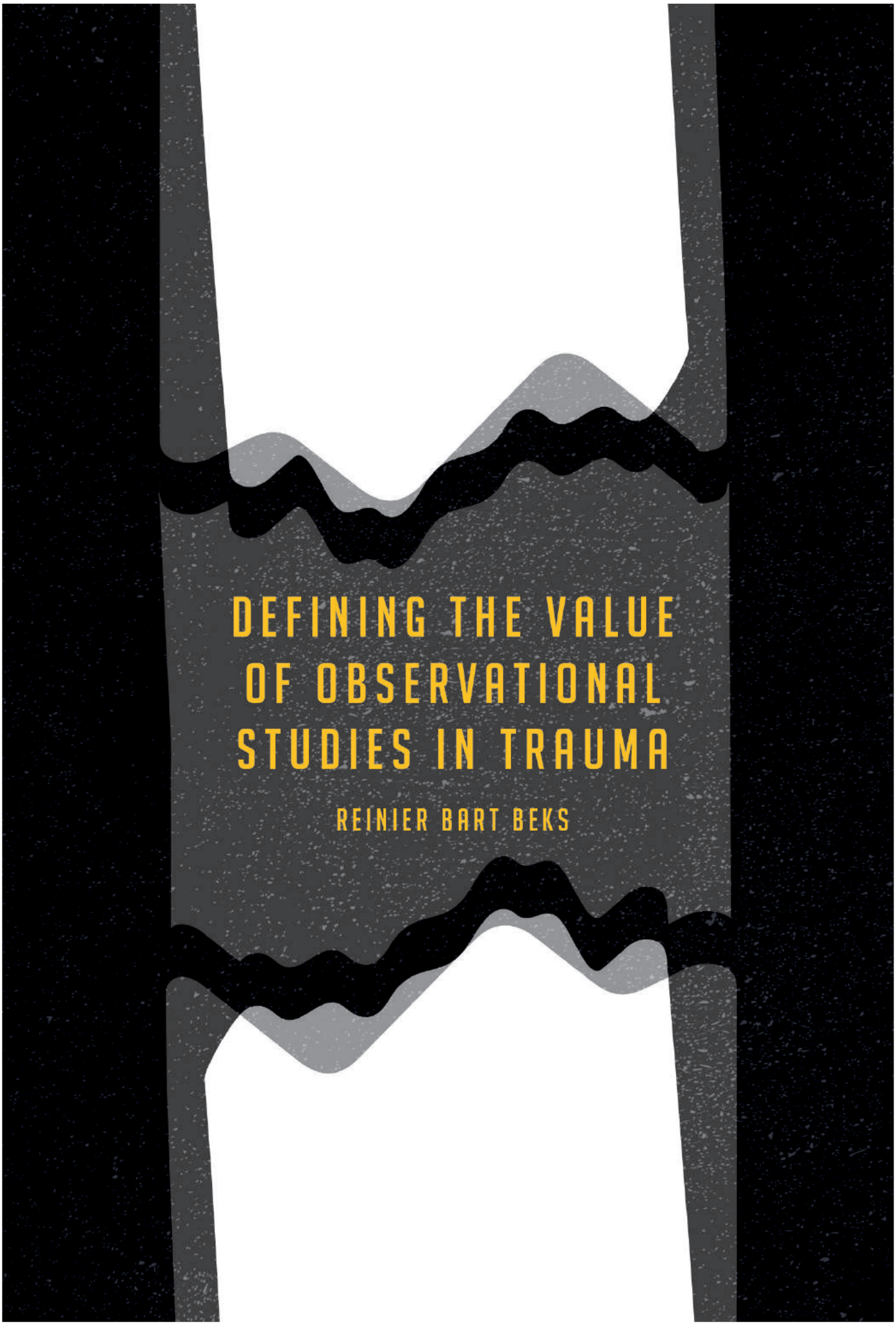

DEFINING THE VALUE OF OBSERVATIONAL STUDIES IN TRAUMA
REINIER BART BEKS

DEFINING THE VALUE OF OBSERVATIONAL STUDIES IN TRAUMA

REINIER BART BEKS

Defining the value of observational studies in trauma

PhD thesis, Utrecht University

Cover design: Andrea Volkers

Printed by: Noovamediaproductions

ISBN: 978-90-9031-663-5

The publication of this thesis was financially supported by:
Chipsoft

Defining the value of observational studies in trauma

Definiëring van de waarde van observationele studies in trauma

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 18 april 2019 des middags te 4.15 uur

door

Reinier Bart Beks
geboren op 24 augustus 1989
te Haren

Promotoren:
Prof. dr. L.P.H. Leenen
Prof. dr. R.H.H. Groenwold

Copromotoren:
Dr. R.M. Houwert
Dr. D. van der Velde

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION AND THESIS OUTLINE

Introduction

Medical decision making generally used to be the result of the subjective interpretation and clinical view of the doctor which was subject to personal believes and other factors. A change started when Alvan Feinstein identified different forms of bias associated with clinical reasoning and wrote about this in his book *Clinical Judgment* in 1968.[1] In 1972, Archie Cochrane and afterwards many others, published on the need for controlled trials to support or abandon medical interventions which ultimately led to the introduction of *evidence based medicine* (EBM).[2] The goal of EBM is to optimize medical decision making based on the best available evidence and has resulted in a clear hierarchy of different research designs.

The randomized placebo controlled trial (RCT) is considered the gold standard for clinical research.[3] In an RCT patients are randomized between a treatment and a control group while both the patient and the physician are blinded for the treatment choice. This results in two similar and comparable groups and consequently observed differences in outcome between both groups can be ascribed to the treatment. In observational studies on treatment effect of medical interventions, there is no randomization and blinding of the patient and physician for the treatment choice is not possible. Therefore, an observational study is considered to be subject to bias and results of this type of study are considered less credible for decision making.

> *"As early as 1703 orders had been issued to Naval surgeons requiring them to keep journals and records of cases, and in 1731 there were further orders about these returns ; but their value to posterity was not always recognized and many were destroyed, including all Cuming's journals."*[8]
>
> Ralph Cuming, a Navel surgeon, was acknowledged to be the originator of performing an amputation of the upper limb in a patient after a war injury in 1808. Unfortunately, the value of his journals was not recognized and few of the records kept by the Navel surgeons have remained. The story of Cuming illustrates an early example of observational research not valued to its full potential in advancing surgical treatment.

Over the years EBM more and more focused on the use of RCTs for treatment recommendation while results from observational studies are considered inferior and less applicable. The RCT has become the standard for pharmacological research and also in the surgical field, RCTs are increasingly performed.[4,5] However, surgical RCTs have been fraught with challenges such as recruitment difficulty, small sample sizes, failed randomization, and inadequate blinding.[4–7] Obstacles for high quality surgical RCTs include patient and surgeon related factors as well as methodological factors.[4] Compared to a more complex working mechanism of medical drugs, surgical treatment is a more tangible concept and consequently patients might have a stronger treatment preference and are less willing to participate in randomization. Preference of the surgeon is another major factor and is influenced by personal experience, culture, and technical expertise for a specific surgical treatment. As a

result, the required state of equipoise to perform an RCT is difficult to reach. What is more, methodological challenges come from the strict in/exclusion criteria and (consequently) the often low patient numbers in surgical RCTs which makes the interpretation and implementation of results more difficult. These challenges are further impeded by a lack of funding.[4] Finally, surgery as opposed to pharmacological treatment, is a combination of pre-, peri-, and postoperative procedures subject to the skill and expertise of the surgical team. One could question to what extent such a complex process is captured with an RCT which aims to control for all influencing variables.

Observational studies have an important role in reporting of adverse events, long term outcomes, and the effects of new surgical techniques. To study treatment effect, results from observational studies as compared to RCTs, are often better generalizable because of less strict inclusion criteria, have more possibilities such as subgroup analysis, and an observational study is easier, faster and less expensive to perform. With these advantages, it is important to know when observational studies might produce credible results as an alternative for RCTs, especially in a world with fast technological advancements, continuous implementation of new techniques and limited financial possibilities.

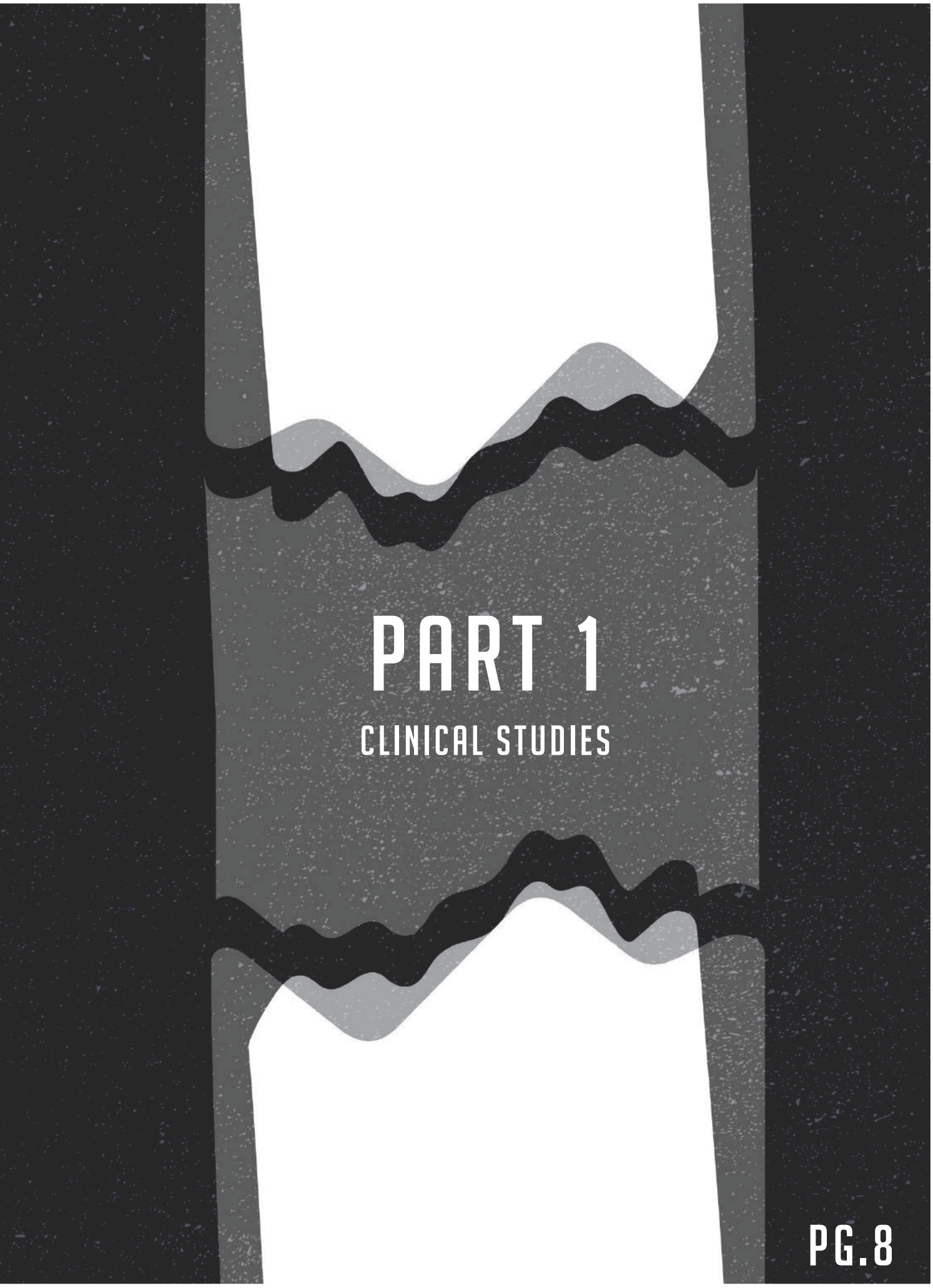
Outline

The overall aim of the this thesis is to evaluate the role of observational studies compared to RCTs in trauma patients with respect to the current hierarchy of research designs.

This thesis consists of three parts. The *first part* covers different research questions in trauma patients for which an observational design is suitable and even indispensable: (rare) complications after surgical treatment for distal biceps tendon repair (chapter 2), long-term follow-up after an existing surgical technique for proximal humerus fractures (chapter 3) or after a new surgical technique for multiple rib fractures (chapter 4), and evaluation of rib fixation compared to (conservative) standard of care (chapter 5).
The *second part* includes methodological considerations of observational studies. A viewpoint on how to value observational studies in trauma patients (chapter 6) is followed by a comparison of results of observational studies and RCTs in patients with proximal humerus fractures, flail chest, and multiple rib fractures (chapter 7 – 8).
The *third part* summarizes differences in observational studies and RCTs and when observational studies can produce valid treatment effects in trauma patients (chapter 9). Finally, a study protocol is presented which aims to function as a template for future research in trauma patients (chapter 10). This thesis concludes with a general discussion and future perspective (chapter 11).

References

1. Feinstein AR. *Clinical Judgment*. R.E. Krieger Pub. Co; 1974.
2. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Services*. Nuffield Provincial Hospitals Trust; 1972.
3. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med*. 2008;5(3):e67. doi:10.1371/journal.pmed.0050067.
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. Sibai T, Carlisle H, Tornetta P. The darker side of randomized trials: recruitment challenges. *The Journal of bone and joint surgery American volume*. 2012;94 Suppl 1(Suppl 1):49-55. doi:10.2106/JBJS.L.00240.
6. Chan S, Bhandari M. The Quality of Reporting of Orthopaedic Randomized Trials with Use of a Checklist for Nonpharmacological Therapies. *The Journal of Bone and Joint Surgery (American)*. 2007;89(9):1970. doi:10.2106/JBJS.F.01591.
7. James MA. Insufficient Post Hoc Statistical Power: A Potential Pitfall of a Well-Designed Randomized Controlled Surgical Trial: Commentary on an article by Geert A. Buijze, MD, PhD, et al.: "Three-Dimensional Compared with Two-Dimensional Preoperative Planning of Corrective Osteotomy for Extra-Articular Distal Radial Malunion. A Multicenter Randomized Controlled Trial". *The Journal of bone and joint surgery American volume*. 2018;100(14):e98. doi:10.2106/JBJS.18.00256.
8. Keevil JJ. Ralph Cuming and the interscapulothoracic amputation in 1808. *The Journal of bone and joint surgery British volume*. 1949;31B(4):589-595.



PART 1

CLINICAL STUDIES

CHAPTER 2

FACTORS ASSOCIATED WITH ADVERSE EVENTS AFTER DISTAL BICEPS TENDON REPAIR OR RECONSTRUCTION.

R.B. Beks, F.M.A.P. Claessen, L.S. Oh, D. Ring, N.C. Chen

Journal of Shoulder and Elbow Surgery. 2016

Abstract

Background

Factors associated with adverse events after distal biceps tendon repair or reconstruction are incompletely understood. This study examined factors associated with adverse events, prevalence of adverse events, and rate of second surgeries after distal biceps repair or reconstruction.

Methods

Three hundred seventy-three adult patients who underwent repair or reconstruction of the distal biceps tendon between January 2002 and March 2015 at one of three area hospitals were analyzed to determine factors associated with adverse events after surgical repair or reconstruction of the distal biceps tendon.

Results

Eighty-two of 373 distal biceps tendon repairs or reconstruction (22%) had an adverse event, 5.3% were major adverse events. In multivariable analysis, a single incision anterior approach and obesity were associated with a higher rate of adverse events. Fifteen patients (18% of patients with an adverse event and 4% of all patients) had a second surgery after distal biceps tendon surgery.

Conclusion

Patients should be counseled that one in five patients will have a minor complication and one in 20 patients will have a major complication after surgery on the distal biceps tendon. The most common adverse event is lateral antebrachial cutaneous neuroapraxia.

Introduction

Untreated distal biceps tendon rupture results in an average loss of 30-40% of supination strength and reduced endurance [9, 16, 17]. Surgical reattachment of the ruptured tendon will improve strength, although not completely in some cases [2, 6, 11]. Typically, repair restores supination strength to about 80-90% of normal [10, 12].

Between 10% and 40% of patients have been reported to experience adverse events after surgical repair reconstruction of the distal biceps tendon [1, 4, 7, 14, 18]. Cain and colleagues reported a 36% complication rate in 198 consecutive patients after distal biceps tendon repair or reconstruction either with a bone tunnel, suture anchor, or cortical button [5]. Major complications such as posterior interosseous nerve palsy, heterotopic ossification and re-rupture occurred in 9% of patients [5]. Minor complications, including superficial infection, wound separation and lateral antebrachial cutaneous neuropathy were more common [5].

Factors associated with adverse events after surgery for the distal biceps tendon are incompletely understood. We used a database of patients treated at three hospitals to study the primary null hypothesis that there are no factors associated with adverse events after surgical repair or reconstruction of the distal biceps tendon. Additionally we addressed the following secondary study questions: (1) what is the prevalence of adverse events after surgical repair or recontruction of the distal biceps tendon and (2) what is the rate of a second surgery for an adverse event.

Material and Methods

Study design, setting and participants

This retrospective study was approved by our institutional review board. We reviewed 409 patients who underwent surgery to address rupture or tendinopathy of the distal biceps tendon between January 2002 and March 2015 at three regional hospitals. Two of these hospitals are level 1 trauma centers and one hospital is a community hospital tied to a level 1 trauma center.

We identified patients using Current Procedural Terminology (CPT) codes for distal biceps tendon rupture (codes: 24340, 24341, 24342).
A multi-institutional Research Patient Data Registry (RPDR) was used to collect data of patients who underwent surgical repair or reconstruction of the distal biceps tendon. RPDR is a centralized clinical data registry that comprises diagnostic codes (International Classification of Diseases, 9th revision code), CPT codes, demographic information (e.g. sex, date of birth, and race), radiology and operative reports, and visit notes.

Thirty-six patients were excluded: 19 did not have adequate documentation; 10 had prior surgical repair of the biceps tendon in another hospital; six patients had biceps injury related

to a traumatic wound, and one patient had debridement of the distal biceps tendon without repair. This resulted in a final cohort of 373 patients.

Outcome measures

We tracked adverse events (e.g. re-rupture, nerve injury, infection, heterotopic ossification) after surgical repair or reconstruction of the distal biceps tendon as described in the medical record. We considered symptomatic heterotopic ossification, re-rupture, deep infection, and motor nerve dysfunction as major adverse events. We considered suture abscesses and radial forearm numbness minor adverse events. Forearm numbness was not subcategorized into injury of specific nerves because of ambiguous visit notes and often unreliable discrimination between different sensory nerves. It is likely that most of these are injuries to the lateral antebrachial cutaneous or – much less likely – the radial sensory nerve.

Explanatory variables

The following patient characteristics were obtained: age, sex, obesity, and smoking. Additionally, we collected the following surgery related factors: hospital, experience of the surgeon after board examination, rupture or tendinopathy, acute or chronic rupture, surgical technique (bone tunnel, suture anchor, cortical button, and tenodesis screw), surgical approach (anterior using one incision, anterior using two incisions, posterior using one incision, or anteroposterior (combined) using two incisions), use of graft, and time from injury to surgery. Obesity and smoking were retrieved using ICD-9 codes. All other variables were obtained through chart review. We included only the first distal biceps tendon repair or reconstruction in cases of bilateral tendon rupture.

The mean age in this cohort was 48 years (range, 20 – 74 years) and included five females (1.3 %). Eighty-two of 373 patients who underwent surgical repair or reconstruction of the distal biceps tendon (22%) had a recorded adverse event (Table I). Fifty-nine patients had nerve injury after surgical repair or reconstruction of the distal biceps tendon (16%). Six patients had a motor palsy including 4 posterior interosseous nerve palsies and two median nerve palsies. All of the nerve palsies recovered within an average of 4.7 months (range, 3 – 7.5 months). Eight patients developed heterotopic ossification restricting motion; four patients were treated using the combined approach, three using an anterior approach, and one from a posterior approach. Follow up was available an average of 34 weeks after surgery (range, 1 – 349 weeks). Fifty-one patients (14%) had less than 2 months evaluation in the medical record. We assumed that those patient had no major adverse events.

Table I. Adverse events per surgical approach after distal biceps tendon surgery (n =373)

	Surgical approach, n			Total,
Adverse event	Anterior	Posterior	Combined	n(%)
Radial side forearm numbness	44	2	7	53 (15)
Superficial infection / suture abscess	7	0	2	9 (2.4)
Heterotopic ossification	3	1	4	8 (2.1)
Posterior interosseous nerve palsy	2	1	1	4 (1.1)
Re-rupture	6	0	0	6 (1.6)
Median nerve palsy	2	0	0	2 (0.54)
Total				82 (22%)

Statistical analysis

Data was described using frequencies and percentages for dichotomous and nominal variables, mean and confidence interval for normally distributed continuous data, and median and interquartile range for non-normally distributed continuous data.

In bivariate analysis the association of adverse events with patient characteristics and surgery related factors was assessed with a student t-test for continuous explanatory variables (age), a Wilcoxon rank-sum test for non-normally distributed explanatory variables (surgeon experience and time till surgery) and ordinal categorical explanatory variables (surgical approach, surgical technique, and hospital), and a Fisher Exact test for dichotomous explanatory variables (sex, alcohol abuse, smoking, obesity, acute or chronic, rupture or tendinopathy, use of graft, and incision).

Factors with $p < 0.10$ in bivariate analysis were entered into a multivariable logistic regression analyses to assess if possible factors were independently associated with adverse events after surgical repair of the distal biceps tendon. All analyses were performed with Stata 13 (StataCorp LP) a two-tailed p value of less than 0.05 was considered significant.

Results

In bivariate analysis, factors associated with adverse events after surgical distal biceps tendon repair or reconstruction were obesity *(p* = 0.038) and surgical approach *(p* = 0.034) (Table II). A multivariable model demonstrated that obesity (odds ratio [OR] 1.88, 95% confidence interval [CI] 1.01 - 3.52, p = 0.048) and the anterior approach using a single incision technique (OR 2.30, 95% CI 1.21 – 4.37, p = 0.011) as compared to the combined approach were independently associated with adverse events (Table III).

In bivariate analysis, there were no factors associated with major adverse events after surgical distal biceps tendon repair or reconstruction (Table IV).

Fifteen patients (18% of patients with an adverse event and 4% of all patients) had a second surgery after distal biceps tendon repair or reconstruction. Seven patients had removal of heterotopic ossification; two had pain and clicking secondary to heterotopic bone, five had limited forearm rotation, and one had concomitant ulnar nerve release. Six patients had a re-rupture of their distal biceps tendon repair or reconstruction; three patients had repair using a cortical button; two patients with a suture anchor, and one patient with a tenodesis screw. Three of six patients with re-rupture had a second repair using a bone tunnel technique and three patients chose nonoperative treatment. Three patients had incision and drainage of an infection. One of the patients had three debatable surgeries on the radial nerve trying to address radial forearm pain after surgery for biceps tendinopathy.

Table II. Bivariate analysis - Factors associated with adverse events after distal biceps tendon surgery

	Adverse events n=82	No adverse events n=291	
Parameter	*Mean (CI)*	*Mean (CI)*	*P value*
Age	47 (45-49)	48 (47-49)	*0.37*
Parameter	*Median (IQR)*	*Median (IQR)*	*P value*
Surgeon experience	7 (1-15)	5 (2-11)	*0.30*
Time till surgery	16.5 (11-48)	18 (10-40)	*0.99*
Parameter	*n (%)*	*n (%)*	*P value*
Men	81 (99)	287 (99)	*0.99*
Smoking*	4 (4.9)	19 (6.5)	*0.80*
Obesity*	19 (23)	39 (13)	*0.038*
Repair within a month	54 (66)	186 (64)	*0.80*
Use of graft	2 (2.4)	11 (3.8)	*0.74*
Rupture (versus tendinopathy)	72 (88)	263 (90)	*0.54*
Surgical approach			
Combined, 2 incisions	14 (17)	93 (32)	*0.034*
Anterior, 1 incision	57 (70)	158 (54)	
Anterior, 2 incisions	7 (8.5)	30 (10)	
Posterior, 1 incision	4 (4.9)	10 (3.4)	
Incisions			
1	61 (74)	168 (58)	*0.007*
>1	21 (26)	123 (42)	
Surgical technique			
Bone tunnel	17 (21)	91 (31)	*0.41*
Suture anchor	24 (29)	70 (24)	
Cortical button + tenodesis screw	21 (26)	70 (24)	
Cortical button	18 (22)	53 (18)	
Tenodesis screw	2 (2.4)	7 (2.4)	
Hospital			
Hospital I	27 (33)	124 (43)	*0.16*
Hospital II	26 (32)	66 (23)	
Hospital III	29 (35)	101 (35)	

CI = Confidence Interval, IQR = Interquartile range

* according to medical records

Table III. Multiple logistic regression analysis: factors associated with adverse events after distal biceps tendon surgery

	Odds ratio (CI)	Standard error	*P* value
Obesity	1.88 (1.01 - 3.52)	0.60	*0.048*
Approach/incision			
Combined / 2-incision	*reference value*	*reference value*	*reference value*
Anterior / 1-incision	2.30 (1.2 - 4.4)	0.75	*0.011*
Anterior / 2-incision	1.47 (0.54 - 4.01)	0.75	*0.45*
Posterior / 1-incision	2.90 (0.79 - 10.56)	1.61	*0.107*

Discussion

The rate of adverse event after repair or reconstruction of the distal biceps tendon is reported to be as high as 40%, however, it is unclear what factors are associated with adverse events [1, 4, 7, 14, 18, 19]. Therefore we aimed to find factors associated with adverse events after surgical repair or reconstruction of the distal biceps tendon. In our study, 22% of patients experienced an adverse event (the vast majority dysfunction of the lateral antebrachial cutaneous nerve), but only 18% of patients with an adverse event had a second surgery. A single incision anterior approach and obesity were independently associated with adverse events after surgical repair or reconstruction of the distal biceps tendon.

This study had a number of limitations. First, adverse events may be underreported in retrospective studies as patients could have been treated for adverse events in a different hospital not used in our database search. Given that dysfunction of the lateral antebrachial cutaneous nerve was the most common adverse event, it is likely that a prospective study examining carefully for dysfunction of this nerve would find higher rates than a retrospective database study. This might have a substantial influence on the results. Second, we included patients with a follow up shorter than three months and assume no adverse events. We base this on prior studies where we find that few operated patients leave our health system. However, this could have resulted in a lower number of reported adverse events. Third, we used CPT codes to identify the initial diagnoses, which is subject to miscoding. A small amount of miscoding is typical for database studies. Finally, it should be noted that the posterior only approach was only used in 14 patients, all with tendinopathy, and the multivariable comparison is underpowered to make meaningful analyses for these patients. However, this study represents a large series of surgically treated distal biceps tendon repairs. A wide variety of surgical approaches and techniques are represented in the data.

Table IV. Bivariate analysis - Factors associated with major adverse events after distal biceps tendon surgery

	Major adverse events n=20	No adverse events n=353	
Parameter	*Mean (CI)*	*Mean (CI)*	*P value*
Age	50 (45-55)	48 (47-49)	*0.39*
Parameter	*Median (IQR)*	*Median (IQR)*	*P value*
Surgeon experience	3 (0-11.5)	5 (2-12)	*0.29*
Time till surgery	34.5 (9.5-93.5)	17 (10-40)	*0.19*
Parameter	*n (%)*	*n (%)*	*P value*
Men	19 (95)	349 (99)	*0.24*
Smoking*	2 (10)	21 (6)	*0.35*
Obesity*	4 (20)	54 (15)	*0.53*
Repair within a month	9 (45)	231 (65)	*0.091*
Use of graft	1 (5.0)	12 (3.4)	*0.52*
Rupture (versus tendinopathy)	16 (80)	319 (90)	*0.13*
Surgical approach			
Combined, 2 incisions	5 (25)	102 (29)	*0.12*
Anterior, 1 incision	9 (45)	206 (58)	
Anterior, 2 incisions	4 (20)	33 (9.4)	
Posterior, 1 incision	2 (10)	12 (3.4)	
Incisions			
1	11 (55)	218 (62)	*0.64*
>1	9 (45)	136 (38)	
Surgical technique			
Bone tunnel	6 (30)	102 (29)	*0.44*
Suture anchor	4 (20)	90 (26)	
Cortical button + tenodesis screw	3 (15)	88 (25)	
Cortical button	6 (30)	65 (18)	
Tenodesis screw	1 (5.0)	8 (2.3)	
Hospital			
Hospital I	11 (55)	140 (40)	*0.33*
Hospital II	5 (25)	87 (25)	
Hospital III	4 (20)	126 (36	

CI = Confidence Interval, IQR = Interquartile range

* according to medical records

Our finding that adverse events after surgical repair or reconstruction of the distal biceps tendon are more than twice as likely in patients treated with a single incision anterior approach than those treated with a two incision approach is in agreement with the study of El-Hawary et al. [8]. The anterior dissection and retraction likely leads to lateral antebrachial cutaneous neuroapraxia and some of the branches may be permanently injured [3, 13, 15]. We speculate that the size of the anterior exposure may relate to the risks of lateral antebrachial cutaneous nerve dysfunction, with smaller incisions less likely to cause sensory disturbances. Obesity – the other factor associated with adverse events – may compromise exposure leading to a larger incision or more retraction on the lateral antebrachial cutaneous nerve. In contrast with earlier studies, surgery for tendinopathy was not associated with adverse events [5, 14].

Our 22% rate of adverse events is comparable to the range of 10 to 40% and the 36% in other large published series [4, 5, 7, 14]. Retrospective data is at risk for underreporting, particularly with relatively adaptable problems such as lateral antebrachial cutaneous nerve dysfunction, therefore the true rate of adverse events might be higher in a prospective cohort. Underreporting could be the result of transient nerve paresthesia not being considered as an adverse event because it may be an expected outcome after surgery for the distal biceps. Also, patients may not notice or mention a small area of numbness, particularly when they are in pain.
Nearly one in five patients with an adverse event had a second surgery, most often for removal of heterotopic ossification. A two-incision technique (combined approach) was not associated with heterotopic ossification in our study[14].

