new rehabilitation protocols and operative techniques, regardless the use of operative or nonoperative treatment. However, it should be noted that both the level of high-quality studies and the study period were arbitrarily chosen.

We found no difference in pooled effect estimates from RCTs and observational studies. This is in line with previous reports showing that differences in effect estimates between RCTs and observational studies are small.[8,9,11,13,15,17] Observational studies, however, have also been associated with an overestimation of treatment effects compared with RCTs.[68,69] Hemkens et al assessed the difference in treatment effect estimates for mortality between observational studies and RCTs.[69] They evaluated 16 observational studies and 36 subsequent RCTs investigating the same clinical questions. Overall, observational studies significantly overestimated the effects of treatment compared with RCTs.[69] This overestimation of treatment effects could be explained by the effects of bias and confounding in observational studies.[70] However, overestimates by observational studies could also be explained by the potential selection bias in RCTs. RCTs require strict conditions such as selection of participants, inclusion/exclusion criteria, randomization method, and outcome measurements. The patient population in daily clinical practice can differ from the often highly selected patient populations in RCTs, which could be the reason for the discrepancy between treatment effects.[71,72] Nevertheless, observational studies increase sample size, which could lead to the evaluation of small treatment effects and infrequent outcome measures. Furthermore, the addition of observational studies might provide insight into a variety of populations and long-term effects. These results could improve the representation of daily clinical practice, with various levels of surgical experience and differences in operative techniques, provided that confounding has been adequately addressed.[12,13] In this meta-analysis, pooled effect estimates obtained from RCTs and observational studies were similar. Several orthopedic trauma meta-analyses including both RCTs and observational studies have shown high quality observational studies to result in similar treatment effects to RCTs.[15-17] These findings indicate that the effect of potentially unmeasured confounding in high quality observational studies seems relatively small, emphasizing the possible benefits of combining different study designs for the evaluation of objective outcome measures after surgical treatment.

Several potential limitations in this review need to be considered. Firstly, results might be influenced by missing articles. However, in addition to the extensive electronic database search, funnel plots did not indicate evidence for publication bias. Three studies could not be obtained in full text, but these articles were all published before 1996.[2,6,27,28] Secondly, the methodological quality of included studies was assessed by the MINORS criteria, which do not differentiate between randomized and non-randomized studies. However, the MINORS criteria were externally validated using RCTs and are able to distinguish adequately between study designs, as well-designed randomized trials score higher than well designed non-randomized studies.[21] The incidence of complications could be affected by the use of different treatment protocols. Five studies mentioned the use of prophylaxis for deep vein thrombosis.[51,53,56,58,59] However, descriptions were not comprehensive and the duration and types of prophylaxis varied widely. Finally, sensitivity analyses for the evaluation of weight bearing status and accelerated rehabilitation were performed using data from both RCTs and observational studies. However, the primary analysis showed no significant difference in effect estimates between the two study designs in terms of re-rupture rate.

Operative treatment of acute Achilles tendon ruptures reduces the risk of re-rupture compared with nonoperative treatment, although the incidence of re-ruptures is low and differences are small (2.3% v 3.9%). Operative treatment results in a higher risk of other complications compared with nonoperative treatment, mostly attributable to the increased risk of infection. Nonoperative treatment might be the preferred treatment for acute Achilles tendon rupture, owing to the higher risk of other complications after operative treatment and the relative small benefit in re-rupture rate. However, patient specific factors should always be taken into consideration and patients should be counselled about the incidence of complications.

Unfortunately, comparison of the literature remains difficult owing to a wide variety of rehabilitation protocols, weight bearing restrictions, treatment modalities, patient oriented outcome measures, and duration of follow-up. The discordance among studies makes comparisons between treatment modalities difficult, indicating a substantial need for further research. We suggest future research to focus on the effect of comorbidities on the success of treatment for Achilles tendon rupture. Studies could compare outcomes according to different age groups and evaluate effects in a variety of populations such as in patients with immunosuppression, diabetes mellitus, increased body mass index, neuropathy, peripheral

vascular disease, and dermatological disorders. Furthermore, future studies should strive to determine the optimal treatment for acute Achilles tendon ruptures on the basis of patients' expectations. Operative treatment is associated with complications inherent to the treatment itself, such as infection. However, athletic people may prefer operative treatment to enhance and expedite their outcomes, whereas a sedentary person with limited functional outcome expectations may prefer nonoperative treatment.[3] We believe that more data are needed for the development of a shared decision making algorithm to guide surgeons and physicians regarding the most appropriate treatment option for each individual patient.

Conclusion

In this meta-analysis, operative treatment of acute Achilles tendon ruptures reduced the risk of re-rupture compared with nonoperative treatment. However, re-rupture rates are low and differences between treatment groups are small, with a risk difference of 1.6%. Operative treatment results in a higher risk of other complications, with a risk difference of 3.3%, mostly due to the increased risk of infection. Patients should be counselled about complications, and the final decision for operative or nonoperative management should be based on patient specific factors and shared decision making. Further research is needed for the development of a shared decision-making algorithm. Moreover, this review emphasizes the potential benefits of adding high quality observational studies in meta-analyses to complement RCTs for the evaluation of objective outcome measures after surgical treatments.

References:

1. Ganestam A, Kallemose T, Troelsen A, Barfod KW. Increasing incidence of acute Achilles tendon rupture and a noticeable decline in surgical treatment from 1994 to 2013. A nationwide registry study of 33,160 patients. *Knee Surg Sports Traumatol Arthrosc.* 2016;24(12):3730-3737.
2. Egger AC, Berkowitz MJ. Achilles tendon injuries. *Curr Rev Musculoskelet Med.* 2017;10(1):72-80.
3. Khan RJ, Carey Smith RL. Surgical interventions for treating acute Achilles tendon ruptures. *Cochrane database Syst Rev.* 2010;(9):CD003674.
4. Jiang N, Wang B, Chen A, Dong F, Yu B. Operative versus nonoperative treatment for acute Achilles tendon rupture: a meta-analysis based on current evidence. *Int Orthop.* 2012;36(4):765-773.
5. Wilkins R, Bisson LJ. Operative versus nonoperative management of acute Achilles tendon ruptures: a quantitative systematic review of randomized controlled trials. *Am J Sports Med.* 2012;40(9):2154-2160.
6. Soroceanu A, Sidhwa F, Aarabi S, Kaufman A, Glazebrook M. Surgical versus nonsurgical treatment of acute Achilles tendon rupture: a meta-analysis of randomized trials. *J Bone Joint Surg Am.* 2012;94(23):2136-2143.
7. Erickson BJ, Mascarenhas R, Saltzman BM, et al. Is Operative Treatment of Achilles Tendon Ruptures Superior to Nonoperative Treatment?: A Systematic Review of Overlapping Meta-analyses. *Orthop J Sport Med.* 2015;3(4):2325967115579188.
8. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med.* 2000;342(25):1878-1886.
9. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med.* 2000;342(25):1887-1892.
10. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane database Syst Rev.* 2014;(4):MR000034.
11. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *J Clin Epidemiol.* 2010;63(3):238-245.
12. Arditi C, Burnand B, Peytremann-Bridevaux I. Adding non-randomised studies to a Cochrane review brings complementary information for healthcare stakeholders: an augmented systematic review and meta-analysis. *BMC Health Serv Res.* 2016;16(1):598.
13. Frieden TR. Evidence for Health Decision Making - Beyond Randomized, Controlled Trials. *N Engl J Med.* 2017;377(5):465-475.
14. Houwert RM, Smeeing DPJ, Ahmed Ali U, Hietbrink F, Kruyt MC, van der Meijden OA. Plate fixation or intramedullary fixation for midshaft clavicle fractures: a systematic review and meta-analysis of randomized controlled trials and observational studies. *J shoulder Elb Surg.* 2016;25(7):1195-1203.
15. Smeeing DPJ, van der Ven DJC, Hietbrink F, et al. Surgical Versus Nonsurgical Treatment for Midshaft Clavicle Fractures in Patients Aged 16 Years and Older: A Systematic Review, Meta-analysis, and Comparison of Randomized Controlled Trials and Observational Studies. *Am J Sports Med.* 2017;45(8):1937-1945.
16. Beks RB, Ochen Y, Frima H, et al. Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. *J shoulder Elb Surg.* 2018;27(8):1526-1534.
17. Beks RB, Peek J, de Jong MB, et al. Fixation of flail chest or multiple rib fractures: current evidence and how to proceed. A systematic review and meta-analysis. *Eur J Trauma Emerg Surg.* October 2018.
18. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000;283(15):2008-2012.
19. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol.* 2009;62(10):1006-1012.
20. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1-34.
21. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg.* 2003;73(9):712-716.
22. Nilsson-Helander K, Thomee R, Silbernagel KG, et al. The Achilles tendon Total Rupture Score (ATRS): development and validation. *Am J Sports Med.* 2007;35(3):421-426.
23. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. *Cochrane Collab.* 2011;0(March):10-11. http://www.cochrane.org/training/cochrane-handbook.
24. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 1997;315(7109):629-634.
25. The Nordic Cochrane Centre. Review Manager. *Cochrane Collab.* 2014:1-43.
26. Lill H, Moor C, Fecht E, Kalbe P, Echtermeyer V. Achilles tendon rupture - Operative or conservative functional treatment? *Aktuelle Traumatol.* 1996;26(3):95-100.
27. Haggmark T, Liedberg H, Eriksson E, Wredmark T. Calf muscle atrophy and muscle function after non-operative vs operative treatment of achilles tendon ruptures. *Orthopedics.* 1986;9(2):160-164.
28. Inglis AE, Scott WN, Sculco TP, Patterson AH. Ruptures of the tendo achillis. An objective assessment of surgical and non-surgical treatment. *J Bone Joint Surg Am.* 1976;58(7):990-993.

29. Moller M, Lind K, Movin T, Karlsson J. Calf muscle function after Achilles tendon rupture. A prospective, randomised study comparing surgical and non-surgical treatment. *Scand J Med Sci Sports*. 2002;12(1):9-16.
30. Olsson N, Nilsson-Helander K, Karlsson J, et al. Major functional deficits persist 2 years after acute Achilles tendon rupture. *Knee Surg Sports Traumatol Arthrosc*. 2011;19(8):1385-1393.
31. Olsson N, Karlsson J, Eriksson BI, Brorsson A, Lundberg M, Silbernagel KG. Ability to perform a single heel-rise is significantly related to patient-reported outcome after Achilles tendon rupture. *Scand J Med Sci Sports*. 2014;24(1):152-158.
32. Westin O, Nilsson Helander K, Gravare Silbernagel K, Moller M, Kalebo P, Karlsson J. Acute Ultrasonography Investigation to Predict Reruptures and Outcomes in Patients With an Achilles Tendon Rupture. *Orthop J Sport Med*. 2016;4(10):2325967116667920.
33. Costa ML, MacMillan K, Halliday D, et al. Randomised controlled trials of immediate weight-bearing mobilisation for rupture of the tendo Achillis. *J Bone Joint Surg Br*. 2006;88(1):69-77
34. Jackson G, Sinclair VF, McLaughlin C, Barrie J. Outcomes of functional weight-bearing rehabilitation of Achilles tendon ruptures. *Orthopedics*. 2013;36(8):e1053-9.
35. Kotnis R, David S, Handley R, Willett K, Ostlere S. Dynamic ultrasound as a selection tool for reducing achilles tendon reruptures. *Am J Sports Med*. 2006;34(9):1395-1400.
36. Bergkvist D, Astrom I, Josefsson P-O, Dahlberg LE. Acute Achilles tendon rupture: a questionnaire follow-up of 487 patients. *J Bone Joint Surg Am*. 2012;94(13):1229-1233.
37. Carden DG, Noble J, Chalmers J, Lunn P, Ellis J. Rupture of the calcaneal tendon. The early and late management. *J Bone Joint Surg Br*. 1987;69(3):416-420.
38. Cukelj F, Bandalovic A, Knezevic J, Pavic A, Pivalica B, Bakota B. Treatment of ruptured Achilles tendon: Operative or non-operative procedure? *Injury*. 2015;46 Suppl 6:S137-42.
39. Ebinesan AD, Sarai BS, Walley GD, Maffulli N. Conservative, open or percutaneous repair for acute rupture of the Achilles tendon. *Disabil Rehabil*. 2008;30(20-22):1721-1725.
40. Fahlstrom M, Bjornstig U, Lorentzon R. Acute Achilles tendon rupture in badminton players. *Am J Sports Med*. 1998;26(3):467-470.
41. Grubor P, Grubor M. Treatment of Achilles tendon rupture using different methods. *Vojnosanit Pregl*. 2012;69(8):663-668.
42. Gwynne-Jones DP, Sims M, Handcock D. Epidemiology and outcomes of acute Achilles tendon rupture with operative or nonoperative treatment using an identical functional bracing protocol. *Foot ankle Int*. 2011;32(4):337-343.
43. Jaakkola JI, Beskin JL, Griffith LH, Cernansky G. Early ankle motion after triple bundle technique repair vs. casting for acute Achilles tendon rupture. *Foot ankle Int*. 2001;22(12):979-984.
44. Lim CS, Lees D, Gwynne-Jones DP. Functional Outcome of Acute Achilles Tendon Rupture With and Without Operative Treatment Using Identical Functional Bracing Protocol. *Foot ankle Int*. 2017;38(12):1331-1336.
45. Miller D, Waterston S, Reaper J, Barrass V, Maffulli N. Conservative management, percutaneous or open repair of acute Achilles tendon rupture: a retrospective study. *Scott Med J*. 2005;50(4):160-165.
46. Nestorson J, Movin T, Moller M, Karlsson J. Function after Achilles tendon rupture in the elderly: 25 patients older than 65 years followed for 3 years. *Acta Orthop Scand*. 2000;71(1):64-68.
47. Rajasekar K, Gholve P, Faraj AA, Kosygan KP. A subjective outcome analysis of tendo-Achilles rupture. *J Foot Ankle Surg*. 2005;44(1):32-36.
48. Renninger CH, Kuhn K, Fellars T, Youngblood S, Bellamy J. Operative and Nonoperative Management of Achilles Tendon Ruptures in Active Duty Military Population. *Foot ankle Int*. 2016;37(3):269-273.
49. van der Linden-van der Zwaag HMJ, Nelissen RGHH, Sintenie JB. Results of surgical versus non-surgical treatment of Achilles tendon rupture. *Int Orthop*. 2004;28(6):370-373.
50. Wang D, Sandlin MI, Cohen JR, Lord EL, Petrigliano FA, SooHoo NF. Operative versus nonoperative treatment of acute Achilles tendon rupture: An analysis of 12,570 patients in a large healthcare database. *Foot Ankle Surg*. 2015;21(4):250-253.
51. Weber M, Niemann M, Lanz R, Muller T. Nonoperative treatment of acute rupture of the achilles tendon: results of a new protocol and comparison with operative treatment. *Am J Sports Med*. 2003;31(5):685-691.
52. Cetti R, Christensen SE, Ejsted R, Jensen NM, Jorgensen U. Operative versus nonoperative treatment of Achilles tendon rupture. A prospective randomized study and review of the literature. *Am J Sports Med*. 1993;21(6):791-799.
53. Keating JF, Will EM. Operative versus non-operative treatment of acute rupture of tendo Achillis: a prospective randomised evaluation of functional outcome. *J Bone Joint Surg Br*. 2011;93(8):1071-1078.
54. Lantto I, Heikkinen J, Flinkkila T, et al. A Prospective Randomized Trial Comparing Surgical and Nonsurgical Treatments of Acute Achilles Tendon Ruptures. *Am J Sports Med*. 2016;44(9):2406-2414.
55. Metz R, Verleisdonk E-JMM, van der Heijden GJ-M-G, et al. Acute Achilles tendon rupture: minimally invasive surgery versus nonoperative treatment with immediate full weightbearing--a randomized controlled trial. *Am J Sports Med*. 2008;36(9):1688-1694.
56. Moller M, Movin T, Granhed H, Lind K, Faxen E, Karlsson J. Acute rupture of tendon Achillis. A prospective randomised study of comparison between surgical and non-surgical treatment. *J Bone Joint Surg Br*. 2001;83(6):843-848.

57. Nistor L. Surgical and non-surgical treatment of Achilles Tendon rupture. A prospective randomized study. *J Bone Joint Surg Am*. 1981;63(3):394-399.
58. Nilsson-Helander K, Silbernagel KG, Thomee R, et al. Acute achilles tendon rupture: a randomized, controlled study comparing surgical and nonsurgical treatments using validated outcome measures. *Am J Sports Med*. 2010;38(11):2186-2193.
59. Olsson N, Silbernagel KG, Eriksson BI, et al. Stable surgical repair with accelerated rehabilitation versus nonsurgical treatment for acute Achilles tendon ruptures: a randomized controlled study. *Am J Sports Med*. 2013;41(12):2867-2876.
60. Twaddle BC, Poon P. Early motion for Achilles tendon ruptures: is surgery important? A randomized, prospective study. *Am J Sports Med*. 2007;35(12):2033-2038.
61. Willits K, Amendola A, Bryant D, et al. Operative versus nonoperative treatment of acute Achilles tendon ruptures: a multicenter randomized trial using accelerated functional rehabilitation. *J Bone Joint Surg Am*. 2010;92(17):2767-2775.
62. Majewski M, Rickert M, Steinbruck K. [Achilles tendon rupture. A prospective study assessing various treatment possibilities]. *Orthopade*. 2000;29(7):670-676.
63. van der Eng DM, Schepers T, Goslings JC, Schep NWL. Rerupture rate after early weightbearing in operative versus conservative treatment of Achilles tendon ruptures: a meta-analysis. *J Foot Ankle Surg*. 2013;52(5):622-628.
64. De Boer AS, Van Lieshout EMM, Van Moolenbroek G, Den Hartog D, Verhofstad MHJ. The effect of time to post-operative weightbearing on functional and clinical outcomes in adults with a displaced intra-articular calcaneal fracture; A systematic review and pooled analysis. *Injury*. February 2018.
65. Smeeing DPJ, Houwert RM, Briet JP, et al. Weight-bearing and mobilization in the postoperative care of ankle fractures: a systematic review and meta-analysis of randomized controlled trials and cohort studies. *PLoS One*. 2015;10(2):e0118320.
66. Eliasson P, Agergaard A-S, Couppe C, et al. The Ruptured Achilles Tendon Elongates for 6 Months After Surgical Repair Regardless of Early or Late Weightbearing in Combination With Ankle Mobilization: A Randomized Clinical Trial. *Am J Sports Med*. 2018;46(10):2492-2502.
67. Yang B, Liu Y, Kan S, et al. Outcomes and complications of percutaneous versus open repair of acute Achilles tendon rupture: A meta-analysis. *Int J Surg*. 2017;40:178-186.
68. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA*. 2001;286(7):821-830.
69. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPA. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. *BMJ*. 2016;352:i493.
70. Boyko EJ. Observational research--opportunities and limitations. *J Diabetes Complications*. 2013;27(6):642-648.
71. Van Spall HGC, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA*. 2007;297(11):1233-1240.
72. Khan AY, Preskorn SH, Baker B. Effect of study criteria on recruitment and generalizability of the results. *J Clin Psychopharmacol*. 2005;25(3):271-275.

Supplementary materials to Chapter 6

Table S1. Search syntax performed last on April 25th, 2018

Database	Syntax
PubMed/MEDLINE (n= 1049)	((((((((((achilles tendon[MeSH Terms]) OR achilles tendon[Title/Abstract]) OR achill*[Title/Abstract]) OR tendoachill*[Title/Abstract]) OR calcaneal*[Title/Abstract]) OR calcanean*[Title/Abstract]) OR calcaneus[Title/Abstract])) AND ((((((rupture[MeSH Terms]) OR Tendon Injuries[MeSH Terms]) OR ruptu*[Title/Abstract]) OR injur*[Title/Abstract]) OR lesion*[Title/Abstract]) OR tear*[Title/Abstract]))))) AND ((((((((((((surgical procedures, operative[MeSH Terms]) OR orthopedics[MeSH Terms]) OR surg*[Title/Abstract]) OR operat*[Title/Abstract]) OR orthop*[Title/Abstract]) OR kessler[Title/Abstract]) OR bunnell[Title/Abstract]) OR krackow[Title/Abstract]) OR ma and griffit[Title/Abstract]) OR achillon[Title/Abstract]) OR tenolig[Title/Abstract]) OR dresden[Title/Abstract]) OR percuta*[Title/Abstract]))) AND ((((((((((((((((Conservative Treatment[MeSH Terms]) OR physical therapy modalities[MeSH Terms]) OR conservative[Title/Abstract]) OR conventional[Title/Abstract]) OR non-operative[Title/Abstract]) OR non operative[Title/Abstract]) OR nonoperative[Title/Abstract]) OR non-surgical[Title/Abstract]) OR non surgical[Title/Abstract]) OR nonsurgical[Title/Abstract]) OR cast*[Title/Abstract]) OR brace*[Title/Abstract]) OR splint*[Title/Abstract]) OR boot*[Title/Abstract]) OR bandage*[Title/Abstract]) OR tape[Title/Abstract]) OR taping[Title/Abstract]))
Embase (n= 1181)	('achilles tendon rupture'/de OR (('achilles tendon'/exp OR 'achilles tendon':ab,ti OR 'achill*':ab,ti OR 'tendoachill*':ab,ti OR 'calcanean*':ab,ti OR 'calcaneus':ab,ti) AND ('rupture'/de OR 'tendon injury'/de OR 'injur*':ab,ti OR 'ruptu*':ab,ti OR 'lesion*':ab,ti OR 'tear*':ab,ti))) AND (('surgery'/de OR 'orthopedic surgery'/de OR 'surg*':ab,ti OR 'operat*':ab,ti OR 'orthop*':ab,ti OR 'kessler':ab,ti OR 'bunnell':ab,ti OR 'krackow':ab,ti OR 'griffit':ab,ti OR 'achillon':ab,ti OR 'tenolig':ab,ti OR 'dresden':ab,ti OR 'percuta*':ab,ti) AND ('conservative treatment'/de OR 'conservative':ab,ti OR 'conventional':ab,ti OR 'non-operative':ab,ti OR 'non operative':ab,ti OR 'nonoperative':ab,ti OR 'non-surgical':ab,ti OR 'non surgical':ab,ti OR 'nonsurgical':ab,ti OR 'cast*':ab,ti OR 'brace*':ab,ti OR 'splint*':ab,ti OR 'boot*':ab,ti OR 'bandage*':ab,ti OR 'tape':ab,ti OR 'taping':ab,ti))
CENTRAL (n= 217)	(Achilles AND rupture)
CINAHL (n= 249)	(((MH achilles tendon OR TI achilles tendon OR AB achilles tendon OR TI achill* OR AB achill* OR TI tendoachill* OR AB tendoachill* OR TI calcaneal* OR AB calcaneal* OR TI calcanean* OR AB calcanean* OR TI calcaneus OR AB calcaneus) AND (MH rupture OR MH tendon injuries OR TI rupt* OR AB rupt* OR TI injur* OR AB injur* OR TI lesion* OR AB lesion* OR TI tear* OR AB tear*)) AND ((MH surgical procedures, operative OR MH orthopedics OR TI surg* OR AB surg* OR TI operat* OR AB operat* OR TI orthop* OR AB orthop* OR TI kessler OR AB kessler OR TI bunnell OR AB bunnell OR TI krackow OR AB krackow OR TI ma and griffit OR AB ma and griffit OR TI achillon OR AB achillon OR TI tenolig OR AB tenolig OR TI dresden OR AB dresden OR TI percuta* OR AB percuta*) AND (MH Conservative Treatment OR MH physical therapy modalities OR TI conservative OR AB conservative OR TI conventional OR AB conventional OR TI non-operative OR AB non-operative OR TI non operative OR AB non operative OR TI nonoperative OR AB nonoperative OR TI non-surgical OR AB non-surgical OR TI non surgical OR AB non surgical OR TI nonsurgical OR AB nonsurgical OR TI cast* OR AB cast* OR TI brace* OR AB brace* OR TI splint* OR AB splint* OR TI boot* OR AB boot* OR TI bandage* OR AB bandage* OR TI tape OR AB tape OR TI taping OR AB taping))

Table S2. Quality assessment according to the MINORS criteria in a meta-analysis of Achilles tendon ruptures

Criteria	Reported and adequate (2)	Reported but inadequate (1)	Not reported (0)
Clearly stated aim	Aim including outcomes reported	Aim reported without outcomes	Not reported
Inclusion consecutive patients	Inclusion/exclusion criteria reported	Unclear description inclusion/exclusion criteria	Not reported
Prospective collection data	Prospective	Not applicable	Not applicable
Appropriate endpoints	Appropriate endpoints to aim study	Endpoints not appropriate to aim study	Not reported
Unbiased assessment	Blinded evaluation of outcomes	Reason not blinding stated	Not reported
Appropriate follow-up	≥ 1 year	< 1 year	Not reported
Loss to follow-up < 5%	≤ 5%	> 5%	Not applicable
Prospective calculation study size	Prospective power-analysis performed	Prospective calculation without power-analysis	Not applicable
Adequate control group	Operative versus nonoperative treatment	Not applicable	Not applicable
Contemporary groups	Study/control group managed during same period	Study/control not managed during same period	Not reported
Baseline equivalence groups	Baseline characteristics described and comparable	Baseline characteristics not comparable	Not reported
Adequate statistical analyses	Statistical analysis described including type of analyses	Inadequate description statistical analysis	Not reported

Items are scored 0 (not reported), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S3. Treatment characteristics of included studies in a meta-analysis of Achilles tendon ruptures

Study	Year	Overall number	Number		Operative method	Operative suture technique	Nonoperative treatment	Duration (weeks) nonoperative	Accelerated fun. Rehab	Weightbearing ≤ 4 weeks
RCTs			**OP**	**NON**						
Cetti et al.	1993	111	56	55	Open	Bunnell	Cast	8	No	No
Keating et al.	2011	80	39	41	Open	Kessler	Cast	10	No	No
Lantto et al.	2016	60	32	28	Open	Krackow	Cast/orthosis	7	No	Yes
Metz et al.	2008	83	42	41	MI	Bunnell	Cast/brace	7	No	Yes
Möller et al.	2001	112	59	53	Open	Kessler	Cast	8	No	Yes
Nilsson-Helander et al.	2010	97	49	48	Open	Kessler	Cast/brace	8	Yes	No
Nistor et al.	1981	107	46	61	Open	Bunnell	Cast	8	N/A	N/A
Olsson et al.	2013	100	49	51	Open	Kessler	Brace	8	No	Yes
Twaddle et al.	2007	50	25	25	Open	Krackow	Cast/orthosis	8	Yes	No
Willits et al.	2010	144	72	72	Open	Krackow	Brace	8	Yes	Yes
Observational studies										
Bergkvist et al.	2012	487	220	267	N/A	N/A	Brace	8	N/A	No
Carden et al.	1987	71	35	36	Open	End-to-end	Cast	8	N/A	No
Costa et al.	2006	96	48	48	Open	End-to-end	Cast/orthosis	12	No	Both
Cukelj et al.	2015	90	60	30	Open/MI (n=30/30)	Lindholm/Ma-Griffith	Cast	8	N/A	No
Ebinesan et al.	2008	63	51	12	Open/MI (n=20/31)	Krackow/Percutaneous	Cast	6-8	No	No
Fahlström et al.	1998	31	22	9	N/A	N/A	Cast	N/A	N/A	N/A
Grubor et al.	2012	42	34	8	Open/MI (n=15/19)	Lindholm/Ma-Griffith	Cast	10	No	No
Gwynne-Jones et al.	2011	363	143	220	Open	Kessler	Cast	8	No	No
Jaakkola et al.	2001	73	35	38	Open	Bunnell	Cast	8-10	No	No
Jackson et al.	2013	80	29	51	Open	Kessler	Boot	8	Yes	Yes
Kotnis et al.	2006	125	67	58	Open/MI (n=37/30)	N/A	Cast	8	Yes	No
Lim et al.	2017	200	99	101	Open	Kessler	Cast/boot	8	No	No
Miller et al.	2005	172	140	32	Open/MI (n=86/54)	Kessler/Percutaneous	Cast	6	No	No
Nestorson et al.	2000	24	14	10	N/A	End-to-end	Cast/brace	8	N/A	N/A
Rajasekar et al.	2005	35	21	14	Open	Kessler	Cast	10-13	No	No
Renninger et al.	2016	57	27	30	Open/MI (N/A)	N/A	N/A	N/A	Yes	N/A
Van der Linden et al.	2004	292	212	80	Open	Bunnell	Cast	12	N/A	Yes
Wang et al.	2015	12570	7625	4945	N/A	N/A	N/A	N/A	No	N/A
Weber et al.	2003	47	24	23	Open	Kessler	Boot	12	No	Yes

OP|NON operative/nonoperative; N/A not available; MI minimally invasive; n number; Fun. Rehab functional rehabilitation

Table S4. Quality assessment of included studies in a meta-analysis of Achilles tendon ruptures

MINORS criteria	**RCTs**	Cetti et al. 1993	Keating et al. 2011	Lantto et al. 2016	Metz et al. 2008	Möller et al. 2001	Nilsson-Helander et al. 2010	Nistor et al. 1981	Olsson et al. 2013	Twaddle et al. 2007	Willits et al. 2010	**Observational studies**	Bergkvist et al. 2012	Carden et al. 1987	Costa et al. 2006	Cukelj et al. 2015	Ebinesan et al. 2008	Fahlström et al. 1998	Grubor et al. 2012	Gwynne-Jones et al. 2011	Jaakkola et al. 2001	Jackson et al. 2013	Kotnis et al. 2006	Lim et al. 2017	Miller et al. 2005	Nestorson et al. 2000	Rajasekar et al. 2005	Renninger et al. 2016	Van der Linden et al. 2004	Wang et al. 2015	Weber et al. 2003
Clearly stated aim		2	2	2	2	2	2	2	2	2	2		2	2	2	2	1	1	1	2	2	2	2	2	2	1	2	2	2	2	2
Inclusion consecutive patients		2	2	2	2	2	2	1	2	2	2		2	0	0	0	2	1	0	2	2	2	2	1	2	1	2	2	2	1	2
Prospective collection of data		2	2	2	2	2	2	2	2	2	2		0	0	2	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0
Appropriate endpoints		1	2	2	2	2	2	2	2	2	2		2	1	2	1	2	1	1	2	1	2	1	2	2	1	1	2	2	1	1
Unbiased assessment endpoints		0	1	0	0	0	1	1	1	1	1		0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
Appropriate follow-up		1	2	2	2	2	2	2	2	1	2		2	2	1	0	0	1	0	0	1	0	1	2	2	0	1	0	0	0	1
Loss to follow-up < 5%		0	2	2	2	2	0	0	2	0	0		0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Prospective calculation size		0	2	2	2	0	2	0	2	2	2		0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Adequate control group		2	2	2	2	2	2	2	2	2	2		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Contemporary groups		2	2	2	2	2	2	2	2	2	2		2	2	2	1	2	2	1	2	2	2	2	0	2	2	1	1	1	0	1
Baseline equivalence of groups		2	2	2	2	2	2	1	2	1	2		1	1	1	1	0	1	0	0	1	0	0	2	2	2	0	2	1	0	1
Adequate statistical analysis		2	2	2	2	2	2	1	2	2	2		1	0	1	0	0	2	0	0	0	0	1	2	1	1	0	2	2	2	2
Total MINORS score		**16**	**23**	**22**	**22**	**20**	**21**	**16**	**23**	**19**	**21**		**14**	**10**	**14**	**7**	**9**	**11**	**5**	**10**	**11**	**12**	**16**	**13**	**15**	**10**	**9**	**13**	**13**	**8**	**12**

Items are scored 0 (not reported/ not applicable), 1 (reported but inadequate) or 2 (reported and adequate). The overall score ranging from 0 to 24 for comparative studies

Table S5. Outcome measures of included studies in a meta-analysis of Achilles tendon ruptures

Study	Re-rupture n (%)		Complication n (%)		ATRS ≤ 1 year (mean, SD or range)		ATRS > 1 year (mean, SD or range)		Return sports (months, SD or range)		Return work (weeks, SD or range)	
RCTs	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**	**OP**	**NON**
Cetti et al. 1993	3 (5)	8 (15)	17 (30)	4 (7)	N/A		N/A		N/A		6.2 (0.5-19)	8.0 (0-52)
Keating et al. 2011	2 (5)	4 (10)	3 (8)	2 (5)	N/A		N/A		7.8 (3-12)	8.0 (4-12)	N/A	
Lantto et al. 2016	1 (3)	4 (14)	1 (3)	0	N/A		N/A		N/A		N/A	
Metz et al. 2008	3 (7)	5 (12)	9 (21)	15 (37)	N/A		N/A		N/A		8.4 (12)	15.4 (16)
Möller et al. 2001	1 (2)	11 (21)	12 (20)	1 (2)	N/A		N/A		N/A		8 (7)	11 (8)
Nilsson-Helander et al. 2010	2 (4)	6 (13)	6 (12)	3 (6)	75 (31-100)*	90 (31-100)*	N/A		N/A		N/A	
Nistor et al. 1981	2 (4)	5 (8)	31 (67)	0	N/A		N/A		N/A		13 (0-30)	9 (0-44)
Olsson et al. 2013	0	5 (10)	8 (16)	2 (4)	82 (20)	80 (23)	N/A		N/A		N/A	
Twaddle et al. 2007	2 (8)	1 (4)	0	0	N/A		N/A		N/A		N/A	
Willits et al. 2010	2 (3)	3 (4)	11 (15)	3 (4)	N/A		N/A		N/A		N/A	
Observational studies												
Bergkvist et al. 2012	6 (3)	19 (7)	5 (2)	6 (2)	N/A		83 (19)	78 (22)	N/A		N/A	
Carden et al. 1987	0	0	N/A		N/A		N/A		8.9 (1.0)	6.5 (0.7)	8.5 (1)	5.4 (1)
Costa et al. 2006	2 (4)	2 (4)	12 (25)	2 (4)	N/A		N/A		N/A		N/A	
Cukelj et al. 2015	0	3 (10)	2 (3)	0	N/A		N/A		N/A		N/A	
Ebinesan et al. 2008	1 (2)	1 (8)	2 (4)	0	N/A		N/A		N/A		N/A	
Fahlström et al. 1998	0	2 (22)	N/A		N/A		N/A		N/A		7 (0-13)	11 (0-26)
Grubor et al. 2012	3 (9)	4 (50)	4 (12)	0	N/A		N/A		N/A		N/A	
Gwynne-Jones et al. 2011	2 (1)	19 (9)	2 (1)	0	N/A		N/A		N/A		N/A	
Jaakkola et al. 2001	0	3 (8)	9 (26)	0	N/A		N/A		N/A		N/A	
Jackson et al. 2013	1 (3)	2 (4)	3 (10)	0	94 (23-100)*	84 (25-100)*	N/A		N/A		N/A	
Kotnis et al. 2006	1 (1)	2 (3)	5 (7)	2 (3)	N/A		N/A		N/A		N/A	
Lim et al. 2017	2 (2)	6 (6)	N/A		N/A		84.8	85.3	N/A		N/A	
Miller et al. 2005	6 (4)	3 (9)	17 (12)	2 (6)	N/A		N/A		8.0	7.5	10.3	10.4
Nestorson et al. 2000	1 (7)	4 (40)	6 (43)	3 (30)	N/A		N/A		N/A		N/A	
Rajasekar et al. 2005	0	1 (7)	2 (10)	0	N/A		N/A		N/A		N/A	
Renninger et al. 2016	1 (4)	2 (7)	0	0	N/A		N/A		N/A		29 (13-52)	35 (21-70)
Van der Linden et al. 2004	10 (5)	4 (5)	24 (11)	1 (1)	N/A		N/A		N/A		N/A	
Wang et al. 2015	160 (2)	119 (2)	259 (3)	54 (1)	N/A		N/A		N/A		N/A	
Weber et al. 2003	1 (5)	4 (17)	5 (21)	2 (9)	N/A		N/A		5.7	5.5	6 (1-12)	4 (0-12)

OP|NON operative/nonoperative; n number; SD standard deviation; N/A not available; ATRS Achilles Tendon Rupture Score; * median

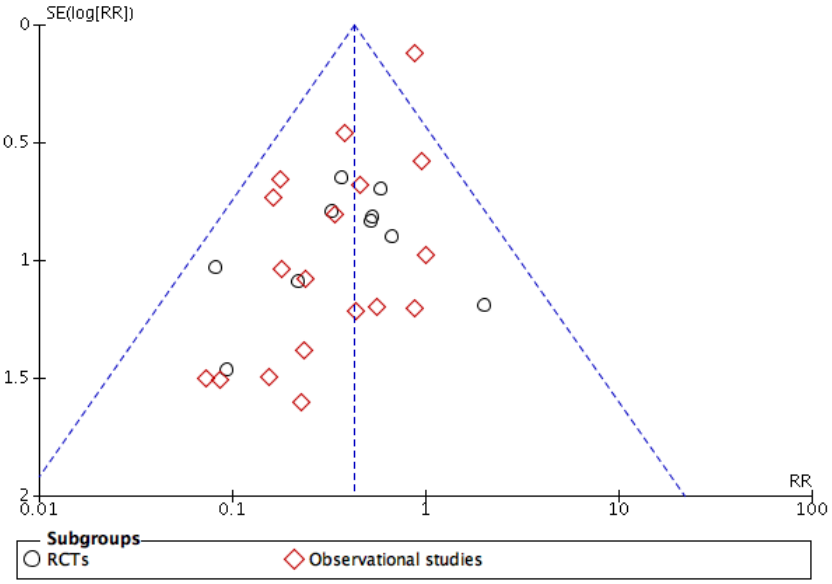

Figure S1. Funnel plot of re-rupture rate in a meta-analysis of Achilles tendon ruptures (RR risk ratio; SE standard error).

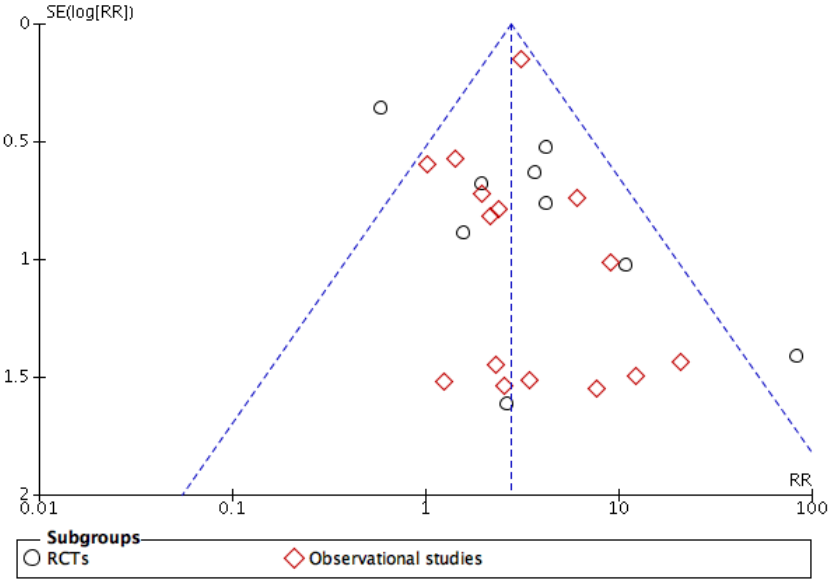

Figure S2. Funnel plot of complication rate in a meta-analysis of Achilles tendon ruptures (RR risk ratio; SE standard error).

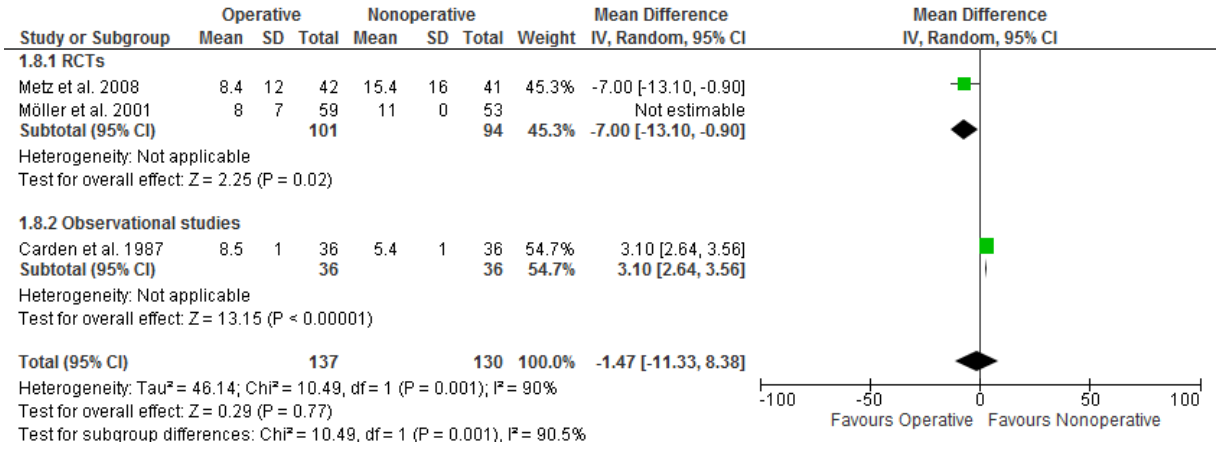

Figure S3. Forest plot of return to work in weeks in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

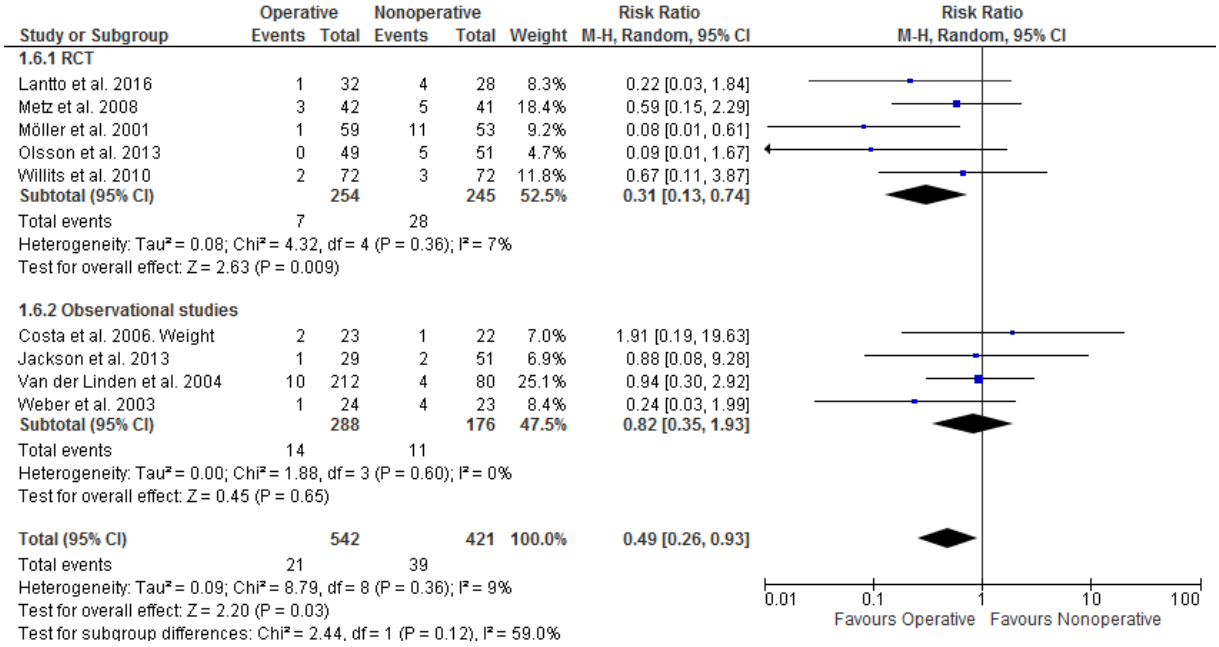

Figure S4. Forest plot of re-rupture rate in studies with ≤ 4 weeks full weightbearing in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

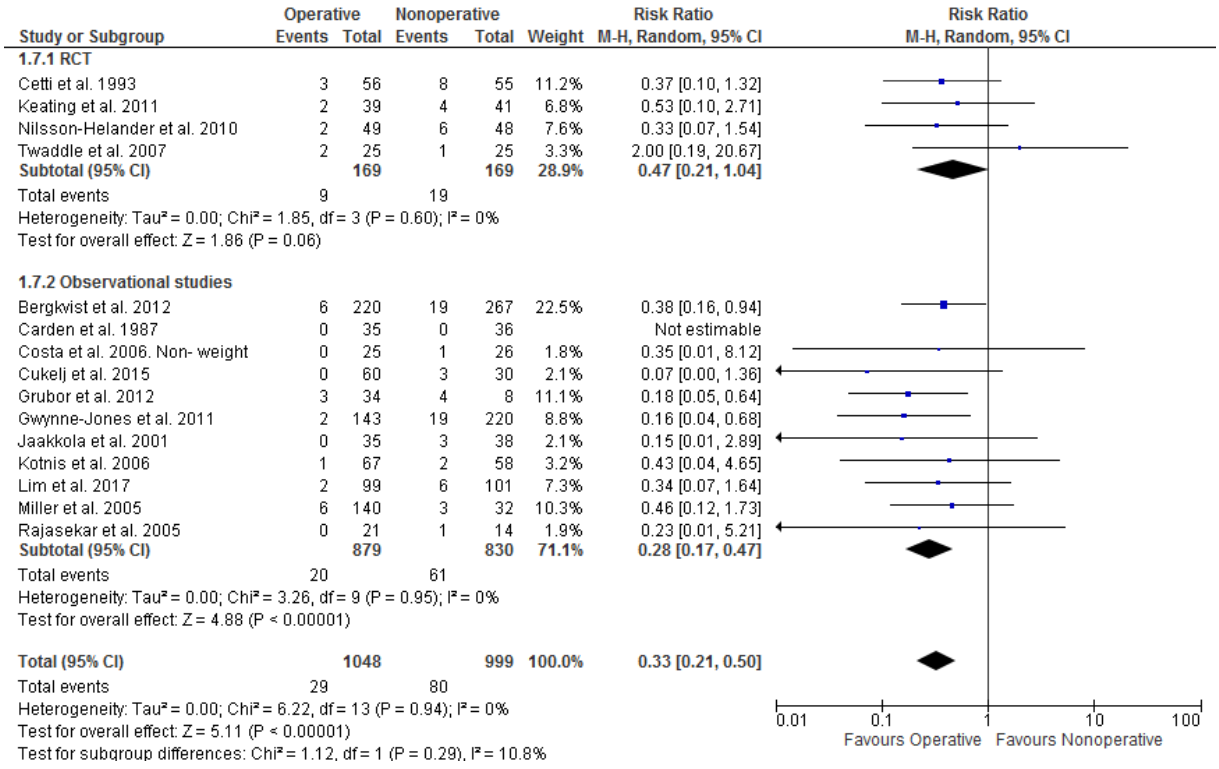

Figure S5. Forest plot of re-rupture rate in studies with > 4 weeks full weightbearing in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

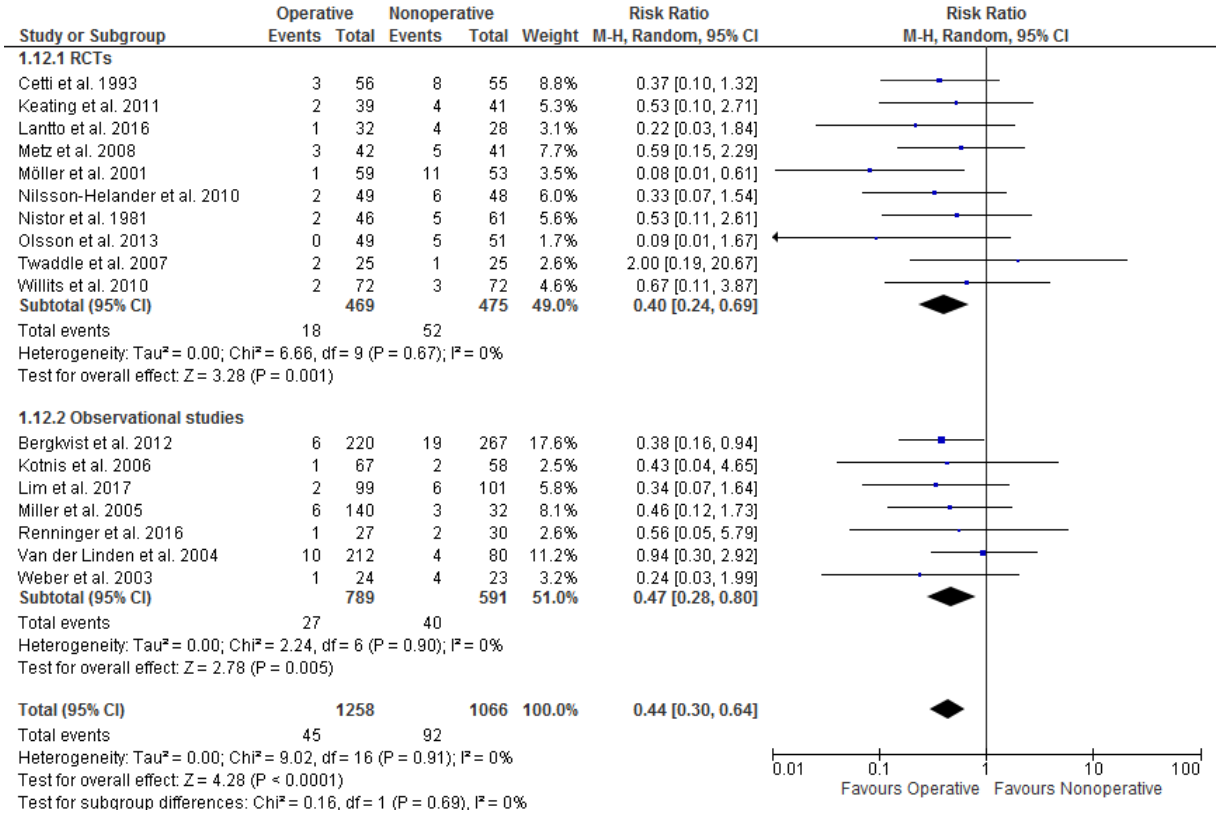

Figure S6. Forest plot of re-rupture rate in high-quality studies in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

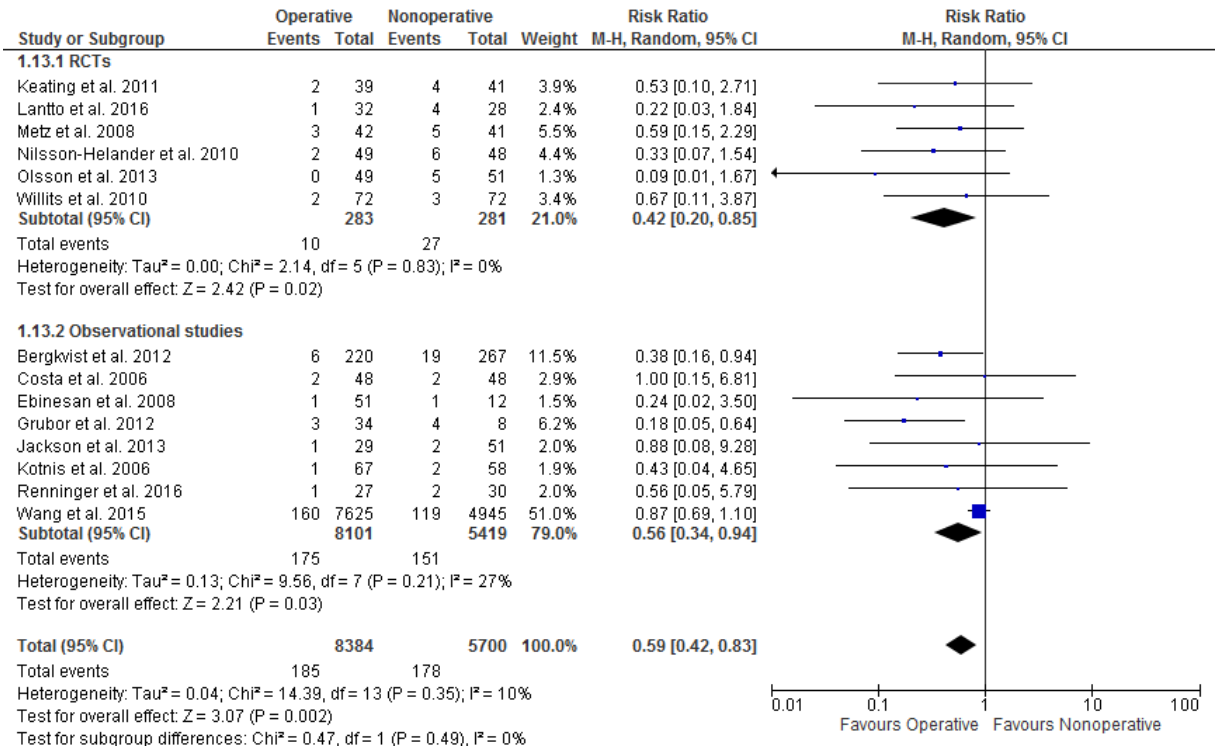

Figure S7. Forest plot of re-rupture rate in studies with a study period after the year 2000 in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

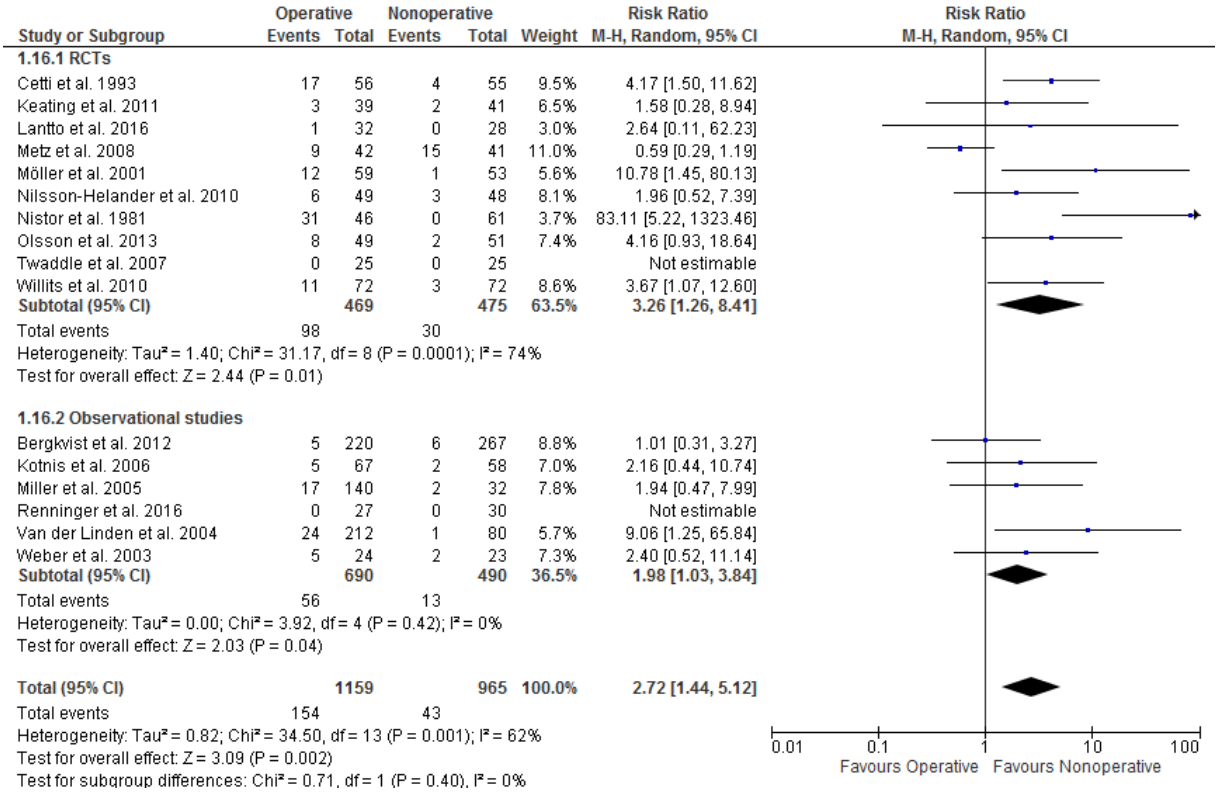

Figure S8. Forest plot of complication rate in high-quality studies in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

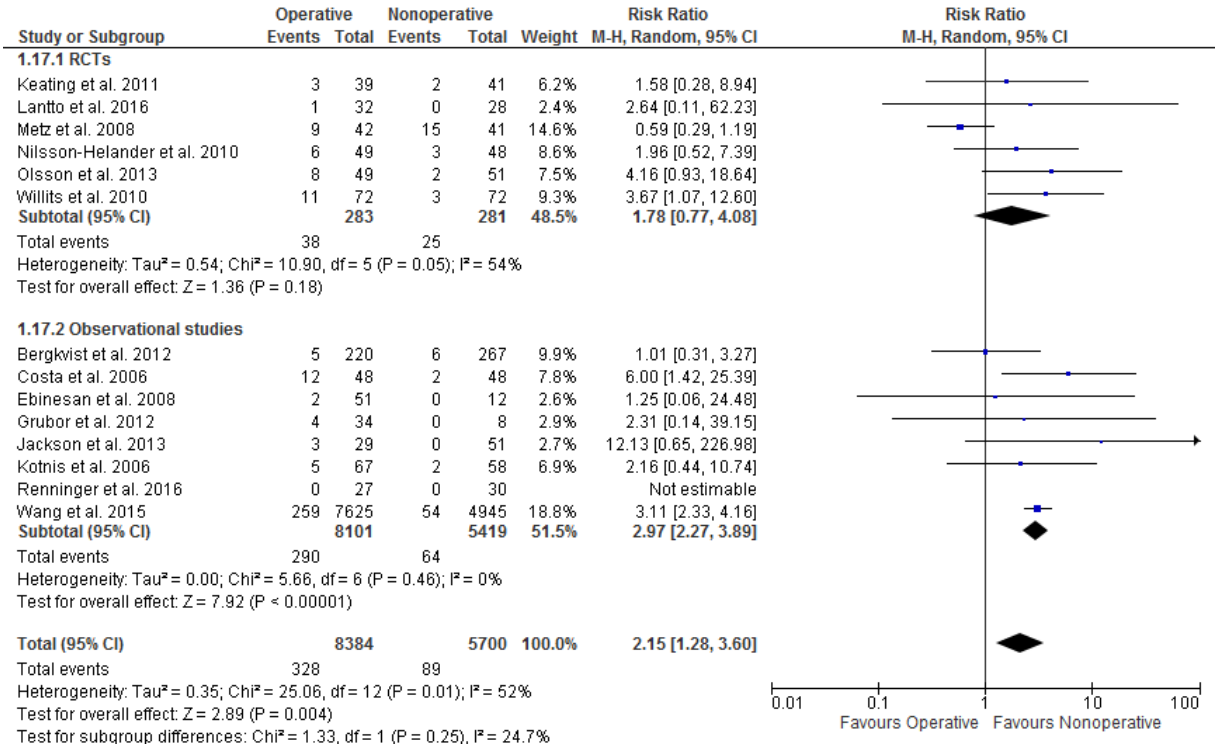

Figure S9. Forest plot of complication rate in studies with a study period after the year 2000 in a meta-analysis of Achilles tendon ruptures. M-H, Mantel-Haenszel.

PART 2

PATIENT-REPORTED OUTCOME MEASURES

CHAPTER 7

Surgical treatment of acute and chronic AC joint dislocations: five-year experience with conventional and modified LARS fixation by a single surgeon

Yassine Ochen, Reinier B. Beks, Benjamin L. Emmink, Philippe Wittich, Detlef van der Velde, R. Marijn Houwert, Jort Keizer

(Journal of Orthopaedics)

Abstract

Background

Different surgical fixation methods are available for the treatment of acromioclavicular (AC) joint dislocations. The aim of this study was to present the results of five years of experience with the Ligament Augmentation and Reconstruction System (LARS) fixation technique by a single surgeon.

Methods

A single-center retrospective cohort study was performed. All patients treated for an AC joint dislocation with LARS fixation by the same surgeon between 2012 and 2016 (n=20) were eligible for inclusion. All these dislocations were unstable injuries, Rockwood type-III or higher, requiring acute or chronic repair. The primary outcome was the QuickDASH score. Secondary outcomes were the Subjective Shoulder Value (SSV), Numerical Rating Scale (NRS) pain score, return to work, complications, and implant removal.

Results

17 patients (85%) were available for final follow-up. The median follow-up was 23 months (IQR; 17─34). The median QuickDASH score was 7 (IQR; 2−18), the median SSV was 90 (IQR; 80−90), and the median NRS pain score was 2 (IQR;1−3). Patients returned to work after a median of 8 weeks (IQR; 6−12). There was no significant difference in functional outcome scores between acute and chronic repair, or between the conventional and modified LARS fixation groups. There were two major complications requiring revision surgery, one ruptured LARS ligament and one case of deep wound infection. Implant removal was performed in one patient.

Conclusion

The LARS ligament fixation technique seems to be effective for the treatment of AC joint dislocations, resulting in good short- and mid-term patient-reported functional outcome. LARS fixation might also be an acceptable treatment option for active patients with symptomatic chronic AC dislocations.

Introduction

Dislocation of the acromioclavicular (AC) joint is a frequently encountered injury, with an incidence of 8.9/100,000 per year and most common in the young to middle-aged population.[1] Furthermore, dislocation of the AC joint represents 9% of injuries in the shoulder region.[2]

AC joint dislocations can be classified according to the Rockwood classification based on the relation to the coracoclavicular (CC) ligament, the deltoid muscle, the trapezius muscle and the direction of dislocation. The Rockwood classification consist of six types, from minor subluxation to complete dislocation. The less severe Rockwood type-I or II AC dislocation are incomplete separations with an intact CC ligament, and generally treated conservatively. The optimal treatment of the most common AC joint dislocation, Rockwood type-III, remains unclear. The Rockwood type-III dislocation involves tears of both the AC and CC ligaments, with 25% to 100% displacement compared with the contralateral side. Operative treatment is recommended for Rockwood type-IV, V, and VI due to severe dislocation and >100% displacement compared with the contralateral side.[2-7] However, the management of AC joint dislocations also depends on a variety of factors, including the patient's level of activity and age.[7]

Different surgical fixation methods are available for the treatment of AC dislocations. However, no consensus has been reached regarding the optimal fixation method. In general, the operative management includes repair of the AC ligament, CC ligament repair or rigid internal fixation of the AC joint.[5-7]

The use of rigid fixation methods is commonly used for the treatment of AC dislocations. However, these implants have been related to complications such as implant irritation, implant dislocation, implant migration and loss of reduction. Moreover, implant removal is often required due to implant-related complications and impingement syndrome.[5,8]

Several synthetic ligament devices such as PDS, the Gore-Tex, Dacron, Tightrope system, carbon fiber and Mersilene tape have been used to overcome the shortcomings of rigid implants.[5,6,8] Complications related to the use of these synthetic devices include ligament failure, incomplete reduction, foreign body reaction, bony erosion, coracoid fractures, and clavicle fractures.[8]

The Ligament Augmentation and Reconstruction System (LARS) is a more recently developed artificial ligament. The LARS fixation act as a non-rigid and extra-articular reinforcement, allowing stabilization and reduction of the CC ligament. The LARS artificial ligament is composed of industrial strength polyester fibers, providing superior strength to the original CC ligaments.[7]

The mid-term results after the use of LARS fixation for the treatment of acute and chronic AC joint dislocations have not been widely studied. The aim of this study was to present five years of experience with the conventional and modified LARS fixation technique by a single surgeon, evaluating functional outcome scores, return to work, and complications.

Methods

Study design

Approval was granted by the Institutional Review Board and informed consent was obtained from all subjects. A retrospective cohort study was performed using data from a level II trauma center. All patients with AC joint dislocations who were treated with LARS fixation by the same surgeon between 2012 and 2016 were eligible for inclusion. Inclusion criteria were: (1) unstable AC joint dislocation, (2) acute and chronic repair, (3) age 18 years or older, (4) LARS fixation, (5) minimum of six months' follow-up, and (6) operated by a single surgeon. Acute repair was defined as AC dislocation treated within eight weeks of injury. Chronic was defined as persisting AC dislocation requiring repair more than eight weeks following injury, despite nonoperative treatment. Data collection was performed by reviewing electronic medical records, operative reports, radiology reports and telephone interviews by an independent research fellow. Electronic medical records were reviewed to collect the following baseline characteristics: age, gender, trauma date, trauma mechanism, time from injury to fixation, Rockwood classification, surgical indication, complications and revision surgery.

Surgical procedure

Operations were performed under general anesthesia with the patient placed in a beach chair position. An incision was made over the AC joint using a sagittal incision. The deltoideus and trapezius muscles were detached from the lateral clavicle. The AC joint was exposed, after which debridement of fibrotic tissue was performed. If necessary, the lateral end of the clavicle was resected to allow adequate anatomical reduction. A guide wire was used to place a LARS ligament

around the base of the coracoid process. Two bony tunnels were drilled through the superior clavicle. A 4-mm hole was drilled from craniodorsal to caudoventral through the clavicle, lateral from the coracoid. Medial from the coracoid, a 4.5mm hole was drilled through the clavicle from cranioventral to caudodorsal. The LARS ligament was fixated in the lateral clavicle drill-hole with an interference screw. Subsequently, reduction was performed under direct visualization by putting the LARS ligament under tension. Following adequate reduction, the medial clavicle site was also fixated with an interference screw. From 2015 onwards, a modified LARS fixation technique was used.[9] Following the fixation with interference screws, excess ligament was passed around the coracoid a second time medial from the coracoid. The LARS ligament was then fixated with a figure eight knot and secured with fiber wire (Figure 1). AC joint reduction and screw placement were checked under fluoroscopic guidance. Finally, the m. deltoideus muscle was reinserted to the lateral clavicle and the fascia was closed in layers.

Postoperative management

Patients were immobilized in a sling for four weeks postoperatively. During this period patients were allowed early active mobilization and performed daily shoulder pendulum exercises. Weight-bearing activities and resisted exercises were not permitted until approval from the treating surgeon. Follow-up visits with control radiographs were scheduled after four weeks (Figure 2), after which patients could start with resisted exercises and physiotherapy. Removal of the LARS ligament and interference screws was not routinely performed.

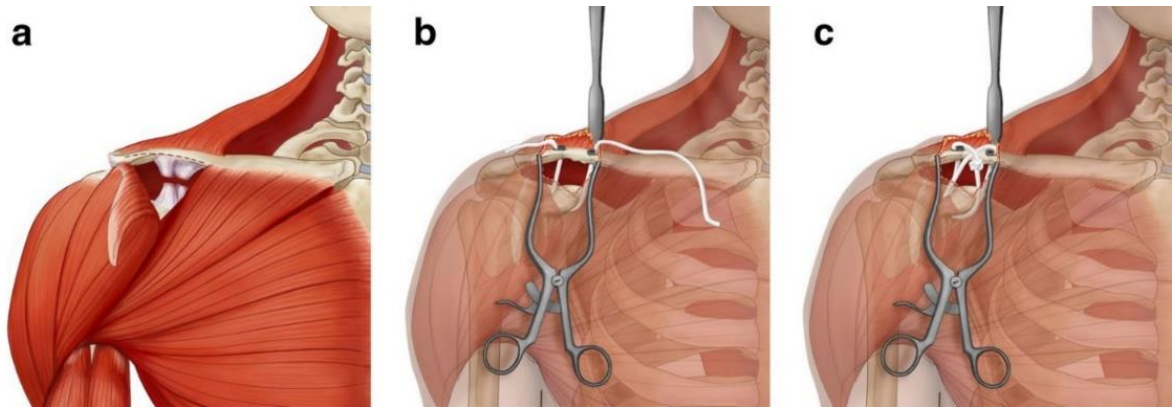

Figure 1. Surgical technique for modified acromioclavicular joint reconstruction with LARS ligament. **A.** Opened trapeziusdeltoid fascia. **B.** Clavicle reduction and LARS fixation on the clavicle. **C.** Figure of eight reconstruction.

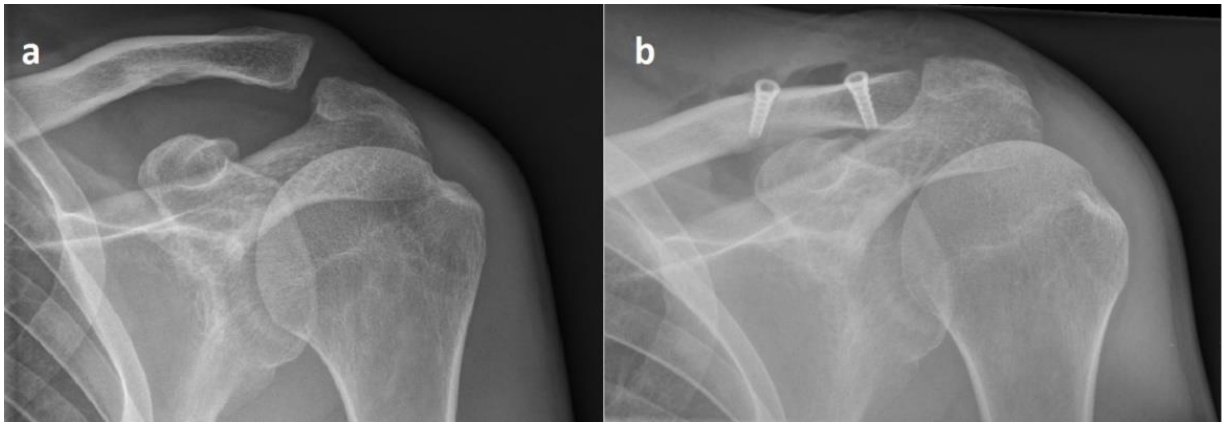

Figure 2. **A.** Anteroposterior view of an acromioclavicular dislocation Rockwood type-III. **B.** Anteroposterior view after coracoclavicular ligament repair with a LARS ligament.

Evaluation

Outcomes were assessed at least 6 months following LARS fixation using the QuickDASH score, Subjective Shoulder Value (SSV), Numerical Rating Scale (NRS) pain score and return to work. The QuickDASH is a validated and shortened version of the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH). The QuickDASH is a patient-reported outcome instrument developed to measure upper extremity disability and symptoms, resulting in a score ranging no disability (0) to most severe disability (100).[10,11] The SSV is a subjective value for shoulder function expressed as a percentage of an uninjured shoulder, which would score 100%.[12,13] The NRS is a 11-point scale to measure pain intensity, ranging from no pain (0) to worst imaginable pain (10).[14] Return to work was defined as the duration in weeks before resuming work.

Statistical analysis

Descriptive results are presented as mean values with standard deviations and range (SD, range), median values with interquartile range (IQR) or absolute numbers and percentages (%). Continuous variables were evaluated using the Mann–Whitney U test. The significance level was defined as a p value <0.05. All statistical analyses were performed using STATA 13.1 (StataCorp LP, TX, USA).

Results

Study population

In total, 20 patients met the inclusion criteria. Two patients could not be contacted, and one patient was not able to answer the questions due to progressive dementia. Resulting in the

Table 1. Baseline characteristics (n=20)

	n (%)
Age [mean, SD]	46 (18)
Gender	
Male	17 (85)
Female	3 (15)
Side injury	
Right	8 (40)
Left	12 (60)
Trauma mechanism	
Sports-related	7 (35)
Bicycle accident	5 (25)
Motor vehicle accident	5 (25)
Fall	3 (15)
Rockwood classification	
III	11 (55)
IV	6 (30)
V	3 (15)
Indication for surgery	
Persistent or progressive pain	12 (60)
Rockwood classification	7 (35)
Patient's request	1 (5)
Repair	
Acute	9 (45)
Chronic	11 (55)
LARS fixation technique	
Conventional	14 (70)
Modified	6 (30)
Follow up in months [median, IQR]	23 (18-34)

n number; SD standard deviation; IQR interquartile range

Table 2. Distribution of Rockwood types according to timing of repair

Rockwood classification	**Acute (n=9)**	**Chronic (n=11)**
III	2 (22%)	9 (82%)
IV	5 (56%)	1 (9%)
V	2 (22%)	1 (9%)

n number

inclusion of 17 patients (85%) for follow-up. The mean age was 46 years (SD 18, range 17-80) and 17 patients (85%) were male (Table 1). In most cases, patients sustained the shoulder injury in a sports-related accident (35%). Eleven patients (55%) sustained an AC dislocation Rockwood type-III, six patients (30%) a Rockwood type-IV, and three patients (15%) a Rockwood type-V. The main indication for operative treatment was persistent or progressive shoulder pain (60%). Acute LARS repair was performed in nine patients (45%). Conventional LARS fixation was performed in 14 patients (70%) and six patients (30%) were treated with the modified LARS fixation technique. The median follow-up was 23 months (IQR; 17─34). The distribution of Rockwood classification and timing of repair is provided in Table 2.

Functional outcome

The median QuickDASH score at final follow-up was 7 (IQR; 2−18), as shown in Table 3. The median SSV was 90 (IQR; 80−90) and the NRS pain score 2 (IQR;1−3). Patients returned to

Table 3. Functional outcome measures (n=17)

	Median (IQR)
QuickDASH	7 (2-18)
SSV	90 (80-90)
NRS	2 (1-3)
Return to work (weeks)	8 (6-12)

n number; IQR interquartile range; SSV Subjective Shoulder Value; NRS Numerical Rating Scale pain score

Table 4. Functional outcome according to timing of repair and LARS fixation technique

	Acute (n=7)	**Chronic (n=10)**	**p-value**
QuickDASH [median, IQR]	2 (0-16)	8 (2-20)	0.183
SSV [median, IQR]	90 (80-100)	90 (75-90)	0.244
NRS [median, IQR]	0 (0-2)	2 (1-3)	0.089
Return to work (weeks) [median, IQR]	8 (6-22)	8 (6-10)	0.456
	Conventional (n=12)	**Modified (n=5)**	
QuickDASH [median, IQR]	5 (1-19)	9 (7-18)	0.632
SSV [median, IQR]	90 (70-93)	90 (90-90)	0.548
NRS [median, IQR]	2 (0-3)	2 (1-2)	0.914
Return to work (weeks) [median, IQR]	8 (6-12)	8 (8-12)	0.668

n number; IQR interquartile range; SSV Subjective Shoulder Value; NRS Numerical Rating Scale pain score

work after a median of 8 weeks (IQR; 6−12). There was no significant difference in functional outcome scores between acute and chronic repair, or between the conventional and modified LARS fixation groups (Table 4).

Complications and hardware removal

There were two patients (10%) with major complications, both requiring revision surgery. One patient who was treated with conventional LARS fixation sustained a rupture of the LARS ligament nine weeks after fixation, revision surgery was performed with repeat LARS ligament fixation. One patient who was treated with the modified LARS fixation technique was re-admitted to the hospital with a deep wound infection requiring incision and drainage. Screw removal due to irritation was performed in one patient (5%).

Discussion

This retrospective study evaluated outcome after conventional and modified LARS ligament fixation of both acute and chronically repaired AC joint dislocations performed by a single surgeon. LARS ligament fixation for the treatment of AC dislocations resulted in good mid-term functional outcome.