Conclusion

Patients should be counseled that one in five patients will have a minor or major complication (most often lateral antebrachial cutaneous nerve dysfunction) and one in 20 patients will have a major complication after repair or reconstruction of the distal biceps tendon. Obese patients seem to be at greater risk. Future studies should address the placement and size of the anterior incision and potential modifications in obese patients as a way to protect the lateral antebrachial cutaneous nerve.

References

1 Austin L, Mathur M, Simpson E, Lazarus M. Variables influencing successful two-incision distal biceps repair. Orthopedics 2009;32:88.

2 Baker BE, Bierwagen D. Rupture of the distal tendon of the biceps brachii. Operative versus non-operative treatment. The Journal of bone and joint surgery American volume 1985;67:414-417.

3 Baratz M, King GJ, Steinmann S. Repair of distal biceps ruptures. The Journal of hand surgery 2012;37:1462-1466. 10.1016/j.jhsa.2012.02.008

4 Bisson L, Moyer M, Lanighan K, Marzo J. Complications associated with repair of a distal biceps rupture using the modified two-incision technique. Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2008;17:67S-71S. 10.1016/j.jse.2007.04.008

5 Cain RA, Nydick JA, Stein MI, Williams BD, Polikandriotis JA, Hess AV. Complications following distal biceps repair. The Journal of hand surgery 2012;37:2112-2117. 10.1016/j.jhsa.2012.06.022

6 Cil A, Merten S, Steinmann SP. Immediate active range of motion after modified 2-incision repair in acute distal biceps tendon rupture. The American journal of sports medicine 2009;37:130-135. 10.1177/0363546508323749

7 Cohen MS. Complications of distal biceps tendon repairs. Sports medicine and arthroscopy review 2008;16:148-153. 10.1097/JSA.0b013e3181824eb0

8 El-Hawary R, Macdermid JC, Faber KJ, Patterson SD, King GJ. Distal biceps tendon repair: comparison of surgical techniques. The Journal of hand surgery 2003;28:496-502. 10.1053/jhsu.2003.50081

9 Freeman CR, McCormick KR, Mahoney D, Baratz M, Lubahn JD. Nonoperative treatment of distal biceps tendon ruptures compared with a historical control group. The Journal of bone and joint surgery American volume 2009;91:2329-2334. 10.2106/JBJS.H.01150

10 Giacalone F, Dutto E, Ferrero M, Bertolini M, Sard A, Pontini I. Treatment of distal biceps tendon rupture: why, when, how? Analysis of literature and our experience. Musculoskeletal surgery 2015;99 Suppl 1:67-73. 10.1007/s12306-015-0360-5

11 Greenberg JA, Fernandez JJ, Wang T, Turner C. EndoButton-assisted repair of distal biceps tendon ruptures. Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2003;12:484-490. 10.1016/S1058274603001733

12 Grewal R, Athwal GS, MacDermid JC, Faber KJ, Drosdowech DS, El-Hawary R et al. Single versus double-incision technique for the repair of acute distal biceps tendon ruptures: a randomized clinical trial. The Journal of bone and joint surgery American volume 2012;94:1166-1174. 10.2106/JBJS.K.00436

13 Keener JD. Controversies in the surgical treatment of distal biceps tendon ruptures: single versus double-incision repairs. Journal of shoulder and elbow surgery /

American Shoulder and Elbow Surgeons [et al] 2011;20:S113-125. 10.1016/j.jse.2010.11.009

14 Kelly EW, Morrey BF, O'Driscoll SW. Complications of repair of the distal biceps tendon with the modified two-incision technique. The Journal of bone and joint surgery American volume 2000;82-A:1575-1581.

15 Miyamoto RG, Elser F, Millett PJ. Distal biceps tendon injuries. The Journal of bone and joint surgery American volume 2010;92:2128-2138. 10.2106/JBJS.I.01213

16 Morrey BF, Askew LJ, An KN, Dobyns JH. Rupture of the distal tendon of the biceps brachii. A biomechanical study. The Journal of bone and joint surgery American volume 1985;67:418-421.

17 Nesterenko S, Domire ZJ, Morrey BF, Sanchez-Sotelo J. Elbow strength and endurance in patients with a ruptured distal biceps tendon. Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2010;19:184-189. 10.1016/j.jse.2009.06.001

18 Nigro PT, Cain R, Mighell MA. Prognosis for recovery of posterior interosseous nerve palsy after distal biceps repair. Journal of shoulder and elbow surgery / American Shoulder and Elbow Surgeons [et al] 2013;22:70-73. 10.1016/j.jse.2012.08.001

19 Van den Bogaerde J, Shin E. Posterior interosseous nerve incarceration with endobutton repair of distal biceps. Orthopedics 2015;38:e68-71. 10.3928/01477447-20150105-92

CHAPTER 3

LONG TERM FOLLOW UP AFTER MIPO PHILOS PLATING FOR PROXIMAL HUMERAL FRACTURES.

H. Frima, C. Michelitsch, R.B. Beks, R.M. Houwert, Y.P. Acklin, C. Sommer

Archives of Orthopaedic and Trauma Surgery. 2018

Abstract

Introduction

Minimally invasive plate osteosynthesis (MIPO) has been described as a suitable technique for the treatment of proximal humerus fractures, but long-term functional results have never been reported. The aim of this study was to describe the long-term functional outcome and implant related irritation after MIPO for proximal humerus fractures.

Methods

A long-term prospective cohort analysis was performed on all patients treated for a proximal humerus fracture using MIPO with a Philos plate (Synthes, Switzerland) between December 2007 and October 2010. The primary outcome was the QuickDASH score. Secondary outcome measures were the Subjective Shoulder Value (SSV), implant related irritation and implant removal.

Results

Seventy-nine out of 97 patients (81%) with a mean age of 59 years were available for follow-up. The mean follow-up was 8.3 years (SD 0.8). The mean QuickDASH score was 5.6 (SD 14). The mean SSV was 92 (SD 11). Forty out of 79 patients (50,6%) had implant removal, and of those, 27/40 (67,5%) were due to implant related irritation. On average, the implant was removed after 1.2 years (SD 0.5). In bivariate analysis there was an association between the AO classification and the QuickDASH (p = 0.008).

Conclusion

Treatment of proximal humerus fractures using MIPO with Philos through a deltoid split approach showed promising results. A good function can be assumed due to the excellent scores of patient oriented questionnaires. However, about one third of the patients will have a second operation for implant removal due to implant related irritation.

Introduction

Proximal humerus fractures are very common and account for 5% of all fractures in the emergency department, with an incidence of 82 per 100,000 people[1-6]. The incidence has an unipolar distribution with a typical patient being relatively fit, female and more than 80 years old[7]. Most patients are treated non-operatively while one out of five will undergo surgery even though no clear benefit of operative treatment has been shown[6,8,9].

The standard approach for osteosynthesis of proximal humerus fractures is the deltopectoral approach, which generally is considered the open approach[1,10-12]. Over the past decade, there has been an increasing interest in minimally invasive plate osteosynthesis (MIPO) of proximal humerus fractures through the deltoid split approach[10,13-19]. Previously reported possible advantages of MIPO are: less soft tissue stripping and a lower risk of injury to the ascending branch of the anterior circumflex humeral artery resulting in lower rates of avascular necrosis (AVN) and shorter operation time[10,15,16,18]. Possible disadvantages are risk of damage to the axillary nerve and, in case of a later shoulder prosthesis, the need for a different second surgical approach[10]. Several studies have reported on the short-term results of this technique[10,14,15,19,20]. Although long-term results of the open approach for proximal humerus fracture treatment have been reported, little is known about the long-term results after MIPO with Philos[12,21].

The aim of this study was to analyze the long-term functional outcome after MIPO with Philos for proximal humerus fractures. Additionally we assessed implant related irritation and implant removal.

Methods

Study design

Between December 2007 and October 2010, 191 patients with a proximal humerus fracture were treated with MIPO through a 'deltoid split' approach in our center using the Philos® system (Synthes, Switzerland). Patients were operated by 16 different surgeons. Two of these surgeons performed 50% of all operations. In 2013 Acklin et al. published prospectively gathered data on the short-term outcome of 97 of these patients available for follow-up[10]. In the current study, this cohort was approached and analyzed again to obtain long-term outcome on these patients. Exclusion criteria were death, a second trauma to the operated arm, inability to answer questions, or absence of written consent. This study was approved by the Cantonal Ethic Committee Zürich (KEK-ZH-Nr. 2017-00428).

Operative procedure and indications

All patients were treated in a MIPO technique. In beach chair position a minimally invasive anterolateral deltoid split approach was performed. After reduction of the humeral head and

non-absorbable suture insertion in the tendons of the rotator cuff, a five hole Philos® plate was inserted. This was done sub-muscular, either percutaneously or with a radiolucent aiming device, under Langenbeck protection to preserve the axillary nerve. The plate was fixed to the humeral head with 4 locking screws and, depending on bone quality, with two to four conventional or locking screws to the shaft. The non-absorbable sutures were then knotted to the plate for additional stabilization and to prevent secondary dislocation.
Postoperatively, patients were allowed immediate active-assisted mobilization without sling immobilization. Abduction of more than 90° was not allowed in the first six weeks.

Indications for operative treatment were a varus displacement of >20°, a valgus displacement of >40°, an increased reclination >30°, a lateral displacement of > ½ diaphyseal diameter, and/or displacement of the major and/or minor tubercle of >5-10mm.

Baseline characteristics and outcome measures
Baseline characteristics were obtained from the prospectively collected data by Acklin et al. [10]. All patients were contacted by phone by an independent study nurse to assess shoulder function using the QuickDASH questionnaire[22] and the Subjective Shoulder Value (SSV)[23]. Implant removal was assessed using the algorithm of Hulsmans et al.[24]. If patients could not be reached after a minimum of five phone call attempts, their contact person and general practitioner were approached for contact details and the internet was searched for an alternative telephone number. A letter was sent to patients who could not be reached by phone, asking the patient to contact us. Patients were considered lost to follow-up if all these attempts were unsuccessful.

The primary outcome measure was shoulder function as measured by the QuickDASH score[22]. The QuickDASH is a validated measure for disability of the arm, shoulder and hand and provides a summative score on a 100-point scale, where a higher score indicates more disability. A QuickDASH score of less than 15 is considered an excellent result and a score of >40 indicates poor shoulder function[25].

Secondary outcome measures were SSV and implant related irritation or implant removal. The SSV is a subjective value for shoulder function determined by the patient after answering the following question: “What is the overall percent value of your shoulder if a completely normal shoulder represents 100%?”, with 100% indicating the best function[23]. The SSV has shown a reliable agreement with the validated Constant Score for measuring shoulder function[26]. Implant removal and implant related irritation were discussed and analysed using the algorithm of Hulsmans et al., developed to analyse the presence of implant related irritation[24]. In addition, all patients were asked if they have had re-operations or were diagnosed with AVN in another hospital.

Statistical analysis

Data were described using frequencies and percentages for dichotomous and categorical variables, mean and standard deviation (SD) for normally distributed continuous data, and median and interquartile range (IQR) for non-normally distributed continuous data. In bivariate analysis, the association patients characteristics with the QuickDASH and SSV were assessed using a Mann-Whitney test for dichotomous variables (age), a Kruskal-Wallis test for ordinal variables (AO classification[27] and trauma mechanism) and a Spearman's rank correlation coefficient for continuous variables (age). A p value < 0.05 was considered significant which was tested using non-parametrical tests. The analyses were performed with SPSS, version 22.0 (IBM Corp., Armonk, NY) for Windows.

Results

Informed consent was obtained from all individual participants included in the study. A total of 79 (81%) patients were available for follow-up and included for analysis (Figure 1). The mean age at the time of accident was 59 (SD ± 13) years and 37 (47%) patients were male (Table 1). The most common trauma mechanism was injury during skiing or snowboarding (51%). There were 16 (20%) type A, 33 (42%) type B, and 30 (38%) type C fractures according to the AO classification[27]. There were no significant differences in age, trauma mechanism, and AO classification of patients available for follow-up as compared to the initial cohort (data not shown). The mean follow-up duration was 8.3 years (SD 0.8).

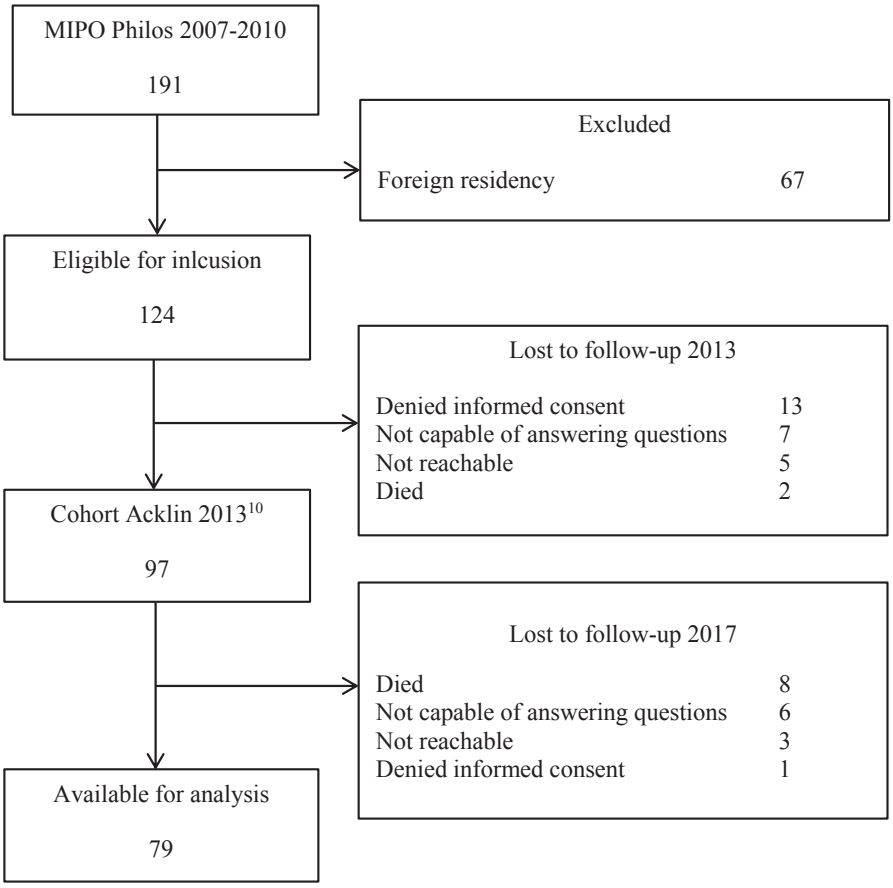


Fig. 1 Flowchart of patient inclusion

The mean QuickDASH score was 5.6 (SD 14) and the mean SSV was 92 (SD 11) (Table 2). A total of 40/79 (50,6%) patients had implant removal on average 1.2 years (SD 0.5) after the initial osteosynthesis (Table 2). Twenty-seven of the 79 (34,2%) patients had implant removal due to implant irritation and 13/79 (16,5%) patients requested implant removal without implant irritation.

Table 1. Baseline characteristics

Variable	Baseline cohort (n=79) n (%)
Age (mean, SD)	59 (13)
Male	37 (47)
ASA	
1	28 (35)
2	49 (62)
3	2 (2.5)
4	0 (0)
Dominant hand side	31 (39)
Trauma mechanism	
Ski / Snowboard	40 (51)
Low energy	29 (37)
Traffic accident	6 (7.6)
Other	4 (5.1)
AO Classification	
A	16 (20)
B	33 (42)
C	29 (38)
Follow-up time in years (mean, SD)	8.3 (0.8)

SD standard deviation

Table 2. Outcome measures

Variable	Mean (SD)	Median (IQR)
Functional outcome		
QuickDASH	5.6 (14)	0 (0 - 4.5)
Subjective Shoulder value	92 (11)	97 (90 - 100)
Implant related irritation / removal (n, %)		
Implant not removed, no irritation	34 (44)	
Implant not removed, irritation but implant removal not neccessary	1 (1.32)	
Implant not removed, irritation, no request for removal due to fear of surgery	2 (2.6)	
Implant not removed, irritation, considering removal	0 (0)	
Implant removed routinely or on patient's request without irritation	13 (17)	
Implant removed due to implant irritation	27 (35)	
Duration till removal of PHILOS plate in years	1.2 (0.5)	

SD standard deviation, IQR interquartile range

Table 3. Bivariate analysis

Variable	QuickDASH			SSV		
	Mean (SD)	Median (IQR)	*P* value	Mean (SD)	Median (IQR)	*P* value
Age continous (coefficient)	0.139		0.223	-0.002		0.988
Age categorical						
Age < 65	3.8 (13)	0 (0 - 2.3)	0.06	94 (8.7)	98 (90 - 100)	0.209
Age > 65	8.4 (16)	2.3 (0 - 6.8)		89 (13)	95 (80 - 100)	
Trauma mechanism						
Ski / Snowboard	2.0 (3.1)	0 (0 - 2.3)		95 (6.6)	99 (90 - 100)	
Low energy	8.5 (17)	0 (0 - 6.8)		89 (13)	95 (80 - 100)	
Traffic accident	20 (33)	5.7 (2.3 - 14)	0.232	83 (16)	90 (75 - 90)	0.155
Other	1.1 (1.3)	1.1 (0 - 2.3)		93 (5.4)	93 (89 - 98)	
AO						
A	0.4 (0.9)	0 (0 - 0)		93 (7.7)	95 (89 - 100)	
B	4.6 (9.4)	0 (0 - 4.5)	0.008	93 (10)	98 (90 - 100)	0.844
C	9.5 (20)	2.3 (0 - 6.8)		91 (12)	97 (85 - 100)	

SD standard deviation, IQR interquartile range

In bivariate analysis, there was a significant difference between AO fracture type and QuickDASH score with a mean of 0.4 (SD 0.9) for type A, 4.6 (SD 9.4) for type B and 9.5 (SD 20) for type C fractures (p = 0.008) (Table 3). There was no association of age or trauma mechanism with the QuickDASH score and also no association of age, trauma mechanism or AO classification with the SSV.

Previously published short-term follow-up of this cohort by Acklin et al. showed that all fractures were healed and no hardware failure occurred on follow-up radiographs. The mean radiological follow-up was 18 ± 6 months. There was a small but significant progression of varus displacement visible on last follow-up radiographs compared to postoperative evaluation (40°±8 and 41°±8; p= 0.015, respectively. Secondary screw perforation occurred in seven (7%) patients on average 7 weeks postoperatively and required operative screw(s) replacement. Four patients (4%) had axillary nerve injury with atrophy of the anterior border of the deltoid muscle, however, without clinical consequences.

Furthermore, eight (8%) patients developed some degree of radiological AVN (grade 3-5) in short term follow-up. With a mean 99 months follow-up, five patients with AVN were available for long-term follow-up. The mean time to diagnosis was 16.7 months, the mean

radiological follow-up was 34 months. They had a mean QuickDASH of 21 (SD 29) and a mean SSV of 72 (SD 12) which was significantly worse compared to patients without diagnosis of AVN in short-term follow-up (p=0.001 and p<0.001, respectively). One patient with a QuickDASH of 73 was offered a reversed arthroplasty but she refused. One patient with AVN grade 5 with a QuickDASH score of 13 and a SSV of 60 is considering a reversed arthroplasty. Of all patients with AVN three had a screw perforation of the head and four had their implant removed. At long-term follow-up there were no new reported cases of AVN based on the interview.
In total, one patient received a reversed arthroplasty. This was because of a symptomatic malunion. The major tubercle was not anatomically reduced and healed with a cranial step. At 80 months follow-up this patient had a QuickDASH of 31 and a SSV of 65. Two other new reported complications occured. Two patients developed a recurrent shoulder dislocation of whom one was operated for a rotator cuff repair.

Discussion

MIPO with Philos has been described as a suitable technique for the treatment of proximal humerus fractures, but long-term functional results have never been reported. The aim of this study was to describe the long-term functional outcome and implant related irritation after MIPO for proximal humerus fractures. In our cohort, we found a very good QuickDASH score and SSV representing an excellent functional outcome at more than 8 years of follow-up after MIPO with Philos for proximal humerus fractures. For this long-term follow-up we used patient reported questionnaires but were not able to obtain an objective clinical and radiological examination. Forty of the 79 patients (50,6%) had implant removal, and of those, 27/40 (67,5%) were due to implant related irritation. We found a mean QuickDASH score of 5.6 and a mean SSV of 92, which can be considered an excellent outcome. In 2013, Acklin et al. reported the one-year follow-up of this cohort and found a Constant score of 75 (SD 11) that corresponded to a shoulder function of 91% compared to the uninjured side. Other studies presenting one or two year follow-ups reported mean DASH scores ranging from 14.5 to 26 after MIPO and 31 to 32 after an open procedure[1,4,14,17,20]. It is still debated whether further improvement of shoulder function is to be expected 12 months after treatment. Hirschmann et al. found only a slight improvement after one year[28]. Other studies published no further improvement at longer follow-up[12,29].

Few studies have reported on long-term follow-up after operative treatment of proximal humerus fractures but almost all studies investigated the open approach. Ockert et al. investigated 43 patients who were operated on using the open approach with a median follow-up of 10 years and reported a mean DASH score of 24[12]. Most patients had an excellent outcome while 16% of the patients were considered to have a poor outcome. Bahrs et al. analysed 77 patients with a mean follow-up of eight years; eight patients were operated on using MIPO and 68 via an open approach[21]. They found a good mean DASH score of 12

with 77% of the patients having an excellent/good result and 23% having a satisfactory or worse result. No difference in Constant score between surgical approaches nor a correlation with the variables age and AO classification was found.

More than half of the patients had their implant removed. The majority (68%) because of implant related irritation and 32% requested implant removal because they did not want the material in their shoulder for the rest of their lives. Our findings are in line with Ockert et al. who reported a 40% implant removal rate[12]. They also found a significant improvement of functional outcome after implant removal. Similarly, another study reported improvement of shoulder function after implant removal among patients with implant related irritation treated with MIPO Philos plating[30]. In our study, we did not have sufficient data to report on improvement of shoulder function after implant removal. However, based on the study of Acklin et al., the high rate of implant removal in our cohort might have been beneficial for the excellent long-term results[30].

Results of different operative treatments should be put into perspective with regard to the conservative treatment for proximal humerus fractures. In a Cochrane review of eight randomized and quasi-randomized controlled trials the authors conclude that there was no evidence that supported the benefit of operative treatment of proximal humerus fractures[8]. But these results have to be interpreted with caution as the results did not cover two-part tuberosity fractures, fractures in young people, high-energy trauma, fracture-dislocations and head splitting fractures. Recently, the five-year follow-up results of the PROFHER trial, the most influential trial leading to conclusions in the Cochrane review, were published in which patients with a proximal humerus fracture were randomized between conservative and operative treatment[29]. In this medium-term follow-up study, the results of 109 patients were reported and no differences in Oxford Shoulder Score and EQ-5D-3L Score were observed. They concluded that there is no evidence that supports the trend of increased surgery for patients with displaced proximal humerus fractures. Nevertheless, there are major shortcomings in this study. First, the study is designed as a superiority study. Only 32% of the screened patients were included (e.g. several patients with clear indication for surgery were excluded). In 11% of cases, fairly inexperienced surgeons (e.g. registrars) performed the operation and 17% were operated on with something other than a plate (e.g. hemi-arthroplasty). So these results raise serious doubts. In addition, Kruithof et al. presented the long-term follow-up of conservatively treated patients with proximal humerus fractures between 2000 and 2013[2]. After exclusion, there was data of 410 patients with a good median DASH score of 6.67 at a follow up of 7.5 years. Sub-analysis revealed a significant better outcome of patients younger than 65 years old at the time of injury. They concluded that long-term functional outcome and quality of life were good in most patients after proximal humeral fractures.

Our study has several limitations that need to be addressed. First, we report on a subgroup of the original cohort that could have led to bias and limited generalizability of the study results. However, compared to the original cohort there were no differences in baseline characteristics in terms of age, trauma mechanism and AO-classification. Second, our hospital is situated in a recreational area in the mountains. Therefore, as compared to other hospitals, our patient population consists of younger and many relatively fit patients who were injured during outdoor sports activities. Therefore, our results might not be applicable to the typical proximal humerus fracture patient (female and > 65 years of age)[7]. Nevertheless, in bivariate analysis there was no association of age or trauma mechanism with functional outcome. Third, the sample size in this study is small and as this was a single center study, no appropriate control group was available. Furthermore, in this long-term follow-up study, we used telephone interviews in order to get sufficient follow-up. Therefore we were not able to perform a clinical examination of the shoulder or obtain long-term radiological follow-up. Consequently, no radiological data is available to report on the actual number of patients that developed AVN or implant failure. It can be argued to what extent radiological grade of AVN translates to limitations experienced by the patient[31], although, it seems that patients diagnosed with radiological AVN do have a worse functional outcome in this cohort. In addition, a possible disadvantage of the deltoid-split approach is second deltopectoral incision should a prosthesis be necessary in the future. However, this occurred only once in our patient sample. Finally, more than half of the patients were operated by two trauma surgeons dedicated to shoulder surgery while 14 different surgeons operated the other patients, which could have resulted in a performance bias. However, bivariate analysis did not show a difference in functional outcome of patients treated by the two surgeons versus patients treated by the 14 other surgeons. This possibly reflects the effect of in-hospital training and standardized procedure with the introduction of the MIPO technique and postoperative protocol in our hospital.

The results of this study can be of guidance when discussing treatment options for a proximal humerus fracture. The challenge for the future will be to determine which patient will benefit from operative treatment and which patient should be treated conservatively. We recommend operative treatment with MIPO for fit and active patients with a displaced proximal humerus fracture.

Conclusion

Treatment of proximal humerus fractures using MIPO with Philos through a deltoid split approach showed promising results. A good function can be assumed due to the excellent scores of patient oriented questionnaires. However, about one third of the patients will have a second operation for implant removal due to implant related irritation.