These findings are in accordance with previous studies that evaluated LARS ligament fixation. Lu et al.[8] treated 24 patients with LARS artificial ligaments, they reported a mean Constant score of 94.5 (SD 9.3) after a follow-up of 36 months (range 6–60). Tiefenboeck et al.[15] presented the

results after a mean of 7.4 years of 47 patients treated with the LARS ligament. They reported good-to-excellent outcome in all patients, with a mean DASH score of 2.6 and Constant score of 93. Giannotti et al.[16] evaluated the use of the LARS artificial ligament in 17 patients, shoulder function was evaluated using the Constant score and Simple Shoulder test after a follow-up ranging from 1 to 41 months. They reported excellent results on both the Constant Score and the Simple Shoulder test for all 17 patients.

Giannotti et al.[16] reported one patients with radiographic enlargement of the clavicular screw tunnels, although, reduction was maintained. Lu et al.[8] concluded LARS fixation can provide immediate stability and allow early shoulder mobilization with good functional results and few complications. They performed follow-up radiographs showing slight loss of reduction in four patients, calcification of the CC ligament in four patients, degenerative changes around the AC joint in two patients and clavicular osteolysis around the screws in one patient. Tiefenboeck et al.[15] reported complications in five patients (11%), with four patients requiring revision surgery. Major complications occurred in three patients consisting of one loss of reduction and two cases of late infection after a mean of 18.6 months. Additionally, implant removal was performed in one patient due to screw pullout and irritation after 36 months. These findings are in line with our study with two patients (10%) requiring revision surgery, due to a ruptured LARS ligament and one case of deep infection. Screw removal due to irritation was performed in one patient. Further research is needed to evaluate the development of potential late complications following LARS fixation.

Previous case series have shown LARS fixation to be an effective fixation method for the treatment of AC joint dislocations. However, previous studies have mainly focused on acute AC dislocations. Lu et al.[8] and Tiefenboeck et al.[15] only treated patients with acute Rockwood type-III or higher AC dislocations. Giannotti et al.[16] evaluated the use of the LARS artificial ligament after both acute and chronic repair, however, they only treated Rockwood type-IV and V dislocations. The current study is the first to evaluate LARS fixation after both acute and chronically repaired for Rockwood type-III or higher AC dislocations. The results indicate LARS fixations to be an effective fixation method for the treatment of both acute and chronic AC joint dislocations. Therefore, LARS fixation is also an acceptable treatment option for active patients with symptomatic chronic Rockwood type-III or higher AC dislocations.

Early repair of AC dislocations has been reported to provide satisfactory results independent of surgical fixation method.[17] However, there is no consensus for the treatment of chronic AC dislocations.[17] Previous studies have shown the treatment of chronic AC dislocations to result is less favorable outcome and higher complication rates compared to acute repair.[18,19] Fraschini et al.[17] previously recommended LARS fixation for the treatment of chronic complete AC dislocations. Fraschini et al.[17] retrospectively compared outcome of 90 patients with chronic Rockwood type-III or higher AC dislocations, 30 patients treated with Dacron vascular prosthesis, 30 patients with LARS ligament and 30 patients with conservative treatment. Their results showed operative treatment resulted in significant better functional outcome compared to conservative treatment. However, treatment with Dacron vascular prosthesis resulted in a higher complications rate (43%) compared to LARS fixation (3%). In the current study, treatment with LARS fixation resulted in good functional outcome in both patients with acute and chronic AC joint dislocations. Unfortunately, accurate comparison of outcomes in the acute and chronic group are not possible due to the small sample size. Further research could focus on the effect early and delayed LARS fixation have on functional outcome following treatment of the AC dislocations.

To our knowledge, this study and the study by Marcheggiani Muccioli et al.[9] are the only two studies to report the use of a modified LARS fixation technique for the treatment of AC dislocations. The modified technique involves passing excess LARS ligament around the coracoid a second time. Thus, creating a figure of eight knot, which is then reinforced and secured with fiber wire. This modification adds an anterior translation force to the construct and increases the resistance to the opposing superior-inferior forces on the clavicle.[20] Following the use of the modified LARS fixation technique we did not encounter any cases of ligament rupture or loss of reduction.

This study has several limitations. First, this was a retrospective case series, outcome following LARS fixation was assessed in a relatively small number of patients without control group. However, the strength of this study is that all patients were treated in a single center by a single surgeon, which reduces the variability of the surgical skill and the clinical results. Second, functional outcome was only assessed by patient-reported outcome measures. However, both the QuickDASH, SSV and NRS are validated and reliable verbal outcome instruments, easily performed by telephone interview. Third, it was not possible to evaluate the degree of clinical

improvement, as preoperative functional outcome scores were not available. Unfortunately, the use of LARS fixation for the treatment AC joint dislocations has not been widely studied and comparison of literature remains difficult due to small sample sizes and inclusion of different types of AC joint dislocations. Therefore, a multicenter study might provide insight into the long-term results following LARS fixation of different Rockwood types, and in different patient populations.

Conclusion

The LARS ligament technique seems to be an effective and safe surgical fixation method for the treatment of AC joint dislocations, resulting in good short- and mid-term patient-reported functional outcome after a median follow-up of 23 months. In addition to acute AC dislocations, LARS fixation might also be an acceptable treatment option for active patients with symptomatic chronic Rockwood type-III or higher AC dislocations.

References:

1. Hindle P, Davidson EK, Biant LC, Court-Brown CM. Appendicular joint dislocations. *Injury.* 2013;44(8):1022-1027.
2. Mazzocca AD, Arciero RA, Bicos J. Evaluation and treatment of acromioclavicular joint injuries. *The American journal of sports medicine.* 2007;35(2):316-329.
3. Beitzel K, Cote MP, Apostolakos J, et al. Current concepts in the treatment of acromioclavicular joint dislocations. *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association.* 2013;29(2):387-397.
4. Chillemi C, Franceschini V, Dei Giudici L, et al. Epidemiology of isolated acromioclavicular joint dislocation. *Emergency medicine international.* 2013;2013:171609-171609.
5. Stucken C, Cohen SB. Management of acromioclavicular joint injuries. *The Orthopedic clinics of North America.* 2015;46(1):57-66.
6. Moatshe G, Kruckeberg BM, Chahla J, et al. Acromioclavicular and Coracoclavicular Ligament Reconstruction for Acromioclavicular Joint Instability: A Systematic Review of Clinical and Radiographic Outcomes. *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association.* 2018;34(6):1979-1995.e1978.
7. Li X, Ma R, Bedi A, Dines DM, Altchek DW, Dines JS. Management of acromioclavicular joint injuries. *The Journal of bone and joint surgery American volume.* 2014;96(1):73-84.
8. Lu N, Zhu L, Ye T, et al. Evaluation of the coracoclavicular reconstruction using LARS artificial ligament in acute acromioclavicular joint dislocation. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA.* 2014;22(9):2223-2227.
9. Marcheggiani Muccioli GM, Manning C, Wright P, Grassi A, Zaffagnini S, Funk L. Acromioclavicular joint reconstruction with the LARS ligament in professional versus non-professional athletes. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA.* 2016;24(6):1961-1967.
10. Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: comparison of three item-reduction approaches. *The Journal of bone and joint surgery American volume.* 2005;87(5):1038-1046.
11. Gummesson C, Ward MM, Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. *BMC musculoskeletal disorders.* 2006;7:44-44.
12. Fuchs B, Jost B, Gerber C. Posterior-inferior capsular shift for the treatment of recurrent, voluntary posterior subluxation of the shoulder. *The Journal of bone and joint surgery American volume.* 2000;82(1):16-25.
13. Gilbart MK, Gerber C. Comparison of the subjective shoulder value and the Constant score. *Journal of shoulder and elbow surgery.* 2007;16(6):717-721.
14. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain.* 2011;152(10):2399-2404.
15. Tiefenboeck TM, Boesmueller S, Popp D, et al. The use of the LARS system in the treatment of AC joint instability - Long-term results after a mean of 7.4 years. *Orthopaedics & traumatology, surgery & research : OTSR.* 2018;104(6):749-754.
16. Giannotti S, Dell'osso G, Bugelli G, Cazzella N, Guido G. Surgical treatment of acromioclavicular dislocation with LARS artificial ligament. *European journal of orthopaedic surgery & traumatology : orthopedie traumatologie.* 2013;23(8):873-876.
17. Fraschini G, Ciampi P, Scotti C, Ballis R, Peretti GM. Surgical treatment of chronic acromioclavicular dislocation: comparison between two surgical procedures for anatomic reconstruction. *Injury.* 2010;41(11):1103-1106.
18. Weinstein DM, McCann PD, McIlveen SJ, Flatow EL, Bigliani LU. Surgical treatment of complete acromioclavicular dislocations. *The American journal of sports medicine.* 1995;23(3):324-331.
19. Rolf O, Hann von Weyhern A, Ewers A, Boehm TD, Gohlke F. Acromioclavicular dislocation Rockwood III-V: results of early versus delayed surgical treatment. *Archives of orthopaedic and trauma surgery.* 2008;128(10):1153-1157.
20. Baxter JA, Phadnis J, Robinson PM, Funk L. Functional outcome of open acromioclavicular joint stabilization for instability following distal clavicle resection. *Journal of orthopaedics.* 2018;15(3):761-764.

CHAPTER 8

Surgical treatment of Neer type II and type V lateral clavicular fractures: comparison of hook plate versus superior plate with lateral extension: a retrospective cohort study

Yassine Ochen, Herman Frima, R. Marijn Houwert, Marilyn Heng, Mark van Heijl, Egbert J. M. M. Verleisdonk, Detlef van der Velde

(European Journal of Orthopaedic Surgery & Traumatology)

Abstract

Background

Different fixation methods are used for treatment of unstable lateral clavicle fractures (LCF). Definitive consensus and guidelines for the surgical fixation of LCF have not been established. The aim of this study was to compare patient-reported functional outcome after open reduction and internal fixation with the clavicle hook plate (CHP) and the superior clavicle plate with lateral extension (SCPLE).

Methods

A dual-center retrospective cohort study was performed. All patients operatively treated for unstable Neer type II and type V LCF between 2011 and 2016, with the CHP (n = 23) or SCPLE (n = 53), were eligible for inclusion. The primary outcome was the QuickDASH score. Secondary outcomes were the numerical rating scale (NRS) pain score, complications, and implant removal.

Results

A total of 67 patients (88%) were available for the final follow-up. There was a significant difference in bicortical lateral fragment size, 15 mm (± 4, range 6–21) in the CPH group compared to 20 mm (± 8, range 8–43) in the SCPLE group ($p \leq 0.001$). There was no significant difference in median QuickDASH score (CHP; 0.00 [IQR 0.0–0.0], SCPLE; 0.00 [IQR 0.0–4.5]; $p = 0.073$) or other functional outcome scores (NRS at rest; $p = 0.373$, NRS during activity; $p = 0.559$). There was no significant difference in median QuickDASH score or other functional outcome scores between Neer type II and type V fractures. There was no significant difference in complication rate, CHP 11% and SCPLE 8% (relative risk 1.26; [95% CI 0.25–6.33; $p = 0.777$]). The implant removal rate was 100% in the CHP group compared to 42% in the SCPLE group (relative risk 2.40; [95% CI 1.72–3.35; $p \leq 0.001$]).

Conclusion

Both the CHP and SCPLE are effective fixation methods for the treatment of unstable LCF, resulting in excellent patient-reported functional outcome and similar complication rates. SCPLE fixation is an effective fixation method for the treatment of both Neer type II and type V LCF. The SCPLE has a lower implant removal rate. Therefore, if technically feasible, we recommend SCPLE fixation for the treatment of unstable LCF.

Introduction

The fracture of the clavicle is frequently encountered in the emergency department, accounting for 2.6–4% of fractures in the adult population. Furthermore, clavicle fractures represent 35–44% of fractures in the shoulder region. Although the majority involve the midshaft, lateral fractures account for 10–30%.[1-6]

Lateral clavicle fractures (LCF) are classified according to Neer based on their relation to the coracoclavicular ligaments.[6,7] Neer types I, III and IV are considered to be stable fractures and are generally treated conservatively. The unstable Neer type II and V fractures account for approximately 10–52% of LCF. Surgical management is recommended for these unstable LCF, as nonoperative treatment results in a 22–50% nonunion rate.[1-6,8,9] Neer type II fractures are unstable due to the detachment of the coracoclavicular ligaments from the medial fragment. Neer type V fractures have a comminuted character, with only an inferior fragment remaining attached to the coracoclavicular ligament.[4,6,7] Fixation of LCF proves to be a challenge as it can be difficult to get a firm hold on small lateral fragments. In addition, opposing forces contribute to considerable displacement of the fracture ends. Therefore, LCF can usually only be stabilized by rigid fixation methods.[4,9] Different surgical fixation methods are available for the treatment of unstable LCF. However, at present, no consensus has been reached regarding the optimal fixation method.

The clavicle hook plate (CHP) is fixated with a small hook under the acromion posterior to the acromioclavicular joint. Complications related to the CHP such as acromial osteolysis, acromion fractures, rotator cuff tears and sub-acromial impingement have been reported.[4,5,10,11] The superior clavicle plate with lateral extension (SCPLE) is a more recently developed locking compression plate. The SCPLE has multiple locking screws on the lateral end, divergently configured to maximize screw purchase on LCF fragments. The SCPLE does not interfere with the acromioclavicular joint and has a relatively low-profile.[12-17] Previous case series have shown the SCPLE to be an effective fixation method for the treatment of unstable Neer type II fractures.[12-17] However, the results after SCPLE fixation of Neer type V fractures have not yet been studied.

Currently, both the CHP and SCPLE are being used for the treatment of LCF. However, definitive consensus and guidelines for the surgical fixation of LCF have not yet been established.

The aim of this study was to retrospectively evaluate patients treated with CHP and SCPLE fixation by comparing patient-reported functional outcome, complication-, and implant removal rates. Our hypothesis was that the SCPLE would result in better functional outcome and would lead to a reduction in complication- and implant removal rates.

Methods

Study design

A retrospective cohort study was performed using data from two level II trauma centers. All patients with an unstable LCF who were treated operatively between January 2011 and June 2016 were eligible for inclusion. Inclusion criteria were: (1) acute LCF, (2) age 18 years or older, (3) Neer type II or type V fracture, (4) fixation with CHP or SCPLE, (5) fixation within 2 weeks of injury, and (6) minimum of one-year follow-up. Exclusion criteria were: (1) history of prior shoulder injuries or (2) neurovascular disorders of the affected shoulder. Data collection was performed by reviewing electronic medical records, operative reports, radiology reports, and telephone interviews by an independent research fellow. Electronic medical records were reviewed to collect baseline characteristics regarding affected shoulder, age, gender, trauma date, trauma mechanism, time from injury to surgery, fixation method, previous shoulder injuries, and lateral fragment size. Lateral fragment size was measured in millimeters (mm) on the anterior–posterior view radiograph. Overall lateral fragment size was defined as the total length of the largest lateral fragment. The largest intact bicortical fragment, which would allow for adequate screw fixation, was considered as the bicortical lateral fragment length. Informed consent was obtained from all subjects, and approval was granted by the institutional review board.

Surgical procedure

Patients were treated by means of open reduction and internal fixation (ORIF) using a CHP (3.5 mm LCP; Depuy Synthes GmbH, Oberdorf, Switzerland) or SCPLE (3.5/2.7 mm LCP; Depuy Synthes GmbH, Oberdorf, Switzerland). Implant selection was based on the surgeon's preference. CHP and SCPLE fixation were performed by several surgeons in both trauma centers. Operations were performed under general anesthesia with the patient placed in a beach chair position. An incision was made using a standard superior approach. The fracture site was exposed preserving as much periosteum as possible. Reduction was performed under direct visualization, and fragments were temporarily fixated using K-wires or reduction forceps. Fracture reduction, implant position, and screw placement were checked under fluoroscopic

guidance. Coracoclavicular ligament repair was not routinely performed. Finally, the fascia and skin were closed in layers.

Clavicle hook plate

In cases of CHP fixation, a small incision was made in the posterior capsule of the acromioclavicular joint to allow sub-acromial hook placement. Trial plates were used to determine correct length and depth. Definitive CHP fixation was completed with the insertion of 3.5 mm angular stable or conventional screws (Figure. 1).

Superior clavicle plate with lateral extension

In cases of SCPLE fixation, there was no involvement of the acromioclavicular joint. A plate with an appropriate length was chosen to allow adequate fixation with 3.5 mm conventional or angular stable screws in the medial fragment and smaller 2.7 mm angular stable screws in the lateral end (Figure. 2).

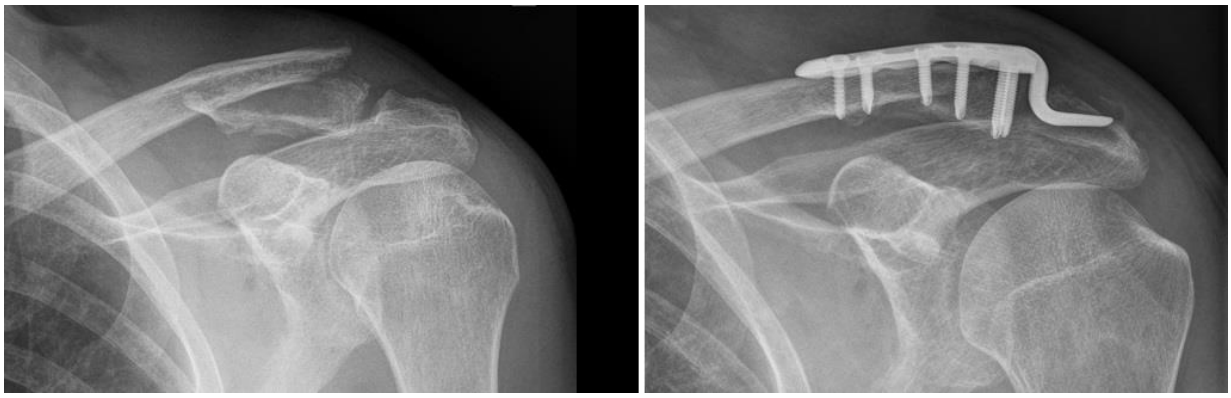

Figure 1. Preoperative radiograph of LCF and postoperative radiograph after CHP fixation.

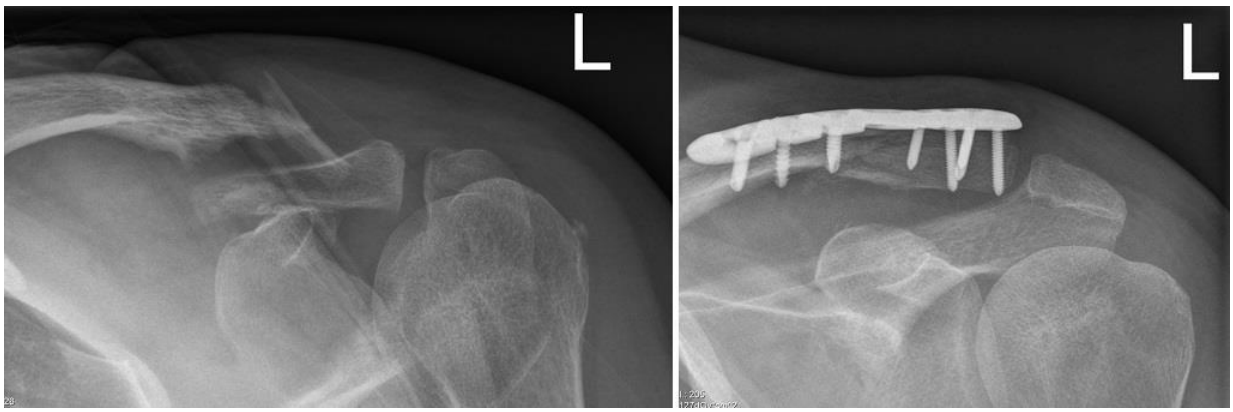

Figure 2. Preoperative radiograph of LCF and postoperative radiograph after SCPLE fixation.

Postoperative management

Both groups received the same postoperative management. Radiographs were taken 1 day postoperatively. Patients were temporarily immobilized in a sling until the pain subsided; early mobilization and active range of motion exercises were allowed when tolerated. Weight-bearing activities and resisted exercises were not permitted until approval from the treating surgeon. Follow-up visits were scheduled at 2, 4, and 12 weeks postoperatively. Additional outpatient visits were scheduled depending on fracture consolidation. Removal of the SCPLE was not routinely performed, as opposed to the CHP where removal was recommended to all patients.

Primary outcome

Functional outcome was assessed at least 12 months following ORIF, using the Dutch language version of the QuickDASH score. The QuickDASH is a validated and shortened version of the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH). The QuickDASH is a patient-reported outcome instrument developed to measure upper extremity disability and symptoms, resulting in a score ranging from 0 (no disability) to 100 (most severe disability).[18,19]

Secondary outcome

Secondary outcomes were the numerical rating scale (NRS) pain score at rest and during activity, complications, revision surgery and implant removal. The NRS is a reliable and commonly used 11-point scale to measure pain intensity, ranging from 0 (no pain) to 10 (worst imaginable pain).[20] Complications included infection, nonunion, mal-union, implant failure, and implant removal-related complications. Infections were subdivided in superficial-skin or deep-wound infection. Superficial infection was defined as redness, swelling, or purulent discharge from the wound that was treated with antibiotics alone. If surgical irrigation and debridement was required, it was considered a deep infection. Nonunion was defined as the absence of fracture consolidation 6 months after surgery. Malunion was defined as a symptomatic deformity of the clavicle. Implant failure was defined as implant displacement, implant breakage, or breakage of screws. Revision surgery was defined as the need for subsequent surgery other than implant removal. Infection and re-fracture following implant removal were considered implant removal-related complications. Implant-related irritation and indication for implant removal were analyzed using a series of questions developed by Hulsmans et al.[21] Responses to these questions allowed categorization of implant removal into (1) routinely or on patient's request without irritation or (2) patient's request due to irritation. Patients with the implant still in situ received a different

series of questions, leading to categorization of why implant was not removed; (1) not experiencing irritation, (2) experiencing irritation but removal not necessary, (3) experiencing irritation but no request for removal due to fear of re-operation, or (4) experiencing irritation, considering removal.

Statistical analysis

Descriptive results are presented as mean values with standard deviations and range (SD, range) and median values with interquartile range (IQR) or absolute numbers and percentages (%). Continuous variables were evaluated using an independent sample t test or Mann–Whitney U test. Categorical variables were compared using the Pearson's Chi-squared test. The Fisher's exact test was used in case of small count sizes. Mean differences and relative risks (RR) were calculated with 95% confidence intervals (CIs). The significance level was defined as a p value < 0.05. All statistical analyses were performed using IBM SPSS Statistics version 24 for Windows (IBM Corp, Armonk, NY).

Results

Study population

A flowchart of the patient cohort is shown in Fig. 3. In total, 76 patients met the inclusion criteria. However, eight patients could not be contacted, and one patient refused participation. This resulted in the inclusion of 67 patients (88%) for analysis. The baseline characteristics are shown in Table 1. The CHP group included 19 patients (28%) compared to 48 patients (72%) in the SCPLE group. The most frequent fracture pattern was Neer type II found in 43 patients (64%). The overall lateral fragment size was 39 mm (SD 12, range 14–83). There was a significant difference in bicortical lateral fragment size, 15 mm (SD 4, range 6–21) in the CPH group compared to 20 mm (SD 8, range 8–43) in the SCPLE group ($p < 0.001$). The mean time from injury to surgery was 6.9 days (SD 3.6, range 0–14). The mean follow-up was 37.5 months (SD 17.9, range 12–76).

Functional outcome

There was no significant difference in functional outcome, as shown in Table 2. The median QuickDASH score in the CHP group was 0.00 (IQR; 0.0–0.0), as opposed to 0.00 (IQR; 0.0–4.5) in the SCPLE group ($p = 0.073$). There were 15 patients (79%) with a QuickDASH score of 0 in the CHP group (range 0–21) compared to 25 patients (52%) in the SCPLE group (range 0–23).

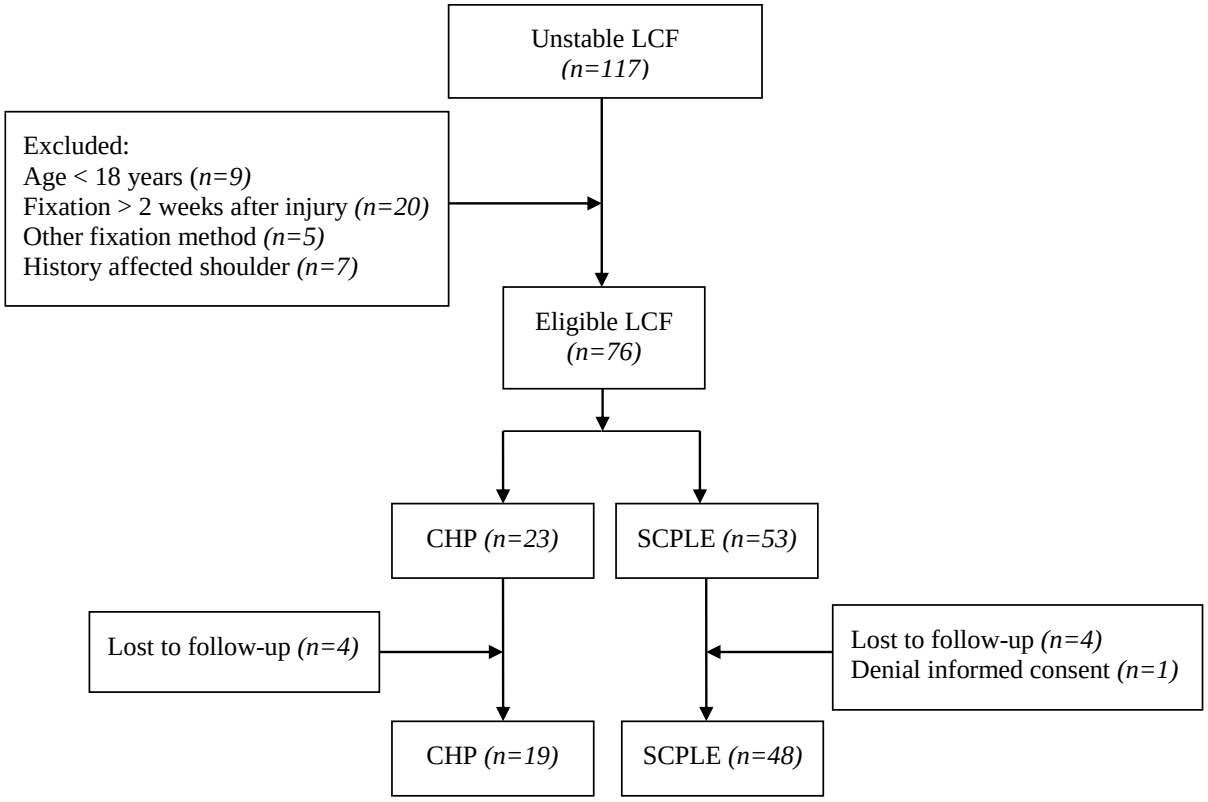

Figure 3. Flowchart representing patient selection for analysis of CHP versus SCPLE for unstable LCF.

The median NRS pain score at rest was 0.00 (IQR; 0.0–0.0) in the CHP group and 0.00 (IQR; 0.0–0.0) in the SCPLE group (p = 0.373). There were 16 patients (84%) with a NRS pain score at rest of 0 in the CHP group (range 0–6) compared to 44 patients (94%) in the SCPLE group (range 0–3). In the CHP group, the median NRS pain score during activity was 0.00 (IQR; 0.0–1.0) compared to 0.00 (IQR; 0.0–2.0) in the SCPLE group (p = 0.559). There were 14 patients (74%) with a NRS pain score during activity of 0 in the CHP group (range 0–8) compared to 30 patients (63%) in the SCPLE group (range 0–7).

Functional outcome according to Neer type

In both treatment groups, there was no significant difference in median QuickDASH score or other functional outcome scores between the Neer type II and type V fractures (Table 3). The median QuickDASH score in the Neer type II group following CHP fixation was 0.00 (IQR; 0.0–2.3), as opposed to 0.00 (IQR; 0.0–0.6) in the Neer type V group (p = 0.623). In the SCPLE group, the median QuickDASH score in the Neer type II group was 0.00 (IQR; 0.0–5.1), as opposed to 2.30 (IQR; 0.0–4.5) in the Neer type V group (p = 0.764).

Table 1. Baseline characteristics

	Overall n (%)*	CHP n (%)*	SCPLE n (%)*	95% CI difference	p-value
Patients	67	19	48		
Age [mean, SD]	43 (14)	42 (17)	43 (12)	-8.29-6.54	0.814
Gender					
Male	54 (81)	13 (68)	41 (85)		0.169
Female	13 (19)	6 (32)	7 (15)		
Side injury					
Left	39 (58)	8 (42)	31 (65)		0.108
Right	28 (42)	11 (58)	17 (35)		
Affected side dominant side					
Yes	27 (40)	10 (53)	17 (35)		0.270
No	40 (60)	9 (47)	31 (65)		
Neer classification					
Type II	43 (64)	13 (68)	30 (63)		0.780
Type V	24 (36)	6 (32)	18 (38)		
Overall lateral fragment (mm) [mean, SD]	39 (12)	37 (12)	40 (12)	-9.39-3.55	0.371
Bicortical lateral fragment (mm) [mean, SD]	19 (7)	15 (4)	20 (8)	-8.40--2.64	<0.001
Time injury to surgery (days) [mean, SD]	6.9 (3.6)	7.5 (3.5)	6.7 (3.6)	-1.15-2.72	0.419
Follow-up (months) [mean, SD]	37.5 (17.9)	31.3 (16.3)	40.0 (18.0)	-18.25-0.77	0.071

* Percentages may not add up to 100 due to rounding.

Table 2. Functional outcome and implant-related complications

	CHP (n=19) n (%)*	SCPLE (n=48) n (%)*	Relative risk (95% CI)	p-value
QuickDASH median [IQR]	0.00 (0.0-0.0)	0.00 (0.0-4.5)		0.073
QuickDASH distribution [range]	0–21	0–23		
0	15 (79)	25 (52)		
0–10	3 (16)	19 (40)		
10–20	0	3 (6)		
20–25	1 (5)	1 (2)		
NRS pain at rest [median, IQR]	0.00 (0.0-0.0)	0.00 (0.0-0.0)		0.373
NRS pain at rest distribution [range]	0–6	0–3		
0	16 (84)	44 (92)		
0–3	2 (11)	3 (6)		
3–6	1 (5)	1 (2)		
NRS pain during activity [median, IQR]	0.00 (0.0-1.0)	0.00 (0.0-2.0)		0.559
NRS pain during activity distribution [range]	0–8	0–7		
0	14 (74)	30 (63)		
0–3	1 (5)	7 (15)		
3–6	2 (11)	8 (17)		
6–8	2 (11)	3 (6)		
Complications	2 (11)	4 (8)	1.26 (0.25-6.33)	0.777
Complication classification				0.929
Implant failure	1 (5)	3 (6)		
Nonunion	1 (5)	1 (2)		
Revision surgery	1 (5)	2 (5)	1.26 (0.12-13.13)	0.999

* Percentages may not add up to 100 due to rounding. QuickDASH score: 0=no disability to 100=most severe disability. NRS pain score: 0=no pain to 10=worst imaginable pain.

Implant removal

Implant removal rates and indications are presented in Table 4. CHP fixation was associated with a significant higher removal rate. CHP removal was, according to protocol, performed in all 19 patients (100%) compared to 20 patients (42%) in the SCPLE group (relative risk 2.40; 95% CI 1.72–3.35; $p < 0.001$). The mean time to removal was 4.3 months (SD 2.2, range 2–10) and 13.6 months (SD 11.5, range 5–50) in the CHP and SCPLE groups, respectively (mean difference

Table 3. Functional outcome according to Neer classification

Neer	Type II	Type V	p-value
CHP n (%)*	13 (68)	6 (32)	
QuickDASH median [IQR]	0.00 (0.0-2.3)	0.00 (0.0-0.6)	0.623
NRS pain score at rest [median, IQR]	0.00 (0.0-0.0)	0.00 (0.0-0.0)	1.000
NRS pain score during activity [median, IQR]	0.00 (0.0-2.0)	0.00 (0.0-2.0)	0.734
SCPLE n (%)*	30 (63)	18 (38)	
QuickDASH median [IQR]	0.00 (0.0-5.1)	2.30 (0.0-4.5)	0.764
NRS pain score at rest [median, IQR]	0.00 (0.0-0.0)	0.00 (0.0-0.0)	0.609
NRS pain score during activity [median, IQR]	0.00 (0.0-3.3)	0.00 (0.0-1.0)	0.999

* Percentages may not add up to 100 due to rounding. QuickDASH score: 0=no disability to 100=most severe disability. NRS pain score: 0=no pain to 10=worst imaginable pain.

Table 4. Implant removal rate and indication

	CHP (n=19) n (%)	SCPLE (n=48) n (%)	Mean difference (95% CI)	Relative risk (95% CI)	p-value
Implant removal	19 (100)	20 (42)		2.40 (1.72-3.35)	<0.01
Reason implant removed					0.695
Routinely/ patient's request, no irritation	3 (16)	5 (25)		0.63 (0.17-2.29)	
Due to irritation	16 (84)	15 (75)		1.23 (0.81-1.55)	
Time to implant removal (months) [mean, SD]	4.3 (2.2)	13.6 (11.5)	-9.3 (-14.76-3.82)		0.002
Status implant not removed					NP
Not experiencing irritation	0	12 (43)			
Irritation, but removal not necessary	0	6 (21)			
Irritation, no removal, fear re-operation	0	5 (18)			
Irritation, considering removal	0	5 (18)			

NP statistical analyses Not Possible because all CHP implants were removed.

− 9.287; 95% CI − 14.757 to 3.817; p = 0.002). In the CHP group, three patients (16%) reported removal without irritation and 16 patients (84%) reported removal due to irritation. There were no cases of implant removal-related complications. In the SCPLE group, 28 patients (58%) did not have the implant removed and 12 patients (43%) reported not to experience irritation.

Complications

Complications were reported in two patients (11%) in the CHP group compared to four patients (8%) in the SCPLE group (relative risk 1.26; 95% CI 0.25–6.33; p = 0.777) (Table 2). Complications in the CHP group consisted of one case of implant failure due to implant displacement and one case of nonunion. Complications in the SCPLE group included three cases of implant failure and one case of nonunion. The implant failures in SCPLE group consisted of two implant displacements and one case of screw breakage. No cases of infection or mal union were observed. In total, there were three patients that needed revision surgery. In the CHP group, one patient received a lateral clavicle resection due to nonunion. Two revision surgeries were performed in the SCPLE group, one due to severe implant displacement and one case of nonunion. The SCPLE implant displacement was treated by repeat SCPLE fixation. The nonunion was treated with temporary K-wires fixation for 9.5 months.

Discussion

There was no significant difference in patient-reported functional outcome or complication rate between CHP and SCPLE fixation. However, the CHP was used more often on fractures with a small lateral bicortical fragment. There was no significant difference in patient-reported functional outcome between Neer type II and type V LCF fractures. Furthermore, there was a significant higher implant removal rate in the CHP group. In the SCPLE group, 57% of patients with the implant still in situ reported varying degrees of implant-related irritation.

Both the SCPLE and CHP result in excellent functional outcome. These findings are in accordance with previous comparative studies. Zhang et al.[22] compared functional outcome of 36 patients with the SCPLE implant to 30 patients with the CHP using the Constant–Murley score and demonstrated no significant difference between groups. Erdle et al.[23] compared the results of 19 patients with CHP and 13 patients with SCPLE fixation, and they reported no significant difference between the groups when using the Constant score, the Oxford shoulder score, and the subjective shoulder value.

In the current study, the bicortical lateral fragment size was significantly smaller in the CHP group. Erdle et al.[23] reported no significant difference in lateral fragment size; however, they did not report whether the intact lateral fragment was bicortical. In the current study, the largest intact bicortical lateral fragment size which would allow for adequate screw fixation was measured. Our results indicate implant selection was influenced by the bicortical lateral fragment size. We recommend further research to focus on lateral fragment size to determine whether lateral fragment size negatively affects functional outcome and complication rates with the use of different implants.

Previous case series have shown the SCPLE to be an effective fixation method for the treatment of unstable Neer type II fractures.[12-17] Zhang et al.[22] treated fractures with a lateral fragment size larger than 2 cm with the SCPLE, and comminuted fractures close to the acromioclavicular joint were treated with the CHP with additional ligament repair. The comparative study by Erdle et al.[23] only included Neer type IIb fractures. To our knowledge this is the first study to evaluate the use of SCPLE fixation for the treatment of Neer type V fractures. In the current study, treatment with SCPLE fixation resulted in good functional outcome in both 30 patients (63%) with Neer type II and 18 patients (38%) with Neer type V fractures. These findings indicate SCPLE fixation

is also an acceptable treatment option for acute Neer type V fractures, despite their comminuted character.

There was no a significant difference in complication rate between CHP and SCPLE fixation, which is in contrast to previous comparative studies. Zhang et al.[22] found a significantly higher complication rate, 23.3% in the CHP group compared to 5.6% in the SCPLE group ($p = 0.04$). However, Zhang et al.[22] included symptomatic hardware as a complication, and they reported three cases (10%) of symptomatic hardware in the CHP group and none in the SCPLE group. Erdle et al.[23] also reported a significantly higher overall prevalence of complications in the CHP cohort (89%) compared to the SCPLE cohort (38%) ($p = 0.014$). Erdle et al.[23] included radiographical proof of persistent acromial osteolysis and posttraumatic acromioclavicular joint arthrosis as complications.

The previous comparative studies included complications such as acromial osteolysis, posttraumatic acromioclavicular joint arthrosis, and sub-acromial impingement syndrome. These complications could be regarded as CHP implant specific. The CHP is fixated with a small hook under the acromion, posterior to the acromioclavicular joint which acts as a lever and maintains fracture reduction. However, this mechanism not only limits abduction of the arm, it may also affect the acromion and induce discomfort. The SCPLE does not interfere with the acromioclavicular joint, which results in the absence of acromial and impingement complications. Furthermore, there are several reports that indicated that these CHP implant-specific complications can resolve after removal.[11,24] Renger et al.[11] evaluated the use of the CHP in 44 patients, and 30 patients (68%) reported implant-related discomfort. Renger et al.[11] found all implant-related complaints and osteolytic defects to disappear after implant removal.