References

1 Jones CB, Sietsema DL, Williams DK (2011) Locked plating of proximal humeral fractures: is function affected by age, time, and fracture patterns? Clin Orthop Relat Res 469 (12):3307-3316. doi:10.1007/s11999-011-1935-6

2 Kruithof RN, Formijne Jonkers HA, van der Ven DJC, van Olden GDJ, Timmers TK (2017) Functional and quality of life outcome after non-operatively managed proximal humeral fractures. Journal of orthopaedics and traumatology : official journal of the Italian Society of Orthopaedics and Traumatology. doi:10.1007/s10195-017-0468-5

3 Launonen AP, Lepola V, Saranko A, Flinkkila T, Laitinen M, Mattila VM (2015) Epidemiology of proximal humerus fractures. Archives of osteoporosis 10:209. doi:10.1007/s11657-015-0209-4

4 Olerud P, Ahrengart L, Soderqvist A, Saving J, Tidermark J (2010) Quality of life and functional outcome after a 2-part proximal humeral fracture: a prospective cohort study on 50 patients treated with a locking plate. J Shoulder Elbow Surg 19 (6):814-822. doi:10.1016/j.jse.2009.11.046

5 Palvanen M, Kannus P, Niemi S, Parkkari J (2006) Update in the epidemiology of proximal humeral fractures. Clin Orthop Relat Res 442:87-92

6 Rangan A, Handoll H, Brealey S, Jefferson L, Keding A, Martin BC, Goodchild L, Chuang LH, Hewitt C, Torgerson D (2015) Surgical vs nonsurgical treatment of adults with displaced fractures of the proximal humerus: the PROFHER randomized clinical trial. Jama 313 (10):1037-1047. doi:10.1001/jama.2015.1629

7 Court-Brown CM, Garg A, McQueen MM (2001) The epidemiology of proximal humeral fractures. Acta orthopaedica Scandinavica 72 (4):365-371. doi:10.1080/000164701753542023

8 Handoll HH, Brorson S (2015) Interventions for treating proximal humeral fractures in adults. The Cochrane database of systematic reviews (11):Cd000434. doi:10.1002/14651858.CD000434.pub4

9 Jawa A, Burnikel D (2016) Treatment of Proximal Humeral Fractures: A Critical Analysis Review. JBJS reviews 4 (1). doi:10.2106/jbjs.rvw.o.00003

10 Acklin YP, Stoffel K, Sommer C (2013) A prospective analysis of the functional and radiological outcomes of minimally invasive plating in proximal humerus fractures. Injury 44 (4):456-460. doi:10.1016/j.injury.2012.09.010

11 Greiner S, Kaab MJ, Haas NP, Bail HJ (2009) Humeral head necrosis rate at mid-term follow-up after open reduction and angular stable plate fixation for proximal humeral fractures. Injury 40 (2):186-191. doi:10.1016/j.injury.2008.05.030

12 Ockert B, Siebenburger G, Kettler M, Braunstein V, Mutschler W (2014) Long-term functional outcomes (median 10 years) after locked plating for displaced fractures of the proximal humerus. J Shoulder Elbow Surg 23 (8):1223-1231. doi:10.1016/j.jse.2013.11.009

13 Acklin YP, Sommer C (2012) Plate fixation of proximal humerus fractures using the minimally invasive anterolateral delta split approach. Operative Orthopadie und Traumatologie 24 (1):61-73. doi:10.1007/s00064-011-0051-9

14 Brunner F, Sommer C, Bahrs C, Heuwinkel R, Hafner C, Rillmann P, Kohut G, Ekelund A, Muller M, Audige L, Babst R (2009) Open reduction and internal fixation of proximal humerus fractures using a proximal humeral locked plate: a prospective multicenter analysis. J Orthop Trauma 23 (3):163-172. doi:10.1097/BOT.0b013e3181920e5b

15 Falez F, Papalia M, Greco A, Teti A, Favetti F, Panegrossi G, Casella F, Necozione S (2016) Minimally invasive plate osteosynthesis in proximal humeral fractures: one-year results of a prospective multicenter study. Int Orthop 40 (3):579-585. doi:10.1007/s00264-015-3069-z

16 Lin T, Xiao B, Ma X, Fu D, Yang S (2014) Minimally invasive plate osteosynthesis with a locking compression plate is superior to open reduction and internal fixation in the management of the proximal humerus fractures. BMC musculoskeletal disorders 15:206. doi:10.1186/1471-2474-15-206

17 Oh HK, Cho DY, Choo SK, Park JW, Park KC, Lee JI (2015) Lessons learned from treating patients with unstable multifragmentary fractures of the proximal humerus by minimal invasive plate osteosynthesis. Arch Orthop Trauma Surg 135 (2):235-242. doi:10.1007/s00402-014-2138-x

18 Sohn HS, Jeon YS, Lee J, Shin SJ (2017) Clinical comparison between open plating and minimally invasive plate osteosynthesis for displaced proximal humeral fractures: A prospective randomized controlled trial. Injury 48 (6):1175-1182. doi:10.1016/j.injury.2017.03.027

19 Sohn HS, Shin SJ (2014) Minimally invasive plate osteosynthesis for proximal humeral fractures: clinical and radiologic outcomes according to fracture type. J Shoulder Elbow Surg 23 (9):1334-1340. doi:10.1016/j.jse.2013.12.018

20 Laflamme GY, Rouleau DM, Berry GK, Beaumont PH, Reindl R, Harvey EJ (2008) Percutaneous humeral plating of fractures of the proximal humerus: results of a prospective multicenter clinical trial. J Orthop Trauma 22 (3):153-158. doi:10.1097/BOT.0b013e3181694f7d

21 Bahrs C, Kuhle L, Blumenstock G, Stockle U, Rolauffs B, Freude T (2015) Which parameters affect medium- to long-term results after angular stable plate fixation for proximal humeral fractures? J Shoulder Elbow Surg 24 (5):727-732. doi:10.1016/j.jse.2014.08.009

22 Beaton DE, Wright JG, Katz JN (2005) Development of the QuickDASH: comparison of three item-reduction approaches. J Bone Joint Surg Am 87 (5):1038-1046. doi:10.2106/jbjs.d.02060

23 Jost B, Pfirrmann CW, Gerber C, Switzerland Z (2000) Clinical outcome after structural failure of rotator cuff repairs. J Bone Joint Surg Am 82 (3):304-314

24 Hulsmans MH, van Heijl M, Houwert RM, Hammacher ER, Meylaerts SA, Verhofstad MH, Dijkgraaf MG, Verleisdonk EJ (2017) High Irritation and Removal Rates After Plate

or Nail Fixation in Patients With Displaced Midshaft Clavicle Fractures. Clin Orthop Relat Res 475 (2):532-539. doi:10.1007/s11999-016-5113-8

25 Angst F, Schwyzer HK, Aeschlimann A, Simmen BR, Goldhahn J (2011) Measures of adult shoulder function: Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) and its short version (QuickDASH), Shoulder Pain and Disability Index (SPADI), American Shoulder and Elbow Surgeons (ASES) Society standardized shoulder assessment form, Constant (Murley) Score (CS), Simple Shoulder Test (SST), Oxford Shoulder Score (OSS), Shoulder Disability Questionnaire (SDQ), and Western Ontario Shoulder Instability Index (WOSI). Arthritis care & research 63 Suppl 11:S174-188. doi:10.1002/acr.20630

26 Gilbart MK, Gerber C (2007) Comparison of the subjective shoulder value and the Constant score. J Shoulder Elbow Surg 16 (6):717-721. doi:10.1016/j.jse.2007.02.123

27 Marsh JL, Slongo TF, Agel J, Broderick JS, Creevey W, DeCoster TA, Prokuski L, Sirkin MS, Ziran B, Henley B, Audige L (2007) Fracture and dislocation classification compendium - 2007: Orthopaedic Trauma Association classification, database and outcomes committee. J Orthop Trauma 21 (10 Suppl):S1-133

28 Hirschmann MT, Fallegger B, Amsler F, Regazzoni P, Gross T (2011) Clinical longer-term results after internal fixation of proximal humerus fractures with a locking compression plate (PHILOS). J Orthop Trauma 25 (5):286-293. doi:10.1097/BOT.0b013e3181f2b20e

29 Handoll HH, Keding A, Corbacho B, Brealey SD, Hewitt C, Rangan A (2017) Five-year follow-up results of the PROFHER trial comparing operative and non-operative treatment of adults with a displaced fracture of the proximal humerus. The bone & joint journal 99-b (3):383-392. doi:10.1302/0301-620x.99b3.bjj-2016-1028

30 Acklin YP, Michelitsch C, Sommer C (2016) Elective implant removal in symptomatic patients after internal fixation of proximal humerus fractures improves clinical outcome. BMC musculoskeletal disorders 17:119. doi:10.1186/s12891-016-0977-z

31 Robinson CM, Khan LA, Akhtar MA (2006) Treatment of anterior fracture-dislocations of the proximal humerus by open reduction and internal fixation. The Journal of bone and joint surgery British volume 88 (4):502-508. doi:10.1302/0301-620x.88b4.17195

CHAPTER 4

LONG-TERM FOLLOW-UP AFTER RIB FIXATION FOR FLAIL CHEST AND MULTIPLE RIB FRACTURES.

R.B. Beks, M.B. de Jong, R.M. Houwert,
A.A.R. Sweet, I.G.J.M. De Bruin, G.A.M. Govaert,
K.J.P. Wessem, R.K.J. Simmermacher, F.Hietbrink,
R.H.H. Groenwold, L.P.H. Leenen

European Journal of Trauma and Emergency Surgery 2018

Abstract

Purpose

Rib fixation for flail chest has been shown to improve in-hospital outcome, but little is known about treatment for multiple rib fractures and long-term outcome is scarce. The aim of this study was to describe the safety, long-term quality of life, and implant related irritation after rib fixation for flail chest and multiple rib fractures.

Methods

All adult patients with blunt thoracic trauma who underwent rib fixation for flail chest or multiple rib fractures between January 2010 and December 2016 in our level-1 trauma facility were retrospectively included. In hospital characteristics and implant removal were obtained via medical records and long-term quality of life was assessed over the telephone.

Results

Of the 864 patients admitted with ≥3 rib fractures, 166 (19%) underwent rib fixation; 66 flail chest patients and 99 multiple rib fracture patients with an ISS of 24 (IQR 18–34) and 21 (IQR 16–29), respectively. Overall, the most common complication was pneumonia (n=58, 35%). Six (9%) patients with a flail chest and 3 (3%) with multiple rib fractures died; only 1 because of injuries related to the thorax. On average at 3.9 years follow-up was obtained from 103 patients (62%); 40 with flail chest and 63 with multiple rib fractures reported an EQ-5D index of 0.85 (IQR 0.62–1) and 0.79 (0.62–0.91), respectively. Forty-eight (48%) patients had implant related irritation and 9 (9%) had implant removal.

Conclusions

We show that rib fixation is a safe procedure and that patients reported a relative good quality of life. Patients should be counseled that after rib fixation approximately half of the patients will experience implant related irritation and about 1 in 10 patients requires implant material removal.

Background

Chest trauma is currently the second leading cause of trauma-related death and multiple rib fractures are the most common injury in these patients.[1] Due to the impact of pulmonary complications, flail chest and multiple rib fractures are still associated with a 10 – 22% mortality rate with increasing rates for every additional rib involved.[2]

Conservative treatment for rib fractures is considered the gold standard and consists of mechanical ventilation (if indicated), pulmonary hygiene, and adequate pain management. In the last century, many different surgical techniques concerning rib fixation were described in literature without becoming common clinical practice. However, due to technical improvements there is a growing popularity of surgical rib fixation which aims to increase stability of the chest, lessen chest wall deformity, and improve pulmonary function.[3]

In a recent meta-analysis the authors recommend rib fixation over conservative treatment for adult patients with flail chest in order to decrease mortality, shorten days on mechanical ventilation, hospital and intensive care length of stay, and decrease incidence of pneumonia and need for tracheostomy.[3] Although rib fixation of patients with flail chest showed promising results, little is known about rib fixation for patients with multiple rib fractures without a flail chest. Furthermore, only few small studies have described the long-term outcome and quality of life after rib fixation.[4–7] Therefore, the aim of this study was to describe the safety, long-term quality of life, and implant related irritation after rib fixation for flail chest or multiple rib fractures.

Methods

Study design and participants

All medical records of patients admitted with rib fractures following blunt thoracic trauma between January 2010 and December 2016 in the University Medical Center Utrecht, a level-1 trauma facility, were retrospectively reviewed. Eligible patients were identified using procedural codes and the Dutch National Trauma Registry. For this study, we included all adult patients with blunt thoracic trauma who underwent rib fixation for flail chest (defined as three or more consecutive ribs fractured in at least two places and clinical signs of paradoxical chest wall movement) or multiple rib fractures (defined as three or more unilateral rib fractures). We did not further distinguish between multiple rib fractures with or without chest deformity due to the retrospective nature of this study. Exclusion criteria were age below 18 years, fewer than three fractured ribs, no availability of an admission CT scan of the chest, and transfer from or to another hospital. Our institutional review board approved a waiver of consent under protocol number 17-914/C.

Indication for surgery

The indication for surgical rib fixation followed from a clinical based algorithm considering several injury and patient specific characteristics as shown in Figure 1. There was a strict

indication for patients with a clinical flail chest (paradoxical breathing). Failure of pain management with tachypnea and dyspnea was considered an indication for surgical rib fixation in patients with multiple rib fractures.

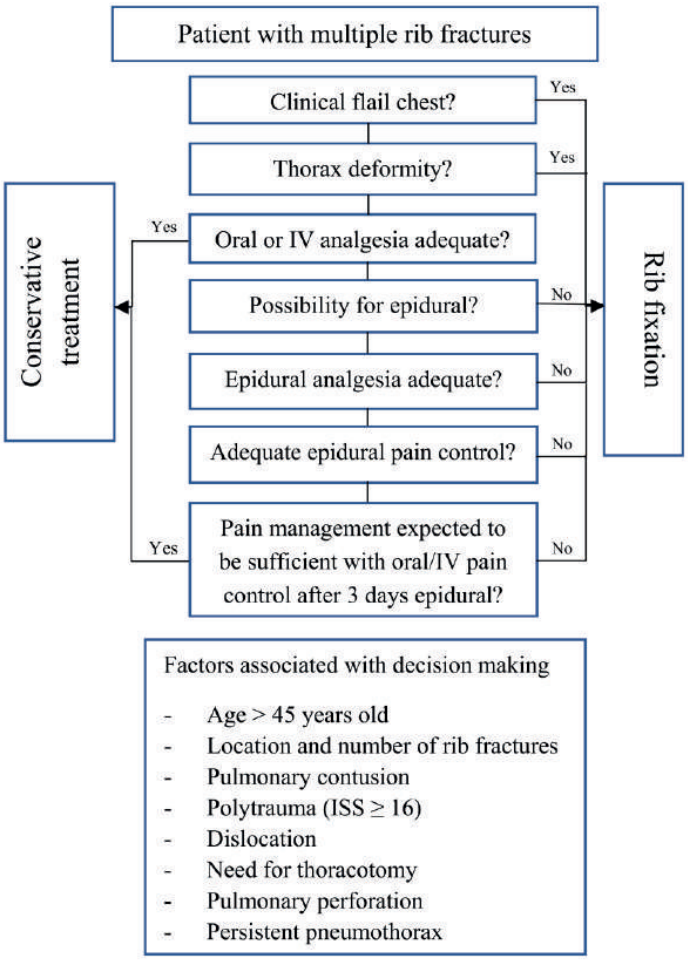


Fig. 1 Clinical treatment algorithm for patients with rib fractures

Patient characteristics at hospital admission

The following characteristics were obtained from medical records based on the recording at admission: age, sex, American Society of Anaesthesiologists (ASA) score, trauma mechanism, Abbreviated Injury Scale (AIS), ISS, Thoracic Trauma Severity Score (TTSS), number of rib fractures, bilateral rib fractures, involvement of the first rib as these are associated with higher impact trauma, rib fractures in the upper/middle/lower third or dorsal side of the thorax, displacement, concomitant injuries as described on the admission CT scan, and blood pH and base excess. The TTSS (range 0 – 25) is a scoring system that helps to predict thorax related complications after thoracic trauma and is based on number of rib fractures, pulmonary contusion, PaO2/FIO2 ratio, pleural involvement, and age.[8] Displacement was defined as a shaft width displacement of the fracture parts in the transversal plane on CT. Dorsal fractures were defined as rib fractures behind the dorsal axillary line.

Surgical procedure and characteristics

All procedures were performed or supervised by senior trauma surgeons experienced with surgical treatment of rib fractures. Preoperative planning of the procedure was done using chest computed tomography (CT) with 3D reconstructions. Preoperative antibiotic prophylaxis (2 grams of Cefazolin) was administered intravenously in all patients. Depending on the site of the fractures, patients were positioned in the supine, lateral or prone position. The surgical approach was performed as described by Taylor.[9] After reduction, internal fixation using the MatrixRIB™ system (Depuy Synthes®, Amersfoort, The Netherlands) was performed. Fixation was preferably done with 3 bicortical screws on each side of the fracture. If plate fixation was not possible due to anatomical boundaries and rib fixation was deemed necessary, splints were used. The number of fixed ribs was at the discretion of the surgeon, and depended on anatomical boundaries and the possibility to regain stability of the chest wall during respiration. Tube thoracostomy was performed in case of pneumothorax or hemothorax at initial presentation or clinical suspicion of pneumothorax during surgery. Postoperative chest radiography was performed in all patients to document surgical result and to rule out early complications. Patients were encouraged to mobilize as soon as possible with the help of physiotherapy and aggressive pain management. All patients had an outpatient department visit six weeks after discharge and were counselled to visit if they experienced any thorax related problems like pain, dyspnoe or irritation.

The following surgery related characteristics regarding rib fixation were extracted from the medical record: time until surgery, duration of surgery, surgical approach, number of ribs fixated, the ratio of fixated ribs to fractured ribs, side of rib fixation, and fixation of dorsal rib fractures.

Short and long-term outcome measures

Short-term outcome measures were hospital length of stay (HLOS), ICU-LOS, duration of invasive mechanical ventilation (IMV), need for tracheostomy, and incidence of surgical complications after rib fixation (e.g. pneumonia, implant related infection, wound infection, and acute respiratory distress syndrome [ARDS]). Pneumonia was defined as having clinical signs (fever, coughing, desaturation) requiring antibiotic treatment, with or without positive cultures. Implant related infection was defined as clinical symptoms (e.g., redness, drainage from surgical wound, fever, pain, elevated CRP, or leukocytes) requiring incision and drainage and intravenous antibiotics following a previously published protocol.[10] ARDS was defined by severe hypoxemia with a PaO2/FIO2 smaller than 100mm Hg.

Long-term outcome measures were quality of life, number of implant removals due to complications of patient complaints, and level of dyspnea. To assess the long-term outcome measures after rib fixation, patients were contacted by phone after a minimum of 12 months of follow-up. The patient's contact person and general practitioner were approached for

additional contact details if patients could not be reached after a minimum of five phone call attempts.

Quality of life was assessed with the EQ-5D-5L, which is a standardized instrument for generic health status measurement.[11] The EQ-5D-index ranges from -0.33 to 1.00 where higher scores indicate better quality of life. The EQ-VAS is a patient's subjective measurement of generic health ranging from 0 and 100, where higher scores represent better subjective health experience. The level of dyspnea was measured with the modified Medical Research Council Dyspnea Scale (mMRC) which is a five-category scale that characterizes the level of dyspnea with physical activity where higher scores corresponds with more dyspnoea.[12] Patients who had implant removal were asked for the reason of removal following the algorithm and definitions as described by Hulsmans et al.[13] Implant removal due to irritation was considered a minimum of six months after rib fixation and after discussing the possible harms and benefits with the patient. Apart from the well-known pitfalls after implant removal in general, the most important pitfall of rib implant removal is the risk of a pneumothorax. Therefore standard chest tube placement should be considered after this procedure. Implant related irritation at the time of the interview was defined as physical complaints which could be attributed to the implant.

Statistical analysis

All analyses were performed separately for the groups of patients with flail chest and the group of patients with multiple rib fractures. Baseline characteristics were presented as median and interquartile range (IQR) for continuous variables, and absolute numbers with percentage for categorical variables. The non-parametric outcome measures were normalized with a cubic transformation for left skewed data and a log transformation for HLOS and ICU-LOS. In bivariate analysis, the association of the HLOS, ICU-LOS, and EQ-5D-index with the baseline characteristics was assessed using linear regression. Variables with a p value of below .05 in this analysis were entered into a multivariable linear regression model to assess their ability to explain the variation in in HLOS, ICU-LOS, and quality of life. Given the small dataset with the high number of potential variables a robustness check of the primary multivariable regression model was performed by means of the least shrinkage and selection operator (LASSO) technique.[14] LASSO performs automatic variable selection by shrinking coefficients and giving a penalty for the number of variables in the model. LASSO is considered a robust and objective alternative for the more regularly performed step wise variable selection for multivariable regression. The two statistical models were compared in terms of the variables that showed a relation with the outcome of interest. All analyses were performed with Stata 13 (StataCorp LP, College Station, TX, USA); a p value of less than .05 was considered significant.

Results

Between 2010 and 2016, in our hospital, a total of 864 patients were admitted with chest trauma resulting in three or more rib fractures. Ultimately, 166 patients (19%) who underwent rib fixation were included for analysis; 67 with flail chest and 99 with multiple rib fractures (Figure 2). Of these, 137 (83%) were treated with plate osteosynthesis, 29 (17%) with a combination of plate osteosynthesis and intramedullary splints, and one only with intramedullary splints. Outcome information, at a minimum of twelve months after rib fixation, was obtained from 103 patients (62%); 40 with flail chest and 63 with multiple rib fractures.

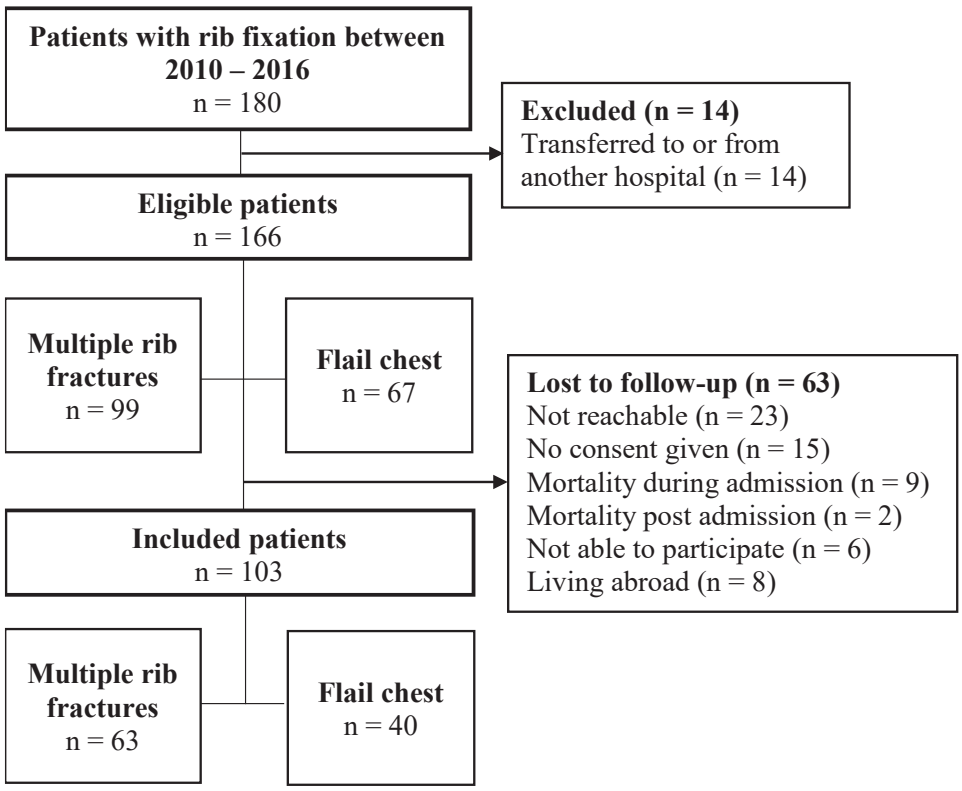


Fig. 2 Flowchart of patient inclusion

Flail chest

The median age of patients with flail chest was 57 (IQR 48 – 69) years and the majority were male (n=52, 78%) (Table 1). The median ISS was 24 (IQR 18 – 34) and the median number of fractured ribs was 10 (IQR 8 – 12). Rib fixation was performed after a median of two (IQR 1 – 3) days and the ratio of fixated ribs to fractured ribs was 0.49 (Table 2).

Among patients with flail chest, the most common complication was pneumonia (n=26, 39%) followed by excess pleural fluid (n=3, 5%) and implant related infection (n=2, 3%) (Table 3).

One patient had a tension pneumothorax perioperatively and required a chest tube. Six (9%) patients died during hospital admission; all were because of concomitant injuries that were not related to the rib fractures. Two patients had an infaust neurological prognosis, one patients died of cardiac failure, one patient developed secondary bacterial meningitis, and one patient with metastasized carcinoma and IC acquired weakness wished no further treatment.

The median HLOS was 19 (11 – 26) days and 44 (66%) patients required ICU admission with a median ICU-LOS 8 (6 – 14) days (Table 4). The median follow-up duration was 3.1 years (IQR 2.4 – 5.1; range 1 – 7.5) and 40 (60%) patients were available for follow-up. The median quality of life as measured with the EQ-5D index at follow-up was 0.85 (IQR 0.62 – 1) with an EQ-VAS of 75 (IQR 63 – 85). Figure 3 shows the proportion of patients reporting problems specified per EQ-5D domain. Twenty-one (53%) patients reported implant related irritation. Five (13%) patients had their implant removed due to irritation on average 1.1 (range 0.64 – 1.6) years after rib fixation. Patients reporting implant related irritation at the time of the interview had a significant lower median EQ-5D index compared to patients without implant related irritation ($z = 2.97$; $p = 0.003$). Eleven patients (28%) reported mild to severe complaints of dyspnea.

The association between each patient characteristic and the outcomes are presented in Appendix 1. In multivariable linear regression, male sex and sternum fracture appeared to be independently associated with the EQ-5D index (Appendix 2). We did not observe an association with HLOS. A higher AIS-head appeared to be associated with ICU-LOS. The associations found in the three multivariable models were also found when applying LASSO, indicating robustness of the models.

Multiple rib fractures

The median age of the 99 patients with multiple rib fractures was 56 (IQR 47 – 64) years and the majority were male (n=82, 82%) (Table 1). The median ISS was 21 (IQR 16 – 29) and the median number of fractured ribs was 7 (IQR 6 – 10). Surgery was performed after a median of two (IQR 1 – 4) days and the ratio of fixated ribs to fractured ribs was 0.52 (Table 2).

Among patients operated on multiple rib fractures, pneumonia was the most common complication (n=32, 32%) followed by excess pleural fluid (n=3, 3%) and implant related infection (n=3, 3%) (Table 3). Two (2%) patients suffered a tension pneumothorax postoperatively and were successfully treated with a chest tube. One (1%) patient needed revision surgery due to two dislocated intramedullary splints resulting in a hemothorax. Three (3%) patients died during hospital admission; one because of respiratory failure possibly associated with the suffered rib fractures and the other two as a result of concomitant injuries not related to the thorax. One had unmanageable infectious episodes from unknown origin and did not want further treatment. One patient had a systemic

inflammatory response with decompensated liver cirrhosis, kidney failure, and developed acute respiratory distress syndrome (ARDS).

The median HLOS was 14 (IQR 10 - 28) days and 44 patients (44%) required ICU admission with a median ICU-LOS of 9 (IQR 2 – 16) days (Table 4). The median follow-up was 4.4 years (IQR 3.4 – 5.9; range 1 – 7.6) and 63 patients (63%) were available for follow-up. The median quality of life as measured with the EQ-5D index at follow-up was 0.79 (IQR 0.62 – 0.91) with an EQ-VAS of 73 (IQR 65 – 80). Figure 3 shows the proportion of patients reporting problems specified per EQ-5D domain. After rib fixation for multiple rib fractures, 28 (44%) of the patients experienced implant related irritation. Four patients (6.3%) had their implant removed due to irritation on average 1.8 (range 0.91 – 4.2) years after rib fixation. Patients reporting implant related irritation at the time of the interview had a significant lower median EQ-5D index compared to patients without implant related irritation ($z = 3.30$; $p = 0.001$). Nine patients (14%) reported mild to serious complaints of dyspnea.

The association between each patient characteristic and the outcomes are presented in Appendix 3. In multivariable regression, we did not observe an association of the EQ-5D index and the baseline characteristics (Appendix 4). A higher AIS-head, AIS-extremities, and AIS-abdomen appeared to be associated with HLOS. A higher AIS-face, AIS-extremities, and base excess appeared to be associated with ICU-LOS. The associations found in the three multivariable models were also found when applying LASSO.

Table 1. Baseline characteristics of patient with rib fixation for flail chest or multiple rib fractures

Variable	Flail chest n = 67	Multiple rib fractures n = 99
Age (median, IQR)	57 (48 - 69)	56 (47 - 64)
Male (n, %)	52 (78)	81 (82)
ASA-score (n, %)		
1-2	57 (92)	82 (84)
> 2	5 (8)	16 (16)
Trauma mechanism (n, %)		
Motor vehicle accident	25 (37)	33 (33)
Fall from height / stairs	17 (25)	29 (29)
Other	25 (37)	37 (37)
AIS (median, IQR)		
Head	0 (0 - 3)	0 (0 - 2)
Face	0 (0 - 0)	0 (0 - 0)
Thorax	4 (3 - 4)	4 (3 - 4)
Abdomen	0 (0 - 2)	0 (0 - 2)
Extremities	2 (0 - 3)	2 (0 - 2)

(table 1 continued)

ISS (median, IQR)	24 (18 - 34)	21 (16 - 29)
TTSS (median, IQR)	13 (11 - 15)	10 (8 - 12)
No. of rib fractures (median, IQR)	10 (8 - 12)	7 (6 - 10)
Bilateral rib fractures (n, %)	26 (39)	34 (34)
First rib fracture (n, %)		
Unilateral	18 (27)	16 (16)
Bilateral	7 (10)	11 (11)
Location rib fracture (n, %)		
Costae 1 - 4	62 (93)	84 (85)
Costae 5 - 8	67 (100)	99 (100)
Costae 9 - 12	46 (69)	60 (61)
Displacement (n, %)	47 (70)	58 (59)
Dorsal fracture (n, %)	59 (88)	67 (68)
Concomitant injuries (n,%)		
Pulmonary contusion	44 (66)	43 (43)
Pneumothorax	50 (75)	66 (67)
Hemothorax	16 (24)	21 (21)
Sternum fracture	7 (10)	16 (16.2)
Blood pH (median, IQR)	7.3 (7.28 - 7.4)	7.4 (7.3 - 7.4)
Base Excess (median, IQR)	-2 (-5 - -1)	-1 (-3.5 - 0.7)

ASA American Society of Anesthesiologists; ISS injury severity score; TTSS Thoracic trauma severity score; AIS abbreviated injury score; IQR interquartile range

Table 2. Surgery related characteristics

Variable	Flail chest n = 67	Multiple rib fractures n = 99
Time until surgery (days, median, IQR)	2 (1 - 3)	2 (1 - 4)
Duration of surgery (minutes, median, IQR)	130 (91 - 155)	98 (71 - 122)
Surgical approach (n, %)		
Anterior	9 (13)	12 (12)
Anterolateral	9 (13)	17 (17)
Posterior	10 (15)	19 (19)
Posterolateral	32 (48)	39 (39)
Combination	7 (10)	12 (12)
No. of ribs fixated (median, IQR)	4 (4 - 6)	4 (3 - 5)
No. of ribs fixated / total ribs fractured (median, IQR)	0.5 (0.36 - 0.6)	0.5 (0.38 - 0.67)
Side of rib fixation (n, %)		
Left	34 (51)	45 (46)
Right	26 (39)	46 (47)
Bilateral	7 (10)	8 (8)
Fixation of dorsal fractures (n, %)	35 (52)	36 (36)

IQR interquartile range; ICU Intensive care unit; IMV invasive mechanical ventilation

Table 3. In hospital complications after rib fixation

In hospital complications	Flail chest (n,%) n = 47	Multiple rib fractures (n,%) n = 53
Pneumonia	26 (39)	32 (32)
Excess pleural fluid	3 (4.5)	3 (3)
Implant related infection	2 (3)	3 (3)
Hemothorax	2 (3)	2 (2)
Pneumothorax	2 (3)	2 (2)
Tension pneumothorax	1 (1)	2 (2)
ARDS	2 (3)	3 (3)
Postoperative bleeding	1 (1.5)	1 (1)
Wound infection	1 (1.5)	0 (0)
Pleural empyema	1 (1.5)	0 (0)
Hematoma	0 (0)	1 (1)
Revision of dislocated splints	0 (0)	1 (1)
In hospital mortality	6 (9)	3 (3)

ARDS acute respiratory distress syndrome

Table 4. Outcome measures after rib fixation for flail chest or multiple rib fractures

	Flail chest	Multiple rib fractures
Short-term outcome measures	n = 67	n = 99
HLOS (days, median, IQR)	19 (11 - 26)	14 (10 - 28)
ICU admission (n,%)	44 (66)	44 (44)
ICU-LOS among those admitted to ICU (days, median, IQR)	8 (6 - 14)	9 (2 - 16)
Number of patient with IMV (n,%)	40 (60)	35 (35)
Duration of IMV among those ventilated (days, median, IQR)	6 (4 - 12)	9 (4 - 16)
Tracheostomy (n, %)	7 (10)	9 (9)
Long-term outcome measures	n = 40	n = 63
EQ-5D index (median, IQR)	0.85 (0.62 - 1)	0.79 (0.62 - 0.91)
EQ VAS (median, IQR)	75 (63 - 85)	73 (65 - 80)
Implant related irritation (n, %)	21 (53)	28 (44)
Implant removed (n, %)	5 (13)	4 (6)
Reason removed (n, %)		
Attributable to implant-related irritation	5 (13)	4 (6)
Patient's wish or surgeon's preference	0 (0)	0 (0)
Status not removed (n, %)		
No irritation	19 (47)	35 (56)
Experiencing irritation, but implant removal not necessary	12 (30)	11 (18)
Experiencing irritation, but no request for removal owing to fear of reoperation	1 (3)	2 (3)
Experiencing irritation, considering removal	3 (8)	10 (16)
Revision implant (n, %)	1 (3)	1 (2)
mMRC (n, %)		
0	17 (43)	31 (49)
1	12 (30)	23 (37)
2	6 (15)	5 (8)
3	4 (10)	3 (5)
4	1 (3)	1 (2)
Follow-up duration in years (median, IQR)	3.1 (2.4 - 5.1)	4.4 (3.4 - 5.9)
Follow-up range duration in years (min, max)	1 - 7.5	1 - 7.6

HLOS hospital length of stay; ICU-LOS intensive care unit length of stay; IQR interquartile range; IMV invasive mechanical ventilation; mMRC modified Medical Research Council

Discussion

In this cohort study of 166 patients admitted to a Dutch level-1 trauma facility the reported quality of life was relatively good after rib fixation for flail chest or multiple rib fractures at a median follow-up of 3.1 and 4.4 years, respectively. A mortality rate of 5% was demonstrated in this cohort. Approximately half of the patients experienced implant related irritation after rib fixation and about 10 percent had the implant material, or part of it, removed due to this irritation. At follow up 15 – 18% of the patients reported mild tot serious complaints of dyspnea as measured with the mMRC.