Implant-related irritation and implant removal were analyzed using the series of questions developed by Hulsmans et al.[21] In the current study, all CHP implants were removed after a mean of 4.3 months, in line with previous studies recommending CHP removal after fracture consolidation.[11] The comparative study by Zhang et al.[22] reported all CHPs were removed compared to 12 SCPLEs (33%). Erdle et al.[23] reported CHP removal was recommended and all CHP implants were removed after a mean period of 4.7 months. In the Erdle et al.[23] study, 77% of SCPLE implants were removed after a mean period of 12.5 months due to local irritation or on patient's explicit request. In the current study, after a minimum of 12 months following

ORIF, 42% of SCPLE implants were removed. Moreover, 43% of the patients with the SCPLE still in situ reported not to experience any irritation.

This study has some limitations. First, the study is limited by the retrospective nature. This study did not include prospective collection of functional and radiological measures during different follow-up times, which would increase the understanding of the impact implants have prior to implant removal. Second, fixation method was based on surgeon's preference, which could cause bias through selection-by-indication. Therefore, different measurements were performed to determine whether lateral fragment size influenced implant selection. Finally, our study is limited by the small number of included patients in the treatment groups. However, this number is in accordance with previous comparative studies. Unfortunately, results after the use of CHP and SCPLE fixation have not yet been widely studied.

To our knowledge, this is the first study to evaluate the use of the CHP and SCPLE, focusing solely on implant selection without major differences in surgical technique or ligament repair. Furthermore, this is the first study to present the results of SCPLE fixation for the treatment of both Neer type II and type V fractures. Unfortunately, comparison of literature remains difficult due to small sample sizes, wide variety of functional outcome scores, definitions and surgical techniques. Therefore, a large multicenter study might provide insight into long-term results following different treatment modalities, influence of different LCF fractures types, and different patient populations.

Conclusion

Both the CHP and SCPLE are effective fixation methods for the treatment of unstable LCF resulting in excellent patient-reported functional outcome and similar complication rates. SCPLE fixation is an effective surgical fixation method for the treatment of both Neer type II and type V LCF. The SCPLE has a lower implant removal rate compared to the CHP. Therefore, if technically feasible, we recommend SCPLE fixation for the treatment of unstable LCF.

References:

1. Nordqvist A, Petersson C. The incidence of fractures of the clavicle. *Clinical orthopaedics and related research.* 1994(300):127-132.
2. Robinson CM. Fractures of the clavicle in the adult. Epidemiology and classification. *The Journal of bone and joint surgery British volume.* 1998;80(3):476-484.
3. Postacchini F, Gumina S, De Santis P, Albo F. Epidemiology of clavicle fractures. *Journal of shoulder and elbow surgery.* 2002;11(5):452-456.
4. Sambandam B, Gupta R, Kumar S, Maini L. Fracture of distal end clavicle: A review. *Journal of clinical orthopaedics and trauma.* 2014;5(2):65-73.
5. Oh JH, Kim SH, Lee JH, Shin SH, Gong HS. Treatment of distal clavicle fracture: a systematic review of treatment modalities in 425 fractures. *Archives of orthopaedic and trauma surgery.* 2011;131(4):525-533.
6. Neer CS, 2nd. Fractures of the distal third of the clavicle. *Clinical orthopaedics and related research.* 1968;58:43-50.
7. Neer CS, 2nd. Fracture of the distal clavicle with detachment of the coracoclavicular ligaments in adults. *The Journal of trauma.* 1963;3:99-110.
8. Rokito AS, Zuckerman JD, Shaari JM, Eisenberg DP, Cuomo F, Gallagher MA. A comparison of nonoperative and operative treatment of type II distal clavicle fractures. *Bull Hosp Jt Dis.* 2002;61(1-2):32-39.
9. Ropars M, Thomazeau H, Huten D. Clavicle fractures. *Orthopaedics & traumatology, surgery & research : OTSR.* 2017;103(1S):S53-S59.
10. Muramatsu K, Shigetomi M, Matsunaga T, Murata Y, Taguchi T. Use of the AO hook-plate for treatment of unstable fractures of the distal clavicle. *Archives of orthopaedic and trauma surgery.* 2007;127(3):191-194.
11. Renger RJ, Roukema GR, Reurings JC, Raams PM, Font J, Verleisdonk EJ. The clavicle hook plate for Neer type II lateral clavicle fractures. *J Orthop Trauma.* 2009;23(8):570-574.
12. Lee SK, Lee JW, Song DG, Choy WS. Precontoured locking plate fixation for displaced lateral clavicle fractures. *Orthopedics.* 2013;36(6):801-807.
13. Andersen JR, Willis MP, Nelson R, Mighell MA. Precontoured superior locked plating of distal clavicle fractures: a new strategy. *Clin Orthop Relat Res.* 2011;469(12):3344-3350.
14. Tiren D, Vroemen JP. Superior clavicle plate with lateral extension for displaced lateral clavicle fractures: a prospective study. *J Orthop Traumatol.* 2013;14(2):115-120.
15. Beirer M, Siebenlist S, Cronlein M, et al. Clinical and radiological outcome following treatment of displaced lateral clavicle fractures using a locking compression plate with lateral extension: a prospective study. *BMC Musculoskelet Disord.* 2014;15:380-380.
16. Fleming MA, Dachs R, Maqungo S, du Plessis JP, Vrettos BC, Roche SJ. Angular stable fixation of displaced distal-third clavicle fractures with superior precontoured locking plates. *J Shoulder Elbow Surg.* 2015;24(5):700-704.
17. Vaishya R, Vijay V, Khanna V. Outcome of distal end clavicle fractures treated with locking plates. *Chinese journal of traumatology = Zhonghua chuang shang za zhi.* 2017;20(1):45-48.
18. Gummesson C, Ward MM, Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. *BMC musculoskeletal disorders.* 2006;7:44-44.
19. Beaton DE, Wright JG, Katz JN. Development of the QuickDASH: comparison of three item-reduction approaches. *The Journal of bone and joint surgery American volume.* 2005;87(5):1038-1046.
20. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain.* 2011;152(10):2399-2404.
21. Hulsmans MHJ, van Heijl M, Houwert RM, et al. High Irritation and Removal Rates After Plate or Nail Fixation in Patients With Displaced Midshaft Clavicle Fractures. *Clinical orthopaedics and related research.* 2017;475(2):532-539.
22. Zhang C, Huang J, Luo Y, Sun H. Comparison of the efficacy of a distal clavicular locking plate versus a clavicular hook plate in the treatment of unstable distal clavicle fractures and a systematic literature review. *International orthopaedics.* 2014;38(7):1461-1468.
23. Erdle B, Izadpanah K, Jaeger M, et al. Comparative analysis of locking plate versus hook plate osteosynthesis of Neer type IIB lateral clavicle fractures. *Archives of orthopaedic and trauma surgery.* 2017;137(5):651-662.
24. Tiren D, van Bemmel AJ, Swank DJ, van der Linden FM. Hook plate fixation of acute displaced lateral clavicle fractures: mid-term results and a brief literature overview. *Journal of orthopaedic surgery and research.* 2012;7:2-2.

CHAPTER 9

Patient-reported and clinical outcomes after proximal humeral fractures

Yassine Ochen, Abhiram R. Bhashyam, Auréline N.A.F. Boot,
Detlef van der Velde, R. Marijn Houwert, Rolf H.H. Groenwold, Marilyn Heng

(Submitted)

Abstract

Background

The aim of this study was to assess Patient-Reported Outcomes Measurement Information System (PROMIS) scores after nonoperative and operative treatment of proximal humeral fractures. Second, to assess the correlation between the PROMIS physical function (PF), PROMIS Upper Extremity (UE), and QuickDASH (the shortened version of the Disabilities of the Arm, Shoulder, and Hand). Third, the association between objective clinical outcome and Patient-reported outcome measures (PROMs) was assessed.

Methods

A retrospective cohort study was performed using data from two American College of Surgeons (ACS) level I trauma centers. All adult patients with proximal humeral fractures who were admitted between January 2016 and March 2018 with a minimum of 3 months clinical or functional follow-up were eligible for inclusion. Functional outcome was assessed using the PROMIS PF, PROMIS UE and/or the QuickDASH. Clinical outcome measures included occurrence of complications, need for subsequent surgery, and degree of flexion and abduction.

Results

In total, 249 patients were included. The mean age was 65.53 years (SD 14.98) and 183 patients (73.5%) were female. Operative treatment was performed in 92 patients (37%). There were no differences in PROMs between nonoperative and operative treated patients, PROMIS PF (adjusted mean difference -1.50; 95% CI -5.20 – 2.19; p= 0.426), PROMIS UE (adjusted mean difference -0.27; 95% CI -6.66; 6.11; p= 0.933), and QuickDASH (adjusted mean difference -2.33; 95% CI -6.72; 2.05; p= 0.299). Also, there was no difference in complication rate (9.6% versus 12%, p= 0.701) and subsequent surgery rate (6.4% versus 12%, p= 0.195). PROMIS PF was associated with PROMIS UE (correlation (r) 0.83; 95% CI 0.76; 0.88; p <0.001), and QuickDASH (r -0.47; 95% CI -0.62; -0.29; p<0.001). The PROMIS UE was also associated with the QuickDASH (r -0.70; 95% CI -0.81; -0.55; p<0.001). The degree of flexion was associated with all PROMs, PROMIS PF (r 0.43; 95% CI 0.26; 0.58; p <0.001), PROMIS UE (r 0.27; 95% CI 0.04; 0.47; p= 0.019), and QuickDASH (r -0.25.; 95% CI -0.44; -0.04; p= 0.021).

Conclusion

There were no significant differences in outcomes between operative and nonoperative treatment of proximal humeral fractures. The PROMIS PF, PROMIS UE, and QuickDASH score showed a high mutual correlation, as well as a moderate correlation with flexion degree. The correlations found in this study suggest the PROMIS PF and PROMIS UE can be considered similarly useful as a measure for evaluating shoulder function after proximal humeral fractures.

Introduction

The fracture of the proximal humerus is a frequently encountered injury, accounting for 5.7% of all fractures. Proximal humeral fractures are the seventh most common fracture type in adults and the third most common fracture type in the elderly population.[1,2] The management of acute proximal fractures consists of operative or nonoperative treatment.[3] The optimal management of these fractures remains a significant challenge and is much debated.[4] Several meta-analyses have been published on the comparison between operative or nonoperative treatment; however, no consensus has been reached regarding the optimal treatment.[5,6]

The lack of treatment consensus is partly due to the wide range of outcome measures presented in different studies.[4] There is no agreement on the optimal outcome measures for the evaluation of proximal humeral fracture treatment, which has made it difficult to compare outcomes across studies. To accurately evaluate patients after treatment, the chosen outcome measures need to have good measurement properties (validity, reliability, and responsiveness).[4] These properties may differ across setting and patient populations. It is therefore important to evaluate different available patient-reported outcome measures (PROMs) in specific patient populations.[4]

To counter these challenges, the Patient-Reported Outcomes Measurement Information System (PROMIS) questionnaires were developed. In the field of orthopedics, the most common used PROMIS is the PROMIS physical function (PF) score. There has been considerable research that has focused on the performance of the PROMIS compared to various legacy "traditional" outcome measures across various conditions. These studies showed that PROMIS correlates well with legacy outcome measures frequently used in orthopedic outcome studies.[7-13] One of those legacy outcome measures is the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire which is used to evaluate various upper extremity disorders and is, although not developed specifically for proximal humerus fractures, the most commonly used PROM for these fractures.[4] However, despite these early promising results, the widespread use of PROMIS has not been adopted in most orthopedic literature. Furthermore, there is a lack of evidence on PROMIS specifically for orthopedic trauma.[11] There is one previous study that has evaluated the correlation between the PROMIS PF and DASH scores in the proximal humeral fracture population, which has shown encouraging results regarding correlation.[13] However, the more recently developed PROMIS Upper Extremity (UE), the PROMIS PF, and the QuickDASH (shortened version of the DASH) have not been specifically assessed in the proximal humeral

fracture population. Furthermore, discordance between objective clinical outcomes and patient-reported outcome measures (PROMs) is a common phenomenon for several UE injuries.[14] Although several studies have evaluated the association of different PROMs of proximal humeral fractures, fewer studies have explored the association between clinical outcome and PROMs.

The aim of this study was to establish benchmark PROMIS data after nonoperative and operative treatment of proximal humeral fractures. Secondly, this study sought to evaluate the correlation between the PROMIS PF, PROMIS UE, and QuickDASH scores. Finally, we aimed to explore the association between objectively measured clinical outcomes and PROMs.

Methods

Study design

A retrospective cohort study was performed using data from two American College of Surgeons (ACS) level I trauma centers. All adult patients with proximal humeral fractures who were treated with between January 2016 and March 2018 were eligible for inclusion. Eligible patients were identified by searching for Current Procedure Terminology (CPT) codes and International Classification of Diseases (ICD) codes. Inclusion criteria were: (1) age 18 years or older, proximal humeral fracture, and (3) minimum of 3 months clinical or functional follow-up. Exclusion criteria were: (1) treatment for fracture at an outside facility, (2) pathologic fracture, or (3) periprosthetic fracture. Baseline demographic characteristics, clinical and functional data were identified using the institutions' detailed Enterprise Data Warehouse and electronic medical records. Informed consent was obtained from all subjects and approval was granted by the Institutional Review Board.

Patient and treatment characteristics

Electronic medical records were reviewed to collect baseline demographic characteristics regarding age, sex, trauma mechanism, open fracture, AO-classification, number of fractures, number of UE fractures, ASA classification, hospital length of stay (HLOS), operative treatment method, and nonoperative treatment method. Injury mechanisms were further subdivided into high-energy trauma (HET) and classified according to the Advanced Trauma Life Support guidelines. HET mechanisms were defined as falls from height, crush injuries, motor vehicle and motorcycle accidents.[15,16] The AO classification was confirmed by two orthopedic surgeons in the treating institutions. Nonoperative treatment methods included closed reduction and sling

treatment. Operative treatment included Open Reduction and Internal Fixation (ORIF), hemiarthroplasty, and reversed shoulder arthroplasty.

PROMs

The collection of PROMs was standardized in both trauma centers since January 2016. Physical function and upper extremity disability were evaluated using the PROMIS physical function (PF) 10a, PROMIS Upper Extremity (UE) 16a, and/or the QuickDASH. Patients completed the questionnaires on a tablet device as part of their routine follow-up visit at the trauma centers. The PROMIS questionnaire assess limitation and difficulty with certain physical activities with scores ranging from 0 to 100, with higher scores representing higher physical function, and a mean score of 50 for the general population of the United States.[7] The QuickDASH is a validated and shortened version of the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH). The QuickDASH is developed to measure UE disability and symptoms, resulting in a score ranging from 0 (no disability) to 100 (most severe disability), and a mean of 10 points reflecting the general US population average.[17] The time from treatment to questionnaire in months was available for all PROMs.

Clinical outcomes

Clinical outcome measures included occurrence of complications, need for subsequent surgery, flexion and abduction degree during clinical follow-up. Complications included nonunion, malunion, arthrosis, infection, dislocation, implant failure, and avascular necrosis. Subsequent surgery was defined as the need for additional surgery other than implant removal. Flexion and abduction were measured in degrees and were assessed by the treating physician during the last outpatient follow-up visit. The time from treatment to last clinical follow-up in months was available for all patients.

Statistical analyses

Descriptive results are presented as mean values with standard deviations (SD) or absolute numbers and percentages (%). The relation between the treatment method (operative or nonoperative) and PROMs was assessed using linear regression analysis. Regression analysis was performed with and without adjustment for baseline information and presented as crude and adjusted mean difference with 95% confidence interval (CI). Adjustment was made for baseline and clinical factors that might confound the relationship between treatment method and

functional outcome including the time from treatment to questionnaire, age, AO classification, and occurrence of multiple UE fractures. Categorical outcomes were compared using the Pearson's chi-squared test. Pearson's correlation (r) was used to assess the relationship between PROMIS PF, PROMIS UE, and QuickDASH to validate our data against previous studies.[7-11] Furthermore, Pearson's correlation (r) was used to assess the relationship between the two physical exam measurements, flexion and abduction degree, and the different PROMs. The significance level was defined as a p value <0.05. All analyses were performed in R version 3.6.1 (R Development Core Team, Released 2013, Vienna, Austria: R Foundation for Statistical Computing).[18]

Results

Study population

In total, 249 patients met the inclusion criteria. The baseline characteristics stratified by treatment method are shown in Table 1. The mean age was 65.53 years (SD 14.98) and 183 patients (73.5%) were female. In most cases, patients sustained a proximal humeral fracture during a low-energy fall (83.9%). Fourteen patients (5.6%) were involved in a HET. Open fractures were sustained by 3 patients (1.2%). The most common AO classification was type B1 (35.7%). There were 6 patients (2.3%) that sustained multiple UE fractures. Nonoperative treatment was performed in 157 patients (63%), with the majority treated with a sling (86.6%). Of the 92 patients (37%) who received operative treatment, the majority was treated with ORIF (61%). The majority of operative patients had an ASA classification of 2 (63%). The mean HLOS was 1.3 days (SD 2.8).

Relation between treatment method and outcomes

There were 26 complications (10.4%). Subsequent surgery was performed in 21 patients (8.4%). The mean flexion was 125 degrees (SD 37.9) and mean abduction was 104.8 degrees (SD 42.2). The mean duration of time from treatment to last clinical follow-up was 8.2 months (SD 6.8). The mean PROMIS PF was 25.4 (SD 10.10), the mean PROMIS UE was 34.95 (SD 14.93), and the mean QuickDASH was 17.02 (SD 11.08).

Clinical and functional outcome measures stratified by treatment method are shown in Table 2. The results of the crude and adjusted regression analyses of the relation between treatment method and PROMs are shown in Table 3. There were no differences in PROMs, crude as well as adjusted, between nonoperative and operative treatment; PROMIS PF adjusted mean

Table 1. Characteristics of 249 proximal humeral fracture patients

	Overall	**Nonoperative**	**Operative**
Patients	249	157 (63.0)	92 (37.0)
Age	64.53 (14.89)	65.43 (15.05)	62.99 (14.55)
Sex (%)			
Female	183 (73.5)	120 (76.4)	63 (68.5)
Male	66 (26.5)	37 (23.6)	29 (31.5)
Trauma mechanism (%)			
Fall-low energy	209 (83.9)	132 (84.1)	77 (83.7)
Fall-high energy	7 (2.8)	5 (3.2)	2 (2.2)
Motor vehicle crash	6 (2.4)	4 (2.5)	2 (2.2)
Motorcycle crash	2 (0.8)	2 (1.3)	0 (0.0)
Bicycle accident	6 (2.4)	3 (1.9)	3 (3.3)
Sports-related	14 (5.6)	8 (5.1)	6 (6.5)
Other	5 (2.0)	3 (1.9)	2 (2.2)
HET (%)			
No	235 (94.4)	147 (93.6)	88 (95.7)
Yes	14 (5.6)	10 (6.4)	4 (4.3)
Open fracture (%)			
No	246 (98.8)	157 (100.0)	89 (96.7)
Yes	3 (1.2)	0 (0.0)	3 (3.3)
AO classification (%)			
A1	41 (16.5)	29 (18.5)	12 (13.0)
A2	24 (9.6)	23 (14.6)	1 (1.1)
A3	14 (5.6)	8 (5.1)	6 (6.5)
B1	89 (35.7)	78 (49.7)	11 (12.0)
B2	16 (6.4)	9 (5.7)	7 (7.6)
B3	19 (7.6)	2 (1.3)	17 (18.5)
C1	12 (4.8)	5 (3.2)	7 (7.6)
C2	21 (8.4)	3 (1.9)	18 (19.6)
C3	13 (5.2)	0 (0.0)	13 (14.1)
Multiple UE fractures (%)			
No	243 (97.6)	154 (98.1)	89 (96.7)
Yes	6 (2.4)	3 (1.9)	3 (3.3)
Number of UE fractures (%)			
1	243 (97.6)	154 (98.1)	89 (96.7)
2	5 (2.0)	2 (1.3)	3 (3.3)
3	1 (0.4)	1 (0.6)	0 (0.0)
ASA classification (%)*			
1			7 (9.2)
2			48 (63.2)
3			21 (27.6)
Operative method (%)			
ORIF			56 (60.9)
Hemiarthroplasty			3 (3.3)
Reversed shoulder arthroplasty			29 (31.5)
Other			4 (4.3)
Nonoperative method (%)			
Closed reduction + sling		20 (12.7)	
Sling		136 (86.6)	
Other		1 (0.6)	
HLOS (days)	1.32 (2.79)	0.56 (1.43)	2.62 (3.86)

Continuous variables presented as mean (SD); HET high-energy trauma; UE upper extremity; ORIF open reduction and internal fixation; * ASA classification, collected for surgical patients, was known for 76 of operative patients (83%); HLOS Hospital length of stay

difference -1.50; 95% CI -5.20; 2.19; p= 0.426), PROMIS UE adjusted mean difference -0.27; 95% CI -6.66; 6.11; p= 0.933), and QuickDASH adjusted mean difference -2.33; 95% CI -6.72 – 2.05; p= 0.299). Also, there were no differences in complication rate (9.6% versus 12%, p= 0.701) and subsequent surgery rate (6.4% versus 12%, p= 0.195) between treatment methods.

Table 2. Clinical and patient-reported outcomes of 249 proximal humeral fracture patients, stratified by treatment

	Overall	Nonoperative	Operative
Clinical outcomes			
Complications (%)			
No	223 (89.6)	142 (90.4)	81 (88.0)
Yes	26 (10.4)	15 (9.6)	11 (12.0)
Complication classification (%)			
Nonunion	9 (34.6)	5 (33.3)	4 (36.4)
Malunion	2 (7.7)	2 (13.3)	0 (0.0)
Arthrosis	2 (7.7)	2 (13.3)	0 (0.0)
Infection	1 (3.8)	1 (6.7)	0 (0.0)
Dislocation	4 (15.4)	2 (13.3)	2 (18.2)
Implant failure	0 (0.0)	0 (0.0)	0 (0.0)
Chronic pain	4 (15.4)	1 (6.7)	3 (11.5)
Avascular necrosis	2 (7.7)	0 (0.0)	2 (18.2)
Other	2 (7.7)	2 (13.3)	0 (0.0)
Subsequent surgery (%)			
No	228 (91.6)	147 (93.6)	81 (88.0)
Yes	21 (8.4)	10 (6.4)	11 (12.0)
Flexion (degrees) (n=168, 68%)	124.96 (37.91)	123.78 (36.21)	126.69 (40.49)
Abduction (degrees) (n=142, 57%)	104.81 (42.23)	101.88 (42.76)	109.05 (41.45)
Time treatment to clinical FU (months)	8.20 (6.82)	6.35 (5.80)	11.36 (7.29)
PROMs			
PROMIS PF (n=164, 66%)	25.37 (10.10)	24.37 (10.78)	27.52 (8.13)
PROMIS UE (n=126, 51%)	34.95 (14.93)	34.16 (15.89)	36.43 (12.98)
QuickDASH (n=120, 48%)	17.02 (11.08)	16.67 (11.22)	17.50 (10.98)
Time treatment to PROMIS PF (months)	7.34 (5.43)	6.33 (4.60)	9.51 (6.42)
Time treatment to PROMIS UE (months)	7.42 (5.13)	6.43 (4.22)	9.27 (6.14)
Time treatment to QuickDASH (months)	6.38 (5.55)	5.49 (4.68)	7.64 (6.41)

Continuous variables presented as mean (SD); FU follow-up; PROMs Patient-Reported Outcome Measures; PROMIS Patient-Reported Outcomes Measurement Information System; PF physical function; UE upper extremity; DASH Disabilities of the Arm, Shoulder, and Hand

Table 3. Differences in patient-reported outcomes, operative and nonoperative treatment of proximal humeral fracture

	Non		Op		Crude				Adjusted			
	Mean	SD	Mean	SD	MD	95%CI		p-value	MD	95%CI		p-value
PROMIS PF	24.37	10.78	27.52	8.13	-3.15	-6.45	0.14	0.063	-1.50	-5.20	2.19	0.426
PROMIS UE	34.16	15.89	36.43	12.98	-2.27	-7.75	3.20	0.417	-0.27	-6.66	6.11	0.933
QuickDASH	16.67	11.22	17.50	10.98	-0.83	-4.86	3.21	0.688	-2.33	-6.72	2.05	0.299

Non Nonoperative; Op Operative; MD Mean difference; PROMs Patient-Reported Outcome Measures; CI Confidence Interval; PROMIS Patient-Reported Outcomes Measurement Information System; PF physical function; UE upper extremity; DASH Disabilities of the Arm, Shoulder, and Hand. Adjustment was made for baseline and clinical factors that might confound the relationship between treatment method and functional outcome including the time from treatment to questionnaire, age, AO classification, and occurrence of multiple UE fractures.

Associations between outcome measures

PROMIS PF was associated with PROMIS UE (r 0.83; 95% CI 0.76; 0.88; p<0.001), and QuickDASH (r -0.47; 95%CI -0.62; -0.29; p<0.001). The PROMIS UE was also associated with the QuickDASH (r -0.70; 95%CI -0.81; -0.55; p<0.001).

PROMIS PF was associated with flexion degree (r 0.43; 95% CI 0.26; 0.58; p<0.001), however, there was no association with abduction degree (r 0.12; 95%CI -0.09; 0.31; p= 0.269). PROMIS UE was associated with flexion degree (r 0.27; 95% CI 0.04; 0.47; p= 0.019), yet there was no association with abduction degree (r -0.14.; 95%CI -0.37; 0.11; p= 0.281). QuickDASH was

negatively associated with flexion degree (r -0.25.; 95% CI -0.44; -0.04; p= 0.021), however, there was no association with abduction degree (r -0.02; 95%CI -0.26; 0.23; p= 0.876).

Discussion

In this observational study of operative versus nonoperative treatment of proximal humeral fractures, no differences in PROMs nor in clinical outcomes were observed. In patients with a proximal humeral fracture, the PROMIS PF, PROMIS UE, and QuickDASH scores showed a high mutual correlation, as well as a moderate correlation with flexion degree.

Several meta-analyses have been inconclusive as to whether operative treatment of proximal humeral fractures is superior to nonoperative treatment, reporting no clinically relevant differences in outcomes between treatment groups.[5,6] The patient characteristics (i.e. age, sex, injury mechanism, fracture classification) of the studies reported in these meta-analyses are comparable to those of the patients included in the current study. However, due to difference in outcome measures and in duration of follow-up, a direct comparison of results is not possible. In accordance with the previous literature this study found no difference in functional outcomes and complication rates between operative and nonoperative treatment. [5,6] Regardless of treatment method, disappointing functional outcomes including residual shoulder pain, limitations in shoulder motion, and decreased quality of life after proximal humeral fractures have been reported.[19,20] Furthermore, the results in this study emphasize the overall poor outcome following proximal humeral treatment. However, this might be due to the relative short follow-up period in this study. There is still debate on whether further improvement of functional outcome after 12 months is expected after humeral fracture treatment.[21] The incidence of proximal humeral fracture rates will continue to increase with the aging population, causing prolonged and severe disability. In the elderly, the effect of UE fractures on functional outcome might continue for many years, and long-term evaluation is needed to accurately assess the efficacy of treatment and rehabilitation.[20]

Previous studies have shown encouraging results regarding correlation of PROMIS questionnaires with legacy outcome scores, responsiveness to treatment, and validity in orthopedics. Gausden et al.[11], evaluated the correlation of PROMIS in upper extremity fracture patients, and found a high correlation between the DASH and PROMIS PF (r -0.76) and PROMIS UE (r -0.79). Morgan et al.[13] evaluated the correlation between the PROMIS PF and

DASH (r -0.66), specifically in proximal humeral fracture patients older than 60 years. In the current study PROMIS PF (r -0.47) and PROMIS UE (r -0.70) also showed a moderate to high correlation with the QuickDASH in the evaluation of proximal humeral fracture patients. To our knowledge the QuickDASH, the validated and shortened version of the DASH, has not been previously compared to the newer PROMIS PF and PROMIS UE. The PROMIS were developed with the goal to provide standardized, valid, and flexible PROMs collection tools, which make them more useful in research and clinical practice, with features that lower response burden and make it possible to seamlessly incorporation them into patients' medical record.[7-13] However, despite early promising results, the widespread use of PROMIS has not been adopted in most orthopedic literature. Furthermore, there is a lack of evidence on PROMIS specifically for orthopedic trauma.[11] The correlations found in this study suggest the PROMIS PF and PROMIS UE can be considered similarly useful as a measure for evaluating shoulder function after proximal humeral fractures.

Clinical physical shoulder examination such as flexion and abduction are commonly used to assess patients with proximal humerus fractures. Because these fractures occur in a heterogenous group, with a wide variety of injury mechanisms, functional demands, and comorbidities, the potential for discordance between clinical outcomes and PROMs exists.[14] For patients with proximal humerus fractures, the relationship between shoulder impairment and PROMs has not been well-described. One previous study by Slobogean et al.[14], quantified the relationship between patient-reported shoulder outcome and objective shoulder impairment using regression models. They found different associations between objective shoulder impairment and the DASH score, with regression analysis suggesting that shoulder impairment (abduction, external rotation, strength) explained 50% of the DASH.[14] Furthermore, Slobogean et al.[14], reported flexion had the best discriminatory ability for identifying normal shoulder function. In the current study we evaluated this relationship and found a moderate association between all PROMs and flexion degree, however not with abduction degree. These results suggest that emphasizing efforts to improve the degree of flexion might be associated with higher PROMs. These results support a comprehensive approach to surgical quality that incorporates both clinical events and PROMs. The current study was limited due to the inclusion of less range of motion and strength measurements compared to the study by Slobogean et al.[14]. However, the study by Slobogean et al.[14], was limited by the small sample size of their cohort, which included 31 patients. For patients with proximal humerus fractures, the relationship between shoulder impairment and PROMIS

scores and the QuickDASH has not yet been well-described. A discordance between shoulder impairment and PROMs has been demonstrated and further work to identify patient, injury, or treatment factors to minimize this discrepancy is still needed. Although these measures of motion provide some degree of information on impairment, they do not provide insight into the perceived functional outcome. Hence, PROMs are used complementary to clinical examination.[14] The combined use of clinical measures of motion and PROMs allow physicians to evaluate the differences between clinical impairment and patient perceived functional outcome.

Potential limitations in this study need to be acknowledged. First, the study is limited by the retrospective nature. However, to our knowledge, with the inclusion of 249 patients, this study is one of largest cohort to establish data on characteristics and outcome after operative and nonoperative treatment, with functional outcome and range of motion (flexion/abduction) known for up till 68% of patients. Second, the different outcome measures had various endpoints. However, inclusion was limited to at least 3 months of clinical or functional follow-up. Furthermore, the time from treatment to clinical follow-up and time from treatment to questionnaire were known and analysis adjusted accordingly. Third, due to the nature of the injury in a trauma setting, it is not possible to compare and adjust PROMs with baseline scores. However, due to the extensive database adjustment was able for several baseline and clinical factors that might confound associations. Finally, results were limited regarding range of motion and strength measurements. We acknowledge that better evidence is lacking and further evaluation using PROMs and different clinical outcomes in proximal humeral fracture patients is needed. However, given the paucity of data regarding PROMIS and QuickDASH scores after proximal humerus fracture treatment, we hope these results will mainly provide benchmark data that can be used for future comparisons.

Conclusion

There was no significant difference in outcomes between operative and nonoperative treatment of proximal humeral fractures. The PROMIS PF, PROMIS UE, and QuickDASH score showed a high mutual correlation, as well as a moderate correlation with flexion degree. To our knowledge the QuickDASH has not been previously compared to the newer PROMIS PF and PROMIS UE in proximal humeral fractures. The correlations found in this study suggest the PROMIS PF and PROMIS UE can be considered similarly useful as a measure for evaluating shoulder function after proximal humeral fractures.

References:

1. Court-Brown CM, Caesar B. Epidemiology of adult fractures: A review. *Injury.* 2006;37:691-697.
2. Palvanen M, Kannus P, Niemi S, Parkkari J. Update in the epidemiology of proximal humeral fractures. *Clin Orthop Relat Res.* 2006;442:87-92.
3. Handoll HH, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev.* 2015(11):Cd000434.
4. Nowak LL, Davis AM, Mamdani M, Beaton D, Kennedy C, Schemitsch EH. A Systematic Review and Standardized Comparison of Available Evidence for Outcome Measures Used to Evaluate Proximal Humerus Fracture Patients. *J Orthop Trauma.* 2019;33(7):e256-e262.
5. Xie L, Ding F, Zhao Z, Chen Y, Xing D. Operative versus non-operative treatment in complex proximal humeral fractures: a meta-analysis of randomized controlled trials. *SpringerPlus.* 2015;4(1):728.
6. Beks RB, Ochen Y, Frima H, et al. Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. *J Shoulder Elbow Surg.* 2018;27(8):1526-1534.
7. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *J Am Acad Orthop Surg.* 2016;24(11):744-749.
8. Cheung EC, Moore LK, Flores SE, Lansdown DA, Feeley BT, Zhang AL. Correlation of PROMIS with Orthopaedic Patient-Reported Outcome Measures. *JBJS reviews.* 2019;7(8):e9.
9. Gagnier JJ. Patient reported outcomes in orthopaedics. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2017;35(10):2098-2108.
10. Kaat AJ, Rothrock NE, Vrahas MS, et al. Longitudinal Validation of the PROMIS Physical Function Item Bank in Upper Extremity Trauma. *J Orthop Trauma.* 2017;31(10):e321-e326.
11. Gausden EB, Levack AE, Sin DN, et al. Validating the Patient Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests for upper extremity fracture care. *J Shoulder Elbow Surg.* 2018;27(7):1191-1197.
12. Hung M, Saltzman CL, Greene T, et al. The responsiveness of the PROMIS instruments and the qDASH in an upper extremity population. *Journal of patient-reported outcomes.* 2017;1(1):12.
13. Morgan JH, Kallen MA, Okike K, Lee OC, Vrahas MS. PROMIS Physical Function Computer Adaptive Test Compared With Other Upper Extremity Outcome Measures in the Evaluation of Proximal Humerus Fractures in Patients Older Than 60 Years. *J Orthop Trauma.* 2015;29(6):257-263.
14. Slobogean GP, Noonan VK, Famuyide A, O'Brien PJ. Does objective shoulder impairment explain patient-reported functional outcome? A study of proximal humerus fractures. *J Shoulder Elbow Surg.* 2011;20(2):267-272.
15. Advanced trauma life support (ATLS®): the ninth edition. *The journal of trauma and acute care surgery.* 2013;74(5):1363-1366.
16. Linder F, Mani K, Juhlin C, Eklöf H. Routine whole body CT of high energy trauma patients leads to excessive radiation exposure. *Scandinavian journal of trauma, resuscitation and emergency medicine.* 2016;24:7.
17. Hunsaker FG, Cioffi DA, Amadio PC, Wright JG, Caughlin B. The American academy of orthopaedic surgeons outcomes instruments: normative values from the general population. *J Bone Joint Surg Am.* 2002;84(2):208-215.
18. Team R. R Development Core Team. *R: A Language and Environment for Statistical Computing.* 2013.
19. Slobogean GP, Johal H, Lefaivre KA, et al. A scoping review of the proximal humerus fracture literature. *BMC Musculoskelet Disord.* 2015;16:112.
20. Hodgson S. Proximal humerus fracture rehabilitation. *Clin Orthop Relat Res.* 2006;442:131-138.
21. Frima H, Michelitsch C, Beks RB, Houwert RM, Acklin YP, Sommer C. Long-term follow-up after MIPO Philos plating for proximal humerus fractures. *Archives of orthopaedic and trauma surgery.* 2019;139(2):203-209.

CHAPTER 10

Association of patient-reported outcomes with clinical outcomes after distal humerus fracture treatment

Abhiram R. Bhashyam, Yassine Ochen, Quirine M.J. van der Vliet,
Luke P.H. Leenen, Falco Hietbrink, R. Marijn Houwert,
George S.M. Dyer, Marilyn Heng

(Journal of the American Academy of Orthopaedic Surgeons)

Abstract

Background

In this study, we assessed the patient-reported outcomes of distal humerus fracture treatment using Patient-Reported Outcomes Measurement Information System (PROMIS) or QuickDASH (Disabilities of the Arm, Shoulder, and Hand) scores and the association between patient-reported outcomes and clinical outcomes.

Methods

We performed a retrospective cohort study of 76 adult patients who sustained an acute distal humerus fracture between 2016 and 2018; 53 patients completed at least one patient-reported outcome measure used to assess physical function (PF) during their routine follow-up care (69.7% response rate). The average time to follow-up patient-reported outcome measure was 10.3 months. Patients completed the PROMIS PF 10a, PROMIS upper extremity (UE) 16a, and/or QuickDASH based on the treating institution/service. In addition, the PROMIS Global (Mental) subscale score was used as a measure of self-rated mental health. To assess clinical outcomes, we measured radiographic union, range of motion, and postoperative complications.

Results

Most fractures were intra-articular (67.9%), and 84.9% were treated surgically. After treatment, 98.1% of fractures united radiographically. By the final follow-up, the average arc of motion was 18° to 122°. Average (SD) PROMIS PF and UE scores were 41.7 (SD 11.1) and 40.8 (SD 12.4), respectively. The average QuickDASH score was 39.4 (SD 26.5). The arc of flexion-extension and PROMIS Global (Mental) score were independently associated with PROMIS PF and PROMIS UE scores.

Conclusion

We found that clinical factors (the arc of flexion-extension) and patient psychological factors (PROMIS Global Mental score) were independently associated with PROMIS measures of PF after distal humerus fracture treatment. These data can be used to contextualize patient outcomes and guide patient expectations.

Introduction

Fractures of the distal humerus account for 2% of fractures in the adult population (approximately 30% of all humeral fractures).[1-3] An increase in the annual incidence of distal humeral fractures has been reported, likely because of a growing older population.[4,5] In general, these injuries are treated surgically with open reduction and internal fixation (ORIF), but some patients may still be managed with nonsurgical treatment.[1]

Although several studies have evaluated clinical outcomes of distal humeral fractures, fewer studies have explored the association between clinical and patient-reported outcome measures (PROMs).[6,7] A recent systematic review identified 109 articles assessing the outcomes of acute distal humeral fracture but found that clinical and PROMs were not consistently reported, making accurate comparison of treatment effectiveness difficult.[8] In addition, the review found that general health surveys were rarely reported and comparison using Patient-Reported Outcomes Measurement Information System (PROMIS) instruments were not possible.[8]

PROMIS instruments are increasingly used to evaluate PROMs for upper extremity (UE) injuries because they can be administered and scored in a standardized manner, allowing for quality assessment across medical and surgical fields.[9,10] In addition, several studies have demonstrated that PROMIS scores correlate with legacy instruments used to measure the PROMs of orthopedic UE trauma patients.[11-13] Few studies have assessed if there is an association between PROMs (e.g. PROMIS instruments, QuickDASH) and clinical outcomes.[14,15] We hypothesized that the variation in PROMIS scores is associated with clinical outcomes. Therefore, in this study, we collected PROMs after distal humerus fracture treatment using PROMIS or QuickDASH scores and then explored the association between PRO and clinical outcomes.

Methods

Study design

This study was approved by our institutional review board. We performed a retrospective cohort study of 85 consecutive adult patients (>18 years old) who received treatment at one of two American College of Surgeons Level 1 Trauma Centers from January 2016 to February 2018 for an acute distal humerus fracture. Starting in January 2016, collection of patient-reported outcome measures (PROMs) was standardized in the orthopedic clinics at both hospitals. Patients were excluded if their injury was initially treated at an outside hospital or if they had a pathologic or

periprosthetic fracture. Patients who had zero follow-up visits (five patients) or were in hospice care (one patient) were also excluded, as were patients treated with total elbow arthroplasty (three patients). From the 76 eligible patients, 53 patients completed at least one follow-up PROMs used to assess physical function (PF)/UE disability (69.7% response rate) with an average follow-up of 10.3 months (Table 1).

PROMs

Patients completed the PROMIS PF 10a, PROMIS UE 16a, and/or the QuickDASH to assess PF and UE disability on a tablet device as part of their routine follow-up visit at the treating institution.[9,11,13,16,17] In addition, the PROMIS Global was completed and the PROMIS Global (Mental) subscale score was used as a measure of self-rate mental health.[18] The PROMIS instrument scores range from 0 to 100 with a mean score of 50 for the general population of the United States (SD of 10).[9] The QuickDASH is an 11-item questionnaire that measures UE-specific disability with higher scores reflecting more severe disability (range of 0 to 100) and a mean of 11 points reflecting the general US population average.[9]

Clinical outcomes

To assess clinical outcomes, we evaluated radiographic union, range of motion, complications (heterotopic ossification and infection), and unplanned return to the operating room. Symptomatic implants were not considered a complication and were recorded separately. The most recently available anterior-posterior and lateral radiographs were evaluated to assess for radiographic union by the treating surgeon (fellowship-trained in orthopedic trauma or hand/UE) and independently by the first author (A.R.B., fifth year orthopedic surgery resident). Range of motion was assessed by the treating surgeon for flexion contracture (i.e. terminal extension), terminal flexion, and the total arc of flexion-extension at the last outpatient follow-up visit. Patients were deemed to have a functional range of motion if their flexion-extension arc was at least 30° to 130°.[19]

Independent variables

Detailed sociodemographic and clinical data were identified for each patient using our institutions' Enterprise Data Warehouse and the electronic medical record (Table1). Because the patients in this study are from a similar geographic area, median income for each patient was abstracted for each patient using the ZIP code of residence based on census data.[20] Primary

health insurance was divided into three categories (private, Medicaid, and Medicare).[21] Distal humerus fractures were classified using the AO-OTA fracture classification by the treating surgeon and independently by the first-author (A.R.B.), and patients with other fractures were classified as "multiple injuries" (binary classification).[22] To mitigate interobserver variability during analysis, all fractures were then grouped as extra-articular (13.A) or intra-articular (partial articular [13.B] and complete articular [13.C]). The energy of injury mechanism was defined according to the Advanced Trauma Life Support guidelines.[23] Patients who did not meet the criteria for high-energy trauma were considered low-energy trauma. Procedures were grouped as closed treatment, ORIF, or ORIF with ulnar nerve transposition (subcutaneous versus submuscular).

Statistical analysis

Baseline characteristics and clinical results between responders and nonresponders were compared using the Fisher exact test for categoric variables and t-test/analysis of variance for continuous variables to assess for response bias. Multivariable linear regression modeling was used to assess the relationship between PROMs and clinical results of distal humerus fracture treatment. To adjust for factors that may confound the relationship between PROMIS PF/PROMIS UE/QuickDASH and clinical outcomes, we used forward stepwise selection to include those patients' sociodemographic and clinical variables that were notable at an alpha level of 0.10.[14] All models were constrained to include the arc of flexion-extension and complications as relevant, independent, and noncollinear clinical outcomes. We also assessed the relationship between PROMIS PF, PROMIS UE, and QuickDASH using simple linear regression to validate our data against previous studies.[9,11,16] P values <0.05 were considered statistically significant. Stata software, version 13.1 (StataCorp), was used for all analyses.

Results

Study population

In this cohort of 53 patients who underwent treatment of a distal humerus fracture and completed PROMs regarding UE function, most patients were women (67.9%) and Caucasian (83%). The average age was 58 years (median: 72 years; range: 22 to 94 years). Most patients carried private (56.6%) or Medicare (37.7%) insurance. The average follow-up was 10.3 (SD 7.1) months. Among all injuries, 13.2% were the result of high-energy trauma, 5.7% were open, and nine patients sustained multiple fractures. Most distal humerus fractures were intra-articular

Table 1. Sociodemographic and clinical characteristics of responders versus nonresponders

	Responders (n=53, 70%)	**Non-responders (n=23, 30%)**	
	No. of Patients (%) or Mean (SD)	**No. of Patients (%) or Mean (SD)**	**p-value**
Sociodemographic characteristics			
Age at injury (years)	54.5 (20.4)	65.1 (19.4)	0.038
Male	17 (32.1)	8 (34.5)	0.509
White race	44 (83.0)	15 (65.2)	0.081
Median income ($)*	93,600 (30,600)	83,600 (30,900)	0.200
Marital status			0.15
Single	25 (48.1)	5 (23.8)	
Married	22 (42.3)	11 (52.4)	
Widowed	4 (7.7)	3 (14.3)	
Divorced	1 (1.9)	2 (9.5)	
Insurance type			0.025
Private	30 (56.6)	6 (26.1)	
Medicaid	3(5.7)	5 (21.7)	
Medicare	20 (37.7)	12 (52.2)	
Injury-related characteristics			
High-energy trauma	7 (13.2)	3 (13.0)	0.648
Open fracture	3 (5.7)	4 (17.4)	0.119
Multiple injuries	9 (17.0)	3 (13.0)	0.477
AO/OTA fracture classification			0.550
A (Extra-articular)	17 (32.1)	10 (43.5)	
B (Partial-articular)	12 (22.6)	53(13.0)	
C (Complete articular)	24 (45.3)	10 (43.5)	
Procedure-related characteristics			
Procedure			0.133
Closed treatment	8 (15.1)	3 (13.0)	
ORIF	22 (41.5)	16 (69.6)	
ORIF + subcutaneous ulnar nerve transposition	8 (15.1)	2 (8.7)	
ORIF + submuscular ulnar nerve transposition	15 (28.3)	2 (8.7)	
UE specialist	24 (45.3)	9 (39.1)	0.405
Inpatient surgery	31 (58.5)	15 (65.2)	0.387
Post-procedure characteristics			
Length of stay (days)	2.3 (2.3)	2.3 (1.8)	0.933
Discharge to rehab	6 (11.3)	5 (21.7)	0.200
Follow-up time (months)	10.3 (7.1)	5.8 (4.2)	0.001

DASH = Disabilities of the Arm, Shoulder, and Hand, HET = high-energy trauma, OTA = Orthopaedic Trauma Association, OR = odds ratio, ORIF = open reduction and internal fixation, UE = upper extremity, * Median income from ZIP code of residence based on 2016 census data

(67.9%), and 84.9% of patients were treated surgically (84.9%). Approximately 45% of patients were treated by an UE specialist (hand or shoulder/elbow fellowship-trained), 58.5% of injuries were treated as inpatient procedures, and only 11.3% of patients were discharged to rehab. Responders and nonresponders were similar in almost all characteristics, except that nonresponders were younger, more likely to be on Medicare/Medicaid, and had shorter follow-up (Table 1).

Clinical results

After treatment, 98.1% of patients demonstrated radiographic union of their distal humerus fracture. By the final follow-up, average flexion contracture was 18°, terminal flexion was 122°, and the average arc of flexion-extension was 105°; 52.8% of patients had a functional range of

Table 2. Clinical outcomes of responders versus nonresponders and patient-reported functional outcome of responders

	Responders	Non-responders	
	No. of Patients (%) or Mean (SD)	**No. of Patients (%) or Mean (SD)**	**p-value**
Clinical outcomes			
Radiographic union	52 (98.1)	23 (100)	0.697
Flexion contracture (degrees)	18 (21)	19 (12)	0.756
Terminal flexion (degrees)	122 (15)	118 (16)	0.331
Arc of flexion-extension	105 (30)	99 (23)	0.422
Functional arc of motion (30-130 degree)	28 (52.8)	9 (40.9)	0.247
Complication	9 (17.0)	2 (8.7)	0.346
Unplanned return to the OR	6 (11.3)	2 (8.7)	0.732
Patient reported functional outcomes			
PROMIS Physical Function 10a	41.7 (11.1)		
PROMIS Global (Physical)	44.7 (11.6)		
PROMIS Global (Mental)	52.2 (10.4)		
PROMIS Upper Extremity 16a	40.8 (12.4)		
QuickDASH	39.4 (26.5)		

PF = physical function, PROMIS = Patient-Reported Outcomes Measurement Information System, UE = upper extremity

motion (at least 30° to 130° flexion-extension arc). Among all patients, nine patients (14.5%) sustained at least one complication (Table 2). Four patients had heterotopic ossification, three patients had an infection, and two patients had a nonunion. Seven patients had symptomatic implants. Clinical results were similar between responders and nonresponders.

PROMs

Average (SD) PROMIS PF and UE scores were 41.7 (SD 11.1) and 40.8 (SD) 12.4, respectively. The average QuickDASH score was 39.4 (SD 26.5) (Table 2). PROMIS PF scores were associated with PROMIS UE scores ($r = 0.84$, $P < 0.001$) and QuickDASH scores ($r = -0.55$, $P = 0.012$).11-13,16 In addition, PROMIS UE scores were associated with QuickDASH scores ($r = 0.87$, $P < 0.001$).

Association of clinical results with PROMs

After controlling for likely confounding variables using multivariable analysis (e.g. age and sex), the arc of flexion and extension (coefficient [95% confidence interval] = 0.13 [0.06, 0.19], $P < 0.001$) and PROMIS Global (Mental) scores (coefficient [95% confidence interval] = 0.79 [0.59, 0.99], $P < 0.001$) were independently associated with PROMIS PF scores. Similar results were observed for PROMIS UE and QuickDASH scores (Table 3, Figure 1, Figure 2).

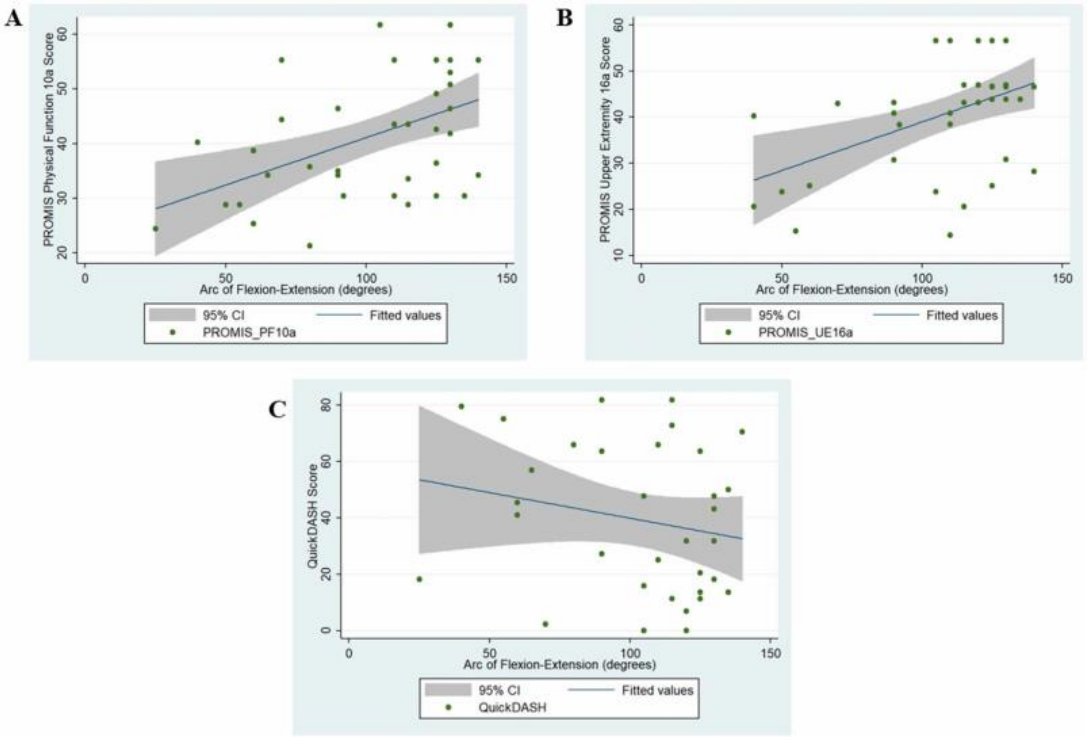

Figure 1. Chart showing the association between functional outcome scores and elbow range of motion (flexion-extension arc); (A) PROMIS PF, (B) PROMIS UE, and (C) QuickDASH. CI = confidence interval, PF = physical function, PROMIS = Patient Reported Outcomes Measurement Information System, UE = upper extremity.

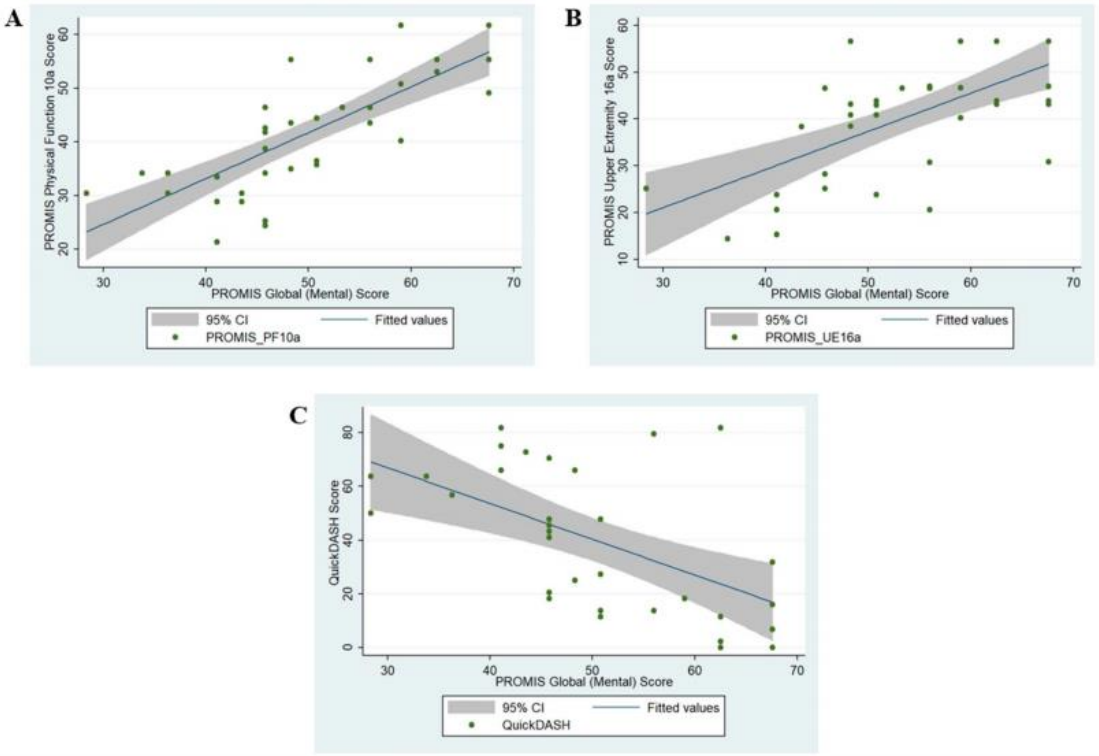

Figure 2. Chart showing the association between functional outcome scores and PROMIS global (mental health) subscale score; (A) PROMIS PF, (B) PROMIS UE, and (C) QuickDASH. CI = confidence interval, PF = physical function, PROMIS = Patient Reported Outcomes Measurement Information System, UE = upper extremity.

Table 3. Multivariable analysis of the association between clinical outcomes and patient-reported functional outcomes adjusted for sociodemographic and clinical factors

	Coefficient	95% CI		p-value	Adjusted R2
PROMIS Physical Function 10a (n=40)					
Arc of Flexion-Extension	0.13	0.06	0.18	<0.001	0.750
Complication	0.25	-4.84	5.35	0.920	
PROMIS Global (Mental)	0.79	0.59	0.99	<0.001	
PROMIS Upper Extremity 16a (n=40)					
Arc of Flexion-Extension	0.15	0.04	0.26	0.007	0.523
Complication	0.86	-8.09	9.81	0.847	
PROMIS Global (Mental)	0.73	0.38	1.09	<0.001	
QuickDASH (n=33)					
Arc of Flexion-Extension	-0.13	-0.40	0.14	0.349	0.349
Complication	-4.17	-25.2	16.9	0.688	
PROMIS Global (Mental)	-1.35	-2.12	-0.57	0.001	

CI = confidence interval, DASH = Disabilities of the Arm, Shoulder, and Hand, PF = physical function, PROMIS = Patient-Reported Outcomes Measurement Information System, UE = upper extremity

Discussion

Historically, clinical (including radiological) outcomes have been used to measure surgical treatment success and quality because they are easily obtained from administrative and clinical records, are easily quantified, and have high face validity.[24] Yet, clinical outcomes do not capture the full patient perspective and multiple recent studies have demonstrated how PROMIS scores can be used to better describe aspects of health status that are reported directly from patients after UE trauma.[9,24] In this study, we present data about the clinical and PROMs after treatment of distal humerus fractures. Our findings demonstrate that the PROMs are associated with clinical outcomes (i.e. range of motion), but each of these sets of metrics has features that are unique and important when evaluating treatment effectiveness.

Although PROMs capture benefits of surgical treatment beyond survival and physiologic markers, the extent to which PROMs are affected by traditionally measured clinical outcomes has remained unclear, especially when using PROMIS scores, abbreviated functional outcome measures (e.g. QuickDASH), or for specific clinical conditions.[24]

In this cohort of distal humerus fractures, the only clinical outcome independently associated with PROMs was the arc of motion (Figure 1). On average, an increase in the arc of flexion-extension of 70° to 80° was associated with an improvement of 8 to 9 points on the PROMIS instruments.[25] This finding is comparable to previous studies which have shown that the arc of motion was related to QuickDASH scores after elbow/wrist trauma.[15,26] In addition, we observed that long-term outcomes (e.g. final arc of motion) were more strongly associated with PROMs than perioperative complications. These findings lend further support to the notion that patients are often satisfied despite adverse or unexpected events and that PROMs likely reflect the durability of clinical outcomes.[14] Our data also suggest that emphasizing efforts to improve the

terminal arc of flexion-extension are likely to be associated with higher PROMs. These results support a comprehensive approach to surgical quality that incorporates both clinical events and self-reported measures of health status.

We also found that the PROMIS Global (Mental) subscale was independently associated with all measures of physical or upper extremity-specific function (Figure 2). On average, increases in PROMIS Global (Mental) subscale scores of 10 to 12 points were associated with 8 to 9 point improvements on PROMIS PF or UE measures.[25] These results are supported by multiple previous studies that have demonstrated how patient mindset may be the most important factor of self-reported outcomes.[18,27]

The importance of patient mental health in the measurement of PROMs presents a plausible explanation for why PROMs are not fully determined by clinical outcomes and, in part, emphasizes the importance of collecting "patient independent" outcome measures. Age, sociodemographic characteristics, or injury-related characteristics were not independently associated with PROMs in our study, although they were in others.[9,18,27] If only PROMs are used when determining financial reimbursement, our results suggest a mechanism by which presurgery mental status may be inappropriately used to select against patients expected to have worse PROMIS PF or UE measures. This further supports the value of a physicians' judgment in the evaluation of outcomes of a care episode.[28]

This study has several limitations. There is a potential for response bias because only 69.7% of eligible patients completed an UE PROM; however, our response rate is similar to other comparable studies and patient/injury characteristics of responders and nonresponders were similar (Table 1).[11,16] Given the retrospective nature of the study, patients had various end points of follow-up, although the effect of this is unclear. The follow-up duration was added to our regression analyses but was omitted in the final multivariable regression models because of the lack of statistically significant association. In addition, not all potential predictors could be assessed. For example, PF before the injury or other patient psychological factors (e.g. PROMIS Pain Interference) may have influenced outcome measures, but these could not be retrieved retrospectively.[29] Finally, some of the lack of influence of clinical outcomes on PROMs may be a limitation of our follow-up. We focused on shorter term PROMs in this study, but future studies should assess this in the long-term, ideally in prospective fashion, because the results may

degenerate over time. Nevertheless, it is reassuring that our analysis recapitulates findings from multiple previous studies.[11-13,15,16]

Conclusion

This study highlights the importance of measuring both clinical and PROMs when evaluating distal humerus fracture treatment effectiveness because each of these metrics is a unique assessor of outcome. Given the paucity of data regarding typical PROMIS or QuickDASH scores after distal humerus fracture treatment, our study also provides benchmark data that can be used for future comparison. Finally, the awareness of factors associated with poorer patient-reported and clinical outcome measures can be used to guide patient expectations and further encourage improvement in range of motion.

References:

1. Anglen J. Distal humerus fractures. *J Am Acad Orthop Surg.* 2005;13(5):291-297.
2. Jupiter JB, Mehne DK. Fractures of the distal humerus. *Orthopedics.* 1992;15(7):825-833.
3. Rose SH, Melton LJ, 3rd, Morrey BF, Ilstrup DM, Riggs BL. Epidemiologic features of humeral fractures. *Clin Orthop Relat Res.* 1982(168):24-30.
4. Palvanen M, Niemi S, Parkkari J, Kannus P. Osteoporotic fractures of the distal humerus in elderly women. *Ann Intern Med.* 2003;139(3):W-w61.
5. Palvanen M, Kannus P, Niemi S, Parkkari J. Secular trends in the osteoporotic fractures of the distal humerus in elderly women. *Eur J Epidemiol.* 1998;14(2):159-164.
6. Wang Y, Zhuo Q, Tang P, Yang W. Surgical interventions for treating distal humeral fractures in adults. *Cochrane Database Syst Rev.* 2013(1):Cd009890.
7. Nauth A, McKee MD, Ristevski B, Hall J, Schemitsch EH. Distal humeral fractures in adults. *J Bone Joint Surg Am.* 2011;93(7):686-700.
8. Zarezadeh A, Mamelson K, Thomas WC, Schoch BS, Wright TW, King JJ. Outcomes of distal humerus fractures: What are we measuring? *Orthopaedics & traumatology, surgery & research : OTSR.* 2018;104(8):1253-1258.
9. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *J Am Acad Orthop Surg.* 2016;24(11):744-749.
10. Makhni EC, Meadows M, Hamamoto JT, Higgins JD, Romeo AA, Verma NN. Patient Reported Outcomes Measurement Information System (PROMIS) in the upper extremity: the future of outcomes reporting? *J Shoulder Elbow Surg.* 2017;26(2):352-357.
11. Gausden EB, Levack AE, Sin DN, et al. Validating the Patient Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests for upper extremity fracture care. *J Shoulder Elbow Surg.* 2018;27(7):1191-1197.
12. Overbeek CL, Nota SP, Jayakumar P, Hageman MG, Ring D. The PROMIS physical function correlates with the QuickDASH in patients with upper extremity illness. *Clin Orthop Relat Res.* 2015;473(1):311-317.
13. Hung M, Voss MW, Bounsanga J, Gu Y, Granger EK, Tashjian RZ. Psychometrics of the Patient-Reported Outcomes Measurement Information System Physical Function instrument administered by computerized adaptive testing and the Disabilities of Arm, Shoulder and Hand in the orthopedic elbow patient population. *J Shoulder Elbow Surg.* 2018;27(3):515-522.
14. Waljee JF, Ghaferi A, Cassidy R, et al. Are Patient-reported Outcomes Correlated With Clinical Outcomes After Surgery?: A Population-based Study. *Annals of surgery.* 2016;264(4):682-689.
15. Lindenhovius AL, Buijze GA, Kloen P, Ring DC. Correspondence between perceived disability and objective physical impairment after elbow trauma. *J Bone Joint Surg Am.* 2008;90(10):2090-2097.
16. Kaat AJ, Rothrock NE, Vrahas MS, et al. Longitudinal Validation of the PROMIS Physical Function Item Bank in Upper Extremity Trauma. *J Orthop Trauma.* 2017;31(10):e321-e326.
17. Hung M, Saltzman CL, Greene T, et al. The responsiveness of the PROMIS instruments and the qDASH in an upper extremity population. *Journal of patient-reported outcomes.* 2017;1(1):12.
18. Stoop N, Menendez ME, Mellema JJ, Ring D. The PROMIS Global Health Questionnaire Correlates With the QuickDASH in Patients With Upper Extremity Illness. *Hand (N Y).* 2018;13(1):118-121.
19. Morrey BF, Askew LJ, Chao EY. A biomechanical study of normal functional elbow motion. *J Bone Joint Surg Am.* 1981;63(6):872-877.
20. Ramsey SD, Bansal A, Fedorenko CR, et al. Financial Insolvency as a Risk Factor for Early Mortality Among Patients With Cancer. *J Clin Oncol.* 2016;34(9):980-986.
21. Shankaran V, Jolly S, Blough D, Ramsey SD. Risk factors for financial hardship in patients receiving adjuvant chemotherapy for colon cancer: a population-based exploratory analysis. *J Clin Oncol.* 2012;30(14):1608-1614.
22. Marsh JL, Slongo TF, Agel J, et al. Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. *J Orthop Trauma.* 2007;21(10 Suppl):S1-133.
23. Advanced trauma life support (ATLS®): the ninth edition. *The journal of trauma and acute care surgery.* 2013;74(5):1363-1366.
24. Waljee JF, Dimick JB. Do Patient-Reported Outcomes Correlate with Clinical Outcomes Following Surgery? *Adv Surg.* 2017;51(1):141-150.
25. Hung M, Tyser A, Saltzman CL, Voss MW, Bounsanga J, Kazmers NH. Establishing the minimal clinically important difference for the PROMIS and qDASH: Level 1 evidence. *Journal of Hand Surgery.* 2018;43(9):S22.

26. Wilcke MK, Abbaszadegan H, Adolphson PY. Patient-perceived outcome after displaced distal radius fractures. A comparison between radiological parameters, objective physical variables, and the DASH score. *J Hand Ther.* 2007;20(4):290-298; quiz 299.
27. Kazmers NH, Hung M, Rane AA, Bounsanga J, Weng C, Tyser AR. Association of Physical Function, Anxiety, and Pain Interference in Nonshoulder Upper Extremity Patients Using the PROMIS Platform. *J Hand Surg Am.* 2017;42(10):781-787.
28. Sadiqi S, Verlaan JJ, Lehr AM, et al. Surgeon Reported Outcome Measure for Spine Trauma: An International Expert Survey Identifying Parameters Relevant for the Outcome of Subaxial Cervical Spine Injuries. *Spine (Phila Pa 1976).* 2016;41(24):E1453-e1459.
29. Jayakumar P, Overbeek CL, Lamb S, et al. What Factors Are Associated With Disability After Upper Extremity Injuries? A Systematic Review. *Clin Orthop Relat Res.* 2018;476(11):2190-2215.

CHAPTER 11

Long-term outcomes after open reduction and internal fixation of bicondylar tibial plateau fractures

Yassine Ochen, Jesse Peek, Michael F. McTague, Michael J. Weaver, Detlef van der Velde, R. Marijn Houwert, Marilyn Heng

(Injury)

Abstract

Background

To establish normative data, long-term patient-reported outcome measures (PROMs) on function and health-related quality of life (HrQoL) after operative treatment of bicondylar tibial plateau fractures. Secondly, to identify risk factors associated with functional outcome and HrQoL.

Methods

We performed a retrospective cohort study at two Level I trauma centers. All adult patients with AO/OTA 41-C or Schatzker V/ VI tibial plateau fractures treated between 2001 and 2016 (n= 450) by open reduction internal fixation (ORIF). The survey was completed by 214 patients (48%). Primary outcome was patient-reported functional outcome assessed with the PROMIS Physical Function (PROMIS PF). Secondary outcomes were HrQoL measured with the EuroQol 5-Dimensions 3-Levels (EQ-5D-3L), infection rate, and total knee arthroplasty (TKA) rate.

Results

Infection occurred in 26 cases (12%) and TKA was performed in 6 patients (3%). The median PROMIS PF scores was 49.8 (IQR;42-54). The median EQ-5D-3L was 0.83 (IQR;0.78-1.0). %). The multivariable regression model revealed female gender, diabetes, and worse HrQoL were correlated with worse functional outcome. The multivariable regression model revealed smoking, diabetes, and the subsequent need for TKA to be correlated with worse HrQoL.

Conclusion

The PROMIS PF and EQ-5D-3L did not reach a minimum clinically important difference. The PROMIS PF items revealed patients had no difficulty in walking more than a mile or climbing a flight of stairs. However, patients were limited in doing vigorous activities and patients should be counseled about the expected long-term outcomes. This study emphasizes the correlation between injury specific functional PROMs and general health measures.

Introduction

Tibial plateau fractures account for approximately 30% of all tibia fractures, and can be classified according to the Schatzker or AO/OTA classification.[1-4] Bicondylar fractures of the tibial plateau, AO/OTA 41-C or the Schatzker V/VI, are complex and severe injuries.[4,5] These bicondylar fractures account for approximately 18% to 39% of all tibial plateau fractures.[5]

The operative management of bicondylar tibial plateau fractures is challenging due to several aspects that need to be addressed including articular reduction, angular stability, coronal alignment, and soft-tissue injuries.[4,5] However, definitive consensus on the operative fixation of bicondylar tibial plateau fractures has not yet been established. Comparisons of the different treatment modalities remains difficult with no fixation method resulting in superior outcomes or associated with a lower risk of complications.[5] The optimal management should be based on patient- and fracture specific characteristics due to the wide range in fracture complexity, severity, and soft-tissue involvement.[4,5] However, the long-term results and functional outcome after operative treatment of bicondylar tibial plateau fractures have not been widely studied. Contributing to the difficulty in choosing the optimal management for bicondylar tibial plateau fractures is the lack of validated patient-reported outcome measures (PROMs).[5]

The PROMIS PF measures are recently developed PROMs, which have been validated in patient populations with orthopedic disorders, and have shown to be psychometrically superior to legacy measures in several key populations.[6] Although increasingly common, PROMIS scores have not been widely used in the evaluation of bicondylar tibial plateau fractures.

The aim of this study was to establish normative data, long-term patient-reported functional outcome and health-related quality of life (HrQoL) after operative treatment of bicondylar tibial plateau fractures. Secondly, this study sought to identify risk factors associated with functional outcome and HrQoL.

Methods

Study design

A retrospective cohort study with follow-up by questionnaire was performed using data from two American College of Surgeons (ACS) level I trauma centers. All adult patients with bicondylar tibial plateau fractures who were treated with ORIF between January 2001 and December 2016

were eligible for inclusion. Eligible patients were identified by searching for Current Procedure Terminology (CPT) codes and International Classification of Diseases (ICD) codes in the institution's Research Patient Data Registry (RPDR). Inclusion criteria were: (1) age 18 years or older, (2) bicondylar tibial plateau fracture (AO/OTA 41-C or Schatzker V/VI), (3) treatment with ORIF, and (4) minimum of 12 months follow-up. Exclusion criteria were: (1) treatment for fracture at an outside facility, (2) pathologic fracture, (3) cognitive impairment, or (4) language other than English. Data collection was performed by reviewing electronic medical records, operative reports, and radiology reports. Eligible patients were invited to participate by a recruitment letter. Questionnaires were administered through telephone interviews or collected online and managed using REDCap (Research Electronic Data Capture). REDCap is a secure, web-based application designed to support data collection for clinical research studies.[7] Informed consent was obtained from all subjects and approval was granted by the Institutional Review Board.

Outcome measures and explanatory variables

Electronic medical records were reviewed to collect baseline demographic characteristics, body mass index (BMI), smoking status, diabetes, Charlson Comorbidity Index (CCI), trauma date, trauma mechanism, time from injury to surgery, fracture and treatment characteristics. BMI was considered if reported within a range of six months prior to or after ORIF. Smoking status was considered positive if the patient was a smoker at the time of fixation. The CCI is a method of categorizing and indexing multiple comorbidities.[8] Injury mechanisms were subdivided into low-energy or high-energy and classified according to the Advanced Trauma Life Support guidelines. High-energy trauma (HET) mechanisms were defined as falls from height, crush injuries, motor vehicle and motorcycle accidents.[9,10] AO/OTA 41-C or Schatzker V/ VI tibial plateau fracture classification was confirmed by two orthopedic surgeons.

The primary outcome measure, patient-reported functional outcome, was assessed at least 12 months following ORIF using the PROMIS Physical Function (PROMIS PF). PROMIS was created to standardize the measurement and reporting of health outcomes to improve patient-reported outcome assessment for research and clinical practice. The PROMIS PF short-form-10 questionnaire consists of ten questions with five response options, assessing limitation and difficulty with certain physical activities, with higher scores representing higher physical function.

The PROMIS PF questionnaire measures the domain of physical functioning, with a mean score of 50 being representative of the general population of the United States.[11,12]

Secondary outcomes were HrQoL measured with the EuroQol 5-Dimensions 3-Levels (EQ-5D-3L) questionnaire, infection rate, and total knee arthroplasty (TKA) rate. The EQ-5D-3L is a five-item questionnaire that measures general health status, with a higher score representing a better quality of life. The EQ-5D-3L includes five dimensions, mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The EQ-5D-3L scores were calculated using a scoring algorithm, with a mean score of 0.88 being representative of the general population of North-American.[13-15] Infections were subdivided in superficial or deep wound infection. Superficial infection was defined as surgical site infection that was treated with antibiotics alone. If surgical irrigation and debridement was required, it was considered a deep infection.

Statistical analysis

Descriptive results are presented as mean values with standard deviations and range (SD, range), median values with interquartile range (IQR) or absolute numbers and percentages (%). Distribution of continuous explanatory and outcome variables were assessed using the Shapiro-Wilk test. Differences in baseline characteristics between responders and non-responders were compared. Continuous variables were evaluated using an independent sample t-test or Mann–Whitney U test. Categorical variables were compared using the Pearson's chi-squared test. The correlations between the PROMIS PF and EQ-5D-3 L outcome measures was assessed using Pearson's correlation coefficient. PROMIS PF and EQ-5D-3 L scores of the study population were compared with the norms for a general North-American population using the independent sample t-test.[12,15] The association between individual predictors and the outcomes measures were determined by bivariate linear regression analyses. Multivariable linear regression analyses were performed to identify factors associated with the outcome measures. To avoid overfitting, the final multivariable linear regression models were selected by forward stepwise regression. In this approach, individual predictors associated with the outcome measure with a p-value <0.1 in the bivariate analyses were included one by one in the multivariable regression model. Predictors no longer significantly associated with the outcome were omitted, only if doing so did not increase the deviance of the model. In the last step, individual predictors initially excluded after bivariate analyses were reincorporated in the multivariable regression model, only if doing so reduced the

overall deviance of the model.[16] The significance level was defined as a p value <0.05. All statistical analyses were performed using STATA® 13.1 (StataCorp LP, TX, USA).

Results

In total, 450 patients met the inclusion criteria. However, 236 patients could not be contacted or refused participation. This resulted in the inclusion of 214 patients (48%) for analysis. The responders were significantly older; 53 years compared with 49 years in the non-responder group (p = 0.004). The responders consisted of less males (50% versus 69%) (p= <0.001), less active smokers (17% versus 28%) (p = 0.005), and more patients with diabetes (11% versus 3%) (p = 0.002) compared with the non-responders. There was no statistical difference in the year of injury between responders 2010 (IQR; 2006–2013) and non-responders 2009 (IQR; 2006–2013) (p = 0.188). The baseline responder characteristics are shown in Table 1.