In our cohort, the mortality rate for patients with flail chest was 9% and for multiple rib fractures 3%; only one death could be directly ascribed as the consequence of the suffered rib fractures. There were three important surgery related complications resulting in a tension pneumothorax; all were successfully treated with a chest tube. The low mortality rate as well as the low number of surgical complications indicate the relative safety of this procedure in this patient cohort. The most frequent complication was pneumonia in 39% of the patients with flail chest and 32% of the patients with multiple rib fractures and is comparable with the existing literature. However, definitions used for pneumonia differ in literature making this outcome measure difficult to compare across studies. The incidence of ARDS was 3% in both groups and was low compared to an ARDS incidence of 13% in a previously published cohort of poly trauma patients, predominantly chest trauma, from our hospital.[15] This low rate of ARDS in our cohort could be attributed to the effects of rib fixation. The rate of implant related infection was 3% in our cohort and was similar to the infection rate reported by Pieracci et al in a similar but smaller cohort.[16]

The duration of mechanical ventilation and ICU-LOS among patients admitted to the ICU in our cohort were comparable or shorter than the three RCTs available on this subject.[17–19] Another interesting finding in our study was that injury severity, as defined by the abbreviated injury scale, in other body regions such as head, face, abdomen, and extremities were associated with a longer HLOS and / or ICU-LOS, while no association was seen with injury severity of the thorax. One explanation could be that rib fixation successfully minimized the impact of chest injury on the outcome measures. ICU-LOS and HLOS are frequently used to measure the success of rib fixation and it should be kept in mind that a small but potential beneficial effect could be masked by associated injury when comparing different treatment strategies for rib fractures. This emphasizes the necessity of sufficient group sizes when comparing treatment strategies in these often heterogeneous group of patients; nonetheless, there is a lack of large patient series in the current literature.

The quality of life in our study, a EQ-5D index of 0.85 for patients with flail chest and 0.79 for patients with multiple rib fractures, is comparable to the Dutch reference population index of 0.87[20] and compared to studies describing different polytrauma cohorts these results were

good.[21–24] There was no difference in quality of life between patients with flail chest and patients with multiple rib fractures as both indices were within the range of the minimal clinically important difference for the EQ-5D (the minimal score difference detectable by the patient).[25,26] Although, one might expect a worse outcome for flail chest patients compared to patients with multiple rib fractures, in this cohort patients with multiple rib fractures had similar injury severity scores which might explain comparability. Caragounis et al presented comparable results after one year follow-up of 45 patients with rib fixation for flail chest and multiple rib fractures with an EQ-5D index of 0.93.[27] Similar results were reported by Mayberry et al in a cohort of fifteen patients after rib fixation.[4] In another study, Campbell et al. reported on quality of life of 20 patients more than one year after rib fixation and showed a lower quality of life as compared to the reference population possibly due to the higher ISS scores in this patients cohort.[6] There was a high number of reported problems per domain ranging from 22 – 60%, with the most substantial limitation experienced in the domain of pain and discomfort. It cannot be extracted from the EQ-5D-5L if the pain is situated in the chest area. Farquhar et al. reported the EQ-5D-5L of 11 patients with rib fixation for flail chest at an unspecified long-term follow-up, and reported a slightly higher number of problems per domain as compared to our results, but also found the highest rate of problems in the domain of pain and discomfort.[28] Although residual pain and chest stiffness are commonly reported in the literature, patient satisfaction is high after rib fixation at long-term follow-up. [5,6,29]

Implant removal after rib fixation is a challenging and time consuming procedure. Due to the angular stable system and soft Titanium, we encountered several technical problems during implant removal. In one case a grinding machine was used to remove plate and screwheads leaving the body of the screws in place. In other cases a diamond drill was used to remove the screwhead from the plate also leaving the screw body behind. Because implant removal is challenging, perforation of the pleura happens easily. Therefore a chest tube should be considered after implant removal.

Two of the three clinical trials in this field performed rib fixation on patients with flail chest who were ventilator dependent without prospect of successful weaning. All three studies had different strict exclusion criteria such as severe injuries to other body systems, head trauma, or patients who did not development acute respiratory failure.[17–19] Because of the heterogenicity in the aforementioned clinical trials, no clear indication for rib fixation has been defined. Also, very few studies have enrolled any substantial number of patients with multiple rib fractures without flail chest making the indication for these patients unknown. We made use of a clinical treatment algorithm (Figure 1) based on previous literature and experience in our hospital, which provides guidance in decision making for both patients with flail chest and patients with multiple rib fractures.

In addition to the right indication, timing of the procedure is of major importance. The main reason for rib fixation is to stabilize the thorax to increase pulmonary mechanics and reduce pain. In a recent published study, Pieracci et al. concluded that early surgical stabilization was indeed associated with favorable outcome.[30] Additionally, they found that late surgical stabilization resulted in a significantly longer operating time for the same type of rib fracture. They hypothesized that this could be ascribed to tissue inflammation resulting in obscured planes and increased bleeding. Therefore, in our hospital, rib fixation is performed according to the treatment algorithm but preferably as early as possible after hospital admission.

The results should be interpreted in the light of several limitations. First, the EQ-5D-5L and mMRC are subjective questionnaires and assess general health and not specifically thorax-related problems. The vast majority of the patients described in this cohort were polytrauma patients, therefore, concomitant injuries but also comorbidities could have influenced the outcome. Second, due to the retrospective nature, this study could be subject to data loss and underreporting of complications. Consequently, no data was available on quality of life of patients before implant removal to objectify any improvement, although no differences were observed after implant removal compared with the rest of the patients. Third, follow-up differed per patient and ranged from 1 – 7.5 years. We assumed that for the majority of patients quality of life will improve most significantly in the first year after trauma and to a lesser extent thereafter, which is supported by our finding that there was no association between follow-up duration and quality of life (Spearman's rho 0.14; $p = 0.164$). Fourth, rib fixation was performed following the incision of a thoracotomy in the earlier years which gradually changed to a more minimal invasive approach in the following years. Nonetheless, there was no correlation between year of surgery and the outcome measures. Finally, the Dutch reference values for the EQ-5D were obtained from the three category EQ-5D version whereas our results were measured using the newer five category version. The additional answer categories provide the possibility for the patient to report milder problems which could have resulted in a higher percentage of reported problems as compared to the available Dutch reference population.

This is the largest study to present the long-term follow-up of patients after rib fixation following a clear clinical treatment algorithm. We show that rib fixation is a save treatment option for both patients with flail chest and patients with multiple rib fractures and that patients report a relatively good quality of life at long-term follow-up as compared to the Dutch reference population. Patients should be counseled that after rib fixation approximately half of the patients will experience implant related irritation and about 1 in 10 patients requires implant material removal due to this irritation. Future studies should focus on further development of the indication for rib fixation and should aim to identify the patient who will benefit most from rib fixation.

Appendices

Appendix 1. Bivariate analysis of the baseline characteristics and the outcome measures of patients with flail chest

Variable	EQ-5D index			HLOS			ICU length of stay		
	coefficient	se	p	coefficient	se	p	coefficient	se	p
Age	0.000	0.001	0.837	0.002	0.006	0.705	0.018	0.010	0.085
Male	0.080	0.033	***0.019***	-0.130	0.206	0.530	0.200	0.366	0.586
ASA-score	0.044	0.090	0.625	0.716	0.312	***0.025***	1.400	0.557	***0.015***
Trauma mechanism	-0.017	0.017	0.332	-0.141	0.098	0.156	-0.102	0.177	0.566
AIS									
Head	-0.007	0.012	0.542	0.088	0.056	0.121	0.275	0.096	***0.005***
Face	0.020	0.017	0.257	0.080	0.109	0.465	0.055	0.195	0.780
Thorax	-0.010	0.021	0.640	0.070	0.115	0.544	0.319	0.200	0.116
Abdomen	0.000	0.011	0.987	0.084	0.065	0.200	0.209	0.114	0.072
Extremities	-0.013	0.011	0.227	0.187	0.061	***0.003***	0.493	0.099	***0.000***
ISS	-0.002	0.001	0.232	0.019	0.007	***0.013***	0.057	0.012	***0.000***
TTSS	-0.005	0.007	0.504	0.022	0.033	0.498	0.193	0.051	***0.000***
No. of rib fractures	-0.002	0.004	0.610	0.059	0.021	***0.007***	0.089	0.038	***0.024***
Bilateral rib fractures	-0.047	0.030	0.121	0.509	0.165	***0.003***	0.678	0.302	***0.028***
First rib fracture	0.005	0.021	0.807	0.260	0.126	***0.043***	0.232	0.227	0.312
Location rib fracture									
Costae 1 - 4	-0.023	0.058	0.693	0.669	0.323	***0.042***	-0.136	0.585	0.817
Costae 5 - 8	NA	NA	NA	NA	NA	NA	NA	NA	NA
Costae 9 - 12	0.011	0.031	0.718	-0.025	0.190	0.897	0.501	0.328	0.132
Displacement	-0.056	0.030	0.071	0.091	0.192	0.636	0.173	0.337	0.610
Dorsal fracture	0.005	0.046	0.920	-0.287	0.268	0.288	-0.374	0.473	0.431
Concomitant injuries									
Lung contusion	-0.009	0.032	0.769	0.329	0.177	0.067	0.487	0.317	0.129
Pneumothorax	-0.030	0.036	0.414	0.068	0.198	0.732	-0.010	0.352	0.977
Hemothorax	0.018	0.034	0.593	-0.149	0.201	0.461	-0.228	0.358	0.526
Sternum fracture	-0.103	0.048	***0.038***	0.472	0.276	0.091	1.255	0.475	***0.010***
Blood pH	0.025	0.142	0.863	-2.167	0.836	***0.012***	-4.863	1.412	***0.001***
Base Excess	0.001	0.004	0.904	-0.083	0.024	***0.001***	-0.174	0.041	***0.000***

Appendix 2. Bivariate analysis of the baseline characteristics and the outcome measures of patients with multiple rib fractures

Variable	EQ-5D index			HLOS			ICU length of stay		
	coefficient	se	p	coefficient	se	p	coefficient	se	p
Age	0.000	0.001	0.767	0.004	0.005	0.414	0.005	0.009	0.600
Male	-0.016	0.044	0.713	-0.301	0.181	0.100	-0.288	0.322	0.373
ASA-score	-0.016	0.049	0.736	0.078	0.191	0.683	0.127	0.334	0.705
Trauma mechanism	0.012	0.018	0.513	-0.094	0.084	0.265	-0.070	0.148	0.639
AIS									
Head	0.008	0.013	0.545	0.220	0.043	***0.000***	0.340	0.077	***0.000***
Face	0.015	0.024	0.516	0.232	0.090	***0.012***	0.522	0.155	***0.001***
Thorax	0.020	0.021	0.362	0.091	0.104	0.382	0.014	0.183	0.941
Abdomen	-0.008	0.011	0.492	0.187	0.048	***0.000***	0.315	0.085	***0.000***
Extremities	0.007	0.013	0.602	0.259	0.049	***0.000***	0.368	0.091	***0.000***
ISS	0.001	0.002	0.488	0.037	0.006	***0.000***	0.056	0.010	***0.000***
TTSS	0.006	0.006	0.308	0.059	0.025	***0.019***	0.122	0.042	***0.005***
No. of rib fractures	-0.001	0.005	0.847	0.047	0.023	***0.042***	0.088	0.040	***0.028***
Bilateral rib fractures	-0.025	0.031	0.424	0.376	0.144	***0.011***	0.627	0.254	***0.015***
First rib fracture	0.005	0.021	0.809	0.111	0.106	0.297	0.252	0.185	0.177
Location rib fracture									
Costae 1 - 4	-0.038	0.042	0.370	0.065	0.205	0.752	-0.273	0.359	0.449
Costae 5 - 8	NA	NA	NA	NA	NA	NA	NA	NA	NA
Costae 9 - 12	-0.026	0.032	0.421	0.171	0.151	0.260	0.164	0.266	0.540
Displacement	-0.007	0.031	0.812	0.119	0.147	0.421	0.397	0.256	0.124
Dorsal fracture	-0.008	0.032	0.802	0.082	0.156	0.603	0.083	0.275	0.765
Concomitant injuries									
Lung contusion	0.009	0.032	0.786	0.255	0.141	0.073	0.547	0.245	***0.028***
Pneumothorax	0.016	0.032	0.619	0.233	0.148	0.121	0.197	0.263	0.456
Hemothorax	0.006	0.038	0.868	0.095	0.173	0.586	0.150	0.304	0.623
Sternum fracture	-0.043	0.040	0.289	-0.054	0.192	0.779	-0.139	0.338	0.682
Blood pH	-0.212	0.208	0.314	-2.102	0.700	***0.003***	-5.739	1.138	***0.000***
Base Excess	-0.003	0.005	0.508	-0.055	0.016	***0.001***	-0.139	0.026	***0.000***

Appendix 3. Multivariable linear regression of the baseline characteristics and the outcome measures of patients with flail chest

model for EQ-5D index	coefficient	se	95% CI		p
Male	0.080	0.031	0.017	0.142	***0.014***
Sternum facture	-0.103	0.045	-0.193	-0.012	***0.027***
model for HLOS					
ASA	0.523	0.293	-0.065	1.111	0.080
AIS extremities	0.120	0.067	-0.014	0.255	0.078
No. of rib fractures	0.017	0.029	-0.040	0.075	0.543
Bilateral fractures	0.192	0.239	-0.287	0.672	0.425
First rib fracture	0.146	0.146	-0.146	0.439	0.320
Base excess	-0.051	0.027	-0.105	0.002	0.060
model for ICU length of stay					
ASA	0.870	0.530	-0.202	1.941	0.109
ISS	-0.008	0.020	-0.048	0.033	0.701
TTSS	0.088	0.060	-0.033	0.209	0.150
AIS extremities	0.226	0.136	-0.050	0.502	0.105
AIS head	0.237	0.107	0.020	0.454	***0.033***
No. of rib fractures	-0.051	0.057	-0.166	0.063	0.370
Bilateral fractures	0.125	0.432	-0.749	0.998	0.774
Sternum facture	0.803	0.740	-0.694	2.301	0.284
Base excess	-0.094	0.053	-0.201	0.012	0.081

Appendix 4. Multivariable linear regression of the baseline characteristics and the outcome measures of patients with multiple rib fractures

model for HLOS	coefficient	se	95% CI		p
TTSS	0.032	0.021	-0.011	0.074	0.143
AIS head	0.133	0.045	0.043	0.223	***0.004***
AIS face	0.112	0.084	-0.054	0.279	0.183
AIS extremities	0.183	0.049	0.086	0.279	***0.000***
AIS abdomen	0.105	0.047	0.013	0.198	***0.026***
No. of rib fractures	-0.001	0.025	-0.051	0.049	0.963
Bilateral fractures	0.036	0.157	-0.277	0.350	0.818
Base excess	-0.030	0.018	-0.066	0.006	0.098
model for ICU-LOS					
AIS head	0.139	0.079	-0.019	0.296	0.083
AIS face	0.363	0.142	0.080	0.647	***0.013***
AIS extremities	0.193	0.083	0.029	0.358	***0.022***
AIS abdomen	0.129	0.079	-0.029	0.287	0.109
TTSS	0.066	0.040	-0.013	0.145	0.100
No. of rib fractures	0.010	0.043	-0.075	0.096	0.811
Pulmonary contusion	0.163	0.229	-0.293	0.618	0.478
Bilateral fractures	-0.132	0.267	-0.664	0.401	0.623
Base excess	-0.113	0.030	-0.173	-0.052	***0.000***

References

1. Vana PG, Neubauer DC, Luchette FA. Contemporary management of flail chest. *The American surgeon*. 2014;80(6):527-535.
2. Bulger EM, Arneson MA, Mock CN, Jurkovich GJ. Rib fractures in the elderly. *J Trauma*. 2000;48(6):1040-1047.
3. Kasotakis G, Hasenboehler EA, Streib EW, et al. Operative fixation of rib fractures after blunt trauma: A practice management guideline from the Eastern Assocaiation for the Surgery of Trauma. *Journal of Trauma and Acute Care Surgery*. 2017;82(3):618-626. doi:10.1097/TA.0000000000001350.
4. Mayberry JC, Kroeker AD, Ham LB, Mullins RJ, Trunkey DD. Long-term morbidity, pain, and disability after repair of severe chest wall injuries. *The American surgeon*. 2009;75(5):389-394.
5. Majercik S, Cannon Q, Granger SR, et al. Long-term patient outcomes after surgical stabilization of rib fractures. *American journal of surgery*. 2014;208(1):88-92. doi:10.1016/j.amjsurg.2013.08.051.
6. Campbell N, Conaglen P, Martin K, Antippa P. Surgical stabilization of rib fractures using inion OTPS wraps-techniques and quality of life follow-up. *Journal of Trauma - Injury, Infection and Critical Care*. 2009;67(3):596-601. doi:10.1097/TA.0b013e3181ad8cb7.
7. Bille A, Okiror L, Campbell A, et al. Evaluation of long-term results and quality of life in patients who underwent rib fixation with titanium devices after trauma. *General Thoracic and Cardiovascular Surgery*. 2013;61(6):345-349. doi:10.1007/s11748-013-0218-4.
8. Pape HC, Remmers D, Rice J, et al. Appraisal of early evaluation of blunt chest trauma: development of a standardized scoring system for initial clinical decision making. *Journal of Trauma*. 2000;49(3):496-504.
9. Taylor BC, French BG, Fowler TT. Surgical approaches for rib fracture fixation. *Journal of Orthopaedic Trauma*. 2013;27(7):e168-e173. doi:10.1097/BOT.0b013e318283fa2d.
10. Hellebrekers P, Leenen LPH, Hoekstra M, Hietbrink F. Effect of a standardized treatment regime for infection after osteosynthesis. *Journal of Orthopaedic Surgery and Research*. 2017;12(1):1-11. doi:10.1186/s13018-017-0535-x.
11. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x.
12. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest*.

1988;93(3):580-586.

13. Hulsmans MHJ, van Heijl M, Frima H, et al. Predicting suitability of intramedullary fixation for displaced midshaft clavicle fractures. *European Journal of Trauma and Emergency Surgery*. October 2017. doi:10.1007/s00068-017-0848-9.

14. Tibshirani R. Regression Shrinkage and Selection via the Lasso Author (s): Robert Tibshirani Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 58 , No . 1 Published by : Wiley for the Royal Statistical Society Stable URL : http://. 2016;58(1):267-288.

15. van Wessem KJP, Leenen LPH. Reduction in Mortality Rates of Postinjury Multiple Organ Dysfunction Syndrome. *Shock*. 2018;49(1):33-38. doi:10.1097/SHK.0000000000000938.

16. Pieracci FM, Lin Y, Rodil M, et al. A prospective, controlled clinical evaluation of surgical stabilization of severe rib fractures. *Journal of Trauma and Acute Care Surgery*. 2016;80(2):187-194. doi:10.1097/TA.0000000000000925.

17. Tanaka H, Yukioka T, Yamaguti Y, et al. Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients. *Journal of Trauma*. 2002;52(4):727-732. doi:10.1097/00005373-200204000-00020.

18. Granetzny A, Abd El-Aal M, Emam E, Shalaby A, Boseila A. Surgical versus conservative treatment of flail chest. Evaluation of the pulmonary status. *Interactive Cardiovascular and Thoracic Surgery*. 2005;4(6):583-587. doi:10.1510/icvts.2005.111807.

19. Marasco SF, Davies AR, Cooper J, et al. Prospective randomized controlled trial of operative rib fixation in traumatic flail chest. *Journal of the American College of Surgeons*. 2013;216(5):924-932. doi:10.1016/j.jamcollsurg.2012.12.024.

20. Hoeymans N, Van Lindert H, Westert GP. The health status of the Dutch population as assessed by the EQ-6D. *Quality of Life Research*. 2005;14(3):655-663. doi:10.1007/s11136-004-1214-z.

21. Wad MS, Laursen T, Fruergaard S, Morgen SS, Dahl B. Survival and health related quality of life after severe trauma – a 15 years follow up study. *Injury*. October 2017. doi:10.1016/J.INJURY.2017.10.001.

22. Ulvik A, Kvåle R, Wentzel-Larsen T, Flaatten H. Quality of life 2-7 years after major trauma. *Acta Anaesthesiologica Scandinavica*. 2008;52(2):195-201. doi:10.1111/j.1399-6576.2007.01533.x.

23. Gross T, Schüepp M, Attenberger C, Pargger H, Amsler F. Outcome in polytraumatized patients with and without brain injury. *Acta Anaesthesiologica Scandinavica*. 2012;56(9):1163-1174. doi:10.1111/j.1399-6576.2012.02724.x.

24. Gunning A, van Heijl M, van Wessem K, Leenen L. The association of patient and trauma characteristics with the health-related quality of life in a Dutch trauma population. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*. 2017;25(1):1-8. doi:10.1186/s13049-017-0375-z.

25. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 2005;14(6):1523-1532.

26. Kvam AK, Fayers PM, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. *European Journal of Haematology*. 2011;87(4):330-337. doi:10.1111/j.1600-0609.2011.01665.x.

27. Caragounis E-C, Fagevik Olsen M, Pazooki D, Granhed H. Surgical treatment of multiple rib fractures and flail chest in trauma: a one-year follow-up study. *World journal of emergency surgery : WJES*. 2016;11:27. doi:10.1186/s13017-016-0085-2.

28. Farquhar J, Almahrabi Y, Slobogean G, et al. No benefit to surgical fixation of flail chest injuries compared with modern comprehensive management: results of a retrospective cohort study. *Canadian journal of surgery Journal canadien de chirurgie*. 2016;59(5):299-303.

29. Bille A, Okiror L, Campbell A, et al. Evaluation of long-term results and quality of life in patients who underwent rib fixation with titanium devices after trauma. *General Thoracic and Cardiovascular Surgery*. 2013;61(6):345-349. doi:10.1007/s11748-013-0218-4.

30. Pieracci FM, Coleman J, Ali-Osman F, et al. A multicenter evaluation of the optimal timing of surgical stabilization of rib fractures. *Journal of Trauma and Acute Care Surgery*. 2018;84(1):1-10. doi:10.1097/TA.0000000000001729.

CHAPTER 5

RIB FIXATION VERSUS CONSERVATIVE TREATMENT FOR FLAIL CHEST AND MULTIPLE RIB FRACTURES AFTER BLUNT THORACIC TRAUMA. A MULTICENTER COHORT STUDY.

R.B. Beks, D. Reetz, M.B. de Jong, R.H.H. Groenwold, F. Hietbrink, M.J.R. Edwards, L.P.H. Leenen, MD, R.M. Houwert, J.P.M. Frölke

European Journal of Trauma and Emergency Surgery. 2018

Abstract

Background

Over the years a trend has evolved towards operative treatment of flail chest although evidence is limited. Furthermore, little is known about operative treatment for patients with multiple rib fractures without a flail chest. The aim of this study was to compare rib fixation based on a clinical treatment algorithm with nonoperative treatment for both patients with a flail chest or multiple rib fractures.

Methods

All patients with ≥3 rib fractures admitted to one of the two contributing hospitals between January 2014 and January 2017 were retrospectively included in this multicenter cohort study. One hospital treated all patients nonoperatively and the other hospital treated patients with rib fixation according to a clinical treatment algorithm. Primary outcome measures were intensive care length of stay and hospital length of stay for patients with a flail chest and patients with multiple rib fractures, respectively. To control for potential confounding, propensity score matching was applied.

Results

A total of 332 patients were treated according to protocol and available for analysis. The mean age was 56 (SD 17) years old and 257 (77%) patients were male. The overall mean Injury Severity Score was 23 (SD 11) and the average number of rib fractures was 8 (SD 4). There were 92 patients with a flail chest, 37 (40%) had rib fixation and 55 (60%) had non-operative treatment. There were 240 patients with multiple rib fractures, 28 (12%) had rib fixation and 212 (88%) had non-operative treatment. For both patient groups, after propensity score matching, rib fixation was not associated with intensive care unit length of stay (for flail chest patients) nor with hospital length of stay (for multiple rib fracture patients), nor with the secondary outcome measures.

Conclusion

No advantage could be demonstrated for operative fixation of rib fractures. Future studies are needed before rib fixation is embedded or abandoned in clinical practice.

Background

Multiple rib fractures are the most common type of thoracic injury and are associated with a high morbidity and mortality, which is to a certain extent due to associated injuries.[1–4] Still, an increased number of rib fractures corresponds to a worse outcome, in part due to respiratory complications resulting from pain and an impaired ventilation capacity.[5–7] Consequently, superinfection leading to pneumonia and prolonged mechanical ventilation are common in patients with chest wall injuries.[2] It is important to distinguish between multiple rib fractures with and without a flail chest, as the latter is associated with an increased mortality rate and significant morbidity.[8–11]

Nonoperative treatment has been the gold standard for the past few decades and is focused on the underlying pulmonary contusion- and rib fracture- associated complications, including pain, atelectasis, and compromised pulmonary hygiene.[4] Over the years, a trend has evolved towards operative treatment of flail chest as physicians aim to improve mortality rates and reduce the prolonged length of stay for these patients. In a recent systematic review, rib fixation in patients with a flail chest was associated with a reduced: intensive care unit length of stay, days on mechanical ventilation, mortality rate, pneumonia rate, and treatment costs, although evidence remains limited.[12] Studies investigating the effect of rib fixation in patients with multiple rib fractures are even more scarce, although two retrospective cohort studies showed promising results.[13,14]

For both flail chest and multiple rib fractures, the indication for surgery is heterogeneously described in the aforementioned studies.[12–14] Therefore, no clear consensus on indication is available based on the current literature. It can be hypothesized that for patients with multiple rib fractures, early fixation might be beneficial. Therefore, the aim of this study was to compare rib fixation based on a clinical treatment algorithm with nonoperative treatment for both patients with a flail chest and patients with multiple rib fractures.

Methods

Study design and participants

All patients with three or more rib fractures admitted to one of the two contributing hospitals between January 2014 and January 2017 were retrospectively included in this multicenter cohort study. Both hospitals are academic tertiary referral centers with a level one trauma facility of similar size. Patients were included if they fulfilled the following criteria: age 18 years and older, blunt thoracic trauma resulting in multiple rib fractures (defined as three or more rib fractures) or a flail chest (defined as three or more consecutive ribs fractured in at least two places and clinical signs of paradoxical chest wall movement), and being alive two days after hospital admission (mean time till surgery). Exclusion criteria were: transfer to another hospital, initial admission in another hospital, no availability of a

computed tomography (CT) scan, and rib fixation more than four days after trauma. Patients were followed from admission until discharge or death.

Eligible patients were identified using procedural codes and the Dutch National Trauma Registry. The non-operative group was formed by all patients with rib fractures admitted to the Radboud University Medical Center where treatment consisted of adequate pain management, supportive mechanical ventilation when indicated, and physiotherapy for breathing exercises according to standard national guidelines. Per protocol every patient with three or more rib fractures was considered for epidural analgesia, if needed this was supported by patient controlled anesthesia (intravenous opioids). Epidural therapy was provided between day 1-5 after trauma. The epidural was removed after 5 days in situ due to the considered risk for infection. The surgical group consisted of all patients who had rib fixation performed in the University Medical Center Utrecht where the same non-operative treatment guidelines were followed, but in addition, rib fixation was considered according to a clinical based algorithm (Figure 1). Pain was arbitrarily defined as a numerical rating scale of 5 or higher during coughing or deep inspiration and if pain was suspected not to decrease over the subsequent days with adequate pain management. It was the subjective decision of the surgeon on call to perform rib fixation.

This study was approved by the institutional review board of the participating centers (METC 17-544/C & 2016-2861).

Surgical procedure rib fixation

All procedures were performed by one of the even senior trauma surgeons experienced in surgical treatment of rib fractures. Rib fixation is performed in this center since 2006. Preoperative planning of the procedure was done using chest computed tomography (CT) with 3D reconstructions. Preoperative antibiotic prophylaxis (2 grams of Cefazolin) was administered intravenously in all patients. Depending on the site of the fractures, patients were positioned in the supine, lateral or prone position and the surgical approach was performed as described by Taylor.[15] In the case of intercostal muscle interposition, debridement was performed. After reduction, internal fixation using the MatrixRIB™ system (Depuy Synthes®, Amersfoort, the Netherlands) was performed. Fixation was preferably done with 3 bicortical screws on each side of the fracture. The number of fixed ribs was at the discretion of the surgeon, and depended upon the possibility to regain stability of the chest wall during respirations. Tube thoracostomy was only performed in the case of clinical suspicion of pneumothorax during surgery. Postoperative chest radiography was performed in all patients to document surgical result and to rule out complications. Patients were allowed to perform their daily activities as soon as possible.