Table 1. Baseline characteristics (n= 214)

	Mean	**SD**
Age (years)	53	13
BMI (n= 198)*	27	(24-30)
	N	%
Gender		
Male	107	50
Female	107	50
Smoking (n= 210)		
Yes	36	17
No	178	83
Diabetes		
Yes	24	11
No	190	89
CCI		
0	76	36
1	56	26
2	44	21
3	17	8
≥ 4	21	10
Fracture side		
Left	111	52
Right	103	48
Open fracture		
Yes	17	8
No	197	78
Mechanism		
Fall - low energy	79	37
Fall - high energy	34	16
Motor vehicle crash	40	19
Motorcycle crash	25	12
Bicycle accident	2	1
Sports-related	19	9
Other	15	7
HET		
Yes	110	51
No	104	49

* Median (IQR); BMI, Body Mass Index; CCI, Charlson Comorbidity Index; HET, High-Energy Trauma; Percentages may not add up to 100 due to rounding

Table 2. Treatment and outcome measures (n=214)

	Median	IQR
Time injury to ORIF (days)	3	1-7
Time external fixation to ORIF (days)	5	4-11
Time injury to infection (days)	15	13-23
Time ORIF to TKA (months)	23	13-29
Time to questionnaire (months)	86	48-134
PROMIS PF	49.8	42-54
EQ-5D-3L	0.83	0.78-1.0
	N	**%**
Time to questionnaire distribution (months)		
12-24	19	9
24-48	35	16
48-72	40	19
>72	120	56
Approaches		
Anterior	48	22
Lateral	115	54
Medial	33	15
Posterior	4	2
Posteromedial	14	7
External fixation		
Yes	61	29
No	153	72
Infection		
No	188	88
Superficial	6	3
Deep	20	9
TKA		
Yes	6	3
No	208	97

ORIF, Open reduction internal fixation; TKA, Total knee arthroplasty; PROMIS PF, Patient-Reported Outcomes Measurement Information System Physical Function; EQ-5D-3L, EuroQol 5-Dimensions 3-Levels; Percentages may not add up to 100 due to rounding.

The mean age at injury for the population of responders was 53 years (SD 13, range 24–89) and 107 patients (50%) were male. Open fractures were sustained by 17 patients (8%). The mechanism of injury was a low-energy fall for 79 patients (37%), with 110 patients (51%) involved in high-energy trauma mechanisms. Treatment and outcome measures are shown in Table 2. The median time from injury to fixation was 3 days (IQR; 1–7) and 61 fractures (29%) were treated with temporizing external fixation. Infection occurred in 26 cases (12%), with 6 superficial infections (3%) and 20 deep infections (9%). TKA was performed in 6 patients (3%) after a median duration of 23 months following ORIF (IQR; 13–29).

PROMs

The questionnaires were completed after a median duration of 86 months from injury (IQR; 48–134) (Table 2). The questionnaires were completed by 120 patients (56%) after more than 72 months following injury, 40 patients (19%) between 48 and 72 months, 35 patients (16%) between 24 and 48 months, and 19 patients (9%) between 12 and 24 months. The mean PROMIS PF score was 47.7 (SD 9.5), significantly lower compared with the mean score of 50 for

Table 3. PROMIS PF items targeting walking or mobility. (n=214)

	Median	IQR
Does your health now limit you in doing vigorous activities, such as running, lifting heavy objects, participating in strenuous sports?	3	2-4
Does your health now limit you in walking more than a mile?	5	3-5
Does your health now limit you in climbing one flight of stairs?	5	4-5
Does your health now limit you in bending, kneeling, or stooping?	4	3-5

PROMIS PF, Patient-Reported Outcomes Measurement Information System Physical Function; Items scored 1 (Cannot do), 2 (Quite a lot), 3 (Somewhat), 4 (Very little), and 5 (Not at all), with higher scores representing higher physical function

Table 4. Multivariable regression analyses

	β regression coefficient	95% CI	p-value
PROMIS PF			
Age	-0.080	-0.168 – 0.006	0.068
Sex			
Male	Ref	–	–
Female	-2.857	-5.034 – -0.680	0.010
BMI	-0.135	-0.310 – -0.039	0.128
Smoking	-0.063	-2.605 – -2.478	0.961
Diabetes	-3.635	-6.576 – -0.694	0.016
Infection	-0.676	-3.568 – 2.216	0.645
TKA	-1.691	-7.168 – 3.785	0.542
EQ-5D-3L	30.910	25.893 – 35.927	<0.001
EQ-5D-3L			
Age	-0.001	-0.003 – 0.000	0.211
Sex			
Male	Ref	–	–
Female	-0.048	-0.104 – 0.007	0.091
Smoking	-0.093	-0.160 – -0.026	0.006
Diabetes Mellitus	-0.086	-0.167 – -0.004	0.038
Infection	-0.057	-0.134 – 0.018	0.139
TKA	-0.212	-0.361 – -0.063	0.005

PROMIS PF, Patient-Reported Outcomes Measurement Information System Physical Function; EQ-5D-3L, EuroQol 5-Dimensions 3-Levels; BMI, Body Mass Index; TKA, Total knee arthroplasty

the reference population score (P= <0.001). The median scores of different items of the PROMIS PF targeting walking or lower extremity mobility are shown in Table 3. The mean EQ-5D-3L was 0.82 (SD 0.2), significantly lower compared with the mean score of 0.88 for the reference population score (p= <0.001). The PROMIS PF and EQ-5D-3L outcome measures were correlated (correlation coefficient = 0.75, p= <0.001). In the bivariate linear regression analyses age, BMI, female gender, diabetes, CCI, time to questionnaire, and TKA were found to be independently associated with worse physical function measured with the PROMIS PF (Supplementary Table 1). Factors associated with lower HrQoL measured by the EQ-5D-3 L score in the bivariate linear regression analyses were age, BMI, female gender, active smoking status, diabetes, CCI, time to questionnaire, and TKA (Supplementary Table 1).

In multivariable regression analyses, female gender (regression coefficient (β) −2.857; 95% Confidence Interval (CI) −5.034 to −0.680; p = 0.010), diabetes (β –3.635; 95% CI −6.576 to −0.694; p = 0.016), and lower HrQoL assessed using the EQ-5D-3L (β 30.910; 95% CI 25.893 –

35.927; p= <0.001) were found to be independently associated with worse physical function measured with the PROMIS PF (Table 4). Factors associated with lower EQ-5D-3 L score in the multivariable regression model were active smoking status (β −0.093; 95% CI −0.160 −0.026; p = 0.006), diabetes (β −0.089; 95% CI −0.167 to −0.004; p = 0.038), and TKA (β −0.212; 95% CI −0.361 to −0.063; p = 0.005).

Discussion

To our knowledge, with the inclusion of 214 patients, this study is the largest cohort to establish normative data and evaluate long-term patient-reported physical function outcome and HrQoL after ORIF of bicondylar tibial plateau fractures. Infection occurred in 26 cases (12%), with 6 superficial infections (3%) and 20 deep infections (9%). TKA following initial ORIF was performed in 6 patients (3%). Both patient-reported physical function and HrQoL were significantly lower compared with the age-by-gender norms for a general North American population. The multivariable regression model revealed female gender, diabetes, and worse HrQoL were correlated with worse functional outcome. The multivariable regression model revealed smoking, diabetes, and the subsequent need for TKA to be correlated with worse HrQoL.

Previous studies have reported impaired functional outcome and HrQoL after bicondylar tibial plateau fractures. Jansen et al.[17] reported the medium-term results of 22 patients with 23 AO/OTA type C fractures of the tibial plateau. They concluded that complex articular tibial plateau fractures continue to have a severe impact on function in the injured knee, with an average Lysholm score of 66.2 and an average KOOS score of 67.84. Timmers et al.[18] presented the results after a mean of 6 years of 82 patients after ORIF of tibial plateau fractures, 46 with Schatzker I–IV and 17 with Schatzker V–VI fractures. They evaluated functional outcome with the KOOS questionnaire and HrQoL using the EuroQol-6D questionnaire. Their overall cohort had a "Fair" functional knee outcome and HrQoL was lower in comparison to the general Dutch population.[18] Rohra et al.[19] presented the functional results of 34 Schatzker type V and VI tibial plateau fractures using The Knee Society Score after treatment with dual plates after a minimum of 3 years. They reported 24 patients (71%) with an Excellent, 8 patients (24%) with Good, 1 patient (3%) with Fair, and 1 patient (3%) with Poor functional Knee Society Scores. Cavallero et al.[20] compared outcomes between locking (n = 29) and nonlocking constructs (n = 27) for the treatment of bicondylar tibial fractures and they reported a PROMIS PF score of 39 and 41,

respectively. Virkus et al.[21] reported a mean PROMIS PF scores of 40, for both 1-stage definitive fixation (n = 28) and 2-stage fixation (n = 24) after initial spanning external fixation for bicondylar tibial fractures.

In the current study, both the PROMIS PF (47 vs. 50) and EQ-5D-3L (0.82 vs. 0.88) were significantly lower compared with the age-by-gender norms for a general North American population. However, these lower scores are likely not clinically significant, as the minimum clinically important difference of the EQ-5D-5 L in patients with hip or knee osteoarthritis among surgical patients has been shown to be 0.32.[22] The PROMIS PF has been shown to have a minimum clinically important difference of 15.98 for patients with knee injuries.[23] In the current study, the PROMIS PF items targeting walking or mobility showed patients had no difficulty in walking more than a mile or climbing flight of stairs. However, patients were limited in doing vigorous activities. Unfortunately, comparison of literature remains difficult due a wide variety of functional outcome scores and the lack of validated patient-centered outcome measures. Further research is needed to focus on specific fracture types to optimize patient-reported outcomes.[5]

The impaired long-term functional outcome and HrQoL in previous studies could be the result of the complexity and severity of bicondylar tibial fractures.[4,5] Bicondylar tibial plateau fractures are usually caused by high-energy trauma mechanisms as a result of motor vehicle collisions, falls from height, motorcycle collisions, and pedestrians being struck by vehicles.[5] These fractures are associated with substantial soft-tissue injuries, and 8% to 43% of bicondylar tibial plateau injuries are presented as open fractures.[5] In our study, open fractures were sustained by only 17 patients (8%), and 110 patients (51%) were involved in high-energy trauma mechanisms. Our results for open fractures and energy of trauma mechanisms showed no association in the bivariate linear regression analyses with long-term physical function or HrQoL.

Operative treatment of bicondylar tibial plateau fracture has been associated with complications such as deep infection, non-union and the need of revision surgery, with an overall high complication rate varying from 28% to 39%.[5] Khatri et al.[24] evaluated 65 patients with Schatzker type V and type VI tibial plateau fractures treated by ORIF. They reported superficial wound infections in 9.2% of patients and 4.6% with deep wound infections. In the current cohort, in accordance with previous literature, infection occurred in 26 cases (12%), with 6 superficial infections (3%) and 20 deep infections (9%). The need for subsequent arthroplasty surgery also

low, with in 6 patients (3%) needing TKA following initial ORIF. However, our study demonstrates that the development of infection and the need for TKA, while suboptimal in the treatment course and recovery, were not associated with worse long-term functional outcome in the multivariable regression model. However, TKA was associated with worse HrQoL in the multivariable regression model. Comparison of literature still remains difficult, with different reports of infection rates, due to the use of different surgical techniques and a variety of approaches.[5]

In the current study, female gender, diabetes, and HrQoL were correlated with worse knee function in the multivariable regression model. To our knowledge, this study is the first to evaluate factors associated with patient-reported functional outcome. In our study, HrQoL was correlated with functional outcome, emphasizing the importance of obtaining both general global health measures and injury specific measures when evaluating outcomes after injuries. A previous study has shown strong correlation between global health measures and injury specific functional scores.[25] Different fracture characteristics of bicondylar tibial plateau fractures contribute to the potential for poor outcome such as associated soft-tissue injury and concomitant injuries. By adjusting injury specific functional outcome measures for general health measures, functional outcome scores might be assessed in the right context when evaluating treatment. This could be important for the evaluating of injuries that occur in the context of high-energy trauma mechanisms, concomitant injuries, and heterogeneous patient populations.

This study has several limitations. First, the study is limited by the retrospective nature. This study did not include prospective collection of functional and radiological measures during standardized follow-up times, which would increase the understanding of the impact treatment and recovery have on patient-reported functional outcome and HrQoL. Second, we were not able to account for all variables that could potentially influence the outcome measures. Therefore, the factors identified to be correlated with our outcome measures should not be considered as the only factors effecting patient-reported functional outcome and HrQoL after ORIF of bicondylar tibial plateau fractures. Third, due to the relative long interval between treatment and follow-up, 91% of patients had a follow up of >24 months, functional outcome scores could be influenced by other conditions, events, or patient factors. Although time to questionnaire was associated with both the PROMIS PF and EQ-5D-3L in the bivariate analyses, there was no association with worse outcome in the multivariable regression model. Fourth, the response rate

was relatively low (48%). However, to our knowledge, with the inclusion of 214 patients, this study is the largest cohort to establish normative data and evaluate long-term outcome of bicondylar tibial plateau fractures focusing on AO/OTA 41-C or Schatzker V/ VI. Fifth, there were several differences in baseline characteristics between the responders and non-responders. The responders consisted of significantly less males, less active smokers, and more patients with diabetes, compared with the non-responders. Therefore, the effect of diabetes on functional outcome and HrQoL might be an overestimation and may not be generalized to all tibial plateau fracture patients. However, the effect of smoking on HrQoL might be an underestimation of the true impact of tobacco use.

Unfortunately, comparison of literature remains difficult due to a wide variety of AO fracture types, operative treatments, approaches, PROMs, and duration to follow-up, indicating a substantial need for further research. We suggest future research to focus on factors that might contribute to the potential for poor outcome such as associated soft-tissue injury and concomitant injuries (poly-trauma). Furthermore, we suggest the prospective collection of functional and radiological measures during standardized follow-up times, which would increase the understanding of the impact treatment and recovery have on patient-reported functional outcome and HrQoL.

Conclusion

Both the PROMIS PF and EQ-5D-3L were lower compared with the age-by-gender norms for a general North American population, however, did not reach a minimum clinically important difference. The PROMIS PF items revealed patients had no difficulty in walking more than a mile or climbing a flight of stairs. However, patients were limited in doing vigorous activities and patients should be counseled about the expected long-term outcomes. Factors that may influence worse functional outcome following ORIF of bicondylar tibia fractures are female gender, diabetes, and patients with lower HrQoL. This study emphasizes the correlation between injury specific functional outcome measures and general health measures. By adjusting injury specific functional outcome measures for general health measures, functional outcome scores might be assessed in the right context.

References:

1. Schatzker J, McBroom R, Bruce D. The tibial plateau fracture. The Toronto experience 1968--1975. *Clinical orthopaedics and related research.* 1979(138):94-104.
2. Kellam JF, Meinberg EG, Agel J, Karam MD, Roberts CS. Introduction: Fracture and Dislocation Classification Compendium-2018: International Comprehensive Classification of Fractures and Dislocations Committee. *Journal of orthopaedic trauma.* 2018;32 Suppl 1:S1-S10.
3. Millar SC, Arnold JB, Thewlis D, Fraysse F, Solomon LB. A systematic literature review of tibial plateau fractures: What classifications are used and how reliable and useful are they? *Injury.* 2018;49(3):473-490.
4. Mthethwa J, Chikate A. A review of the management of tibial plateau fractures. *Musculoskeletal surgery.* 2018;102(2):119-127.
5. Lee AK, Cooper SA, Collinge C. Bicondylar Tibial Plateau Fractures: A Critical Analysis Review. *JBJS reviews.* 2018;6(2):e4-e4.
6. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *The Journal of the American Academy of Orthopaedic Surgeons.* 2016;24(11):744-749.
7. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics.* 2009;42(2):377-381.
8. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *American journal of epidemiology.* 2011;173(6):676-682.
9. Advanced trauma life support (ATLS(R)): the ninth edition. *The journal of trauma and acute care surgery.* 2013;74(5):1363-1366.
10. Linder F, Mani K, Juhlin C, Eklöf H. Routine whole body CT of high energy trauma patients leads to excessive radiation exposure. *Scandinavian journal of trauma, resuscitation and emergency medicine.* 2016;24:7.
11. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Medical care.* 2007;45(5 Suppl 1):S3-S11.
12. Cella D, Riley W, Stone A, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *Journal of Clinical Epidemiology.* 2010;63(11):1179-1194.
13. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Annals of medicine.* 2001;33(5):337-343.
14. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Medical care.* 2005;43(3):203-220.
15. Fryback DG, Dunham NC, Palta M, et al. US norms for six generic health-related quality-of-life indexes from the National Health Measurement study. *Medical care.* 2007;45(12):1162-1170.
16. Shankaran V, Jolly S, Blough D, Ramsey SD. Risk factors for financial hardship in patients receiving adjuvant chemotherapy for colon cancer: a population-based exploratory analysis. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2012;30(14):1608-1614.
17. Jansen H, Frey SP, Doht S, Fehske K, Meffert RH. Medium-term results after complex intra-articular fractures of the tibial plateau. *Journal of orthopaedic science : official journal of the Japanese Orthopaedic Association.* 2013;18(4):569-577.
18. Timmers TK, van der Ven DJC, de Vries LS, van Olden GDJ. Functional outcome after tibial plateau fracture osteosynthesis: a mean follow-up of 6 years. *The Knee.* 2014;21(6):1210-1215.
19. Rohra N, Suri HS, Gangrade K. Functional and Radiological Outcome of Schatzker type V and VI Tibial Plateau Fracture Treatment with Dual Plates with Minimum 3 years follow-up: A Prospective Study. *Journal of clinical and diagnostic research : JCDR.* 2016;10(5):RC05-10.
20. Cavallero M, Rosales R, Caballero J, Virkus WW, Kempton LB, Gaski GE. Locking Plate Fixation in a Series of Bicondylar Tibial Plateau Fractures Raises Treatment Costs Without Clinical Benefit. *Journal of orthopaedic trauma.* 2018;32(7):333-337.
21. Virkus WW, Caballero J, Kempton LB, Cavallero M, Rosales R, Gaski GE. Costs and Complications of Single-Stage Fixation Versus 2-Stage Treatment of Select Bicondylar Tibial Plateau Fractures. *Journal of orthopaedic trauma.* 2018;32(7):327-332.
22. Bilbao A, Garcia-Perez L, Arenaza JC, et al. Psychometric properties of the EQ-5D-5L in patients with hip or knee osteoarthritis: reliability, validity and responsiveness. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation.* 2018;27(11):2897-2908.
23. Hung M, Bounsanga J, Voss MW, Saltzman CL. Establishing minimum clinically important difference values for the Patient-Reported Outcomes Measurement Information System Physical Function, hip disability and osteoarthritis outcome score for joint reconstruction, and knee injury and osteoarthritis ou. *World journal of orthopedics.* 2018;9(3):41-49.
24. Khatri K, Lakhotia D, Sharma V, Kiran Kumar GN, Sharma G, Farooque K. Functional Evaluation in High Energy (Schatzker Type V and Type VI) Tibial Plateau Fractures Treated by Open Reduction and Internal Fixation. *International scholarly research notices.* 2014;2014:589538-589538.
25. Turchin DC, Schemitsch EH, McKee MD, Waddell JP. Do foot injuries significantly affect the functional outcome of multiply injured patients? *Journal of orthopaedic trauma.* 1999;13(1):1-4.

Supplementary materials to Chapter 11

Table S1. Bivariate linear regression analyses

	PROMIS PF			EQ-5D-3L		
	β	**95% CI**	**p-value**	**β**	**95% CI**	**p-value**
Age	-0.191	-0.284 – -0.098	<0.001	-0.002	-0.004 – -0.001	0.007
BMI	-0.440	-0.692 – -0.187	0.001	-0.005	-0.011 – -0.001	0.039
Sex Female	-4.955	-7.427 – -2.482	<0.001	-0.061	-0.113 – -0.0098	0.020
Smoking	-2.975	-6.388 – 0.437	0.087	-0.09	-0.160 – -0.023	0.009
Diabetes	-7.306	-11.313 – -3.300	<0.001	-0.102	-0.185 – -0.019	0.016
CCI	-2.072	-3.021 – -1.124	<0.001	-0.023	-0.043 – -0.003	0.024
Open fracture	0.586	-4.140 – 5.3124	0.807	0.061	-0.034 – 0.156	0.209
HET	2.138	-0.407 – 4.684	0.099	0.031	-0.020 – 0.083	0.236
Time injury to ORIF (days)	0.015	-0.046 – 0.077	0.632	0.001	-0.001 – 0.001	0.455
Time ex-fix to ORIF (days)	0.079	-0.304 – 0.463	0.680	0.001	-0.006 – 0.009	0.695
Time injury to infection (days)	0.031	-0.033 – 0.096	0.287	0.001	-0.001 – 0.002	0.318
Time ORIF to TKA (months)	-0.005	-0.337 – 0.326	0.964	-0.001	-0.009 – 0.007	0.818
Time to questionnaire (months)	0.001	0.001 – 0.001	0.004	0.001	0.001– 0.001	0.023
Approach	0.012	-0.764 – 0.789	0.975	0.002	-0.013 – 0.017	0.792
Ex-fix	-2.385	-5.215 – 0.443	0.098	-0.029	-0.087 – 0.028	0.317
TKA	-10.252	-17.868 – -2.635	0.009	-0.227	-0.381 – -0.073	0.004
Infection	-3.099	-6.990 – 0.096	0.287	-0.076	-0.155 – 0.002	0.058

PROMIS PF, Patient-Reported Outcomes Measurement Information System Physical Function; EQ-5D-3L, EuroQol 5-Dimensions 3-Levels; BMI, Body Mass Index; CCI, Charlson Comorbidity Index; HET, High-Energy Trauma; ORIF, Open reduction internal fixation; Ex-fix externa fixation; TKA, Total knee arthroplasty; β regression coefficient

CHAPTER 12

Validation of PROMIS Physical Function for evaluating outcome following acute Achilles tendon rupture

Yassine Ochen, Daniel Guss, R. Marijn Houwert, Jeremy T. Smith, Christopher W. DiGiovanni, Rolf H.H. Groenwold, Marilyn Heng

(Orthopaedic Journal of Sports Medicine)

Abstract

Background

There is increased demand for valid, reliable, and responsive patient-reported outcome measures (PROMs) to evaluate patients with an Achilles tendon rupture, but not all PROMs currently in use are reliable and responsive for this condition. The primary aim of this study was to evaluate the measurement properties of the more recent Patient Reported Outcomes Measurement Information System Physical Function (PROMIS PF) compared to different PROMs used in patients with an acute Achilles tendon rupture.

Methods

A retrospective cohort study with follow-up by questionnaire was performed using data from two academic centers. All adult patients with an acute Achilles tendon rupture between June 2016 and June 2018 with a minimum of 12 months follow-up were eligible for inclusion. Functional outcome was assessed using the Patient-Reported Outcomes Measurement Information System Physical Function (PROMIS PF) Computerized Adaptive Test (CAT), Foot and Ankle Ability Measure (FAAM) Activities of Daily Living (ADL), FAAM Sports, and Achilles Tendon Total Rupture Score (ATRS). Pearson's correlation (r) was used to assess the correlations between outcome measures. Absolute and relative floor and ceiling effects were calculated.

Results

In total, 103 patients were included. The mean age was 44.7 years (range 19-77) and 76 patients (74%) were male. A total of 82 patients (80%) underwent operative repair while the remainder, 21 patients (20%) underwent nonoperative management. The mean time between treatment to collection of PROMs was 25.3 months (range 15-36). The mean PROMIS PF was 55.4 (SD 9.2), FAAM ADL 92.9 (SD 12.2), FAAM Sports 77.7 (SD 22.9), and ATRS 83 (SD 19.4). The ATRS was correlated with FAAM ADL (r 0.80; 95%CI 0.72; 0.86; $p<0.001$) and FAAM Sports (r 0.86; 95%CI 0.80; 0.90; $p<0.001$). The PROMIS PF was correlated with the FAAM ADL (r 0.66; 95%CI 0.53; 0.75; $p<0.001$), FAAM Sports (r 0.65; 95%CI 0.53; 0.75; $p<0.001$), and ATRS (r 0.69; 95%CI 0.58; 0.78; $p<0.001$). The PROMIS PF did not show absolute floor or ceiling effects (0%). The FAAM ADL (35.9%), FAAM Sports (15.8%), and ATRS (20.4%) had substantial absolute ceiling effects.

Conclusion

The PROMIS PF, FAAM ADL, and FAAM Sports all showed a moderate to high mutual correlation with the ATRS. Notably, however, of all these measures only PROMIS PF avoided substantial floor and ceiling effects. The results of this study suggest the PROMIS PF CAT can be considered a valid, reliable and perhaps the most responsive tool to evaluate patient outcomes after treatment of an Achilles tendon rupture.

Introduction

The incidence of Achilles tendon ruptures is rising not only among young patients, but also among an increasingly aging, but active, population.[1] The role of operative versus nonoperative management remains controversial, but determining the most effective solution for any given patient depends upon patient-reported outcome measurement (PROM) tools that are able to reliably evaluate the success of a chosen clinical treatment strategy.[2-6]

The Achilles Tendon Total Rupture Score (ATRS) is the most commonly used PROM to evaluate outcomes after the treatment of an acute Achilles rupture because it was the first validated, injury-specific PROM.[6-8] The Foot and Ankle Ability Measure (FAAM) is used to evaluate a myriad of lower extremity disorders, and has also been shown to have substantial content relevance to patients with Achilles tendon disorders.[9,10] The more recently developed Patient-Reported Outcomes Measurement Information System (PROMIS) provides a comprehensive set of questionnaires and, critically, items can be administered as a "Computerized Adaptive Test" (CAT) to limit the number of questions that a patient must answer to attain a score. The PROMIS Physical Function (PROMIS PF) CAT has shown to be an excellent method for measuring outcomes for patients with foot and ankle injuries.[11,12] While all the aforementioned instruments are currently employed to evaluate treatment of lower extremity conditions, the correlation between the validated ATRS, FAAM, and PROMIS PF CAT scores in patients with Achilles tendon ruptures has not been evaluated.[13-15]

The primary aim of this retrospective study was to evaluate the validity, reliability and responsiveness of the PROMIS PF, the FAAM, and the ATRS measurement tools in patients with an acute Achilles tendon ruptures.

Methods

Study design

All adult patients who presented to two academic medical centers with an acute Achille tendon rupture between 2016 and 2018 were eligible for inclusion. Eligible patients were identified by searching for Current Procedure Terminology (CPT) codes and International Classification of Diseases (ICD) codes in the institution's Research Patient Data Registry (RPDR). Inclusion criteria were: (1) acute Achilles tendon rupture, (2) 18 years or older, (3) minimum of 12 months follow-up. Exclusion criteria were: (1) treatment for Achilles re-rupture, (2) cognitive impairment,

(3) and language other than English. Data collection was performed by reviewing electronic medical records and, after Institutional Review Board approval, eligible patients were invited to participate in the study by a recruitment letter. Questionnaires were collected online and managed using Research Electronic Data Capture (REDCap).[16]

Patient and treatment characteristics

Electronic medical records and collected REDCap questionnaires were reviewed to collect baseline demographic characteristics regarding age, sex, smoking status, other surgery on the affected leg since initial Achilles treatment, trauma mechanism, Charlson Comorbidity Index (CCI), operative treatment method, nonoperative treatment method, and the time from treatment to questionnaire. Smoking status was subdivided into current, former, and never smoker. The CCI is a method of categorizing and indexing multiple comorbidities.[17] Operative treatment included open and minimally invasive/percutaneous surgery. The operative stitch technique was recorded if noted in the operative report and included Bunnell, Kessler, Krackow, End-to-end, Lindholm/Ma-Griffith, and Kessler/Percutaneous. Immobilization methods used included the use of a cast, boot, or splint. The time from initiation of treatment to the start of rehabilitation was collected. Full weight bearing status was divided in less than 4 weeks and 4 weeks or greater. The use of a functional rehabilitation protocol was recorded (e.g. gradual reduction of plantar flexion, self-administered exercise program, or formal physiotherapy) as well as the use of an accelerated rehabilitation protocol (start early range of motion less than 3 weeks).

PROMs

The collection of PROMs was performed electronically and included the PROMIS Physical Function (PF) v2.0 Computerized Adaptive Test (CAT), FAAM Activities of Daily Living (ADL), FAAM Sports, and the ATRS. The PROMIS questionnaires evaluate the limitations of daily activities, pain, and physical activities, with scores ranging from 0 to 100, with higher scores representing higher function, and a mean score of 50 for the general population of the United States.[18] The PROMIS PF CAT was developed using item response theory to maximize efficient administration from a calibrated items bank of 124 question, a minimum number of 4 items must be answered in order to receive a score. The minimal clinical important difference (MCID) of the PROMIS PF CAT is 16 points. The FAAM is developed to assess physical function for individuals with foot and ankle related disabilities, items are scored on a 5-point Likert scale from "no difficulty at all" to "unable to do", scores are transformed to percentage scores, with higher

scores represent higher levels of functioning.[10] The scores for the FAAM ADL and Sports are regarded valid and generated when subjects complete 90% or more of the items. The MCID of the FAAM ADL/Sports has been reported to be 8 and 9 points, respectively. The ATRS is an instrument developed specifically for measuring outcome after treatment for Achilles tendon ruptures, with items graded on a 11-point Likert scale according to level of limitations and/or difficulties from "major limitations" to "no limitations, with a score of 100 indicating no symptoms and full function.[7] The scores for the ATRS are regarded valid and generated when subjects complete 80% or more of the items. The ATRS has a reported MCID of 10 points. Currently, the ATRS has been identified as the most appropriate PROM to evaluate the management of Achilles tendon ruptures, and thus considered to be the primary comparator.[8,19,20] The time from treatment to questionnaire in months was available for all PROMs. Patients completed all the PROMs questionnaire electronically at the same time, and completed the minimum valid answers required to compute the scores.

Statistical analysis

Descriptive results were presented as mean values with standard deviations and range (SD, range), median values with interquartile range (IQR), or absolute numbers and percentages (%). Pearson's correlation (r), with 95% confidence interval (CI), was used to assess the relationship between the PROMIS PF, FAAM ADL, FAAM Sports, and the ATRS. Correlation coefficients of 0.3 or less were considered weak, 0.31 to 0.39 as moderate-weak, 0.40 to 0.60 as moderate, 0.61 to 0.69 as moderate-high, and larger than 0.70 as high.[21] Additionally, floor and ceiling effect were assessed for all PROMs. Absolute floor was defined as the percentage of patients with the absolute lowest possible PROM score, and absolute ceiling as the percentage with the absolute highest possible PROM score. Relative floor was defined as the percentage of patients that reported the lowest PROM score in the cohort, and relative ceiling as the percentage with the highest PROM score reported in the cohort. Floor or ceiling effects are considered to be substantial if more than 15% of patients achieve the lowest or highest possible score, respectively.[22] The required sample size for studies assessing measurement properties has been advocated to be a sample size of at least 50 patients.[22] The significance level was defined as a p value <0.05. All analyses were performed in R version 3.6.1 (R Development Core Team, Released 2013, Vienna, Austria: R Foundation for Statistical Computing).[23]

Results

Study population

In total, 305 patients met the inclusion criteria, of whom 179 patients (59%) did not respond, and 23 patients (8%) refused participation. This led to the final inclusion of 103 patients (overall response rate of 34%). The different PROMs questionnaires were completed by patients at the same timepoint. The mean time from treatment to PROMs completion was 25.3 months (range 15–36). The patient characteristics, stratified by treatment method, are presented in Table 1.

Treatment method

In total, 82 patients (80%) underwent operative repair. The treatment characteristics are shown in Table 2. Open surgery was performed in 69 patients (86%), with the "Krackow" as most used stitch technique (41%). The median duration of operative treatment to start of rehabilitation was 2.0 weeks (IQR 2.0-5.5). Nonoperative treatment was performed in 21 patients (20%), with "Boot" (48%) used as most common method. The mean duration of nonoperative treatment to start of rehabilitation was 4.5 weeks (IQR 3.0-9.3). The treatment characteristics are shown in Table 2.

Table 1. Characteristics of 103 Achilles tendon rupture patients.

	Overall	**Operative**	**Nonoperative**
Patients	103	82	21
Age injury	44.7 (14.6, 19-77)	42.3 (12.9, 19-74)	54.1 (17.2, 25-77)
Sex (%)			
Male	76 (73.8)	63 (76.8)	13 (61.9)
Female	27 (26.2)	19 (23.2)	8 (38.1)
Smoking (%)			
Current	8 (7.8)	6 (7.3)	2 (9.5)
Former	14 (13.6)	11 (13.4)	3 (14.3)
Never	81 (78.6)	65 (79.3)	16 (76.2)
Other surgery on leg since Achilles treatment (%)	7 (6.8)	6 (7.3)	1 (4.8)
Trauma mechanism (%)			
Sports-related	89 (86.4)	73 (89.0)	16 (76.2)
Ground level fall	5 (4.9)	4 (4,9)	1 (4.8)
Fall from height	1 (1.0)	0 (0.0)	1 (4.8)
Twisting motion	5 (4.9)	3 (3.7)	2 (9.5)
Other	3 (2.9)	2 (2.4)	1 (4.8)
CCI index overall	2.0 (1.3, 1-7)	1.8 (1.0, 1-5)	2.7 (1.8, 1-7)
Treatment to PROMs (months)	25.3 (5.7, 15-36)	25.9 (5.7, 15-36)	23.2 (5.1, 16-31)

Continuous variables presented as mean (SD, range); CCI Charlson Comorbidity Index;
PROMs patient-reported outcome measures

Table 2. Treatment characteristics of 103 Achilles tendon rupture patients.

Operative treatment	
Patients	82
Operative method (%) (n=80)	
Open surgery	69 (86.2)
Minimally invasive/percutaneous	11 (13.8)
Operative stitch technique (%) (n=59)	
Bunnell	1 (1.7)
Kessler	3 (5.1)
Krackow	24 (40.7)
End-to-end	1 (1.7)
Lindholm/Ma-Griffith	20 (33.9)
Kessler/Percutaneous	9 (15.3)
Other	1 (1.7)
Immobilization method (%)	
Cast	3 (3.7)
Boot	0 (0.0)
Splint	79 (96.3)
Full weight-bearing status (%) (n=80)	
<4 weeks	11 (13.8)
≥4 weeks	69 (86.2)
Time from treatment to rehabilitation (weeks) (n=79)	2.0 (2.0−5.5)
Functional rehabilitation protocol (%) (n=81)	79 (97.5)
Accelerated rehabilitation protocol (%) (n=80)	45 (56.2)
Nonoperative treatment	
Patients	21
Nonoperative method (%)	
Cast	4 (19.0)
Boot	10 (47.6)
Splint	7 (33.3)
Full weight-bearing status (%) (n=18)	
<4 weeks	3 (16.7)
≥4 weeks	15 (83.3)
Time from treatment to rehabilitation (weeks) (n=16)	4.5 (3.0−9.3)
Functional rehabilitation protocol (%) (n=17)	17 (100.0)
Accelerated rehabilitation protocol (%) (n=16)	4 (25.0)

Continuous variables presented as median (IQR)

PROMs measurement properties

The overall mean PROMs results were PROMIS PF 55.4 (SD 9.2), FAAM ADL 92.9 (SD 12.2), FAAM Sports 77.7 (SD 22.9), and ATRS 83.0 (SD 19.4). The PROMs stratified by treatment method are shown in Table 3. The mutual correlations between the different PROMs are presented in Table 4. The ATRS showed a high correlation with the FAAM ADL (r 0.80; 95%CI 0.72; 0.86; p<0.001) and with the FAAM Sports (r 0.86; 95%CI 0.80; 0.90; p<0.001). PROMIS PF showed a moderate-high correlation with the FAAM ADL (r 0.66; 95%CI 0.53; 0.75; p<0.001), FAAM Sports (r 0.65; 95%CI 0.53; 0.75; p<0.001), and ATRS (r 0.69; 95%CI 0.58; 0.78; p<0.001). The floor and ceiling effects for the PROMs are presented in Table 5. The PROMIS PF did not show absolute floor or ceiling effects (0%). The FAAM ADL (35.9%), FAAM Sports (15.8%), and ATRS (20.4%) had significant absolute ceiling effects. There were no substantial changes in relative floor and ceiling effects compared to the absolute floor and ceiling effects.

Table 3. Patient-reported outcome measures of 103 Achilles tendon rupture patients

	Overall	Operative	Nonoperative
Patients	103	82	21
PROMIS PF	55.4 (9.2)	56.4 (9.1)	51.5 (8.7)
FAAM ADL	92.9 (12.2)	93.6 (11.6)	90.3 (14.5)
FAAM Sports	77.7 (22.9)	78.7 (22.6)	73.5 (24.1)
ATRS	83.0 (19.4)	83.9 (19.5)	79.6 (19.5)

Continuous variables presented as mean (SD); PROMIS Patient-Reported Outcomes Measurement Information System; PF physical function; FAAM Foot and Ankle Ability Measure; ADL Activities of Daily Living; ATRS Achilles Tendon Total Rupture Score

Table 4. Correlations between patient-reported outcome measures of Achilles tendon ruptures

	PROMIS PF	FAAM ADL	FAAM Sports	ATRS
PROMIS PF	–	0.66 (0.53–0.75)	0.65 (0.53–0.75)	0.69 (0.58–0.78)
FAAM ADL		–	0.68 (0.56–0.77)	0.80 (0.72–0.86)
FAAM Sports			–	0.86 (0.80–0.90)
ATRS				–

Correlations are presented as Pearson's correlation with 95% confidence interval (CI); PROMIS Patient-Reported Outcomes Measurement Information System; PF physical function; FAAM Foot and Ankle Ability Measure; ADL Activities of Daily Living; ATRS Achilles Tendon Total Rupture Score. For all correlations $p < 0.001$.

Table 5. Patient-reported outcome measures floor and ceiling effects of Achilles tendon ruptures

	Absolute floor (%)	Absolute ceiling (%)	Relative floor (%)	Relative ceiling (%)
PROMIS PF	0	0	1.0	2.9
FAAM ADL	0	35.9	1.0	35.9
FAAM Sports	0	15.8	1.0	15.8
ATRS	0	20.4	1.0	20.4

Absolute floor/ceiling effect defined as percentage (%) with absolute lowest/highest possible PROM score; Relative floor/ceiling defined as percentage (%) with lowest/highest PROM score reported in cohort; PROMIS Patient-Reported Outcomes Measurement Information System; PF physical function; FAAM Foot and Ankle Ability Measure; ADL Activities of Daily Living; ATRS Achilles Tendon Total Rupture Score

Discussion

In this cohort study of both the operative and nonoperative functional treatment outcome of Achilles tendon ruptures, the FAAM ADL, FAAM Sports, and PROMIS PF all showed a moderate to high mutual correlation with the ATRS. Of these measures, however, it should be noted that only the PROMIS PF CAT avoided substantial floor as well as ceiling effects.

The overall PROMs results from this study demonstrate good to excellent long-term functional outcome following Achilles tendon treatment. The correlation between the ATRS, FAAM, and PROMIS scores in patients with an Achilles tendon rupture has not been previously evaluated. In this study, the ATRS showed a moderate to high correlation with the FAAM ADL, FAAM Sports, and PROMIS PF. The ATRS is an injury specific PROMs, which has been evaluated and found to be valid, reliable, and responsive. It has also been confirmed and validated in several languages, and many currently consider it the most appropriate PROM to evaluate the management of Achilles tendon ruptures.[8,19,20] Ganestam et al.[24] reported the ATRS in 90 patients with a follow up between 2 and 24 months, and showed a moderately strong criterion validity,

with a ceiling effect of 8%. However, the test–retest variability showed poor reliability, raising questions regarding the use of the ATRS for repeated assessments of individual patients.[24] Kearney et al.[8] evaluated 64 patients, and reported the ATRS demonstrated high internal consistency and responsiveness, with a celling effect of 11% at 9 months follow-up. The ceiling effect of the ATRS (20.4%) in this study was higher compared to previous reports, which could be due to the shorter follow-up in these studies. Functional outcome is likely to continue to improve with longer follow-up, and the previously reported ceiling effects might have underestimated the actual ceiling effects at long-term follow-up.

The FAAM is used to evaluate a variety of lower extremity disorders, and also demonstrated substantial content relevance to patients with Achilles tendon disoders.[9,10] While the FAAM has shown to be a reliable, responsive, and valid measure of physical function in various lower extremity disorders, it was validated in 164 individuals with a broad range of musculoskeletal disorders, with only 2 patients sustaining an Achilles tendon rupture. Reb et al.[9] evaluated the relevance of the FAAM specifically in 75 patients with Achilles tendon disease after a mean of 4 months (range 0-24 months) and concluded a substantial content relevance, however, ceiling effects were apparent for the Sports subscale (42.7%). Subgroup analysis was performed based on treatment groups with ceiling effects for the Sports subscale among nonoperative patients (22%), and ceiling effects for the ADL (21%) and Sports (54%) subscales among operative patients.[9] These ceiling effects are similar to the results presented in the present study.

The PROMIS has shown to be an excellent method for measuring outcomes for patients with foot and ankle surgery.[13,14] The PROMIS PF CAT was developed using item response theory to maximize efficient administration from a calibrated items bank of 124 questions and has been shown to results in equally high reliability and less ceiling effects in the assessment of general orthopedic trauma patients.[25] Hung et al.[11] evaluated the performance of the PROMIS PF CAT specifically for adult patients with common disorders of the foot and ankle, and was found to be an excellent method for measuring outcomes, with a good coverage (floor effect 0%, ceiling effect 0.32%), and an average test administration time of 47 seconds.[11] Hung et al.[12] recently also reported the responsiveness of both the PROMIS CAT and FAAM Sports instruments in the orthopedic foot and ankle population, which including 785 patients and found both to be sensitive and responsive to changes in patient-reported functional health. However, the study included a variety of 43 different disorders without further specifying whether they included

Achilles tendon ruptures.[12] They stated that further assessment of the responsiveness of the PROMIS and FAAM Sports instruments within specific conditions and across different populations is recommended.[12]

The PROMIS questionnaires were developed with the goal of providing standardized, valid, and flexible PROMs collection tools with features that lower response burden and make it possible to seamlessly incorporate them into patients' medical record.[13-15,26] Papuga et al.[26] recently explored the implementation of PROMIS CAT tools with 23,813 patient during outpatient clinics visits and reported an average time to completion of 3.5 minutes. There was no significant change in registration times for new patients, showing the implementation of PROMIS to be effective; results, moreover, could be imported directly into the electronic medical record in real time for use during the clinical visit.[26] Ho et al.[27] assessed whether preoperative PROMIS PF CAT scores were predictive of functional improvement after operative treatment in foot and ankle patients. They found that patients with scores below 29.7 were likely to improve with surgery, whereas patients with scores above 42 were unlikely to improve.[27] Cutoff values like these could help guide surgeons regarding the most appropriate treatment option for each individual patient. Future research could focus on reporting the prognostic cutoff values of the PROMIS PF CAT scores for Achilles tendon ruptures. The PROMIS instruments are already being used in the evaluation of Achilles tendon disorders.[28,29]

PROMs are increasingly used in orthopedic trauma care to evaluate patient-oriented health status in clinical care and research, to assess cost-effectiveness, and, more recently, to influence reimbursement decisions. Today, however, a number of different outcome measures are still used in different Achilles tendon studies.[19,20] Despite early promising results, PROMIS has not been adopted in most orthopedic literature. The performance of the PROMIS compared to various legacy "traditional" outcome measures has been evaluated across various conditions which have shown the PROMIS to correlate well with traditional outcome measures used in orthopedic studies.[13-15] Validity, reliability, and responsiveness are properties that define the clinical relevance of any outcome instrument, and establishing usefulness is an ongoing process meant to substantiate utility under various conditions and populations.[20] These properties may differ across settings and patient populations. It is therefore important to continue to evaluate different available PROMs in specific patient populations. The correlations found in this study suggest that perhaps PROMIS PF CAT ought to be considered the most useful measure for evaluating

patients with an Achilles tendon rupture—particularly when compared to the use of FAAM or ATRS. The PROMIS PF CAT tool did not show substantial floor or ceiling effects, while both the FAAM and the ATRS showed substantial (>15%) ceiling effects. Such effects suggest that extreme items are missing in the upper end of the FAAM and ATRS outcome instruments, indicating limited content validity. Therefore, patients with the highest possible score cannot be distinguished, thus reducing reliability. Furthermore, responsiveness is limited as functional changes cannot be measured in these patients.[22] This might presents challenges in studies that explore and evaluate improvement in the more athletic patients with Achilles tendon ruptures.

Potential limitations in this study need to be acknowledged. First, the study is limited by the nature of the injuries in a trauma setting, hence we were unable to collect PROMs at the time of injury to allow for baseline comparison. Second, only 103 patients contacted (34%) ultimately participated in this study, potentially creating selection bias among those that did agree. Third, the study might be limited by the order in which the PROMs were administered to patients during the electronically collection. Randomization of the order of the PROMs questionnaires may have eliminated the potential effects of survey fatigue. Finally, the different PROMs were only evaluated in the English language.

Conclusion

The PROMIS PF, FAAM ADL, and FAAM Sports all showed a moderate to high mutual correlation with the ATRS. However, of these measures, only the PROMIS PF CAT avoided substantial floor or ceiling effects. The results of this study strongly suggest that PROMIS PF CAT be considered perhaps the most valid and responsive tool for evaluating function after Achilles tendon rupture—regardless of whether patients are treated operatively or non-operatively.

References:

1. Ganestam A, Kallemose T, Troelsen A, Barfod KW. Increasing incidence of acute Achilles tendon rupture and a noticeable decline in surgical treatment from 1994 to 2013. A nationwide registry study of 33,160 patients. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA.* 2016;24(12):3730-3737.
2. Wilkins R, Bisson LJ. Operative versus nonoperative management of acute Achilles tendon ruptures: a quantitative systematic review of randomized controlled trials. *Am J Sports Med.* 2012;40(9):2154-2160.
3. Jiang N, Wang B, Chen A, Dong F, Yu B. Operative versus nonoperative treatment for acute Achilles tendon rupture: a meta-analysis based on current evidence. *International orthopaedics.* 2012;36(4):765-773.
4. Khan RJ, Carey Smith RL. Surgical interventions for treating acute Achilles tendon ruptures. *Cochrane Database Syst Rev.* 2010(9):Cd003674.
5. Ochen Y, Beks RB, van Heijl M, et al. Operative treatment versus nonoperative treatment of Achilles tendon ruptures: systematic review and meta-analysis. *Bmj.* 2019;364:k5120.
6. Egger AC, Berkowitz MJ. Achilles tendon injuries. *Current reviews in musculoskeletal medicine.* 2017;10(1):72-80.
7. Nilsson-Helander K, Thomeé R, Silbernagel KG, et al. The Achilles tendon Total Rupture Score (ATRS): development and validation. *Am J Sports Med.* 2007;35(3):421-426.
8. Kearney RS, Achten J, Lamb SE, Parsons N, Costa ML. The Achilles tendon total rupture score: a study of responsiveness, internal consistency and convergent validity on patients with acute Achilles tendon ruptures. *Health Qual Life Outcomes.* 2012;10:24.
9. Reb CW, Saini SS, Stenson JF, et al. Content Relevance of the Foot and Ankle Ability Measure in Patients With Achilles Tendon Diseases. *Foot Ankle Spec.* 2018;11(3):217-222.
10. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot & ankle international.* 2005;26(11):968-983.
11. Hung M, Baumhauer JF, Latt LD, Saltzman CL, SooHoo NF, Hunt KJ. Validation of PROMIS ® Physical Function computerized adaptive tests for orthopaedic foot and ankle outcome research. *Clin Orthop Relat Res.* 2013;471(11):3466-3474.
12. Hung M, Baumhauer JF, Licari FW, Bounsanga J, Voss MW, Saltzman CL. Responsiveness of the PROMIS and FAAM Instruments in Foot and Ankle Orthopedic Population. *Foot & ankle international.* 2019;40(1):56-64.
13. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *J Am Acad Orthop Surg.* 2016;24(11):744-749.
14. Cheung EC, Moore LK, Flores SE, Lansdown DA, Feeley BT, Zhang AL. Correlation of PROMIS with Orthopaedic Patient-Reported Outcome Measures. *JBJS reviews.* 2019;7(8):e9.
15. Gagnier JJ. Patient reported outcomes in orthopaedics. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2017;35(10):2098-2108.
16. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics.* 2009;42(2):377-381.
17. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J Epidemiol.* 2011;173(6):676-682.
18. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Medical care.* 2007;45(5 Suppl 1):S3-s11.
19. Kearney RS, Achten J, Lamb SE, Plant C, Costa ML. A systematic review of patient-reported outcome measures used to assess Achilles tendon rupture management: what's being used and should we be using it? *British journal of sports medicine.* 2012;46(16):1102-1109.
20. Spennacchio P, Vascellari A, Cucchi D, Canata GL, Randelli P. Outcome evaluation after Achilles tendon ruptures. A review of the literature. *Joints.* 2016;4(1):52-61.
21. Shoukri MM, Cihon C. *Statistical methods for health sciences.* CRC press; 1998.
22. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
23. Team R. R Development Core Team. *R: A Language and Environment for Statistical Computing.* 2013.
24. Ganestam A, Barfod K, Klit J, Troelsen A. Validity and reliability of the Achilles tendon total rupture score. *The Journal of foot and ankle surgery : official publication of the American College of Foot and Ankle Surgeons.* 2013;52(6):736-739.
25. Hung M, Stuart AR, Higgins TF, Saltzman CL, Kubiak EN. Computerized Adaptive Testing Using the PROMIS Physical Function Item Bank Reduces Test Burden With Less Ceiling Effects Compared With the Short Musculoskeletal Function Assessment in Orthopaedic Trauma Patients. *J Orthop Trauma.* 2014;28(8):439-443.

26. Papuga MO, Dasilva C, McIntyre A, Mitten D, Kates S, Baumhauer JF. Large-scale clinical implementation of PROMIS computer adaptive testing with direct incorporation into the electronic medical record. *Health systems (Basingstoke, England)*. 2018;7(1):1-12.
27. Ho B, Houck JR, Flemister AS, et al. Preoperative PROMIS Scores Predict Postoperative Success in Foot and Ankle Patients. *Foot & ankle international*. 2016;37(9):911-918.
28. Okoroha KR, Ussef N, Jildeh TR, et al. Comparison of Tendon Lengthening With Traditional Versus Accelerated Rehabilitation After Achilles Tendon Repair: A Prospective Randomized Controlled Trial. *Am J Sports Med*. 2020;48(7):1720-1726.
29. Bernstein DN, Anderson MR, Baumhauer JF, et al. A Comparative Analysis of Clinical Outcomes in Noninsertional Versus Insertional Tendinopathy Using PROMIS. *Foot Ankle Spec*. 2019;12(4):350-356.

PART 3

OPPORTUNITIES IN TRAUMA RESEARCH

CHAPTER 13

Summary and general discussion

Summary

Challenges in trauma research

In the field of orthopedic trauma surgery, great progress has been made over the last decades which has improved patient care and enhanced the likelihood of surviving serious injury. As a result, focus has shifted from patient survival to improving quality of life and reducing the burden of nonfatal injury. However, research-based advances that improve these outcomes for patients with orthopedic injuries have been constrained. The aims of this thesis were to provide insight into the value of different study designs which evaluate the effects of medical interventions for trauma patients in everyday clinical practice and to assess the use of patient-reported outcome measures (PROMs) as an integrated part of research practice for the assessment of quality of life after nonfatal trauma injury.

PART 1: Value of observational studies

Research based advances in orthopedic trauma are constrained by a lack of high-quality studies for the evaluation of interventions. Although randomized controlled trials (RCTs) are considered the highest level of evidence for such evaluations, this design might not always be ethical, feasible, or necessary to address a specific surgical research question. These challenges are more apparent in the field of trauma surgery, with acute and urgent life-threatening situations. These factors lead to the discontinuation of a large number of surgical trials, and has restricted the translation of study results to routine patient care.[1-3] These challenges have led to a growing debate on the need of RCTs for the evaluation of surgical interventions, and whether well-designed observational studies might complement and add valuable information to results from RCTs.[1,2]

For the assessment of outcomes following pharmaceutical treatment, RCTs are considered superior compared to observational studies. Randomization prevents confounding, which may arise when treatments are selectively prescribed to patients who would potentially benefit. Blinding, of patients and treating physicians, prevents differential changes in health care behavior, and efforts can be made to ensure that assessors of the outcome are blinded for the received treatment. Furthermore, in RCTs efforts can be made to enhance the adherence to the received treatment. However, several aspects of surgical treatments limit the application of these design features.[1,4] Surgical RCTs are commonly described by three types of comparisons. The Type 1 comparison evaluates pharmaceutical treatment in surgical patients, which account for 75% of

surgical trials. For this type of comparison, the traditional RCT appears the most suitable study design. The Type 2 comparison looks into different operative techniques, whereas the Type 3 comparison evaluates operative versus nonoperative treatment. The conduct of Type 2 studies leads to specific challenges as operations are complex procedures with learning curves, varying levels of surgical experience across surgeons, and differences in application of surgical techniques. These challenges are also encountered in Type 3 studies. Moreover, Type 3 studies are challenged by patient and surgeon preference due to the large difference in adverse effects between the operative and nonoperative treatment options, and the irreversibility of operative treatment.[4]

Although it is clear that randomization, concealment of allocation, and blinding are not possible in observational studies, the extent to which these factors impact the validity of a study may differ based on the specific clinical field and research questions.[2] In daily practice, the allocation of surgical interventions can sometimes be close to a random process, possibly improving the validity of observational study designs in research of surgical interventions. Particularly studies of acute surgical treatments might be less sensitive to confounding when the treatment option depends on surgeon preference but not on individual patient characteristics.[5] In such cases, one can speculate that groups of patients who underwent different surgical treatments might be rather similar (except for the treatment option).[5] In **Chapter 2**, we assessed the potential value of routinely collected data on elective operative interventions with two studies (Type 2 studies) on total hip arthroplasty. Our findings support the viewpoint that, in specific cases, the groups of patients who undergo different orthopedic operative interventions indeed appear to be comparable with respect to pre-operative patient characteristics. Therefore, observational studies comparing these operative interventions could be valuable to study comparative effectiveness, in addition to RCTs.[5] The data used in this study came from the nationwide Dutch Arthroplasty Register (LROI), a prospective longitudinal cohort containing high-quality data. Hence, the phenomena observed in this study are not necessarily to be expected in any other observational study. It does, however, provide support that there are cases in which observational studies of operative treatment options are viable and provide valuable information.

In orthopedic trauma research, well-designed observational studies might complement and add valuable information to results from RCTs, or –arguably– could even be used instead of RCTs.[2] Previous studies have looked into the differences in effect estimates from observational studies

and RCTs.[6-10] Although there are many examples in which results of observational studies concur with those of RCTs, obviously one cannot conclude that results will always be the same. We performed several meta-analyses, which included both RCTs and observational studies. All these meta-analyses evaluated outcome, comparing operative and nonoperative treatments (Type 3 studies), for frequently encountered orthopedic trauma topics. For all comparisons made, there seemed to be clinical equipoise regarding treatment choice. In all the meta-analyses we performed, the pooled effect estimates obtained from RCTs and observational studies were similar. In **Chapter 3**, we compared operative with nonoperative management of displaced proximal humeral fractures. We hypothesized that including observational studies in this meta-analysis would lead to more robust conclusions without impairing the quality of the results. This study demonstrated that the findings were indeed consistent across study designs with respect to different outcome measures. Furthermore, by including studies of both designs, this meta-analysis is currently the largest on this topic. This increase in patient numbers made it possible to perform the first meta-analysis in which subgroup analysis for Neer 3-part and 4-part fractures were possible. In **Chapter 4**, the aim was to compare operative with nonoperative treatment for humeral shaft fractures. The optimal management of these fractures was (and probably still is) a topic of debate, despite two previously published reviews.[11,12] Both reviews focused on RCTs only. Because of the lack of RCTs and the existence of observational studies only at the time, both reviews did not perform any meta-analysis and concluded that the superiority of any treatment option could not be determined. However, by combining evidence from RCTs and observational studies, we were able to include information regarding 1.412 patients. This study showed that satisfactory results can be achieved with nonoperative as well as operative management; however, operative treatment reduced the risk of nonunion compared with nonoperative treatment. The study described in **Chapter 5** aimed to compare functional, clinical, and radiologic outcomes after operative and nonoperative treatment of distal radius fractures. Several meta-analyses had been published on the comparison between operative and nonoperative treatment. However, these meta-analyses had focused specifically on elderly patient populations, aged 60 years or older, and found no difference in functional outcome between treatment groups.[13-15] Nevertheless, the international rate of operative treatment of distal radius fractures had been increasing, despite higher cost and limited evidence of improved functional outcome to support this practice.[16] The addition of observational studies in this meta-analysis increased the sample size and heterogeneity in patient characteristics, which lead to the possibility of evaluating treatment effects across age groups. The findings of this study suggest that

operative treatment might be more effective and have a greater impact on the health and well-being of the younger, non-elderly patients, whereas among elderly there was no difference in functional outcome and a higher complication rate following operative treatment. These results will help in the decision-making process of clinicians treating non-elderly patients with a distal radius fracture, in this often relatively healthy and still working age group. In **Chapter 6** the aim was to compare re-rupture and complication rates after operative and nonoperative treatment of Achilles tendon ruptures. Several meta-analyses of RCTs only, had shown that operative treatment significantly reduces the risk of tendon re-rupture compared with nonoperative treatment. However, operative treatment led to a substantial increase in other complications.[17-20] Despite the results of these previous meta-analyses, the use of operative treatment had declined over the past decade as a result of multiple studies showing comparable results between both treatments.[21,22] This study showed that operative treatment of Achilles tendon ruptures indeed reduces the risk of re-rupture compared with nonoperative treatment, and operative treatment also results in a higher risk of other complications. However, with the addition of observational studies resulting in the inclusion of an additional 14.918 patients, our results showed that differences between treatment groups for re-rupture and complications rates were smaller than previously presumed. Furthermore, the sensitivity analyses including studies with a study period after the year 2000 showed that the differences between treatment groups were even smaller. These findings indicate an overall reduction in complications after treatment of Achilles tendon ruptures due to the development of new rehabilitation protocols, operative techniques, and nonoperative treatment modalities.

Meta-analyses are valuable tools for the assessment of differences in treatment effects. Throughout **PART 1**, we encountered many cases in which the sole focus of including RCTs in previous meta-analyses had restricted the translation of studies to routine patient care. The focus on RCTs alone had made it difficult to perform different subgroup analyses. Inclusion of observational studies, however, made it possible to investigate patient subgroups such as different fracture classifications, age groups or study periods. Some reviews did not perform any meta-analysis at all and concluded that superiority of treatment could not be determined because of the lack of RCTs. Moreover, by the fixation on inclusion of RCTs alone in the eligibility criteria, some meta-analyses have lost sight of the generalizability and context of the evaluated treatments. Some meta-analyses focused so much on only including RCTs, some included studies with study periods going as far back as the year 1973, and by doing so overlooked the

development of new operative techniques and nonoperative treatment modalities during the recent decades.

PART 2: Value of patient-reported outcome measures

Traditionally, trauma research primarily focused on clinical and radiological outcomes, and thus overlooked the quality of life of surviving patients.[3] In modern day clinical practice, the recovery of trauma patients is tracked from injury through prehospital care, acute care, and rehabilitation. However, there is still a lack of understanding of the degree of recovery, the time needed, and the extent to which those who suffered injuries will experience lifelong disability.[23]

Standardized outcome measures and routine collection of PROMs are needed to monitor and assess present and new treatment approaches and to support evidence-based care.[3] Despite the advances and use in routine care, there are still substantial challenges regarding implementation and standardization of PROMs. Given the paucity of data regarding PROM scores for trauma patients, this thesis aimed to provide benchmark data that can be used for future comparison. In **Chapter 7**, we demonstrated that the recently developed Ligament Augmentation and Reconstruction System (LARS) technique seemed to be an effective and safe fixation method for the treatment of AC joint dislocations, resulting in good patient-reported functional outcome. The results presented in **Chapter 8** suggest that the newer Superior Clavicle Plate with Lateral Extension (SCPLE) is an effective fixation method for the treatment of lateral clavicle fractures. In both chapters we used the QuickDASH score, a validated PROM instrument developed to measure upper extremity disability and symptoms. Furthermore, we limited the timeframe of PROM completion to establish mid-term functional outcome. In both chapters, we were challenged by the lack of previous functional PROM findings. Unfortunately, comparison of literature remains difficult due to small sample sizes and a wide variety of (non-validated) functional outcome scores that are being used. Outcome studies are, especially when evaluating new medical interventions, mainly single center, include small samples, or do not apply a minimum follow-up period, including patients with limited follow-up (<6 months). Multicenter studies that evaluate PROMs at multiple time points after injury are lacking, even though this information is important for establishing the long-term burden of injury to provide information about prognosis and guide treatment decisions. These challenges may delay implementation of promising interventions and underline the importance of implementation of standardized quality measurement for the orthopedic trauma population. The results of the study described in

Chapter 9 suggest that the newer Patient-Reported Outcomes Measurement Information System (PROMIS) questionnaires can be considered similarly useful as a measure for evaluating shoulder function after proximal humeral fractures. Given the paucity of data regarding PROMIS scores after proximal humeral fracture treatment, the results of this study are meant to provide benchmark data that can be used for future comparisons. **Chapter 10** also provides benchmark data that can be used for future comparison in patients with a distal humeral fracture. Furthermore, this chapter highlights the importance of measuring both clinical and patient-reported outcomes when evaluating distal humeral fracture treatment. In **Chapter 11**, we evaluated the operative management of bicondylar tibial plateau fractures, which had not been widely studied. Contributing to the difficulty in choosing the optimal management for bicondylar tibial plateau fractures is the absence of validated PROMs. In this chapter, we established normative data and long-term functional PROMIS scores after operative treatment of bicondylar tibial plateau fractures. The results of this chapter also emphasize the importance of obtaining both general global health measures and injury specific measures when evaluating outcomes after injuries. In **Chapter 12**, we evaluated the measurement properties of different PROMs used in patients with an acute Achilles tendon rupture. This chapter showed that even an injury specific PROM, which has been evaluated and found to be valid, reliable, and responsive and many currently consider as the most appropriate, can have limited content validity with respect to assessing long-term function and distinguishing between patients.

The use of PROMs in orthopedic trauma has rapidly increased over the last decades and it is expected that this trend will continue. However, despite the advances and use in routine care, there are still substantial challenges regarding implementation and standardization of PROMs. Throughout **PART 2**, we encountered many of these challenges including the reliability and precision of the instruments used to capture the outcome of interest.[24-26] Standardized quality and outcome measurements have been found difficult to implement for the orthopedic trauma patient population, since there is an almost innumerable combination of injuries secondary to different trauma mechanisms and circumstances.[27] Studies are required to determine which specific outcome measures to institutionalize, how to modify them for different injury and patient specific factors.[27] These challenges include the reliability and precision of the instruments used to capture the outcome of interest. Previous orthopedic studies, evaluating similar conditions, have used a variety of different "traditional" legacy PROMs, making it difficult to compare results. Moreover, the completion of the previous legacy measures can be burdensome

and time consuming. Hence, the challenge is how to compare outcomes score between groups and studies, and how to increase effectiveness while reducing administration time and lowering responder burden for PROMs.[24-26]

General discussion

Opportunities in trauma research

Study designs

The design of an orthopedic trauma RCT is challenged by ethical consideration, the urgent nature of treatment, learning curves for operative procedures, belief in existing treatments, inability to blind surgeons and patients, differences between surgical sites, the inherent heterogeneity of each operative procedure, the related costs, exclusion of high-risk populations, and study duration. Observational studies might, therefore, also have a role in improving the value and best available evidence in orthopedic trauma care.[2,5,28] Given the challenges and obstacles to perform a surgical RCT, compared to an observational study, it is important to understand to what extent differences between these designs impact study results. In general, RCTs allows for blinding and strict compliance, which are relevant for Type 1 studies. However, for Type 2 studies, the role of blinding seems smaller as in clinical practice blinding is usually not possible, and considerations of compliance with an operative procedure are often irrelevant. For Type 3 studies, these factors are dependent on the field, research question, and study design. Confounding and bias should not be presumed to be universally present in all observational surgical studies.

By including observational studies in meta-analyses, the analyzed patients may be more representative of patients encountered in daily clinical practice, which tends to improve generalizability of results. In a health care system with growing financial burden the relative low cost and feasibility also underline the possible added value of observational studies. Our findings support the viewpoint that, in specific cases, one could argue that groups of patients who undergo different orthopedic surgical interventions in practice, are comparable with respect to pre-operative patient characteristics, and therefore results of such observational studies would be valuable to use when assessing comparative effectiveness, in addition to results of RCTs. From a methodological perspective, this thesis emphasizes the potential benefits of observational studies in orthopedic trauma research. We hope these findings will help fuel the debate on the often-used hierarchical structure of research designs in the evaluation of outcomes following surgical

treatment. Furthermore, we hope this thesis will challenge future researcher to look beyond the hierarchy of research designs and consider alternative designs.

Recommendations for surgical treatments are influenced by training, which is affected by cultural and regional differences. In a prospective parallel study design, two countries, two hospitals, or even two surgeons of different "schools of thought" can be compared by evaluating outcomes of patient populations in daily clinical practice, where disagreement exists on the preferred treatment option, and clinical equipoise regarding treatment choice seems to exist. This design requires the collection of the same outcomes during standardized follow-up periods. In such cases, one can speculate that the groups of patients who underwent different surgical treatments might be rather similar (except for the treatment option within the "school of thought"). Another observational study designs could be a parallel cohort study with blinded inclusion based on clinical equipoise.[29] In this design eligibility of each patients is assessed retrospectively by an expert panel of orthopedic surgeons from different medical centers who are blinded for the received treatment. Patients are included if the majority of experts disagree on the suggested treatment method. This will lead to two comparable groups, where there actually exists clinical disagreement on the optimal treatment management. A third alternative design could be a study which involves both randomized and observational arms, as both patients and surgeons can have a strong preference for a certain treatment. As an example, we mention a study that evaluated the efficacy of surgical stabilization of rib fracture.[30] This study was faced with these challenges, and decided to offer randomization as well as observational follow-up. Randomization was declined by nearly 80% of subjects, yet no differences were observed between subjects who chose for the different options.[30]

A framework for the design of observational comparative effectiveness studies is the so-called target trial emulation in which observational studies are considered to attempt to emulate a (target) RCT.[31-33] Observational studies that aim to guide clinical decisions could be evaluated with respect to how well they emulate the intended target trial. Target trial emulation is the application of RCT design principles to guide the analysis of observational data. The aim of the emulation is to improve the quality of the observational study, when RCTs are not available, ethical, or feasible. Through the specification of the eligibility criteria, treatment options, treatment allocation, outcome measures, causal contrast, and follow-up of the intended target trial, one can explore which design and analytic options are appropriate. The emulation of the

target trial and the causal framework can guide researchers to identify and avoid unnecessary bias and provides an explicit manner to express concessions that need to be made in the observational study. However, we are usually not able to emulate the basic elements of target trials such as randomization, concealment of allocation, and blinding. There might not be enough baseline confounders to appropriately control for confounding (i.e., emulate randomization), hence, alternative analytic approaches should be considered. However, these might fundamentally change the treatment options and eligibility criteria. The treatment options might not be adequately defined based on available measures and knowledge, specifying the ill-defined target trial forces the researchers to see and acknowledge these concessions. We might realize that the concessions in the target trial we are trying to emulate are too big, forcing us to redefine the research questions or look into other sources of data. We might still pursue the study, despite not being able to emulate the target trial, however, aware of the concessions that are being made. In this framework, both of these endpoints are effective in improving the quality of epidemiologic research.[31-33]

The studies presented in this thesis illustrate that observational studies are indeed inherently different from RCTs. Observational studies are performed when we are faced with the challenges of conducting a RCT in the presence of practical or ethical constraints. The limitations of observational studies are potentially still present and in the light of few alternative options we need to keep improving these studies, because researchers will keep using observational data to guide clinical decisions. Given the challenges and obstacles to perform a surgical RCT compared to an observational study it is important to understand to what extent differences between these designs impact study results. This thesis shows that particularly studies of surgical treatments, might be less sensitive to confounding if treatment preference or "allocation" to treatment is not dependent on patient characteristics. It is up to the researchers of such studies to provide the arguments to substantiate the claim that treatment groups are expected to be comparable and why a particular research question could be answered using an observational study design. We believe that, with detailed planning and conduct, an observational study can be of substantial added value.

Patient-reported outcome measures

Validity, reliability, and responsiveness are properties that define the clinical value of any outcome instrument, and establishing usefulness is an ongoing process meant to substantiate

utility under different conditions and in different populations.[34] The field of orthopedic trauma needs to evolve towards the systematic standardization and evaluation of PROMs. Systematic reviews on PROMs that are currently in use, and the evidence of their measurement properties, are needed. Furthermore, there is still much work to be done to determine how physicians can use PROMs to track the recovery of patients from injury through acute care, hospitalization, and rehabilitation. In addition, there needs to be determined what the most efficient and effective methods are for implementation PROMs into everyday clinical practice. This involves determining when and how to use PROMs to help clinical decisions-making and how to present these results in electronic medical records. Recommendations are needed to establish uniform timeframes (pre-treatment, early-, mid-, and long-term) for the completion of PROMs. The standardization of these timeframes will enable clear data collection and would make it possible to aggregate outcomes in regional and national registries, which would allow for evaluation of infrequent injuries, the evaluation of small treatment effects, and infrequent outcome measures.

The next step would be to increase the effectiveness of measuring different health outcomes, while reducing administration time and lowering responder burden for PROMs.[24-26] Currently, the PROMIS toolbox provides a comprehensive set of questionnaires and items that can be administered as a "Computerized Adaptive Test" (CAT). PROMIS have developed CAT for various outcome instruments, which uses algorithms with item response theory, to optimize questionnaires depending on previous responses, thus increasing effectiveness, reducing administration time and lowering responder burden.[35] This thesis has highlighted some of the advantages of the PROMIS tools. The major challenge is the recognition and implementation of these tools by the orthopedic trauma community. Despite promising results, PROMIS has not been adopted in most orthopedic literature, yet we hope that researchers and clinicians will also start using PROMIS when designing studies. These studies could provide benchmark data and provide validation of the PROMIS tools with the already used legacy measures. Further studies are still needed to evaluate the value of these outcome measures across different orthopedic populations, to define measurement properties, responsiveness to change over time, and determine clinically important differences.

This thesis presented the first results of PROMs studies, in which we encountered many challenges in the evaluation and optimization of PROM use in orthopedic trauma research. Different benchmark studies were performed, and the effects of new and existing medical

interventions were evaluated. Ultimately, the utilization of PROMs data will require careful planning and commitment, to determine how to interpret and thus utilize this data. However, it is important that the field of orthopedic trauma actively engages in the shift from physician-reported to patient-reported outcomes, to ensure that achieving high value for patients becomes the overarching goal of health care.

References:

1. Chapman SJ, Shelton B, Mahmood H, Fitzgerald JE, Harrison EM, Bhangu A. Discontinuation and non-publication of surgical randomised controlled trials: observational study. *Bmj.* 2014;349:g6870.
2. Beks RB, Bhashyam AR, Houwert RM, et al. When observational studies are as helpful as randomized trials: Examples from orthopedic trauma. *The journal of trauma and acute care surgery.* 2019;87(3):730-732.
3. Balogh ZJ, Reumann MK, Gruen RL, et al. Advances and future directions for management of trauma patients with musculoskeletal injuries. *Lancet (London, England).* 2012;380(9847):1109-1119.
4. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *Bmj.* 2002;324(7351):1448-1451.
5. Houwert RM, Beks RB, Dijkgraaf MGW, et al. Study methodology in trauma care: towards question-based study designs. *Eur J Trauma Emerg Surg.* 2019.
6. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *Jama.* 2001;286(7):821-830.
7. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPJb. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. 2016;352.
8. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *The New England journal of medicine.* 2000;342:1878-1886.
9. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England journal of medicine.* 2000;342:1887-1892.
10. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *Journal of clinical epidemiology.* 2010;63:238-245.
11. Gosler MW, Testroote M, Morrenhof JW, Janzing HM. Surgical versus non-surgical interventions for treating humeral shaft fractures in adults. *Cochrane Database Syst Rev.* 2012;1:Cd008832.
12. Clement ND. Management of Humeral Shaft Fractures; Non-Operative Versus Operative. *Arch Trauma Res.* 2015;4(2):e28013.
13. Song J, Yu A-X, Li Z-H. Comparison of conservative and operative treatment for distal radius fracture: a meta-analysis of randomized controlled trials. *International journal of clinical and experimental medicine.* 2015;8:17023-17035.
14. Ju J-H, Jin G-Z, Li G-X, Hu H-Y, Hou R-X. Comparison of treatment outcomes between nonsurgical and surgical treatment of distal radius fracture in elderly: a systematic review and meta-analysis. *Langenbeck's archives of surgery.* 2015;400:767-779.
15. Chen Y, Chen X, Li Z, Yan H, Zhou F, Gao W. Safety and Efficacy of Operative Versus Nonsurgical Management of Distal Radius Fractures in Elderly Patients: A Systematic Review and Meta-analysis. *The Journal of hand surgery.* 2016;41:404-413.
16. Armstrong KA, von Schroeder HP, Baxter NN, Zhong T, Huang A, McCabe SJ. Stable rates of operative treatment of distal radius fractures in Ontario, Canada: a population-based retrospective cohort study (2004-2013). *Canadian Journal of Surgery.* 2019(1488-2310 (Electronic)).
17. Khan RJ, Carey Smith RL. Surgical interventions for treating acute Achilles tendon ruptures. *Cochrane Database Syst Rev.* 2010(9):Cd003674.
18. Jiang N, Wang B, Chen A, Dong F, Yu B. Operative versus nonoperative treatment for acute Achilles tendon rupture: a meta-analysis based on current evidence. *International orthopaedics.* 2012;36(4):765-773.
19. Soroceanu A, Sidhwa F, Aarabi S, Kaufman A, Glazebrook M. Surgical versus nonsurgical treatment of acute Achilles tendon rupture: a meta-analysis of randomized trials. *J Bone Joint Surg Am.* 2012;94(23):2136-2143.
20. Wilkins R, Bisson LJ. Operative versus nonoperative management of acute Achilles tendon ruptures: a quantitative systematic review of randomized controlled trials. *Am J Sports Med.* 2012;40(9):2154-2160.
21. Ganestam A, Kallemose T, Troelsen A, Barfod KW. Increasing incidence of acute Achilles tendon rupture and a noticeable decline in surgical treatment from 1994 to 2013. A nationwide registry study of 33,160 patients. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA.* 2016;24(12):3730-3737.
22. Egger AC, Berkowitz MJ. Achilles tendon injuries. *Current reviews in musculoskeletal medicine.* 2017;10(1):72-80.
23. Gabbe BJ, Simpson PM, Harrison JE, et al. Return to Work and Functional Outcomes After Major Trauma: Who Recovers, When, and How Well? *Annals of surgery.* 2016;263(4):623-632.
24. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *J Am Acad Orthop Surg.* 2016;24(11):744-749.
25. Cheung EC, Moore LK, Flores SE, Lansdown DA, Feeley BT, Zhang AL. Correlation of PROMIS with Orthopaedic Patient-Reported Outcome Measures. *JBJS reviews.* 2019;7(8):e9.
26. Gagnier JJ. Patient reported outcomes in orthopaedics. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2017;35(10):2098-2108.
27. Bhashyam AR, van der Vliet QMJ, Ochen Y, et al. Injury-related variation in patient-reported outcome after musculoskeletal trauma: a systematic review. *Eur J Trauma Emerg Surg.* 2019.
28. Rouleau DM. Designing clinical studies in orthopedic traumatology. *Orthopaedics & traumatology, surgery & research : OTSR.* 2018;104(1s):S1-s7.