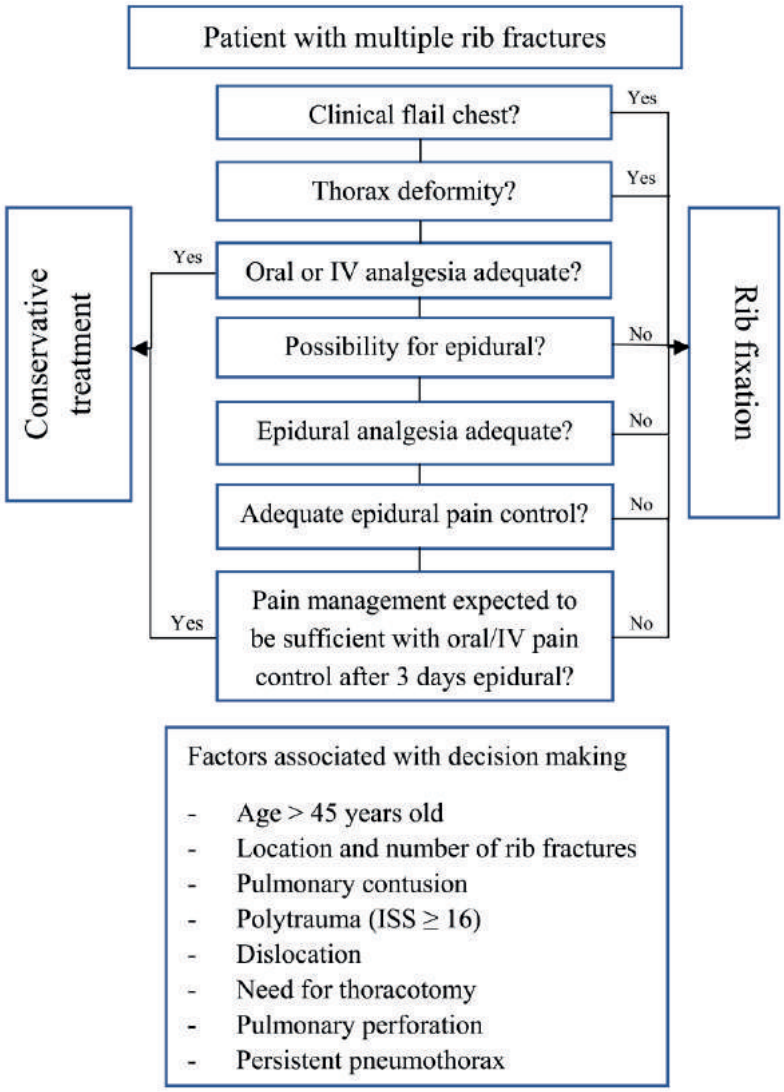


Fig. 1 Clinical treatment algorithm for patients with rib fractures

Baseline characteristics

Data for the following baseline characteristics were extracted from medical records: age, sex, American Society of Anesthesiologists (ASA) score, trauma mechanism, Injury Severity Score (ISS), thoracic trauma severity score (TTSS), abbreviated injury scale (AIS) head, AIS face, AIS thorax, AIS abdomen, AIS extremities, number of rib fractures, bilateral rib fractures, concomitant injuries (pulmonary contusion, pneumothorax, hemothorax, sternum fracture) as recorded on the admission CT scan, and first available blood pH and base excess. Additionally, for the surgical group: duration until surgery in days, duration of surgery in minutes, and number of surgically fixated rib fractures. The ISS is a measure (range 0 – 75) of the severity of traumatic injury and is calculated by adding the square of the three highest AIS scores. The AIS is a standardized anatomical-based coding system ranging from zero to five to classify the severity of traumatic injury per body region. The AIS is registered in the Dutch National Trauma Registry by trained data managers based on radiology reports from

admission CT scans and medical records. The TTSS is a score (range 0 – 25) based on number of rib fractures, pulmonary contusion, PaO_2/FiO_2 ratio, age, and pleural involvement and helps to predict outcome after thoracic trauma.[16–18] Rib fractures, pulmonary contusion, pneumothorax, and hemothorax were assessed on the admission CT scan.

Outcome measures

In line with previous trial reports, the primary outcome measure for patients with a flail chest was intensive care unit length of stay (ILOS) and for patients with multiple rib fractures, hospital length of stay (HLOS). For both patient groups, secondary outcome measures were duration of invasive mechanical ventilation (IMV), duration of epidural analgesia, pneumonia, need for tracheostomy and in hospital mortality. Pneumonia was defined as having clinical signs (fever, coughing, desaturation) requiring antibiotic treatment, with or without positive cultures. Additionally, we assessed in hospital complications after rib fixation.

Statistical analysis

All analyses were stratified by patient group, i.e., performed separately for patients with a flail chest and patients with multiple rib fractures. Baseline characteristics were presented as proportions for categorical variables, mean and standard deviation (SD) for normally distributed continuous variables, and median and interquartile range (IQR) for non-normally distributed continuous variables. Differences in distributions of baseline characteristics between the study groups were quantified by means of standardized differences and statistical tests (t-test for normally distributed continuous data, Mann-Whitney test for non-normally distributed data, and Chi square test for categorical data).[19]

We applied multiple imputation (25 times) to impute missing values for ASA (2.1% [7/332]), TTSS (20% [67/332]), AIS head (0.6% [2/332]), pulmonary contusion (0.6% [2/332]), pH (9.0% [30/332]), and base excess (9.0% [(30/332]). Multiple imputation was performed using the mice() algorithm in R.[20]

To control for potential confounding, propensity score (PS) matching was applied. To minimize the effects of selection bias, we matched patients from the 'operating' center with patients from the 'nonoperative' center, based on all baseline characteristics. First, a PS model was fitted using logistic regression analysis, with rib fracture fixation as the dependent variable, and age, sex, ASA-score, trauma mechanism, ISS, TTSS, AIS head, AIS face, AIS thorax, AIS abdomen, AIS extremities, number of rib fractures, bilateral rib fractures, concomitant injuries, blood pH, and base excess were included as covariates in the model. We performed 2:1 nearest neighbor matching, with a maximum caliper of 0.2 of the standard deviation of the logit of the PS using the Matchit() algorithm in R.[21] After matching, the balance in the distributions of baseline characteristics between the study groups were quantified using standardized differences, where a standardized difference < 0.1 is generally

accepted as indicating fair balance of confounders between the matched treatment groups (i.e. successful matching).[19]

In the primary analysis, for patients with a flail chest, we estimated the relation between rib fracture fixation and ILOS by means of linear regression analysis. For patients with multiple rib fractures, we estimated the relation between rib fracture fixation and HLOS by means of linear regression analysis. Secondary analyses focused on the relation of rib fracture fixation with duration of IMV and duration of epidural analgesia using linear regression analysis. The relation between rib fixation, pneumonia, tracheostomy, and in hospital mortality was assessed by means of a logistic regression analysis.

A two-tailed p-value less than 0.05 was considered significant. All analyses were performed using R v3.4.1.[22]

Results

A total of 332 patients were available for analysis (Figure 2). The overall mean age was 56 (SD 17) years old and 257 (77%) patients were male (Table 1). Most patients were injured in a motor vehicle accident or after a fall from height resulting on average in 8 (SD 4) rib fractures and an overall mean ISS of 23 (SD 11).

Of the 92 patients with a flail chest, 37 (40%) had rib fixation and 55 (60%) had non-operative treatment (Figure 2). For the flail chest population, surgically treated patients had a lower AIS head and a higher blood pH. Among the 240 patients with multiple rib fractures, 28 (12%) had rib fixation and 212 (88%) had non-operative treatment. In this group, surgical patients had a significantly lower AIS head, higher AIS thorax, higher AIS abdomen, and higher number of rib fractures (Table 1).

The median time until surgery was one day (IQR 1-2)(Table 2). The median number of surgically fixated rib fractures for patients with a flail chest was 5 (4-6) and for patients with multiple rib fractures was 4 (IQR 3-5). Four (6%) patients were treated with both plate osteosynthesis and intramedullary splints; two patients with flail chest and two with multiple rib fractures. Nine (14%) patients had a postoperative complication. Two patients had a persistent postoperative pneumothorax and were treated with a chest tube. Two patients developed pleural empyema requiring video assisted thoracoscopic surgery to evacuate the empyema. One patient had a postoperative tension pneumothorax and was treated with a chest tube. One patient had a hemothorax and required a thoracotomy to evacuate the hematoma. One patient had excess pleural fluid and was treated with a chest tube. One patient had a hematoma near the surgical incision and needed surgical debridement of the old hematoma. And one patient had a deep infection near the osteosynthesis material and was successfully treated with antibiotics.

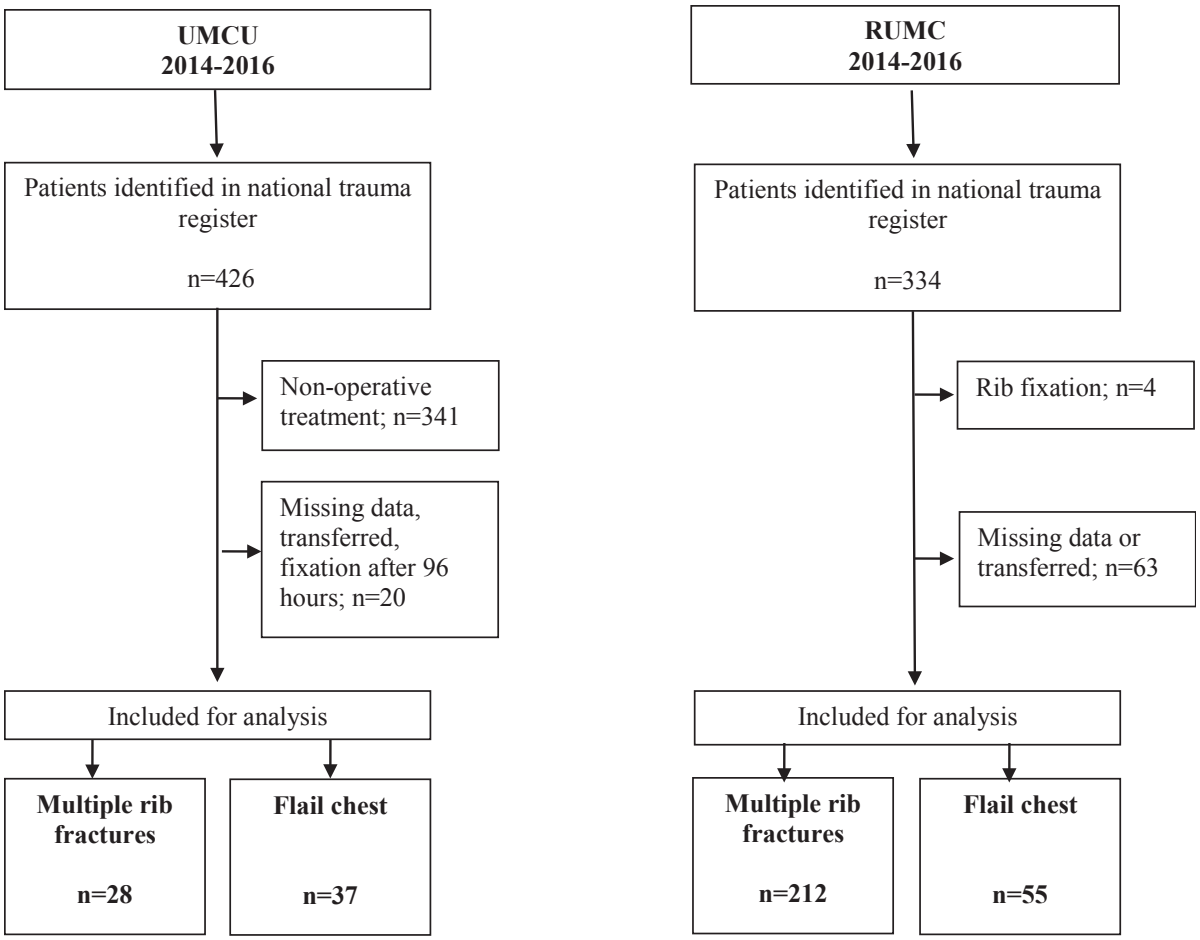


Fig. 2 Flowchart showing inclusion of patients for analysis."UMCU" University Medical Center Utrecht; "RUMC" Radboud University Medical Center

After propensity score matching, for patients with a flail chest there was no association of rib fixation and ILOS (CI -13.9 – 8.5, p = 0.638) and the secondary outcome measures (Table 3). For patients with multiple rib fractures there was no association between rib fixation and HLOS (confidence interval [CI] -0.6 – 13.6, p = 0.074) and the secondary outcome measures (Table 4).

Table 1. Baseline characteristics of patients with flail chest or multiple rib fractures before propensity score matching

Variable	Multiple Rib Fractures					Flail Chest				
	Surgery (n=28)	Non-operative (n=212)	*P* value	SMD unmatched cohort	SMD matched cohort	Surgery (n=37)	Non-operative (n=55)	*P* value	SMD unmatched cohort	SMD matched cohort
Age (mean ± SD)	57 (13)	55 (18)	0.587	0.122	0.037	59 (13)	57 (17)	0.630	0.106	0.075
Male (n, %)	24 (86)	162 (76)	0.386	0.239	0.051	29 (78)	42 (76)	1.000	0.048	0.110
ASA-score (n, %)			0.363	0.270	0.040			0.195	0.452	0.029
1-2	26 (93)	179 (84)				37 (100)	50 (91)			
> 2	2 (7)	33 (16)				0 (0)	5 (9)			
Trauma mechanism (n, %)			0.661	0.181	0.081			0.439	0.273	0.175
Motor vehicle accident	8 (29)	74 (35)				14 (38)	14 (26)			
Fall from height / stairs	11 (39)	66 (31)				10 (27)	19 (35)			
Other	9 (32)	72 (34)				13 (35)	22 (40)			
ISS (mean ± SD)	21 (9.4)	21 (9.6)	0.933	0.017	0.077	25 (12)	30 (13)	0.054	0.418	0.050
TTSS (mean ± SD)	10 (3.6)	9.3 (2.9)	0.237	0.242	0.101	13.1 (3.1)	13 (2.9)	0.951	0.014	0.087
AIS (median, IQR)										
Head	0 (0 - 1)	1 (0 - 3)	**0.008**	0.576	0.088	0 (0 - 3)	3 (0.3 - 4)	**0.001**	0.722	0.050
Face	0 (0 - 0)	0 (0 - 0)	0.138	0.269	0.053	0 (0 - 1)	0 (0 - 0.5)	0.796	0.077	0.069
Thorax	3 (3 - 4)	3 (3 - 3)	**0.010**	0.567	0.017	3 (3 - 4)	4 (3 - 4)	0.149	0.301	0.102
Abdomen	2 (0 - 2)	0 (0 - 0)	**<0.001**	0.585	0.016	0 (0 - 2)	0 (0 - 0.5)	0.238	0.244	0.118
Extremities	1 (0 - 2.3)	2 (0 - 2)	0.770	0.079	0.040	2 (0 - 3)	2 (0 - 3)	0.442	0.164	0.097
No. of rib fractures (median, IQR)	7 (6 - 10.3)	6 (4 - 8)	**0.001**	0.643	0.096	10 (8 - 12)	9 (7 - 13)	0.405	0.049	0.059
Bilateral rib fractures (n, %)	10 (36)	69 (33)	0.904	0.067	0.041	14 (38)	25 (46)	0.610	0.155	0.037

(table 1 cotinued)

Concomitant injuries (n, %)										
Pulmonary contusion	11 (39)	99 (47)	0.561	0.159	0.017	22 (60)	38 (69)	0.467	0.202	0.094
Pneumothorax	19 (68)	117 (55)	0.285	0.263	0.058	27 (73)	47 (86)	0.226	0.311	0.071
Hemothorax	3 (11)	37 (18)	0.529	0.195	0.059	9 (24)	22 (40)	0.182	0.340	0.041
Sternum fracture	5 (18)	24 (11)	0.491	0.186	0.071	5 (14)	9 (16)	0.938	0.080	0.053
Blood pH (mean ± SD)	7.34 (0.12)	7.32 (0.08)	0.308	0.173	0.046	7.33 (0.1)	7.27 (0.13)	0.034	0.475	0.036
Base excess (mean ± SD)	-2.4 (5.6)	-3.1 (3.5)	0.367	0.154	0.052	-2.6 (3.5)	-4.7 (6.2)	0.063	0.426	0.068

SMD standardized mean difference; ASA American Society of Anesthesiologists; ISS injury severity score; TTSS thoracic trauma severity score; AIS abbreviated injury score; SD standard deviation; IQR interquartile range

Table 2. Surgery-related characteristics and in-hospital complications

Variable	Multiple rib fracture n=28	Flail chest n=37
Duration until rib fixation in days (median, IQR)	1 (1-2)	1 (1-2)
Duration of surgery in minutes (mean ± SD)	130 (83)	148 (64)
Number of surgically-fixated rib fractures (median, IQR)	4 (3-5)	5 (4-6)
Ratio surgically-fixated ribs and total number of rib fractures	0.54	0.50
In-hospital complications after surgical rib fixation (n, %)		
Pneumothorax	0 (0)	2 (5.4)
Tension pneumothorax	1 (3.6)	0 (0)
Pleural empyema	0 (0)	2 (5.4)
Excess pleural fluid	1 (3.6)	0 (0)
Infection of osteosynthesis material	1 (3.6)	0 (0)
Hematoma	0 (0)	1 (2.7)
Hemothorax	0 (0)	1 (2.7)

Table 3: Regression analysis assessing the influence of rib fixation for a flail chest after propensity score matching.

Continuous variables	Rib fixation for flail chest Median (IQR) Surgery	Non-operative	δ	95% CI	SE	*P* value
Duration of ICU stay in days	6 (0-13)	2 (0-8)	-2.7	-13.9 - 8.5	5.721	0.638
Duration of IMV in days	3 (0-9)	0 (0-7)	-2.3	-11.6 - 7.0	4.750	0.624
Duration of epidural analgesia in days	0 (0-3)	2 (0-7)	-1.2	-3.4 - 1.0	1.116	0.290
Duration of hospital stay in days	21 (11-31)	11 (8-18)	1.9	-14.3 - 18.0	8.240	0.820
	n (%)		*OR*	*95% CI*	*SE*	*P value*
Pneumonia	4.8 (23)	5.6 (20)	1.1	0.2 - 5.8	0.826	0.871
Tracheostomy	2.6 (12)	3.5 (13)	NA	NA NA	NA	NA
In hospital mortality	2.2 (10)	3.3 (12)	NA	NA - NA	NA	NA

δ indicates the difference in mean outcome value between rib fixation and non-operative treatment; SE standard error; OR odds ratio

ICU intensive care unit; IMV invasive mechanical ventilation; CI confidence interval; IQR interquartile range; NA no answer

Table 4: Regression analysis assessing the influence of rib fixation for multiple rib fractures after propensity score matching.

Outcome variable	Rib fixation for multiple rib fractures Median (IQR) Surgery	Non-operative	δ	95% CI	SE	*P* value
Duration of ICU stay in days	0 (0-11)	1 (0-2)	1.6	-3.5 - 6.7	2.600	0.530
Duration of IMV in days	0 (0-9)	0 (0-1)	2.4	-2.8 - 7.6	2.637	0.365
Duration of epidural analgesia in days	0 (0-4)	0 (0-3)	-0.1	-1.9 - 1.7	0.917	0.939
Duration of hospital stay in days	12 (9-23)	10 (6-16)	6.5	-0.6 - 13.6	3.636	0.074
	n (%)		*OR*	*95% CI*	*SE*	*P value*
Pneumonia	7.4 (34)	5 (14)	3.2	0.8 - 13.9	0.743	0.114
Tracheostomy	1.7 (7.8)	0.7 (2)	NA	NA - NA	NA	NA
In-hospital mortality	0 (0)	3.3 (9.1)	NA	NA - NA	NA	NA

δ indicates the difference in mean outcome value between rib fixation and nonoperative treatment; SE standard error; OR odds ratio

ICU intensive care unit; IMV invasive mechanical ventilation; CI confidence interval; IQR interquartile range; NA no answer

Discussion

We compared rib fixation with nonoperative treatment for both flail chest and multiple rib fractures. After propensity score matching, adjusting for all anticipated confounding variables, rib fixation for a flail chest was not associated with differences in ILOS or the other outcome measures. Neither did we find a difference in HLOS for rib fixation in patients with multiple rib fractures, nor for the other outcome measures.

In our study there was no association between rib fixation and the primary and secondary outcome measures as compared to nonoperative treatment for patients with a flail chest. Three RCTs have been published on this subject. The first was from Tanaka et al. who studied 37 patients (18 surgical, 19 non-operative) with a flail chest unable to wean from mechanical ventilation and performed surgery on average seven days after admission; they excluded patients with severe head trauma, spinal injury, and no development of respiratory failure.[23] Granetzny et al. compared 40 patients (20 surgical, 20 non-operative) with a flail chest and performed surgery 24 to 36 hours after intensive care admission; they excluded patients with disturbed consciousness after head trauma, fractures of the upper three ribs, and severe associated trauma to other systems.[24] Marasco et al. studied 46 patients (23 surgical, 23 non-operative) with a flail chest who were ventilator dependent without prospect of successful weaning within 48 hours and performed surgery on average 4.6 days after admission; they excluded patients of 80 years old and older, spinal injury, open fractures, and a Glasgow Coma Scale of <10 at the scene or on admission.[25] All three studies reported a significant decrease in DMV and ILOS. One possible explanation for these contrasting results as compared to our study might be the more restrictive inclusion criteria used in the aforementioned studies. In our study, all patients with multiple rib fractures or a flail chest were studied, including patients with head trauma or other severe injuries. Less strict inclusion criteria will result in a more diverse patient selection and will increase the generalizability of the results; however, it could also have diminished the effect of rib fixation in an already heterogeneous patient group.

Interestingly, the ILOS of both the surgical (median 6 days; mean 8.9 days) and non-operative group (median 3 days; mean 10.5 days) in our cohort were lower as compared to Tanaka et al. (surgical:16.5; non-operative: 26.8 days), Granetzny et al. (surgical: 9.6 and non-operative: 14.6 days), and Marasco et al. (surgical: 13.5 and non-operative: 18.7 days).[23–25] Also the DMV in our entire cohort was lower as compared to the published RCTs.

In the current literature, only one study compared rib fixation with non-operative treatment for patients with multiple rib fractures without a flail chest. In a retrospective study with 124 patients, Qiu et al. reported a significantly shorter HLOS after rib fixation for multiple rib fractures as compared to non-operative treatment (11.1 days vs 15.9 days; p=0.013) and also found lower pneumonia rates (4.6% vs 17%; p = 0.025).[13] Fitzgerald et al. performed a cohort

study of patients 65 years old and older with more than one rib fracture, but did not report the number of patients with a flail chest.[14] In that study, rib fixation resulted in a decrease in mortality and respiratory complications compared to non-operative treatment. Khandelwal et al. presented a study with 67 patients (38 surgical, 29 non-operative) with only two patients with a flail chest in the surgical group.[26] They found a significant reduction in pain intensity and early return to work after rib fixation.

Few studies have reported on complication rates after rib fixation. Of the published trials only Granetzny et al. reported a complication rate of 35% including pneumonia and mortality.[24] Other complications were empyema (5%), mediastinitis (10%), wound infection (10%), and chest wall deformity (5%). In another prospective study, Pieracci et al. reported an infection rate of 3% after rib fixation but did not report on other complications.[27] In our study, nine (14%) of the surgically treated patients had a postoperative complication.

The results of this study should be interpreted considering several limitations. The retrospective design of the study might have affected the outcome measures due to the effects of data loss and under reporting. Although a clinical algorithm was used to select suitable patients for surgical treatment in UMCU, the final decision was made by the attending surgeon-on-call which is a potential selection bias. Pain is the most important indication for rib fixation in our clinical based treatment algorithm. However, due to the retrospective design of this study, we were unable to compare pain scores and interventions for pain treatment. Therefore, we might have missed this potential beneficial effect of rib fixation. Instead, we used HLOS as a surrogate marker for treatment success, but this outcome measure might have been influenced by other factors such as intensive care treatment, ventilation modalities and logistic issues with patient transfer and could therefore have diminished differences in treatment effect. Additionally, we did not use a scoring system or other determinant for rib fixation other than the clinical algorithm (Figure 1). However, we are confident to have included all the potential factors associated with decision making in the operating center in our propensity score matched model. Finally, there is still no good fracture classification to distinguish between fracture type and location. It is speculated that lateral and lower rib fractures are more painful due to increased mobility of the fracture parts. Fracture classification could influence success of rib fixation and this should be investigated in future studies.

Even though this study is one of the largest studies reporting on this subject, the number of included patients is still relatively small and was possibly too small to detect relatively small yet clinically meaningful differences. Furthermore, as part of the between-hospital comparison and due to clinical practice, there were differences in the baseline criteria between the surgical group and the non-operatively treated group. However, using a propensity score model, we were able to successfully match on all measured baseline characteristics eliminating possible confounding due to measured patient characteristics. As

with any observational study, our results are potentially biased by unmeasured confounding (e.g. subjective indications for surgery, pain scores and fracture classification), be it that we believe we have included most confounders in our analysis and the potential impact of unmeasured confounding therefore seems limited.

The University Medical Hosptial Utrecht was the first hospital in the Netherlands to perform rib fixation for patients with flail chest and multiple rib fractures. With more than seven years of experience, rib fixation has become an established procedure with a univocal clinical-based treatment algorithm, with its main focus on clinical signs of flail chest and pain. Nevertheless, no benefit could be demonstrated in this population with rib fractures who received early operative fixation in their clinical course. Therefore results of this study, combined with the limited existing evidence and the substantial costs of surgical treatment, emphasize the need for future studies before rib fixation is embedded or abandoned in clinical practice, but also to identify specific patient groups who would benefit from rib fixation. These studies should focus on optimization of the indication and describe long-term outcome after rib fixation.

References

1. Bulger EM, Arneson MA, Mock CN, Jurkovich GJ. Rib fractures in the elderly. *J Trauma*. 2000;48(6):1040-1047.

2. Ziegler DW, Agarwal NN. The morbidity and mortality of rib fractures. *The Journal of trauma*. 1994;37(6):975-979.

3. Lin FC-F, Li R-Y, Tung Y-W, Jeng K-C, Tsai SC-S. Morbidity, mortality, associated injuries, and management of traumatic rib fractures. *Journal of the Chinese Medical Association*. 2016;79(6):329-334. doi:10.1016/j.jcma.2016.01.006.

4. Vana PG, Neubauer DC, Luchette FA. Contemporary management of flail chest. *The American surgeon*. 2014;80(6):527-535.

5. Testerman GM. Adverse Outcomes in Younger Rib Fracture Patients. *Southern Medical Journal*. 2006;99(4):335-339. doi:10.1097/01.smj.0000203815.29757.d3.

6. Brasel K, Guse C, Layde P, Weigelt J. Rib fractures: relationship with pneumonia and mortality. *Critical Care Medicine*. 2006;34(6):1642-1646.

7. Flagel BT, Luchette FA, Reed RL, et al. Half-a-dozen ribs: The breakpoint for mortality. *Surgery*. 2005;138(4):717-725. doi:10.1016/j.surg.2005.07.022.

8. Cannon RM, Smith JW, Franklin GA, Harbrecht BG, Miller FB, Richardson JD. Flail chest injury: are we making any progress? *Am Surg*. 2012;78(4):398-402.

9. Dehghan N, De Mestral C, McKee MD, Schemitsch EH, Nathens A. Flail chest injuries: A review of outcomes and treatment practices from the national trauma data bank. *Journal of Trauma and Acute Care Surgery*. 2014;76(2):462-468. doi:10.1097/TA.0000000000000086.

10. Liman ST, Kuzucu A, Tastepe AI, Ulasan GN, Topcu S. Chest injury due to blunt trauma. *European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery*. 2003;23(3):374-378.

11. Clark GC, Schecter WP, Trunkey DD. Variables affecting outcome in blunt chest trauma: flail chest vs. pulmonary contusion. *The Journal of trauma*. 1988;28(3):298-304.

12. Kasotakis G, Hasenboehler EA, Streib EW, et al. Operative fixation of rib fractures after blunt trauma: A practice management guideline from the Eastern Assocaiation for the Surgery of Trauma. *Journal of Trauma and Acute Care Surgery*. 2017;82(3):618-626. doi:10.1097/TA.0000000000001350.

13. Qiu M, Shi Z, Xiao J, Zhang X, Ling S, Ling H. Potential Benefits of Rib Fracture Fixation in Patients with Flail Chest and Multiple Non-flail Rib Fractures. *The Indian journal of*

surgery. 2016;78(6):458-463. doi:10.1007/s12262-015-1409-2.

14. Fitzgerald MT, Ashley DW, Abukhdeir H, Christie DB. Rib fracture fixation in the 65 years and older population: A paradigm shift in management strategy at a Level I trauma center. *Journal of Trauma and Acute Care Surgery*. 2017;82(3):524-527. doi:10.1097/TA.0000000000001330.

15. Taylor BC, French BG, Fowler TT. Surgical approaches for rib fracture fixation. *Journal of Orthopaedic Trauma*. 2013;27(7):e168-e173. doi:10.1097/BOT.0b013e318283fa2d.

16. Pape HC, Remmers D, Rice J, et al. Appraisal of early evaluation of blunt chest trauma: development of a standardized scoring system for initial clinical decision making. *Journal of Trauma*. 2000;49(3):496-504.

17. Aukema TS, Beenen LFM, Hietbrink F, Leenen LPH. Validation of the Thorax Trauma Severity Score for mortality and its value for the development of acute respiratory distress syndrome. *Open Access Emergency Medicine*. 2011;3:49-53. doi:10.2147/OAEM.S22802.