29. Stadhouder A, Buskens E, de Klerk LW, et al. Traumatic thoracic and lumbar spinal fractures: operative or nonoperative treatment: comparison of two treatment strategies by means of surgeon equipoise. *Spine (Phila Pa 1976).* 2008;33(9):1006-1017.
30. Pieracci FM, Leasia K, Bauman Z, et al. A multicenter, prospective, controlled clinical trial of surgical stabilization of rib fractures in patients with severe, nonflail fracture patterns (Chest Wall Injury Society NONFLAIL). *The journal of trauma and acute care surgery.* 2020;88(2):249-257.
31. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol.* 2016;183(8):758-764.
32. Hernan MA. The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data. *American journal of public health.* 2018;108(5):616-619.
33. Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. *Eur J Epidemiol.* 2017;32(6):473-475.
34. Spennacchio P, Vascellari A, Cucchi D, Canata GL, Randelli P. Outcome evaluation after Achilles tendon ruptures. A review of the literature. *Joints.* 2016;4(1):52-61.
35. Hung M, Stuart AR, Higgins TF, Saltzman CL, Kubiak EN. Computerized Adaptive Testing Using the PROMIS Physical Function Item Bank Reduces Test Burden With Less Ceiling Effects Compared With the Short Musculoskeletal Function Assessment in Orthopaedic Trauma Patients. *J Orthop Trauma.* 2014;28(8):439-443.

PART 4

APPENDICES

DUTCH SUMMARY

(Nederlandse samenvatting)

Uitdagingen in trauma onderzoek

Op het gebied van de traumachirurgie is de afgelopen decennia grote vooruitgang geboekt, waardoor de patiëntenzorg is verbeterd en de kans op het overleven van ernstig letsel is vergroot. Als gevolg hiervan is de focus verschoven van de overleving van de patiënt naar het verbeteren van de kwaliteit van leven en het verminderen van de lasten van niet-dodelijk letsel. Op onderzoek gebaseerde resultaten die deze uitkomsten voor patiënten evalueren, zijn echter beperkt. Het doel van dit proefschrift was om inzicht te geven in de waarde van verschillende onderzoeksopzetten die de effecten van medische interventies voor traumapatiënten in de dagelijkse klinische praktijk evalueren, en om het gebruik van patiënt-gerapporteerde uitkomstmaten (PROMs) te beoordelen als geïntegreerd onderdeel van trauma onderzoek voor het beoordelen van de kwaliteit van leven.

DEEL 1: Waarde van observationele studies

Op onderzoek gebaseerde vooruitgang in de traumachirurgie wordt beperkt door een gebrek aan kwalitatief hoogwaardig onderzoek naar de effecten van interventies. Hoewel gerandomiseerd gecontroleerd onderzoek (RCTs) als de gouden standaard wordt beschouwd voor dergelijke evaluaties, is deze onderzoeksopzet niet altijd ethisch verantwoord, haalbaar of noodzakelijk om een specifieke chirurgische onderzoeksvraag te beantwoorden. Deze uitdagingen zijn meer zichtbaar op het gebied van de traumachirurgie, met acute en urgente levensbedreigende situaties. Deze factoren hebben geleid tot het stopzetten van een groot aantal chirurgische RCTs en hebben de vertaling van onderzoeksresultaten naar routinematige patiëntenzorg beperkt.[1-3] Deze uitdagingen hebben geleid tot een steeds grotere discussie over de noodzaak van RCTs voor de evaluatie van chirurgische ingrepen, en de vraag of goed opgezette observationele studies van meerwaarde kunnen zijn naast resultaten van RCTs.[1,2]

Voor de beoordeling van uitkomsten na farmaceutische behandelingen worden RCTs als superieur beschouwd in vergelijking met observationele studies. Randomisatie voorkomt confounding, welke kan optreden wanneer behandelingen selectief worden toegewezen aan patiënten die er mogelijk het meest baat bij hebben. Blindering van patiënten en behandelend artsen voorkomt veranderingen in ziektegedrag en er kunnen inspanningen worden geleverd om ervoor te zorgen dat beoordelaars van uitkomsten blind zijn voor de toegewezen behandeling. Bovendien kunnen in RCTs inspanningen worden geleverd om de therapietrouw bij de toegewezen behandeling te verbeteren. Verschillende aspecten van chirurgische behandelingen beperken echter de toepassing van deze onderzoeksopzet.[1,4] Chirurgische RCTs worden

gedefinieerd door drie typen vergelijkingen. De Type-1 vergelijking evalueert farmaceutische behandelingen bij chirurgische patiënten, dit omvat 75% van de chirurgische RCTs. Voor deze vergelijking lijkt de traditionele RCT de meest geschikte onderzoeksopzet. De Type-2 vergelijking evalueert verschillende operatietechnieken, terwijl de Type-3 vergelijking operatieve versus niet-operatieve behandelingen evalueert. Het uitvoeren van Type-2 onderzoek leidt tot specifieke uitdagingen, aangezien operaties complexe procedures zijn met leercurves, verschillende niveaus van chirurgische ervaring en verschillen in toepassing van chirurgische technieken. Deze uitdagingen komen ook voor in Type-3 onderzoek. Bovendien wordt Type-3 onderzoek bemoeilijkt door de voorkeur van patiënt en chirurg vanwege het grote verschil in consequenties tussen de operatieve en niet-operatieve behandelopties en de onomkeerbaarheid van operatieve behandelingen.[4]

Hoewel het duidelijk is dat randomisatie en blindering van toewijzing, behandelend artsen en beoordelaars niet mogelijk is in observationele studies, kan de mate waarin deze factoren de validiteit van een studie beïnvloeden verschillen op basis van het specifieke klinische veld en onderzoeksvraag.[2] In de dagelijkse praktijk kan de toewijzing van bepaalde chirurgische interventies in de buurt komen van een willekeurig proces, wat mogelijk de validiteit van een observationele studieopzet verbetert. Met name studies naar acute chirurgische behandelingen zijn mogelijk minder gevoelig voor confounding wanneer de gekozen behandelingsoptie afhankelijk is van de voorkeur van de chirurg, maar niet van individuele patiëntkenmerken.[5] In dergelijke gevallen kan men speculeren dat groepen patiënten die verschillende chirurgische behandelingen ondergaan, mogelijk op elkaar lijken (afgezien van de behandeling).[5] In **Hoofdstuk 2** hebben we de potentiële meerwaarde van routinematig verzamelde data over electieve operatieve interventies geëvalueerd, doormiddel van twee studies (Type-2 studies) naar totale heupartroplastiek. Onze bevindingen ondersteunen het standpunt dat, in specifieke gevallen, groepen patiënten die verschillende orthopedische operatieve ingrepen ondergaan inderdaad vergelijkbaar lijken te zijn wat betreft preoperatieve patiëntkenmerken. Daarom kunnen observationele studies die deze operatieve interventies vergelijken, naast RCTs, waardevol zijn om de effecten van interventies te bestuderen.[5] De data die in deze studie zijn gebruikt, zijn afkomstig van de Nederlandse Landelijke Registratie Orthopedische Implantaten (LROI), een prospectief longitudinaal cohort met hoogwaardige data. Daarom zijn de bevindingen die in deze studie zijn waargenomen niet noodzakelijk te verwachten in andere observationele studies. Het biedt echter wel ondersteuning dat er wel degelijk gevallen zijn waarin

observationele studies naar operatieve behandelingsopties meerwaarde hebben en waardevolle informatie kunnen opleveren.

In traumachirurgisch onderzoek kunnen goed opgezette observationele studies de resultaten van RCTs aanvullen, of –mogelijk– zelfs worden gebruikt in plaats van RCTs.[2] Eerdere studies hebben gekeken naar de verschillen in effectschattingen uit observationele studies en RCTs.[6-10] Hoewel er vele voorbeelden zijn waarin resultaten van observationele studies overeenkomen met die van RCTs, kan logischerwijs niet worden geconcludeerd dat de resultaten altijd hetzelfde zullen zijn. We hebben verschillende meta-analyses uitgevoerd, waarbij zowel RCTs als observationele studies werden geïncludeerd. Al deze meta-analyses evalueerden de uitkomsten van operatieve en niet-operatieve behandelingen (Type-3 studies) voor veel voorkomende traumachirurgische onderwerpen. Bij alle vergelijkingen leek er sprake van klinisch equipoise (als men werkelijk niet weet wat de beste interventie is) met betrekking tot de keuze voor een bepaalde behandeloptie. In alle meta-analyses die we hebben uitgevoerd, waren de gepoolde effectschattingen verkregen uit resultaten van RCTs en observationele studies vergelijkbaar. In **Hoofdstuk 3** vergeleken we de operatieve met niet-operatieve behandeling van gedisloceerde proximale humerus fracturen. Onze hypothese was dat het opnemen van observationele studies in deze meta-analyse zou leiden tot meer robuuste conclusies, zonder de kwaliteit van de resultaten te beïnvloeden. Deze studie toonde aan dat de bevindingen tussen de onderzoeksopzetten inderdaad consistent waren met betrekking tot verschillende uitkomstmaten. Bovendien is, door studies van beide onderzoeksopzetten op te nemen, deze meta-analyse momenteel de grootste op dit onderwerp. Deze toename van sample size maakte het mogelijk om de eerste meta-analyse uit te voeren met subgroep analyse voor Neer 3-part en 4-part fracturen. In **Hoofdstuk 4** was het doel om de operatieve en niet-operatieve behandeling van humerusschacht fracturen te vergelijken. De optimale behandeling van deze fracturen was (en is waarschijnlijk nog steeds) een onderwerp van discussie, ondanks twee eerder gepubliceerde reviews.[11,12] Beide reviews waren alleen gericht op de inclusie van RCTs. Vanwege het ontbreken van RCTs en het enkel bestaan van observationele studies op dat moment, hebben beide reviews geen meta-analyse uitgevoerd en geconcludeerd dat de superioriteit van de behandelopties niet kon worden vastgesteld. Door de resultaten uit zowel RCTs als observationele studies te combineren, konden we data van 1.412 patiënten includeren. Deze studie toonde aan dat goede uitkomsten kunnen worden bereikt met zowel niet-operatieve als operatieve behandeling; operatieve behandeling verminderde echter het risico op nonunion in vergelijking met niet-

operatieve behandeling. De studie beschreven in **Hoofdstuk 5** was gericht op het vergelijken van functionele, klinische en radiologische uitkomsten na operatieve en niet-operatieve behandeling van distale radius fracturen. Over de vergelijking tussen operatieve en niet-operatieve behandeling zijn verschillende meta-analyses gepubliceerd. Deze meta-analyses waren echter specifiek gericht op oudere patiëntenpopulaties van 60 jaar of ouder en vonden geen verschil in functionele uitkomst tussen beide behandelopties.[13-15] Desalniettemin was internationaal het percentage van operatieve behandeling van distale radius fracturen gestegen, ondanks hogere kosten en beperkt bewijs van verbeterde functionele uitkomst om deze behandeling te ondersteunen.[16] De toevoeging van observationele studies in deze meta-analyse vergrootte de sample size en heterogeniteit in patiëntkenmerken, wat leidde tot de mogelijkheid om de effecten van beide behandelingen over verschillende leeftijdsgroepen te evalueren. De bevindingen van deze studie suggereren dat operatieve behandeling mogelijk effectiever is en een grotere impact heeft op de functionele uitkomst van jongere patiënten, terwijl er onder ouderen geen verschil was in functionele uitkomst en een hoger percentage complicaties na operatieve behandeling. Deze resultaten zullen artsen helpen bij de besluitvorming om te opereren bij jongere patiënten met een distale radius fractuur, in deze relatief vaak gezonde en nog werkende leeftijdsgroep. In **Hoofdstuk 6** was het doel om het aantal re-rupturen en complicaties te vergelijken na operatieve en niet-operatieve behandeling van achillespeesrupturen. Verschillende meta-analyses bestaande uit alleen RCTs hadden aangetoond dat operatieve behandeling het risico op re-ruptuur significant vermindert in vergelijking met niet-operatieve behandeling. Operatieve behandeling leidde echter tot een aanzienlijke toename van andere complicaties.[17-20] Ondanks de resultaten van deze eerdere meta-analyses was het aantal operatieve behandelingen het afgelopen decennium afgenomen als resultaat van meerdere studies die vergelijkbare resultaten lieten zien tussen beide behandelingen.[21,22] Deze studie toonde aan dat operatieve behandeling van achillespeesrupturen inderdaad het risico op een re-ruptuur verlaagt in vergelijking met niet-operatieve behandeling, en operatieve behandeling leidt ook tot een hoger risico op andere complicaties. Echter, met de toevoeging van observationele studies die resulteerden tot de inclusie van data van nog eens 14.918 patiënten, toonden onze resultaten aan dat de verschillen tussen de behandelopties voor een re-ruptuur en complicaties kleiner waren dan eerder werd aangenomen. Bovendien lieten de subgroep analyses met studies na het jaar 2000 zien dat de verschillen tussen behandelopties nog kleiner waren. Deze bevindingen duiden op een algehele vermindering van complicaties na behandeling van achillespeesrupturen als gevolg van de ontwikkeling van nieuwe revalidatieprotocollen, operatietechnieken en niet-operatieve behandelopties.

Meta-analyses zijn waardevolle instrumenten om de effecten van interventies te evalueren. In **DEEL 1** kwamen we veel gevallen tegen waarbij de focus van eerdere meta-analyses om enkel RCTs te includeren de translatie van onderzoeksresultaten naar routinematige patiëntenzorg had beperkt. De focus om enkel RCTs te includeren maakte het onmogelijk om verschillende subgroep analyses uit te voeren. Door observationele studies te includeren, was het echter mogelijk om subgroepen van patiënten te onderzoeken, zoals verschillende fractuur classificaties, leeftijdsgroepen of studieperiodes. Sommige reviews voerden helemaal geen meta-analyse uit en concludeerden dat de superioriteit van de behandelopties niet kon worden vastgesteld vanwege het ontbreken van RCTs. Bovendien hebben sommige meta-analyses de generaliseerbaarheid en context van de geëvalueerde behandelingen volledig uit het oog verloren door te fixeren op de inclusie van alleen maar RCTs. Sommige meta-analyses waren zo gericht op het opnemen van RCTs, waarbij sommige meta-analyses studies includeerde met studieperiodes die teruggingen tot het jaar 1973, en gingen daarmee voorbij aan de ontwikkeling van nieuwe operatietechnieken en niet-operatieve behandelopties van de afgelopen decennia.

DEEL 2: Waarde van patiënt-gerapporteerde uitkomstmaten

Van oudsher was trauma onderzoek primair gericht op klinische en radiologische uitkomsten, waardoor de kwaliteit van leven van patiënten over het hoofd werd gezien.[3] In de moderne klinische praktijk wordt het herstel van traumapatiënten gevolgd van letsel tot preklinische zorg, acute zorg en revalidatie. Er is echter nog steeds een gebrek aan inzicht in de mate van herstel, de benodigde tijd, en de mate waarin patiënten levenslange invaliditeit ervaren.[23]

Gestandaardiseerde uitkomstmaten en routinematige verzameling van PROMs zijn nodig om huidige en nieuwe behandelmethoden te monitoren en te evalueren, om zodoende evidence-based medicine te ondersteunen.[3] Ondanks de vooruitgang en het gebruik van PROMs in de routinematige zorg, zijn er nog steeds aanzienlijke uitdagingen met betrekking tot implementatie en standaardisatie. Gezien het gebrek aan data over PROM scores voor traumapatiënten, beoogde dit proefschrift referentiewaarden te verschaffen die kunnen worden gebruikt voor toekomstige vergelijkingen. In **hoofdstuk 7** hebben we aangetoond dat de recent ontwikkelde Ligament Augmentation and Reconstruction System (LARS) techniek een effectieve en veilige fixatiemethode lijkt te zijn voor de behandeling van AC-luxaties, wat resulteerde in een goede patiënt-gerapporteerd functionele uitkomst. De resultaten gepresenteerd in **Hoofdstuk 8** suggereren dat de nieuwere Superior Clavicle Plate with Lateral Extension (SCPLE) een

effectieve fixatiemethode is voor de behandeling van laterale claviculafracturen. In beide hoofdstukken hebben we de QuickDASH score gebruikt, een gevalideerde PROM score die is ontwikkeld om invaliditeit en symptomen van de bovenste extremiteit te meten. Bovendien hebben we de periode van voltooiing van PROM scores beperkt om de functioneel uitkomst op de middellange termijn vast te stellen. In beide hoofdstukken werden we uitgedaagd door het ontbreken van eerdere gerapporteerde functionele PROM resultaten. Helaas blijft het vergelijken van literatuur moeilijk vanwege de kleine sample size en grote verscheidenheid aan (niet-gevalideerde) functionele uitkomstsmaten die worden gebruikt. PROMs studies zijn, vooral bij het evalueren van nieuwe medische interventies, voornamelijk monocentrum, omvatten kleine sample size, passen geen minimale follow-up periode toe, of includeren patiënten met een beperkte follow-up (<6 maanden). Multicenter-onderzoeken die PROMs op meerdere tijdstippen evalueren ontbreken, hoewel deze data belangrijk zijn voor het vaststellen van de lange termijn lasten van letsels, om informatie te kunnen geven over prognoses en om behandelbeslissingen te begeleiden. Deze uitdagingen kunnen de implementatie van veelbelovende interventies vertragen en onderstrepen het belang van de implementatie van gestandaardiseerde uitkomstmaten voor de traumapopulatie. De resultaten van de studie beschreven in **Hoofdstuk 9** suggereren dat de nieuwere Patient-Reported Outcomes Measurement Information System (PROMIS) vragenlijsten kunnen worden toegepast als uitkomstmaat voor het evalueren van de schouderfunctie na proximale humerus fracturen. Gezien het gebrek aan data over PROMIS scores na proximale humerus fracturen, zijn de resultaten van deze studie bedoeld om referentiewaarden te verschaffen die kunnen worden gebruikt voor toekomstige vergelijkingen. **Hoofdstuk 10** biedt ook referentiewaarden die kunnen worden gebruikt voor toekomstige vergelijking bij patiënten met een distale humerus fractuur. Verder benadrukt dit hoofdstuk het belang van het meten van zowel klinische als patiënt-gerapporteerde uitkomsten bij het evalueren van behandelingen voor distale humerus fracturen. In **Hoofdstuk 11** hebben we de operatieve behandeling van bicondylaire tibiaplateau fracturen geëvalueerd, welke eerder nog niet uitgebreid was beschreven. Bijdragend aan het feit dat er geen overeenstemming is over de optimale behandeling voor bicondylaire tibiale plateau fracturen is het ontbreken van gevalideerde PROMs. In dit hoofdstuk hebben we normatieve data en lange termijn functionele PROMIS scores vastgesteld na operatieve behandeling van bicondylaire tibiaplateau fracturen. De resultaten van dit hoofdstuk benadrukken ook het belang van het verkrijgen van zowel globale kwaliteit van leven als letsel specifieke uitkomstmaten bij het evalueren van verwondingen. In **Hoofdstuk 12** hebben we de testeigenschappen geëvalueerd van verschillende PROMs die worden gebruikt bij patiënten met

een acute achillespeesruptuur. Dit hoofdstuk toonde aan dat zelfs een letsel specifieke PROM die is gevalideerd, responsief lijk te zijn en door velen momenteel als de meest geschikt functionele uitkomstmaat wordt beschouwd, een beperkte inhoudelijke validiteit kan hebben met betrekking tot responsiviteit en het beoordelen van lange termijn functie.

Het gebruik van PROMs in trauma onderzoek is de afgelopen decennia snel toegenomen en de verwachting is dat deze trend zich zal voortzetten. Ondanks de vooruitgang en het gebruik van de PROMs in de routinematige patiëntenzorg zijn er echter nog steeds aanzienlijke uitdagingen met betrekking tot implementatie en standaardisatie. Tijdens de studies beschreven in **DEEL 2** kwamen we veel van deze uitdagingen tegen, waaronder de betrouwbaarheid en precisie van de instrumenten die werden gebruikt om uitkomsten vast te leggen.[24-26] Gestandaardiseerde kwaliteit van leven en functionele uitkomstmaten bleken moeilijk te implementeren voor de traumapopulatie, aangezien er een bijna ontelbare combinatie is van verwondingen die secundair zijn aan verschillende traumamechanismen en omstandigheden.[27] Er zijn meer studies nodig om te bepalen welke specifieke uitkomstmaten moeten worden geïnstitutionaliseerd en hoe deze kunnen worden aangepast voor verschillende letsels en patiënt specifieke factoren.[27] Deze uitdagingen omvatten de betrouwbaarheid en precisie van de instrumenten die worden gebruikt om de gewenste uitkomst vast te leggen. Eerdere trauma studies, waarin vergelijkbare aandoeningen werden geëvalueerd, hebben van oudsher een verscheidenheid aan "traditionele" PROMs gebruikt, waardoor het moeilijk is om resultaten te vergelijken. Bovendien kan de voltooiing van de eerdere "traditionele" uitkomstmaten omslachtig en tijdrovend zijn. Daarom is de uitdaging hoe uitkomstmaten tussen verschillende patiëntenpopulaties en studies kunnen worden vergeleken, en hoe de betrouwbaarheid en efficiëntie van de PROMs kan worden vergroot terwijl de responslast voor patiënten wordt verlaagd.[24-26]

Referenties:

1. Chapman SJ, Shelton B, Mahmood H, Fitzgerald JE, Harrison EM, Bhangu A. Discontinuation and non-publication of surgical randomised controlled trials: observational study. *Bmj.* 2014;349:g6870.
2. Beks RB, Bhashyam AR, Houwert RM, et al. When observational studies are as helpful as randomized trials: Examples from orthopedic trauma. *The journal of trauma and acute care surgery.* 2019;87(3):730-732.
3. Balogh ZJ, Reumann MK, Gruen RL, et al. Advances and future directions for management of trauma patients with musculoskeletal injuries. *Lancet (London, England).* 2012;380(9847):1109-1119.
4. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *Bmj.* 2002;324(7351):1448-1451.
5. Houwert RM, Beks RB, Dijkgraaf MGW, et al. Study methodology in trauma care: towards question-based study designs. *Eur J Trauma Emerg Surg.* 2019.
6. Ioannidis JP, Haidich AB, Pappa M, et al. Comparison of evidence of treatment effects in randomized and nonrandomized studies. *Jama.* 2001;286(7):821-830.
7. Hemkens LG, Contopoulos-Ioannidis DG, Ioannidis JPJb. Agreement of treatment effects for mortality from routinely collected data and subsequent randomized trials: meta-epidemiological survey. 2016;352.
8. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *The New England journal of medicine.* 2000;342:1878-1886.
9. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *The New England journal of medicine.* 2000;342:1887-1892.
10. Abraham NS, Byrne CJ, Young JM, Solomon MJ. Meta-analysis of well-designed nonrandomized comparative studies of surgical procedures is as good as randomized controlled trials. *Journal of clinical epidemiology.* 2010;63:238-245.
11. Gosler MW, Testroote M, Morrenhof JW, Janzing HM. Surgical versus non-surgical interventions for treating humeral shaft fractures in adults. *Cochrane Database Syst Rev.* 2012;1:Cd008832.
12. Clement ND. Management of Humeral Shaft Fractures; Non-Operative Versus Operative. *Arch Trauma Res.* 2015;4(2):e28013.
13. Song J, Yu A-X, Li Z-H. Comparison of conservative and operative treatment for distal radius fracture: a meta-analysis of randomized controlled trials. *International journal of clinical and experimental medicine.* 2015;8:17023-17035.
14. Ju J-H, Jin G-Z, Li G-X, Hu H-Y, Hou R-X. Comparison of treatment outcomes between nonsurgical and surgical treatment of distal radius fracture in elderly: a systematic review and meta-analysis. *Langenbeck's archives of surgery.* 2015;400:767-779.
15. Chen Y, Chen X, Li Z, Yan H, Zhou F, Gao W. Safety and Efficacy of Operative Versus Nonsurgical Management of Distal Radius Fractures in Elderly Patients: A Systematic Review and Meta-analysis. *The Journal of hand surgery.* 2016;41:404-413.
16. Armstrong KA, von Schroeder HP, Baxter NN, Zhong T, Huang A, McCabe SJ. Stable rates of operative treatment of distal radius fractures in Ontario, Canada: a population-based retrospective cohort study (2004-2013). *Canadian Journal of Surgery.* 2019(1488-2310 (Electronic)).
17. Khan RJ, Carey Smith RL. Surgical interventions for treating acute Achilles tendon ruptures. *Cochrane Database Syst Rev.* 2010(9):Cd003674.
18. Jiang N, Wang B, Chen A, Dong F, Yu B. Operative versus nonoperative treatment for acute Achilles tendon rupture: a meta-analysis based on current evidence. *International orthopaedics.* 2012;36(4):765-773.
19. Soroceanu A, Sidhwa F, Aarabi S, Kaufman A, Glazebrook M. Surgical versus nonsurgical treatment of acute Achilles tendon rupture: a meta-analysis of randomized trials. *J Bone Joint Surg Am.* 2012;94(23):2136-2143.
20. Wilkins R, Bisson LJ. Operative versus nonoperative management of acute Achilles tendon ruptures: a quantitative systematic review of randomized controlled trials. *Am J Sports Med.* 2012;40(9):2154-2160.
21. Ganestam A, Kallemose T, Troelsen A, Barfod KW. Increasing incidence of acute Achilles tendon rupture and a noticeable decline in surgical treatment from 1994 to 2013. A nationwide registry study of 33,160 patients. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA.* 2016;24(12):3730-3737.
22. Egger AC, Berkowitz MJ. Achilles tendon injuries. *Current reviews in musculoskeletal medicine.* 2017;10(1):72-80.
23. Gabbe BJ, Simpson PM, Harrison JE, et al. Return to Work and Functional Outcomes After Major Trauma: Who Recovers, When, and How Well? *Annals of surgery.* 2016;263(4):623-632.
24. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for Orthopaedic Outcomes Measurement. *J Am Acad Orthop Surg.* 2016;24(11):744-749.
25. Cheung EC, Moore LK, Flores SE, Lansdown DA, Feeley BT, Zhang AL. Correlation of PROMIS with Orthopaedic Patient-Reported Outcome Measures. *JBJS reviews.* 2019;7(8):e9.
26. Gagnier JJ. Patient reported outcomes in orthopaedics. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society.* 2017;35(10):2098-2108.
27. Bhashyam AR, van der Vliet QMJ, Ochen Y, et al. Injury-related variation in patient-reported outcome after musculoskeletal trauma: a systematic review. *Eur J Trauma Emerg Surg.* 2019.

ACKNOWLEDGEMENTS

(Dankwoord)

If I have seen further, it is by standing on the shoulders of Giants,
(Isaac Newton, (1642-1727), English mathematician, astronomer, and physicist)

Geachte prof. dr. R.H.H. Groenwold, beste Rolf, bedankt voor, zoals je het recent noemde, het hoogtepunt van mijn academische leven. Vorig jaar heb ik je mogen volgen naar Leiden waar ik heb genoten van onze gedachtewisselingen, je inzichten en je bevlogen manier van onderwijzen. Dankzij jouw passie voor de epidemiologie en de manier waarop je deze kennis op anderen kunt overbrengen, heb ik mij als volwaardig onderzoeker kunnen ontwikkelen. Bovenal wil ik je bedanken voor je inzet en enthousiasme gedurende de totstandkoming van dit proefschrift.

Geachte prof. dr. L.P.H Leenen, uw tomeloze inzet om de traumazorg in Nederland te blijven verbeteren is bewonderenswaardig. Ik wil u graag bedanken voor de begeleiding de afgelopen jaren en kijk ernaar uit in de toekomst vaker met u van gedachten te kunnen wisselen over de inrichting van de traumazorg.

Geachte dr. R.M. Houwert, beste Marijn, elke reis begint met een eerste stap en het is alweer vijf jaar geleden dat jij mij de kans gaf om de eerste stap voor dit proefschrift te zetten. Hierbij kreeg ik altijd een grote mate van vrijheid, waarbij je mij met de juiste vragen aan het denken wist te zetten. Ik wil je graag bedanken voor je onuitputbare optimisme en toewijding gedurende dit promotietraject. Niet alleen op wetenschappelijk gebied, maar ook in leiderschap en daadkracht. De kansen die je mij tijdens het gehele promotietraject hebt geboden, hebben geleid tot waar ik nu sta.

Geachte dr. D. van der Velde, beste Detlef, het is een eer om jou mijn copromotor te mogen noemen. Je hebt mij de afgelopen jaren niet alleen wetenschappelijk begeleid, maar ook persoonlijk, waardoor ik mijzelf als mens heb kunnen ontwikkelen. Bedankt voor je bevlogenheid, klinische blik en uitstekende gevoel voor humor, welke je niet alleen een uitzonderlijke copromotor maken, maar bovenal een goede vriend.

Dear dr. Marilyn Heng, I admire your academic leadership, while being a surgeon at a world-leading orthopaedics department. You were always willing to discuss projects, read papers, and revise manuscripts with astonishing speed and precision. Your academic guidance and enthusiastic encouragement have elevated this thesis to a higher level and will keep inspiring me to move forward.

Dear co-authors, I would like to thank you for the collaboration, which made this thesis possible.

Beste mede onderzoekers, bedankt voor alle hulp en memorabele escapades tijdens congressen. Beste Reinier en Quirine, het is een eer dat ik jullie onderzoekslijnen heb mogen voortzetten. Beste Jesse, waar eerdergenoemde de basis hebben gelegd, heb jij dit proefschrift vormgegeven. Het is alweer 2 jaar geleden dat we samen als "the Dutchies" deel uitmaakten van het team van dr. Heng en ik ben blij dat ik tegenwoordig als collega in het Anton weer kan genieten van je droge humor.

MGH Orthopaedic Trauma team, Marian, Scott, Karen, Elizabeth, Madeline, Mira, and Abhiram, thanks for making me feel at home in Boston. A special thanks goes to Michael McTague, the driving force behind this team.

Dear colleagues of the department of Clinical Epidemiology and C7-81 roomies, thanks for showing me the world of causal inference, prediction modeling, missingness, multiple testing, and cookies on Friday. Dear Tamara, thanks for all your help during the final stages with the thesis formalities.

Beste Bostonians, Appletoni's, Baltimoreans, in het bijzonder Joeky, Rens, Noortje, Sjors, Sabine, Reinier, Michiel, Paul, Olivier, Quirina en Vincent, bedankt voor de onvergetelijke tijd! De weekenden zeilen, kamperen, hiken, tripjes naar New York, Las Vegas, avonden in Club Café en zelfs het hangen aan de gewichten in de gym waren een welkome afwisseling na een week in de cubicles in het MGH.

Beste heren van de Societas, Jaarclub Bockereyder -Thierry in het bijzonder-, Boni en Huize Alba, mijn investering in dit proefschrift ging meer dan eens ten koste van tijd met jullie. Desalniettemin hebben jullie de afgelopen jaren wel gezorgd voor de nodige ontspanning en mooie herinneringen. Ik koester onze vriendschappen en jullie gevraagde of ongevraagde meningen. Tsjoch!

Beste paranimfen, Joeky en Joep, bedankt voor het vervullen van deze rol. Het is een eer om jullie aan mijn zijde te hebben bij de verdediging van dit proefschrift.

Lieve pap en mam, dank voor jullie oneindige liefde en steun. De afgelopen jaren waren niet altijd makkelijk, maar ik vertrouw erop dat onze band hierdoor uiteindelijk sterker wordt. Jullie opvoeding heeft ons van veerkracht en volharding voorzien, waardoor wij ons op deze wijze hebben kunnen ontwikkelen. Ik heb grote bewondering voor de manier waarop jullie altijd hebben gestreden voor en gehamerd op het belang van onze opleidingen, zonder dat jullie zelf de kans hebben gehad om onderwijs te genieten. Zonder jullie steun was ik nooit zo ver gekomen.

Tenslotte, lieve Melanie, bedankt voor je liefde en aanmoediging. Mijn onstuimigheid weet je altijd te bedaren met geduld, begrip en steun. Zonder jou had ik het zeker niet gehaald. Met grote interesse en aanstekelijk enthousiasme heb je elk stuk in dit proefschrift gelezen en gereviseerd. Ik kijk enorm uit naar onze volgende avonturen.

LIST OF PUBLICATIONS

1. Beks RB, **Ochen Y**, Frima H, et al. Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. *Journal of shoulder and elbow surgery.* 2018;27:1526-1534.

2. **Ochen Y**, Beks RB, van Heijl M, et al. Operative treatment versus nonoperative treatment of Achilles tendon ruptures: systematic review and meta-analysis. *BMJ (Clinical research ed).* 2019;364:k5120.

3. **Ochen Y**, Frima H, Houwert RM, et al. Surgical treatment of Neer type II and type V lateral clavicular fractures: comparison of hook plate versus superior plate with lateral extension: a retrospective cohort study. *European journal of orthopaedic surgery & traumatology : orthopedie traumatologie.* 2019;29(5):989-997.

4. **Ochen Y**, Emmink BL, Kootstra TJM, Bosman WM, Wittich P, Keizer J. Nonunion of the Capitate due to Delayed Diagnosis of an Isolated Fracture: A Case Report. *JBJS case connector.* 2019;9(4):e0382.

5. Bhashyam AR, van der Vliet QMJ, **Ochen Y**, et al. Injury-related variation in patient-reported outcome after musculoskeletal trauma: a systematic review. *Eur J Trauma Emerg Surg.* 2020;46(4):777-787.

6. **Ochen Y**, Beks RB, Emmink BL, et al. Surgical treatment of acute and chronic AC joint dislocations: Five-year experience with conventional and modified LARS fixation by a single surgeon. *Journal of orthopaedics.* 2020;17:73-77.

7. **Ochen Y**, Peek J, McTague MF, et al. Long-term outcomes after open reduction and internal fixation of bicondylar tibial plateau fractures. *Injury.* 2020;51(4):1097-1102.

8. van der Vliet QMJ, **Ochen Y**, McTague MF, et al. Long-term outcomes after operative treatment for tibial pilon fractures. *OTA International.* 2019;2(4):e043.

9. Bhashyam AR, **Ochen Y**, van der Vliet QMJ, et al. Association of Patient-reported Outcomes With Clinical Outcomes After Distal Humerus Fracture Treatment. 2020;4(2):e19.00122.

10. van de Wall BJM, **Ochen Y**, Beeres FJP, et al. Conservative vs. operative treatment for humeral shaft fractures: a meta-analysis and systematic review of randomized clinical trials and observational studies. *J Shoulder Elbow Surg.* 2020;29(7):1493-1504.

11. **Ochen Y**, Peek J, van der Velde D, et al. Operative vs Nonoperative Treatment of Distal Radius Fractures in Adults: A Systematic Review and Meta-analysis. *JAMA Netw Open.* 2020;3(4):e203497.

12. Peek J, **Ochen Y**, Saillant N, et al. Traumatic rib fractures: a marker of severe injury. A nationwide study using the National Trauma Data Bank. *Trauma Surg Acute Care Open. 2*020;5(1):e000441.

13. **Ochen Y**, Gademan MGJ, Nelissen RGHH, et al. Potential value of observational studies of surgical orthopedic intervention:two studies of total hip arthroplasty. *Submitted.*

14. **Ochen Y**, Bhashyam AR, van der Velde D, et al. Patient-reported and clinical outcomes after proximal humeral fractures. *Submitted.*

15. **Ochen Y**, Guss D, Houwert RM, et al. Validation of PROMIS Physical Function for evaluating outcome following acute Achilles tendon rupture. *The Orthopaedic Journal of Sports Medicine.*

CURRICULUM VITAE

Yassine Ochen was born on December 6th, 1990 in Utrecht, the Netherlands, where he was raised together with his two brothers and two sisters. In 2009 he graduated from Gymnasium at the St. Bonifatiuscollege in Utrecht and started Medical School at Utrecht University. During his studies he was an active member of several student associations.

In his final years of Medical School Yassine gained increasing interest in the surgical field during his clinical internships at the Department of Surgery of the St. Antonius Hospital in Nieuwegein. It was during this period that he met dr. R.M. Houwert and dr. D. van der Velde, who offered him a chance to join the trauma research group under the guidance of prof. dr. L.P.H Leenen at the Department of Surgery of the University Medical Center in Utrecht. He also worked on several projects to assess the value of different study designs, during an internship at the department of Clinical Epidemiology of the Leiden University Medical Center, under the guidance of prof. dr. R.H.H. Groenwold.
He graduated from Medical School in 2018 and continued his work as a PhD candidate during a one-year research fellowship at the department of Orthopaedic Surgery of the Massachusetts General Hospital in Boston. There he worked on different projects concerning the international validation of Patient-Reported Outcome Measures in orthopaedic trauma, under the supervision of dr. Marilyn Heng, Assistant Professor at the department of Orthopaedic Surgery, Harvard Medical School. After his return to the Netherlands in 2019, he worked as a Clinical Research Coordinator at the Department of Surgery of the St. Antonius Hospital in Nieuwegein.
As a PhD candidate, he was offered the chance to present his work at several national and international conferences. His first publication was featured in an article by Reuters News: "Mix of risk, benefit in Achilles tendon surgery", following the publication in the BMJ: "Operative treatment versus nonoperative treatment of Achilles tendon ruptures: systematic review and meta-analysis". He was rewarded several research grants for his projects and fellowships. He also wrote the study protocol for the LADON study, an international multicenter prospective cohort study, to evaluate proximal humeral fracture management in the Netherlands and Switzerland.
In September 2019, he continued his work as a PhD candidate and followed the training to become an epidemiologist (Epidemioloog-B) at the Clinical Epidemiology department in Leiden, under the supervision of prof. dr. R.H.H. Groenwold.

Yassine currently works as a surgical resident, not in training, at the Department of Surgery of the St. Antonius Hospital in Nieuwegein, under the supervision of dr. D. Boerma.