18. Martinez Casas I, Amador Marchante MA, Paduraru M, Fabregues Olea AI, Nolasco A, Medina JC. Thorax Trauma Severity Score: Is it reliable for Patient's Evaluation in a Secondary Level Hospital? *Bulletin of emergency and trauma*. 2016;4(3):150-155.

19. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res*. 2011;46(3):399-424. doi:10.1080/00273171.2011.568786.

20. Buuren S van, Groothuis-Oudshoorn K. mice :BMultivariate Imputation by Chained Equations in *R*. *Journal of Statistical Software*. 2011;45(3):1-67. doi:10.18637/jss.v045.i03.

21. Ho DE, Imai K, King G, Stuart EA. MatchIt :BNonparametricBPreprocessingBforBParametricB Causal Inference. *Journal of Statistical Software*. 2011;42(8):1-28. doi:10.18637/jss.v042.i08.

22. R Core Team. R: A Language and Environment for Statistical Computing. 2017.

23. Tanaka H, Yukioka T, Yamaguti Y, et al. Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients. *Journal of Trauma*. 2002;52(4):727-732. doi:10.1097/00005373-200204000-00020.

24. Granetzny A, Abd El-Aal M, Emam E, Shalaby A, Boseila A. Surgical versus conservative treatment of flail chest. Evaluation of the pulmonary status. *Interactive Cardiovascular and Thoracic Surgery*. 2005;4(6):583-587. doi:10.1510/icvts.2005.111807.

25. Marasco SF, Davies AR, Cooper J, et al. Prospective randomized controlled trial of

operative rib fixation in traumatic flail chest. *Journal of the American College of Surgeons*. 2013;216(5):924-932. doi:10.1016/j.jamcollsurg.2012.12.024.

26. Khandelwal G, Mathur RK, Shukla S, Maheshwari A. A prospective single center study to assess the impact of surgical stabilization in patients with rib fracture. *International journal of surgery (London, England)*. 2011;9(6):478-481. doi:10.1016/j.ijsu.2011.06.003.

27. Pieracci FM, Lin Y, Rodil M, et al. A prospective, controlled clinical evaluation of surgical stabilization of severe rib fractures. *Journal of Trauma and Acute Care Surgery*. 2016;80(2):187-194. doi:10.1097/TA.0000000000000925.

PART 2

METHODOLOGICAL CONSIDERATIONS

CHAPTER 6

ADDED VALUE OF OBSERVATIONAL STUDIES IN SURGERY: THE HIERARCHICAL STRUCTURE OF STUDY DESIGNS REQUIRES A MORE REFINED APPROACH.

R.B. Beks, R.M. Houwert, R.H.H. Groenwold

Nederlands tijdschrift voor geneeskunde. 2017

Abstract

The randomised placebo-controlled trial (RCT) is the gold standard for the evaluation of medical interventions. Observational studies, on the other hand, usually do not get much credit. For studies investigating surgical interventions this does not always seem entirely justified. A more refined approach might be needed for the often-used hierarchical structure of research designs. Instead of a strict separation of results from RCTs and other designs, results of the different designs should rather be regarded as complementary to each other when evaluating surgical interventions in traumatology.

Samenvatting

De gerandomiseerde placebogecontroleerde trial (RCT) geldt als de gouden standaard voor het evalueren van medische interventies. Observationeel onderzoek daarentegen krijgt doorgaans minder waardering. Voor chirurgisch onderzoek lijkt dit echter niet altijd terecht en daarom is nuancering nodig van de veelgebruikte hiërarchische structuur van onderzoeksopzetten. In plaats van een strikte scheiding aan te brengen tussen de resultaten van RCT's en die van andere studieopzetten zouden we voor het evalueren van operatieve interventies de resultaten van verschillende onderzoeksopzetten veel meer als complementair aan elkaar moeten zien.

De gerandomiseerde placebogecontroleerde trial (RCT) geldt als de gouden standaard voor het evalueren van medische interventies.[1] Patiënten worden gerandomiseerd tussen een interventie- en controlegroep waarbij zowel de patiënt als de behandelaar wordt geblindeerd voor de behandelkeuze. Zo ontstaan 2 vergelijkbare groepen en kunnen waargenomen verschillen in gezondheidsuitkomsten tussen deze onderzoeksgroepen worden toegeschreven aan de interventie.
In observationeel onderzoek naar de effecten van medische interventies vindt geen randomisatie plaats en is blindering niet mogelijk. Hierdoor is deze onderzoeksopzet gevoeliger voor bias, bijvoorbeeld door confounding, en krijgt observationeel onderzoek doorgaans minder waardering.[2,3]

In de loop der jaren heeft zich een hiërarchie van niveaus van bewijskracht ('levels of evidence') ontwikkeld waarin observationeel onderzoek ver onder de RCT wordt ingeschaald. Toch zijn de resultaten van observationeel onderzoek vaak beter te generaliseren door de minder strikte inclusiecriteria, zijn er meer mogelijkheden tot subgroepanalyse en is het onderzoek makkelijker, sneller en voor minder geld uit te voeren dan een RCT.

3 typen onderzoeksvragen

Voor de evaluatie van interventies in de chirurgische kliniek kunnen 3 typen onderzoeksvragen worden onderscheiden.[4]

Geneesmiddel vs. controlemiddel Type 1-onderzoek richt zich op farmacologische interventies bij chirurgische patiënten, waarbij een bepaald geneesmiddel wordt vergeleken met een placebo of een ander geneesmiddel. Om het risico op bias te minimaliseren is randomisatie een vereiste; blindering van de deelnemer, de behandelaar of beiden is optioneel. Een voorbeeld is een recente RCT waarin het effect van farmacologische tromboseprofylaxe werd vergeleken met dat van placebo bij patiënten die een artroscopie van de knie hadden ondergaan en bij patiënten met onderbeengips.[5]

Operatie vs. controleoperatie Type 2-onderzoek vergelijkt verschillende operatieve interventies, waarbij patiënten die een bepaalde ingreep ondergaan worden vergeleken met een controlegroep die een alternatieve operatie of een 'sham'-operatie krijgt. Bij een sham-operatie doorloopt de patiënt hetzelfde traject als de onderzoeksgroep, inclusief anesthesie en gelijke chirurgische incisie, maar zonder de daadwerkelijke ingreep. In tegenstelling tot een farmacologische interventie omvat de operatieve interventie een traject van pre-, peri- en postoperatieve handelingen. Bovendien is elke operatieve interventie een complexe interventie, wat betekent dat de uitkomst ook afhankelijk is van de vaardigheid en expertise van het chirurgische team. Bij type 2-onderzoek gaat het doorgaans niet om de werkzame component van de interventie, zoals bij type 1-onderzoek wel het geval is, maar om het volledige pakket aan handelingen. Type 2-onderzoek is minder gevoelig voor confounding als

verwijzing naar de chirurg een willekeurig proces is en er geen aanwijzingen zijn dat de behandelvoorkeur van de chirurg sterk onderhevig is aan patiëntkenmerken.[6]

Operatie vs. geen operatie Type 3-onderzoek vergelijkt een operatieve nterventie met een niet-operatieve interventie, bijvoorbeeld operatie versus conservatief beleid bij patiënten met een fractuur. Net als bij type 2-onderzoek is de operatieve interventie complex. Maar type 3-onderzoek heeft mogelijk een groter risico op confounding dan type 2-onderzoek doordat de patiëntkenmerken een grotere invloed kunnen hebben op het besluit tussen een operatie of conservatief beleid.

RCT in chirurgie

Voor het minimaliseren van confounding is het belangrijk 2 gelijkwaardige groepen te vormen door het willekeurig toewijzen van behandelingen aan de patiënten (randomiseren). Dit is mogelijk in trials met een relatief laagcomplexe interventie, zoals het toedienen van medicatie (type 1-trials).
Een operatie is per definitie een hoogcomplexe handeling, waarbij een goed resultaat afhangt van de expertise van het chirurgische team. Afhankelijk van hoe er wordt gerandomiseerd kan een situatie ontstaan waarin een chirurg een complexe handeling moet uitvoeren die minder binnen zijn expertise valt doordat een alternatieve behandeling de standaard is in zijn praktijk. Dit bemoeilijkt een eerlijke vergelijking van verschillende interventies. Gekoppeld aan technische expertise hebben chirurgen vaak een sterke behandelvoorkeur ontwikkeld, waardoor er onder hen weinig animo kan zijn voor deelname aan een RCT waarin zij verschillende operatieve behandelingen zullen moeten toepassen.[4,7] Een alternatief is het randomiseren van patiënten naar behandelaars, die vervolgens alle aan hen toegewezen patiënten opereren met de techniek waarin zij het bekwaamst zijn.

Ook patiënten kunnen een sterke voorkeur hebben voor een bepaalde behandeling en willen daarom niet dat de keuze voor een behandeling afhangt van toeval (randomisatie). Dit geldt met name voor type 3-trials omdat de beide behandelopties enorm verschillen en de chirurgische interventie onomkeerbaar is. Dit maakt patiëntvoorkeur een belangrijk obstakel voor chirurgische RCT's.[4,8] Een voorbeeld hiervan is een type 3-RCT waarin fractuurfixatie van proximale humerusfracturen werd vergeleken met een conservatief beleid, maar waaraan meer dan de helft van de gevraagde patiënten weigerde deel te nemen.[9]

In een chirurgische RCT kan het blinderen van een patiënt lastig zijn. Een oplossing hiervoor is de eerdergenoemde sham-operatie. Dit type operatie kent ethische bezwaren, omdat het gepaard gaat met operatierisico's voor een patiënt zonder dat daar mogelijk gunstige effecten tegenover staan. Bovendien is het de vraag of het onderzoek zich richt op een specifieke werkzame component – waarbij blindering noodzakelijk lijkt – of dat er 2 verschillende strategieën worden vergeleken, inclusief de effecten die het gevolg zijn van

kennis over de gevolgde strategie – waarbij blindering juist overbodig is. Het is wenselijk de beoordelaar van de uitkomstmaat wel te blinderen, om op die manier het risico op zogenaamde beoordelaarsbias ('observer bias') te minimaliseren.
Tot slot is het uitvoeren van gerandomiseerd onderzoek met klinische patiënten duur en zijn middelen vaak slechts in beperkte mate beschikbaar. Hierdoor is de follow-upduur in RCT's vaak beperkt vergeleken met bijvoorbeeld retrospectief observationeel onderzoek.

Observationeel onderzoek in chirurgie

Observationele studies binnen de chirurgie zijn waardevol voor de evaluatie van gunstige effecten van operatieve behandelingen en het in kaart brengen van complicaties. De inclusiecriteria van observationeel onderzoek zijn minder selectief dan die van RCT's en daardoor zijn de patiëntencohorten groter en representatiever. Dit betekent dat de resultaten van observationeel onderzoek doorgaans meer van toepassing zijn op de dagelijkse praktijk. Een groter aantal patiënten geeft ook meer inzicht in zeldzame uitkomsten en biedt de mogelijkheid tot subgroepanalyses. Vanzelfsprekend valt of staat de waarde van de resultaten van een observationele studie, net als bij een RCT, met de kwaliteit van de gegevens en een juiste vraagstelling.

Het belangrijkste nadeel van observationeel onderzoek is de gevoeligheid voor bias. Dit risico op bias maakt dat bijvoorbeeld een type 1-onderzoeksvraag vrijwel altijd moet worden beantwoord met een RCT.

Type 2- en type 3-onderzoeksvragen bieden meer plaats voor observationeel onderzoek. Dit wordt ondersteund door een recente meta-analyse naar 2 operatieve behandeltechnieken, plaatfixatie en intramedullaire fixatie, bij patiënten met een midschacht-claviculafractuur waarin zowel de resultaten van RCT's als die van observationeel onderzoek werden meegenomen; dit is dus type 2-onderzoek.[10] Voor de primaire uitkomstmaat, die bestond uit 'non-union' en heroperatie, liet observationeel onderzoek inderdaad hetzelfde behandeleffect zien als de RCT's (tabel 1).[10] Uit een andere meta-analyse naar operatief versus conservatief beleid (type 3-onderzoek) bij patiënten met een midschacht-claviculafractuur kwam ook naar voren dat observationeel onderzoek tot dezelfde resultaten leidt als RCT's (tabel 2).[11]

De meerwaarde van observationeel onderzoek blijkt uit een analyse naar een minder harde uitkomstmaat als terugkeer naar werk. Waar de RCT's te klein waren om harde conclusies te kunnen trekken met betrekking tot deze uitkomstmaat, bevatten de observationele onderzoeken voldoende deelnemers om een gunstig effect van operatief ingrijpen op terugkeer naar werk te kunnen concluderen. Dit laat de toegevoegde waarde van schaalvergroting zien door inclusie van observationeel onderzoek, terwijl invloeden van bias minimaal lijken te zijn. Zoals gezegd is het voor observationeel onderzoek naar chirurgische

interventie bij traumatologiepatiënten van belang dat de verwijzing naar de chirurg een willekeurig proces is en er geen aanwijzingen zijn dat de behandelvoorkeur van de chirurg sterk beïnvloedt wordt door patiëntkenmerken. Wanneer niet aan beide voorwaarden wordt voldaan, is er een groot risico op bias en daardoor op verkeerde effectschattingen.

Een voorbeeld hiervan is de vergelijking van de plaatsing van een prothese versus osteosynthese bij patiënten met een heupfractuur. Observationeel onderzoek overschatte het mortaliteitsrisico na behandeling met een prothese met 40% vergeleken met RCT's.[12] In de observationele onderzoeken kozen chirurgen echter sneller voor een prothese bij zwakkere en oudere patiënten, wat een deel van dit gevonden effect verklaart.

Tot slot

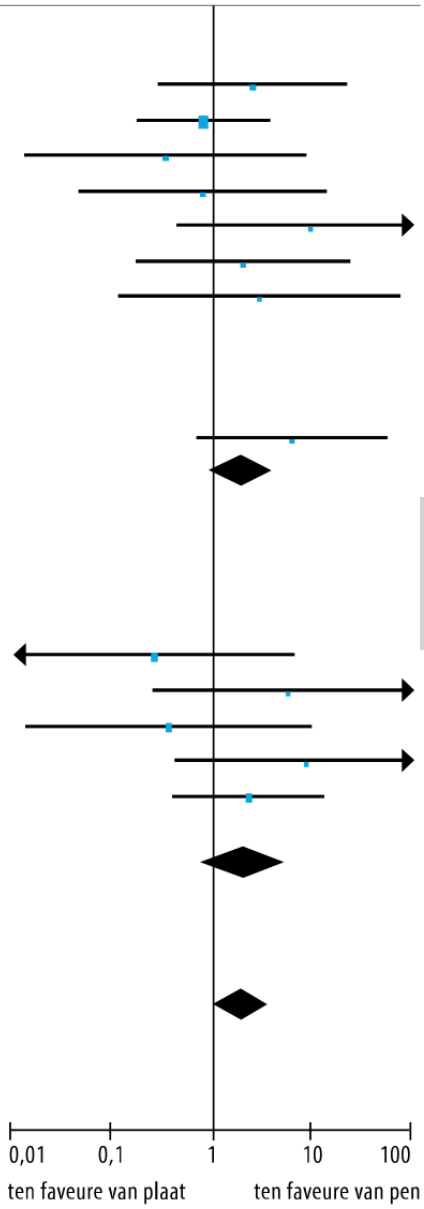
Waarom hebben we observationeel onderzoek nodig voor de evaluatie van operatieve behandelingen? De relatief lage kosten en praktische uitvoerbaarheid van dit type onderzoek maken het mogelijk naar de lange-termijnuitkomsten van grote representatieve groepen of subgroepen van patiënten te kijken, terwijl het risico op confounding voor type 2- en type 3-onderzoek relatief klein kan zijn.

Eén aspect hebben we echter nog niet benoemd en dat is het vakmanschap dat nodig is om te opereren en de creativiteit die ontwikkeling van nieuwe chirurgische technieken en behandelmethoden mogelijk maakt. De creativiteit die sommige chirurgen bezitten laat zien wat er allemaal mogelijk is. Dat is niet te vangen in een RCT, maar wel in observationeel onderzoek. De techniek en vaardigheid van de chirurgen die deelnemen aan RCT's zijn niet altijd representatief voor de vaardigheden van chirurgen in de dagelijkse praktijk. De vertaalslag van de resultaten van een RCT naar de dagelijkse praktijk valt daardoor niet altijd goed te maken.

Voor het evalueren van operatieve interventies zit de kracht in het combineren van de resultaten van RCT's én observationeel onderzoek. In plaats van een strikte scheiding aan te brengen tussen de resultaten van verschillende onderzoeksopzetten zouden we voor het evalueren van type 2- en type 3-interventies de resultaten van observationeel onderzoek en die van RCT's veel meer als complementair aan elkaar moeten zien.

TABEL 1 Resultaten van meta-analyse naar risico op grote re-interventie na behandeling voor midschacht-claviculafractuur met plaatfixatie of plaatsing van intramedullaire pen (type 2 onderzoek)[11]

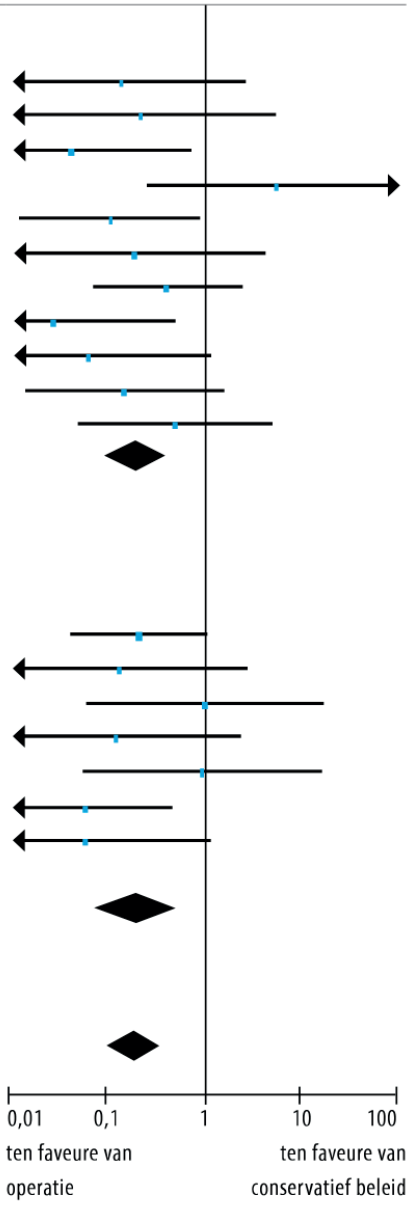
	plaat		pen			
studie of subgroep	**re-interventie; n**	**totaal; n**	**re-interventie; n**	**totaal; n**	**gewicht; %**	**OR (95%-CI)**
cohort						
Bohme	6	53	1	20	8,2	2,43 (0,27-21,5)
Fu	3	50	4	53	23,1	0,78 (0,17-3,68)
Jones	0	24	1	25	9,1	0,33 (0,01-8,59)
Kleweno	1	18	1	14	6,7	0,76 (0,04-13,4)
Lee 2008	2	32	0	56	2,1	9,26 (0,43-199)
Narsaria	2	34	1	32	6,1	1,94 (0,17-22,5)
Saha	1	37	0	34	3,2	2,84 (0,11-72,0)
Tarng	0	32	0	25		
Thyagarajan	0	17	0	17		
Wenninger	0	29	0	33		
Wijdicks	5	43	1	47	5,4	6,05 (0,68-54,1)
subtotaal	20	369	9	356	64,0	1,86 (0,87-4,00)
heterogeniteit: χ^2: 4,94, df = 7 (p = 0,67); I^2: 0%						
test voor totaal effect: Z: 1,60 (p = 0,11)						
RCT						
Andrade	0	33	1	26	10,4	0,25 (0,01-6,49)
Assobhi	2	19	0	19	2,8	5,57 (0,25-124)
Ferran	0	15	1	17	8,7	0,35 (0,01-9,38)
Lee 2007	3	30	0	32	2,7	8,27 (0,41-167)
Van der Meijden	4	58	2	62	11,4	2,22 (0,39-12,6)
Zehir	0	21	0	24		
subtotaal	9	176	4	180	36,0	1,92 (0,70-5,25)
heterogeniteit: χ^2: 3,90, df = 4 (p = 0,42); I^2: 0%						
test voor totaal effect: Z: 1,27 (p = 0,21)						
totaal	29	545	13	536	100	1,88 (1,02-3,46)
heterogeniteit: χ^2: 8,85, df = 12 (p = 0,72); I^2: 0%						
test voor totaal effect: Z: 2,04 (p = 0,04)						
test voor subgroepverschil: χ^2: 0,00, df = 1 (p = 0,97); I^2: 0%						



OR = oddsratio.

TABEL 2 Resultaten van meta-analyse naar risico op non-union na behandeling voor midschacht-claviculafractuur met osteosynthese of conservatief beleid (type 3 onderzoek)[12]

	operatie		conservatief			
studie of subgroep	**non-union; n**	**totaal; n**	**non-union; n**	**totaal; n**	**gewicht; %**	**OR (95%-BI)**
cohort						
Althausen	0	66	4	83	4,0	0,13 (0,01-2,51)
Böhme	0	73	1	47	3,3	0,21 (0,01-5,29)
Daniilidis	0	81	9	70	4,2	0,04 (0,00-0,70)
Grassi	2	40	0	40	3,7	5,26 (0,24-113)
Jones	1	49	7	41	7,5	0,10 (0,01-0,86)
Jubel	0	20	2	20	3,6	0,18 (0,01-4,01)
Khorami	2	35	4	30	11,0	0,39 (0,07-2,32)
Kulshrestha	0	45	8	28	4,1	0,03 (0,00-0,48)
Smekal	0	60	6	52	4,1	0,06 (0,00-1,08)
Thyagarajan	1	34	3	17	6,3	0,14 (0,01-1,48)
Van der Ven	1	39	3	58	6,5	0,48 (0,05-4,82)
subtotaal	7	542	47	486	58,2	0,19 (0,09-0,40)
heterogeniteit: Tau^2: 0,00; χ^2: 9,86, df = 10 (p = 0,45); I^2: 0%						
test voor totaal effect: Z: 4,28 (p < 0,001)						
RCT						
COTS	2	62	7	49	13,1	0,20 (0,04-1,01)
Chen	0	30	3	30	3,8	0,13 (0,01-2,61)
Judd	1	29	1	28	4,3	0,96 (0,06-16,2)
Melean	0	35	4	41	3,9	0,12 (0,01-2,26)
Mirzatolooei	1	26	1	24	4,3	0,92 (0,05-15,6)
Robinson	1	86	16	92	8,3	0,06 (0,01-0,43)
Virtanen	0	26	6	25	4,0	0,06 (0,00-1,07)
Witzel	0	35	0	33		
subtotaal	5	329	38	322	41,8	0,17 (0,07-0,43)
heterogeniteit: Tau^2: 0,00; χ^2: 4,81, df = 6 (p = 0,57); I^2: 0%						
test voor totaal effect: Z: 3,78 (p < 0,001)						
totaal	12	871	85	808	100	0,18 (0,10-0,33)
heterogeniteit: Tau^2: 0,00; χ^2: 14,66, df = 17 (p = 0,62); I^2: 0%						
test voor totaal effect: Z: 5,71 (p < 0,001)						
test voor subgroepverschil: χ^2: 0,01, df = 1 (p = 0,90); I^2: 0%						



OR = oddsratio; COTS = Canadian Orthopaedic Trauma Society.

Referenties

1. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med*. 2008;5(3):e67. doi:10.1371/journal.pmed.0050067.
2. Cole GD, Francis DP. Trials are best, ignore the rest: safety and efficacy of digoxin. *BMJ (Clinical research ed)*. 2015;351(August):h4662. doi:10.1136/bmj.h4662.
3. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*. 1996;312(7040):1215-1218.
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. van Adrichem RA, Nemeth B, Algra A, et al. Thromboprophylaxis after Knee Arthroscopy and Lower-Leg Casting. *N Engl J Med*. 2016. doi:10.1056/NEJMoa1613303.
6. Wijdicks FJ, Houwert M, Dijkgraaf M, et al. Complications after plate fixation and elastic stable intramedullary nailing of dislocated midshaft clavicle fractures: a retrospective comparison. *Int Orthop*. 2012;36(10):2139-2145. doi:10.1007/s00264-012-1615-5.§
7. van der Meijden OA, Houwert RM, Hulsmans M, et al. Operative treatment of dislocated midshaft clavicular fractures: plate or intramedullary nail fixation? A randomized controlled trial. *J Bone Joint Surg Am*. 2015;97(8):613-619. doi:10.2106/JBJS.N.00449.
8. Solomon MJ, McLeod RS. Should we be performing more randomized controlled trials evaluating surgical operations? *Surgery*. 1995;118(3):459-467.
9. 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. doi:10.1001/jama.2015.1629.
10. 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. doi:10.1016/j.jse.2016.01.018.
11. 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. doi:10.1177/0363546516673615.
12. Bhandari M, Tornetta P, Ellis T, et al. Hierarchy of evidence: Differences in results between non-randomized studies and randomized trials in patients with femoral neck fractures. *Archives of Orthopaedic and Trauma Surgery*. 2004;124(1):10-16. doi:10.1007/s00402-003-0559-z.

CHAPTER 7

OPERATIVE VS. CONSERVATIVE TREATMENT OF PROXIMAL HUMERUS FRACTURES. A SYSTEMATIC REVIEW AND META-ANALYSIS OF OBSERVATIONAL STUDIES AND RANDOMISED CONTROLLED TRIALS.

R.B. Beks, Y. Ochen, H. Frima, D.P.J. Smeeing,
O. van der Meijden, T.K. Timmers, D. van der Velde,
M. van Heijl, L.P.H. Leenen,
R.H.H. Groenwold, R.M. Houwert

Journal of Shoulder and Elbow Surgery. 2018

Abstract

Background

There is no consensus on the choice of treatment for displaced proximal humerus 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 (RCT) and observational studies.

Methods

The databases of MEDLINE, Embase, CENTRAL, and CINAHL were searched on September 5th 2017 for studies comparing operative versus nonoperative treatment of proximal humerus 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.

Result:

Twenty-two studies were included; seven 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; I2=69%). Major reinterventions occurred more often in the operative group. Pooled effects of RCTs were similar compared to pooled effects of observational studies for all outcome measures.

Conclusions

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

Introduction

The proximal humerus fracture is the third most common fracture seen in the elderly with an incidence of 82 per 100,000 person years with an annual increasing rate of 13.7% per year over the last 33 years.[1–3] The typical patient is a female aged 65 or over.[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 (ORIF) 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] Additionally, there is no consensus whether surgery is beneficial for the older patient with a displaced proximal humerus fracture.

There is increasing scientific evidence which demonstrates that meta-analyses of both high quality observational studies and RCT's 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 generalizability of the results due to an increase in patient numbers available for (meta-) analysis.

The aim of this systematic review and meta-analysis is (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 for proximal humerus fractures does not improve functional outcome as compared to nonoperative treatment and (2) including observational studies in this meta-analysis will lead to more robust conclusions without decreasing quality of the results.

Methods

This systematic review and meta-analysis followed guidelines published by PRISMA and MOOSE.[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 (RBB, YO) independently searched MEDLINE, Embase, CENTRAL, and CINAHL databases on September 5th, 2017, for studies comparing operative and nonoperative treatment of proximal humerus fractures. The search syntax is provided in Appendix 1. Both

RCTs and observational studies were included. After screening title and abstract of identified records, studies were independently assessed based on full-text. Eligibility criteria were proximal humerus fracture; operative versus nonoperative treatment; reporting of functional outcomes, and complications. Exclusion criteria were language other than English, Dutch, or German; no availability of full-text; inclusion of patients younger than 18 years old; letters, meeting proceedings, and case reports; external osteosynthesis as operative treatment. Disagreement over eligibility was resolved by discussion with a third reviewer (RMH). References of 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 paper was available.

Data Extraction

Data extraction was done independently by two reviewers (RBB, YO) with a data extraction file. The following data were extracted: first author, journal, year of publication, study period, study design, country/countries in which the study was/were 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. Major reintervention was defined as an additional, initially unplanned, surgery for implant failure, deep infection, symptomatic nonunion, subacromial impingement, or avascular necrosis. Planned implant removal was not considered a major reintervention. Fjalestad et al. reported additional follow up of previously published date which were merged with the original article for this meta-analysis.[13,14]

Quality Assessment

Two reviewers (RBB, HF) 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 Appendix 2. Disagreements were resolved by involving a third reviewer (RMH).

Outcome Measures

The primary outcome measure was physical function as measured by the absolute Constant-Murley[16] score 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.

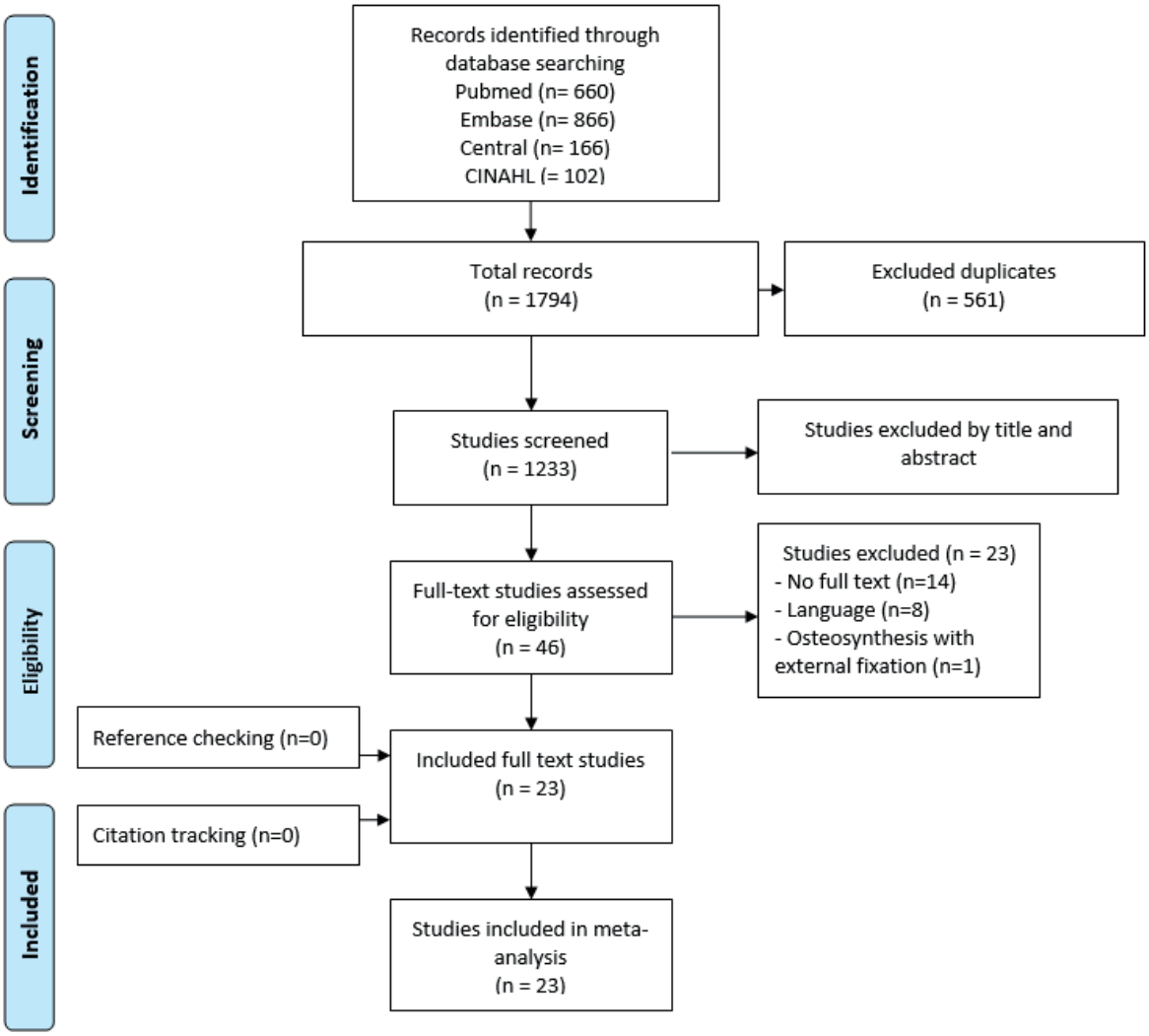
Sensitivity Analyses

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 orthopaedic 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 was treated with a locking plate. Furthermore, a sensitivity analysis was done 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.



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 Appendix 3. The MINORS criteria for unbiased assessment of study end-points 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.

All studies but one included displaced proximal humerus 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 humerus 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 (Appendix 4).[13,14,37,39,40,42,23,25–28,32–34] In patients with a proximal humerus 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 (Figure 1 and 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%)(Appendix 5). 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%)(Appendix 6).

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

Study	Study design	Country	Fracture classification	Treatment groups	Number of patients	Follow-up (months)	Age (years, range or ± SD)	Female/ Male
Boons 2012[4]	RCT	Netherlands	Neer 4-part	Operative: arthroplasty	25	12	76,4 (5,6)	24/1
				Nonoperative: Sling	25	12	79,9 (7,7)	23/2
Fjalestad 2012-14*[10,11]	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 2011[31]	RCT	Sweden	Neer 4-part	Operative: hemiarthroplasty	27	24	75,8 (58-90)	23/4
				Nonoperative: Sling	28	24	77,5 (60-92)	24/4
Olerud 2011b[32]	RCT	Sweden	Neer 3-part	Operative: LP	30	24	72,9 (56-92)	24/6
				Nonoperative: Sling	29	24	74,9 (58-88)	24/5
Rangan 2015[35]	RCT	England	Neer 2,3,4-part	Operative: PHN(n=4), LP(n=90), TB(n=1), arthroplasty(n=10), screw(n=2), other(n=2)	125	24	66,6 (11,8)	97/28
				Nonoperative: Sling or hanging cast	125	24	65,43 (12,09)	95/30
Stableforth 1984[43]	RCT	England	Neer 4-part	Operative: arthroplasty	16	Overall (range): 18-144	65,6 (52-88)	12/4
				Nonoperative: Sling	16		70,1 (60-85)	13/3
Zyto 1997[49]	RCT	Sweden	Neer 3,4-part	Operative: Tension band	20	50	73 (7,5)	18/2
				Nonoperative: Sling	20	50	75 (6,7)	17/3
Court-Brown 2001[9]	PC	Scotland	Neer 2-part	Operative: PHN + tension band fixation	18	12	73	NR
				Nonoperative: Sling	31	12	78	
Hauschild 2013[17]	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[21]	PC	Italy	Neer 2,3,4-part	Operative: K-wire	23	Overall: 86	73,92 (6,01)	Total: 38/13
				Nonoperative: Sling	19		77,47 (6,95)	
Noureai 2014[29]	PC	Iran	Neer 2,3,4-part	Operative: LP, Tension band, K-wire	57	12	Total: 52,9 (15,0)	Total: 70/44
				Nonoperative: Sling	57	12		

(table 1 continued)

Fjalestad 2005[12]	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	Total: 70 (25-95)	Total: 50/20
				Nonoperative: Sling	55	12		
Ilchman 1998[20]	RC+PC	Sweden / Swiss	Neer 3,4-part	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[3]	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[46]	RC	Netherlands	Neer 4,5,6 fracture	Operative: PHN	27	16	64,6 (27-87)	NR
				Nonoperative: Sling	17	69	69,4 (35-84)	
Hageman 2016[15]	RC	Netherlands	Neer 2,3,4-part	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[23]	RC	Germany	Neer 4,5,6 fracture	Operative: LP(n=2), Screws + cerclage(n=7), K-wire(n=4)	13	82	52,5 (14,7)	NR
				Nonoperative: Sling	9	76	52,7 (11,5)	
Lange 2016[24]	RC	Germany	Neer 2,3,4-part	Operative: PHN	41	Overall: 55	69,1 (37-88)	35/6
				Nonoperative: hanging cast or dessault dressing	41		68,9 (42-93)	35/6
Okike 2015[30]	RC	US	Neer 2,3,4-part	Operative: LP	61	Overall: 40	total: 76,9	109/37
				Nonoperative: Sling	146			46/15
Roberson 2017[36]	RC	US	Neer 3,4-part	Operative: reversed arthroplasty	20	53	Overall: 71 (Range 52-88)	19/1
				Nonoperative: Sling	19	29		15/4
Sanders 2011[38]	RC	Australia	Neer 2,3,4-part	Operative: LP	17	37	58 (14)	9/8
				Nonoperative: Sling	18	42	64 (15)	12/6
Tamimi 2015[45]	RC	Canada	Neer 2,3,4-part	Operative: PHN(n=19),LP(n=44),K-wire(n=25)	88	26	Total: 65,3 (15,2)	Total: 57/31
				Nonoperative: Sling	25	28		

*Fjalestad 2012[11] and 2014[10] 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 2. Subgroup & sensitivity analyses of studies included in a meta-analysis of proximal humerus fractures comparing operative to nonoperative treatment

Analysis description	Constant score			Major reintervention			Nonunion			Avascular necrosis		
	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												
RCT	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

Bold: significant pooled effect ($p < 0.05$); 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

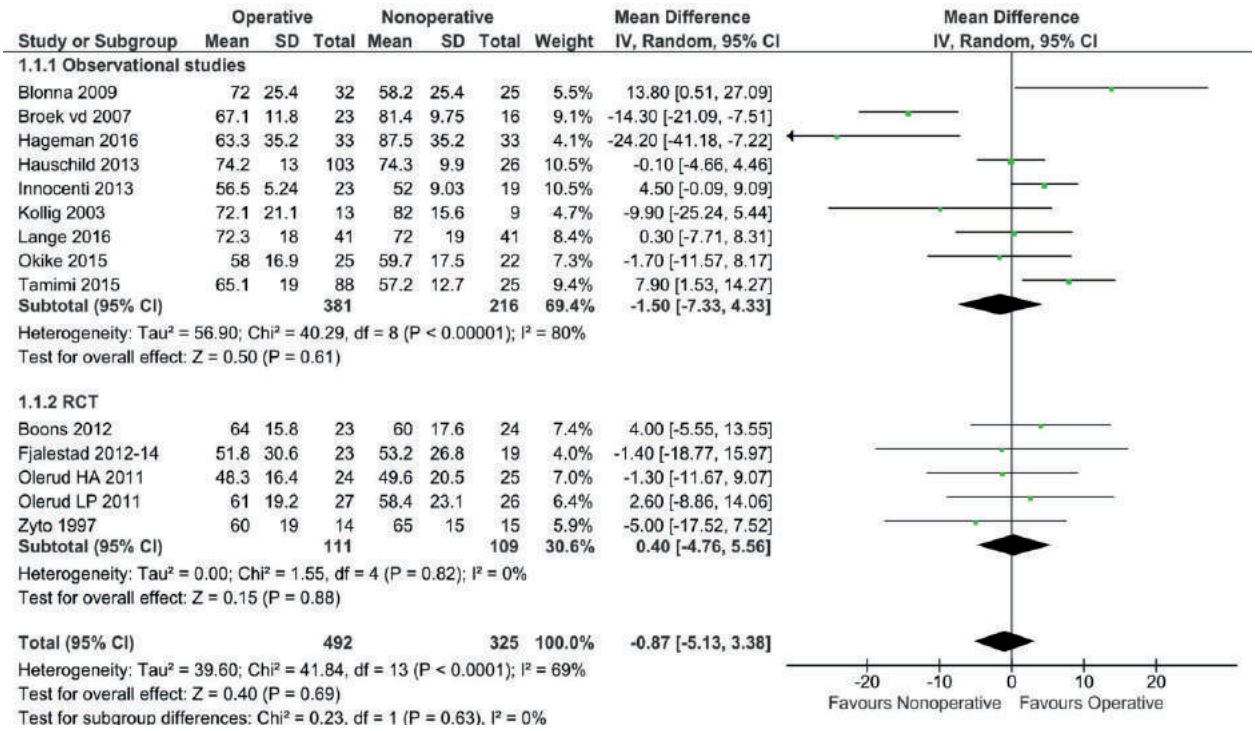


Figure 2. Functional outcome as measured with the Constant-Murley score in a systematic review of proximal humerus fractures comparing operative to nonoperative treatment

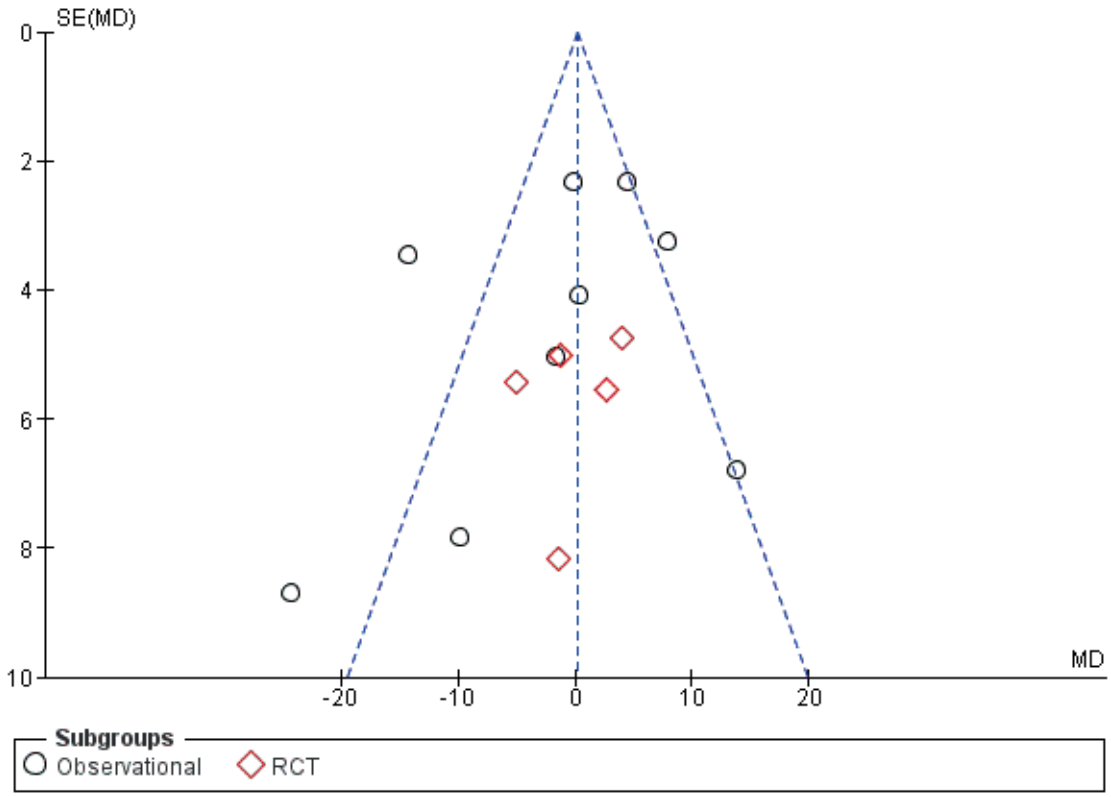


Figure 3. Funnel plot of studies included in a meta-analysis reporting Constant-Murley scores after operative or nonoperative treatment of proximal humerus fractures. (MD mean difference; SE standard error)

Major reinterventions

Fifteen studies (68%, n=938) reported on major reinterventions (Appendix 4).[13,14,34–36,39–41,23,25,27–30,32,33] Two studies had no major reintervention in either treatment arm 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%)(Appendix 7). Utilizing different methods of incorporating studies in the meta-analysis with zero event data in one or both arms yielded similar results (Appendix 8). 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 to 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

Thirteen studies (59%) reported on nonunion (Appendix 4). Operative treatment of proximal humerus fractures resulted in fewer nonunions compared to nonoperative treatment with a RR of 0.45 (CI, 0.23 – 0.89; P=0.02; I^2=0%)(Appendix 9). When stratified by study design, both subgroups showed a similar, non-significant, pooled effect (Table 2).

Avascular necrosis

Thirteen studies (59%) reported on avascular necrosis (Appendix 4). There was no difference in the rate of avascular necrosis between operative and nonoperative treatment for proximal humerus fractures with a RR of 1.24 (CI, 0.87 – 1.77; P=0.24; I^2=24%)(Appendix 10). When stratified by study design, observational studies showed a higher risk of avascular necrosis for the operative group compared to 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 humerus 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 to nonoperative treatment. This the largest meta-analysis in the current literature by including both RCTs and observational studies.

Compared to nonoperative treatment, there is no improved functional outcome after operative treatment for displaced proximal humerus fractures, which confirms findings from previous meta-analyses.[6,43] A recent systematic review of displaced proximal humerus 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] Demographic patient characteristics of the PROFHER trial are comparable to 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 to 21% of patients in this meta-analysis. Therefore, compared to 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 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 orthopaedic 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 humerus fractures. This is the first review to show significantly more severe complications requiring surgical re-intervention after operative treatment of displaced proximal humerus 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 humerus 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 are 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, open fractures, fractures of skeletal immature patients, and 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 feel 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 also improve generalizability of our results. We also demonstrated that findings were consistent across study designs with respect to different outcome measures. Although no subgroup analysis could be performed on elderly patients > 65 years old, the mean age of all patients in this review was 68 years old 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 humerus fractures.

Conclusions

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

Appendices

Appendix 1. Search Syntax

Date of search: March 30th, 2017

- Searchstring 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])

- Searchstring 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)

- Searchstring CENTRAL (The Cochrane Library) (n=166)

humerus AND fracture AND (proximal OR neck OR sub capital OR subcapital)

- Searchstring 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)

Appendix 2. MINROS assessment criteria

Criteria	2	1	0
A clearly stated aim	Aim or hypothesis including outcomes have been reported	Aim or hypothesis have been reported without a clear outcome	Not reported
Inclusion of consecutive patients	Explicit inclusion and exclusion criteria have been reported	Unclear or poor description inclusion and exclusion criteria have been reported	Not reported
Prospective collection of data	Prospective	Retrospective	Not reported
Endpoints appropriate to the aim of the study	Outcomes are appropriate to the aim of the study	Outcomes are not appropriate to the aim of the study	Not reported
Unbiased assessment of the study endpoint	Blind evaluation of objective outcomes and double-blind evaluation of subjective outcomes	One or more outcomes have been blinded	No blinding / not reported
Follow-up period appropriate to the aim of the study	≥ 1 year	< 1 year	Not reported
Loss to follow-up less than 5%	≤ 5%	> 5% and ≤ 20%	Not reported / >20%
Prospective calculation of the study size	Power analysis has been performed	Explanation for the number of included patients without a power analysis	Not reported / not performed
An adequate control group	Plate or intramedullary fixation compared with a conservative treatment	Not applicable	Not reported
Contemporary groups	Study group and controls have been managed during the same time period	Study group and controls have not been managed during the same time period	Not reported / unclear discription
Baseline equivalence of groups	Baseline characteristics have been described for both groups and are comparable	Baseline characteristics have not been described thoroughly or are not comparable	Not reported
Adequate statistical analyses	Statistical analysis has been described including the type of test	Inadequate statistical analysis	Not reported

Appendix 3. Quality assessment of all included studies in a systematic review of proximal humerus fractures comparing operative to nonoperative treatment

Criteria	Blonna 2009	Boons 2012	vd Broek 2007	Court-Brown	Fjalestad 2005	Fjalestad 2012-	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	Tamimi 2015	Zyto 1997
A 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
Endpoints appropriate to the aim of the study	2	2	2	2	1	2	2	2	1	2	2	2	1	2	2	2	2	2	2	1	2	2
Unbiased assessment of the study endpoint	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	0	1	0	0	1
Follow-up period appropriate to the aim of the study	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 less than 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 of the study size	0	2	0	0	0	1	0	1	0	0	0	0	0	0	1	1	2	0	0	0	0	0
An 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 analyses	2	2	0	2	2	2	2	2	1	2	0	2	1	2	2	2	2	2	2	0	2	2
Total quality score MINORS	**18**	**22**	**13**	**18**	**14**	**22**	**17**	**21**	**14**	**19**	**12**	**16**	**13**	**16**	**21**	**21**	**22**	**16**	**18**	**16**	**14**	**21**

Appendix 4. Outcome measures in a systematic review of proximal humerus fractures comparing operative to 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				
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		

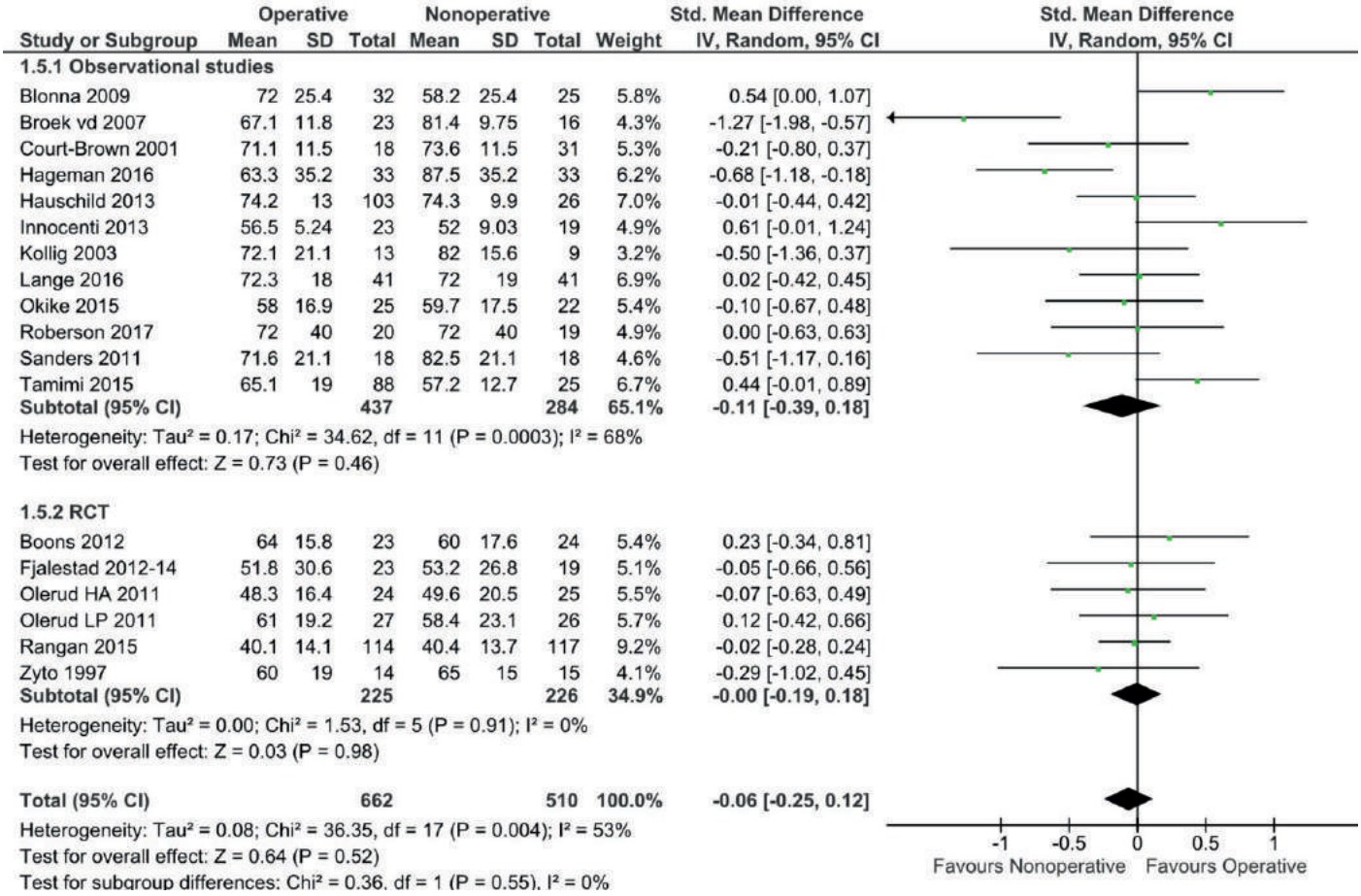
(appendix 4 continued)

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			

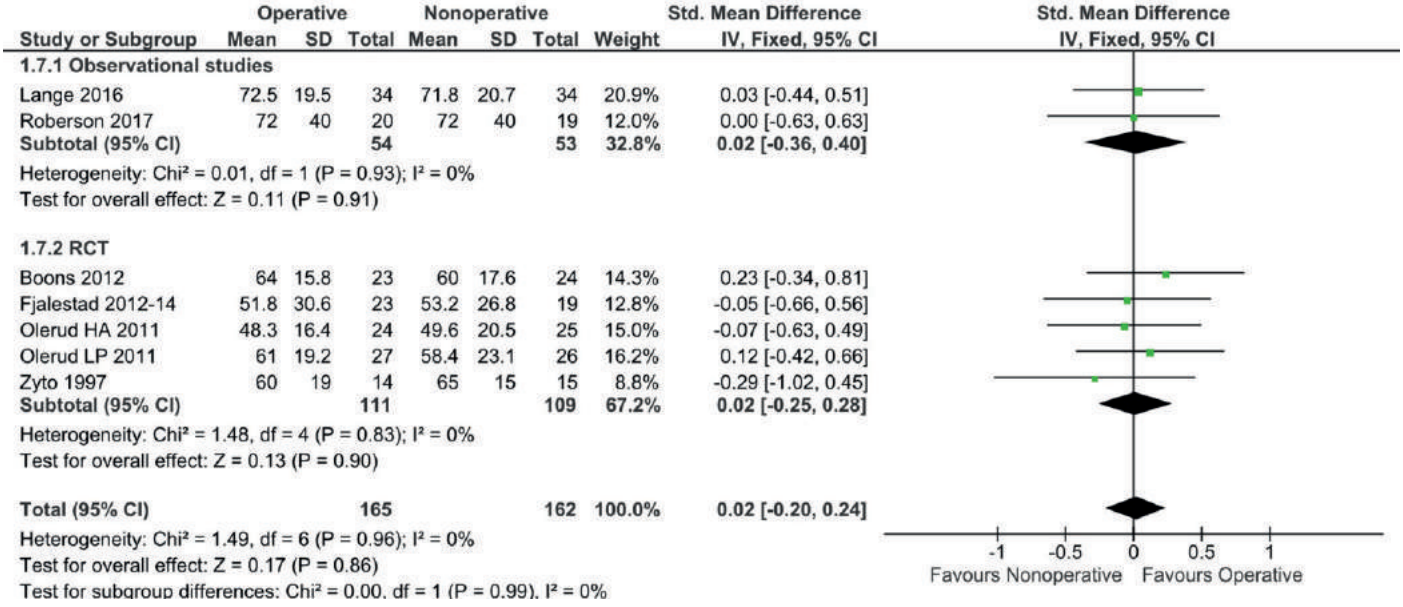
*In this analysis Fjalestad 2012 and 2014 were seen as one study as both studies describe the same patient cohort

AVN avascular necrosis; NR not reported; SD standard deviation

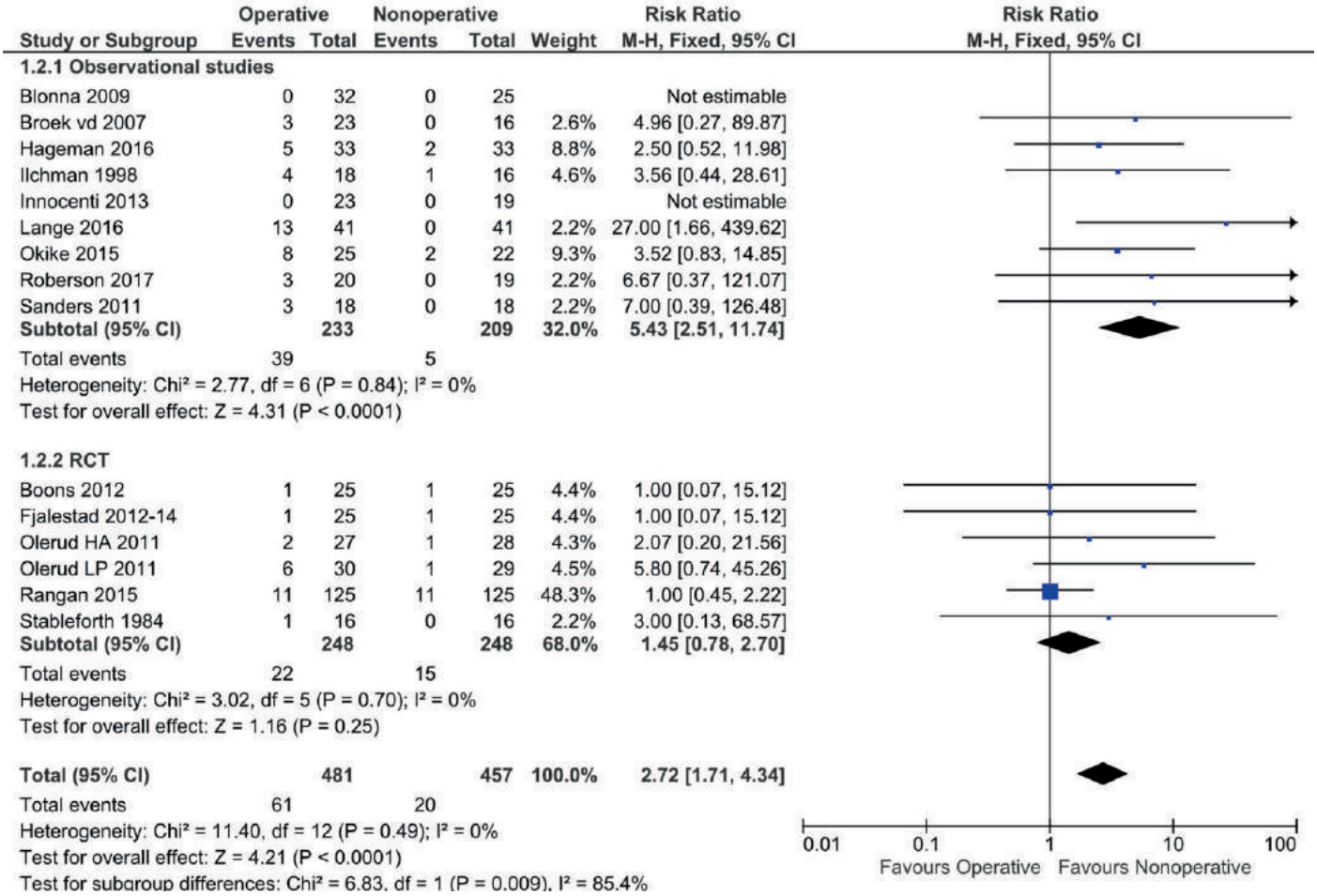
Appendix 5. Standardized mean difference of functional outcome scores in a systematic review of proximal humerus fractures comparing operative to nonoperative treatment.



Appendix 6. 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 to nonoperative treatment.



Appendix 7. Revision surgery in a systematic review of proximal humerus fractures comparing operative to nonoperative treatment



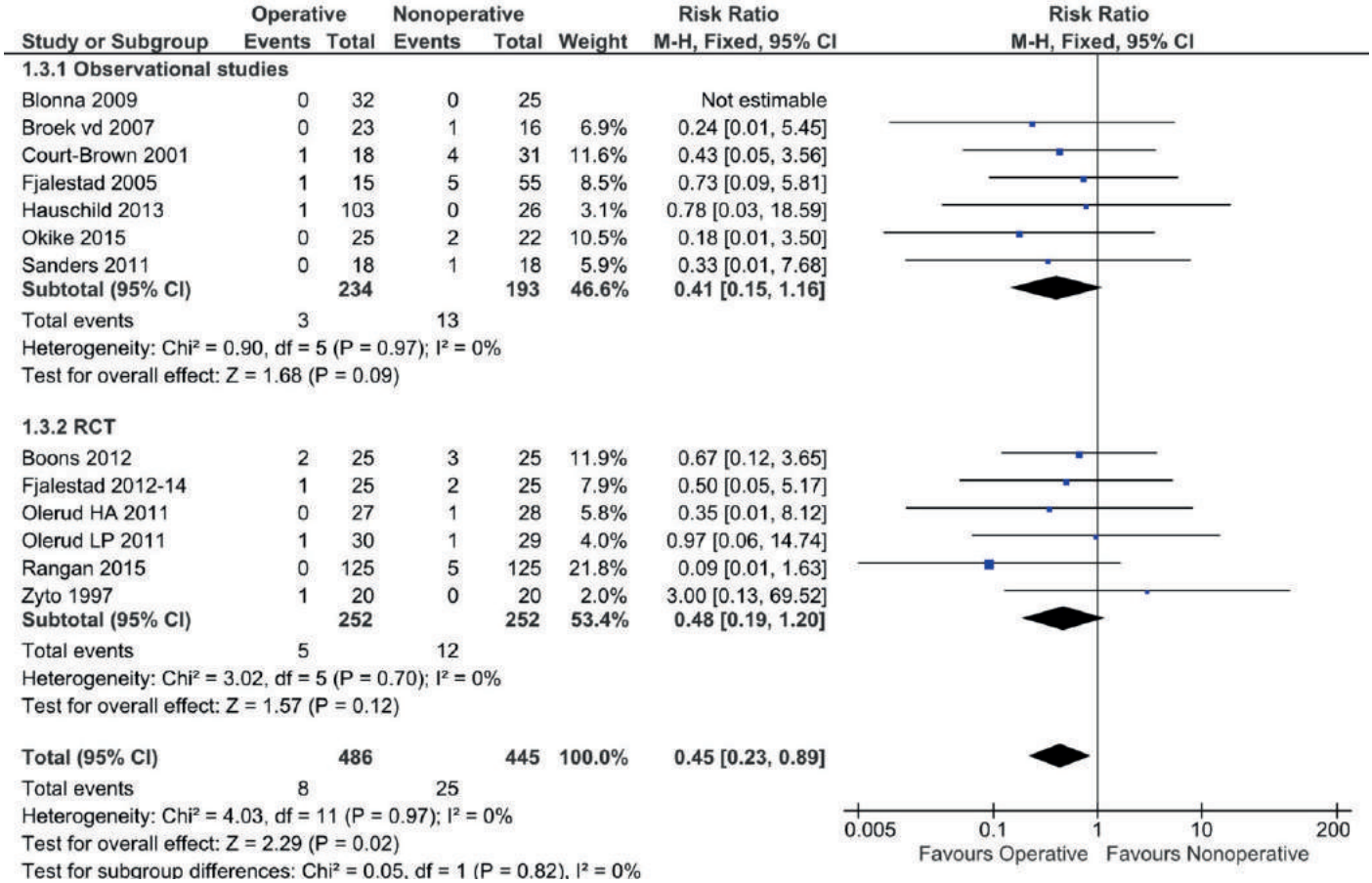
Appendix 8. Impact of different methods to handle zero-event data in a meta-analysis of operative versus nonoperative treatment of proximal humerus fractures and major reintervention

Method	Observational studies OR (95% CI)	RCT OR (95% CI)	Total 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)

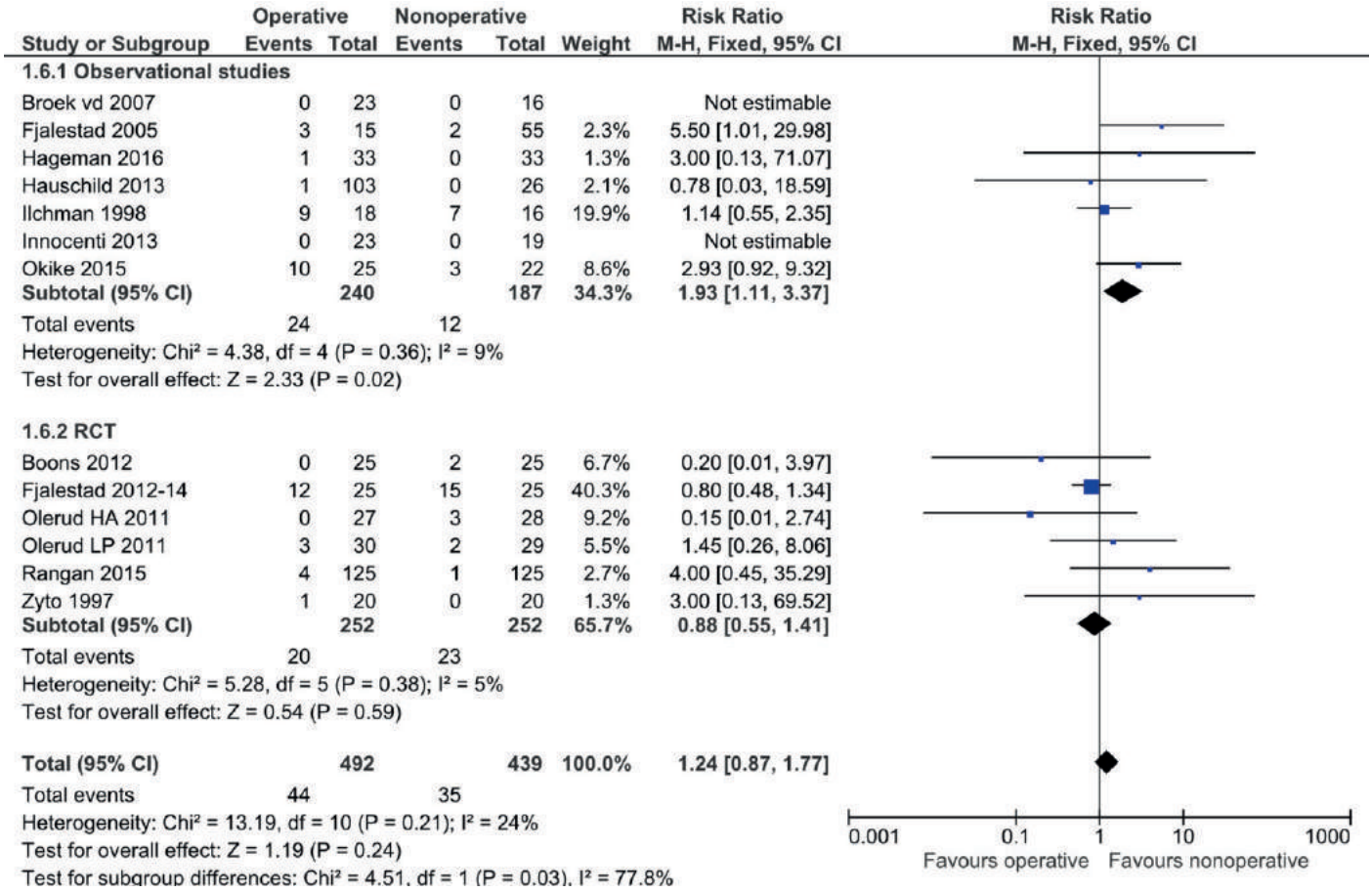
* Method used in meta-analysis; OR odds-ratio; CI confidence interval

In a model with correction 0.5 is added to every table of the 2x2 table

Appendix 9. Nonunion in a systematic review of proximal humerus fractures comparing operative to nonoperative treatment



Appendix 10. Avascular necrosis in a systematic review of proximal humerus fractures comparing operative to nonoperative treatment



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. doi:10.1016/j.otsr.2012.05.013.

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. doi:10.1007/s11657-015-0209-4.

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. doi:10.2106/JBJS.RVW.O.00003.

6. Handoll HH, Brorson S. Interventions for treating proximal humeral fractures in adults. *Cochrane Database Syst Rev*. 2015;(11):CD000434. doi:10.1002/14651858.CD000434.pub4.

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. doi:10.1177/0363546516673615.

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. doi:10.1016/j.jse.2016.01.018.

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. doi:10.1016/j.jclinepi.2009.04.005.

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. doi:10.1002/14651858.MR000034.pub2.

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. doi:10.1371/journal.pmed.1000100.

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. doi:10.1001/jama.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. doi:10.1097/BOT.0b013e31821c2e15.

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. doi:10.1007/s00590-013-1403-y.

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. doi:10.1016/j.jse.2007.06.022.

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. doi:10.1067/mse.2002.127096.

19. Neer CS. Displaced Proximal Humeral Fractures. *J Bone Joint Surg Am*. 1970;52(6):1077-1089. doi:00003086-200601000-00014 [pii].

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. doi:10.1177/0363546516673615.

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. doi:10.1002/sim.2528.

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. doi:10.1007/s40520-013-0063-4.

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. doi:10.4103/2277-9175.146385.

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. doi:10.1016/j.injury.2015.05.049.

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). doi:10.5812/atr.37423.

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. doi:10.1016/s0020-1383(16)30848-8.

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. doi:10.1016/j.jse.2016.10.013.

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. doi:10.1302/0301-620X.79B3.7419.

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. doi:10.1016/j.jse.2011.04.016.

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. doi:10.1016/j.jse.2010.12.018.

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. doi:10.1007/s11999-012-2531-0.

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. doi:10.1001/jama.2015.1629.

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. doi:10.1055/s-2003-37764.

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. doi:10.1007/s00402-005-0803-9.

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. doi:10.1007/s00402-006-0250-2.

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. doi:10.1016/j.jse.2009.02.004.

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. doi:10.1016/j.jse.2011.01.025.

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. doi:10.1007/s00402-013-1798-2.

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. doi:10.1186/s40064-015-1522-5.

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. doi:10.1007/s00068-016-0720-3.

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. doi:10.1056/NEJMra1614394.

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. doi:10.3109/09638288.2013.829529.

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. doi:10.1007/s11999-012-2454-9.

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. doi:10.5312/wjo.v6.i10.838.

CHAPTER 8

FIXATION OF FLAIL CHEST OR MULTIPLE RIB FRACTURES: CURRENT EVIDENCE AND HOW TO PROCEED. A SYSTEMATIC REVIEW AND META-ANALYSIS.

R.B. Beks, J. Peek, M.B. de Jong, K.J.P. van Wessem,
C.F. Öner, F. Hietbrink, L.P.H. Leenen,
R.H.H. Groenwold, R.M. Houwert

European Journal of Trauma and Emergency Surgery 2018

Abstract

Purpose

The aim of this systematic review and meta-analysis was to present current evidence on rib fixation and to compare effect estimates obtained from randomized controlled trials (RCTs) and observational studies.

Methods

MEDLINE, Embase, CENTRAL, and CINAHL were searched on June 16th 2017 for both RCTs and observational studies comparing rib fixation versus nonoperative treatment. The MINORS criteria were used to assess study quality. Where possible, data were pooled using random effects meta-analysis. The primary outcome measure was mortality. Secondary outcome measures were hospital length of stay (HLOS), intensive care unit length of stay (ILOS), duration of mechanical ventilation (DMV), pneumonia, and tracheostomy.

Results

Thirty-three studies were included resulting in 5874 patients with flail chest or multiple rib fractures: 1255 received rib fixation and 4619 nonoperative treatment. Rib fixation for flail chest reduced mortality compared to nonoperative treatment with a risk ratio of 0.41 (95CI 0.27, 0.61, $p<0.001$, $I2=0\%$). Furthermore, rib fixation resulted in a shorter ILOS, DMV, lower pneumonia rate, and need for tracheostomy. Results from recent studies showed lower mortality and shorter DMV after rib fixation but there were no significant differences for the other outcome measures. There was insufficient data to perform meta-analyses on rib fixation for multiple rib fractures. Pooled results from RCTs and observational studies were similar for all outcome measures, although results from RCTs showed a larger treatment effect for HLOS, ILOS, and DMV compared to observational studies.

Conclusions

Rib fixation for flail chest improves short-term outcome, although the indication and patient subgroup who would benefit most remain unclear. There is insufficient data regarding treatment for multiple rib fractures. Observational studies show similar results compared with RCTs.

Introduction

Rib fractures are very common in patients with thoracic trauma and nowadays still associated with significant morbidity and mortality due to the underlying injuries to the lung and heart resulting in more pulmonary complications.[1–4] Compared to multiple rib fractures, flail chest is associated with a worse outcome due to a higher incidence of respiratory compromise and concomitant injuries.[5,6]

A combination of adequate pain control, respiratory assistance, and physiotherapy is considered the gold standard in management of rib fractures.[3] Over the past decades, there has been a growing interest in rib fixation for flail chest and for multiple rib fractures, however, there is no consensus regarding the indication and patient selection for rib fixation.

In the field of (orthopaedic) trauma surgery, there is increasing scientific evidence that inclusion of observational studies could add value to meta-analyses without decreasing quality of the results.[7–10] Adding observational studies result in larger sample sizes and might enable the evaluation of small treatment effects, subgroups, and infrequent outcome measures while also providing information about the generalizability of the results.[11]

The aim of this systematic review and meta-analysis was (1) to present current evidence on outcome after rib fixation compared to nonoperative treatment for both flail chest and multiple rib fractures and (2) to compare effect estimates obtained from RCTs and observational studies.

Methods

This review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines.[12,13] A published protocol for this review does not exist. Ethical committee approval did not apply to this study.

Search Strategy and Eligibility Criteria

A structured literature search was conducted in MEDLINE, Embase, CENTRAL and CINAHL on June 16th, 2017 for both randomized controlled trials (RCTs) and observational studies comparing operative to nonoperative treatment of traumatic rib fractures. The search was not restricted by publication date, language, or other limits. The full search syntax is provided in Appendix 1.

All obtained studies from the literature search were independently screened for eligibility based on title and abstract by two reviewers (RBB, JP). Exclusion criteria were animal studies, abstracts of conferences, case-reports, reviews, inclusion of patients younger than 18 years

old, and studies written in another language than English, French, Dutch or German. Disagreement regarding study selection was resolved by discussion with a third reviewer (RMH). References of included studies were manually screened and citation tracking was conducted using Web of Science to identify additional relevant studies.

Data Extraction

Data were extracted by two independent reviewers (RBB, JP), using a data extraction file. Extracted data included first author, year of publication, study period, study design, country, fracture type, number of fractured ribs, number of included patients, number of patients with flail chest or multiple rib fractures (according to the definition used by the original study), age, gender, type of operative treatment, type of nonoperative treatment, duration of follow-up, loss to follow-up, Injury Severity Score (ISS), Abbreviated Injury Scale (AIS), Glasgow Coma Scale (GCS), hemothorax, pneumothorax, pulmonary contusion, type of implant in operative group, mortality during hospitalization, hospital length of stay (HLOS), intensive care unit length of stay (ILOS), duration of mechanical ventilation (DMV), incidence of pneumonia, need for tracheostomy, complications, revision surgery, and implant removal.

Outcome Measures

The primary outcome measure was mortality during hospitalization. Secondary outcome measures were HLOS, ILOS, DMV, incidence of pneumonia, need for tracheostomy, complications, revision surgery, and implant removal.

Quality Assessment

The Methodological Index for Non-Randomized Studies (MINORS) score was used to assess the included studies.[14] The MINORS is a critical appraisal instrument developed to assess the methodological quality of observational surgical studies. Other quality assessment tools focus on a specific study design while the MINORS is externally validated on RCTs and is therefore a suitable instrument for meta-analyses of different study designs. The MINORS score ranges from 0 – 24 and a higher score reflects better quality. Studies were independently assessed by two reviewers (RBB, JP) using the MINORS criteria and disagreement was resolved by discussion with a third reviewer (RMH). Additional details on the MINORS criteria and scoring system are set out in Appendix 2.

Statistical Analysis

Statistical analyses were performed using Review Manager (RevMan, Version 5.3.5 Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). Data were converted to a mean with standard deviation (SD) using different methods as described in the Cochrane Handbook for Systematic Reviews of Interventions.[15]

Different studies based on the same patient cohort were included only once in the analysis.[16,17] Studies reporting on specific patient subgroups were split and included

separately for meta-analysis, provided sufficient information was reported; Qiu et al. distinguished between the presence or absence of a flail chest and Voggenreiter et al. made subgroups based on the presence or absence of pulmonary contusion.[18,19] Results from both RCTs and observational studies were pooled in the primary analysis.

Meta-analysis was performed if outcome measures of two or more studies were available. For continuous outcome measures, the inverse variance weighted random effects model was used to estimate the pooled difference in the outcome measure for fixation vs. no fixation, with corresponding 95% confidence interval (CI). For dichotomous outcomes, we applied the Mantel-Haenszel method and pooled results are presented as risk ratios (RR) with 95% CI. Heterogeneity between studies was assessed by visual inspection of the forest plots and by estimating statistical measure for heterogeneity, i.e., the I^2 statistic. Inspection of a funnel plot of the study-specific difference in the primary outcome measure against its standard error was done to detect potential publication bias. A two-sided p-value < 0.05 was considered statistically significant.

Subgroup and Sensitivity Analyses

In subgroup analysis, we stratified by study design and pooled effects of RCTs were compared with pooled effects of observational studies. 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 orthopaedic trauma surgery studying both study designs.[8,10,20] To assess the impact of improvement in intensive care management over time, we performed a sensitivity analysis including only studies published in the last 5 years. Different methods were used to include studies with zero events in one or both arms of the outcome measure. To assess the sensitivity of the analyses to the choice of the method of analysis, also the crude methods, DerSimonian Laird method with correction, the inverse variance with and without correction for zero event data, and the Peto method were applied and results were compared for consistency.[21]

Results

Search

The flowchart of the literature search is presented in Figure 1. Ultimately, 33 studies were included [16–19,22–50]. There were three RCTs, two prospective cohort studies, 14 retrospective cohort studies, and 14 case-control studies.

Patient characteristics

The studies included for meta-analysis included 5874 patients; 1255 received rib fixation and 4619 received nonoperative treatment. In the majority of the studies (n=20), patients were surgically treated with plates (Table 1&2). Other surgical methods were K-wires and Judet or Atkin struts. Nonoperative treatment consisted generally of 'best medical treatment' and

included adequate pain management, lung physiotherapy and respiratory support. The weighted average age was 52.9 years and 73% of patients were male. The weighted average of the number of rib fractures was 6.9 in the rib fixation group and 6.0 in the nonoperative group with a weighted mean ISS of 21.2 and 22.4, respectively.

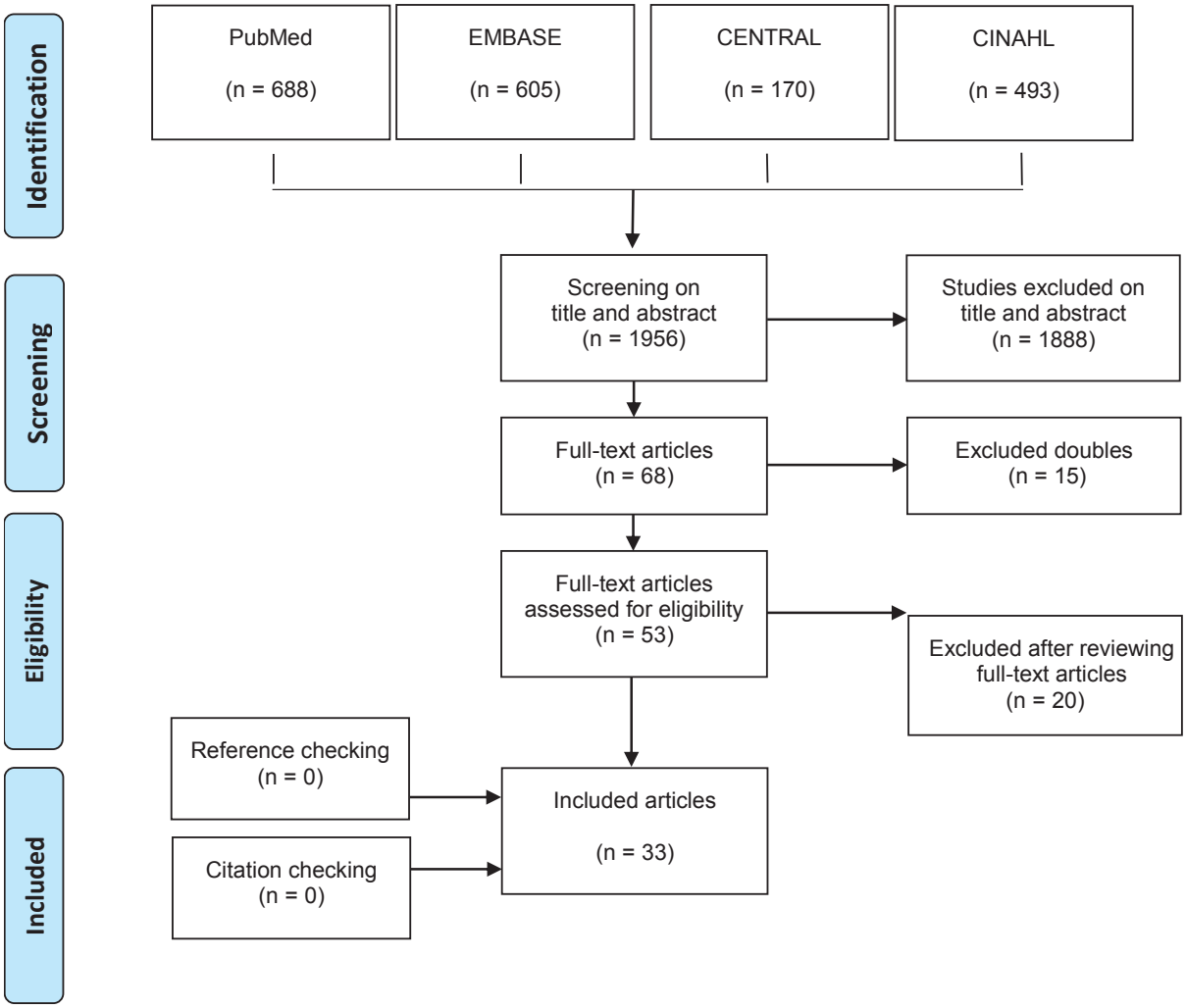


Fig. 1 Flow chart of the literature search

Quality assessment

The average MINORS score of the included studies was 15.4 (SD 2.7; range 9 – 21). The MINORS score for RCTs was 20 (SD 1.0; range 19 – 21) and for observational studies 14.9 (SD 2.4; range 9 – 21). An overview of the study specific MINORS score are provided in Appendix 3.

Mortality

Twenty-five studies (n=4826) reported on mortality (Appendix 4).[18,19,34,36–44,22,45–50,23,25,27,28,30,32,33] Rib fixation resulted in a significant reduction of mortality compared to

nonoperative treatment with a risk ratio (RR) of 0.41 (95%CI 0.27, 0.61, $p < 0.001$, $I^2 = 0\%$) (Figure 2). Different methods of incorporating studies in the meta-analysis with zero event data in one or both arms yielded similar results (Appendix 5). When stratified by study design, RCTs showed a RR 0.57 (95%CI 0.13, 2.52, $p = 0.46$, I^2=0%) vs. RR 0.40 (95%CI 0.26, 0.60, $p < 0.001$, $I^2 = 0\%$) in observational studies (Table 3). Figure 3 shows a funnel plot of the odds ratio and standard error of the included studies using the mortality rate; there was no important asymmetry observed.

Hospital stay length of stay

Twenty-one studies (n=4770) reported on length of hospital stay (Appendix 4).[16,17,37–45,47,23,50,51,25,26,31–35] Rib fixation did not result in a significant reduction of HLOS compared to nonoperative treatment with a mean difference of -1.46 days (95%CI -4.31, 1.39, $p = 0.32$, $I^2 = 96\%$) (Appendix 6). When stratified by study design, the pooled mean difference of RCTs (-8.33 days; 95%CI -14.6, -2.1; $p < 0.001$, $I^2 = 46\%$) was greater compared to observational studies (-0.77; 95%CI -3.72, 2.18; $p = 0.61$, $I^2 = 97\%$) (Table 3).

ICU length of stay

Twenty-six studies (n=4520) reported on length of ICU stay (Appendix 4).[16,17,31–33,35–41,18,42–44,47,50,51,22–26,28,30] Rib fixation resulted in a significant reduction of ILOS compared to nonoperative treatment with a mean difference of -2.0 (95%CI -3.61, -0.38, $p = 0.02$, $I^2 = 85\%$) (Appendix 7). When stratified by study design, RCTs showed a greater difference compared to observational studies (Table 3).

Duration of mechanical ventilation

Twenty-seven studies (n=2063) reported on duration of mechanical ventilation (Appendix 4).[16,17,28,30–32,35–40,18,41,42,45–47,49–51,19,22–27] Rib fixation resulted in a significant reduction of days on mechanical ventilation compared to nonoperative treatment with a mean difference of -4.01 (95%CI -5.58, -2.45, $p < 0.001$, $I^2 = 91\%$) (Appendix 8). When stratified by study design, RCTs showed a greater difference compared to observational studies (Table 3).

Pneumonia

Twenty-five studies (n=4485) reported on the incidence of pneumonia (Appendix 4).[16,17,31–33,37–39,41–44,18,47,50,51,19,22,24–26,28,30] Rib fixation resulted in a significant reduction of pneumonia compared to nonoperative treatment with a risk ratio of 0.59 (95%CI 0.42, 0.83, $p = 0.002$, $I^2 = 79\%$) (Appendix 9). When stratified by study design both subgroups showed similar results (Table 3).

Tracheostomy

Fourteen studies (n=1541) reported on the need of tracheostomy (Appendix 4).[16,17,36–38,45,50,18,22,25,26,28,30,32,34] Rib fixation resulted in a significant reduction of tracheostomies compared to nonoperative treatment with a risk ratio of 0.59 (95%CI 0.36, 0.90, p = 0.01, I^2 = 72%) (Appendix 10). When stratified by study design both subgroups showed similar results (Table 3).

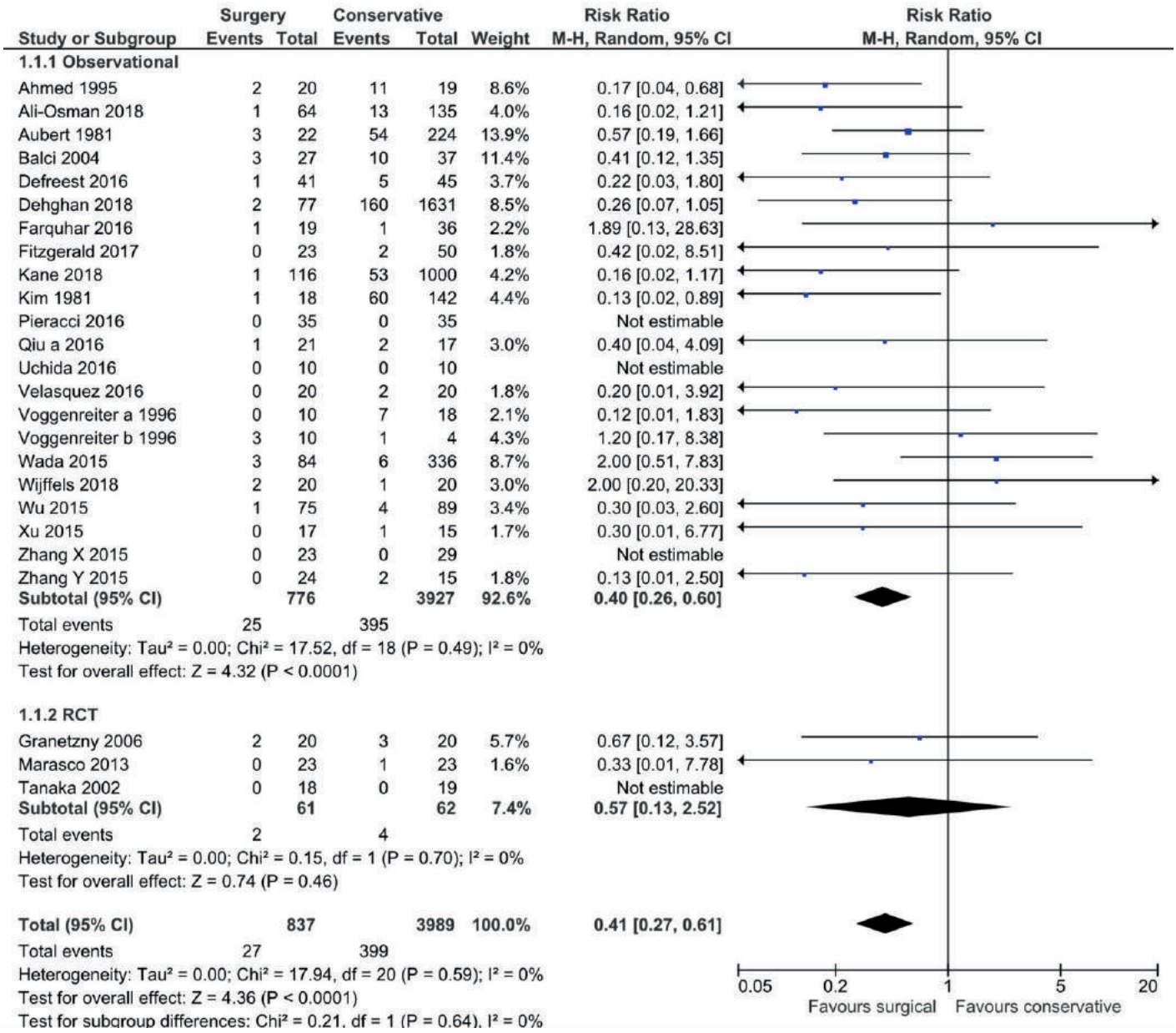


Fig. 2 Mortality in a systematic review of rib fractures comparing operative to nonoperative treatment

Other outcome measures

Nine studies (n=1174) reported on implant removal; five studies reported zero events and four studies reported implant removal ranging from 1.5% to 4.9% (Appendix 4).[17,26,28,36–38,40,45,48] Eleven studies reported on wound infection; five studies reported zero events and six studies reported a wound infection rate ranging from 1.7% to 25%.[18,23,24,26–30,46] Other short and/or long term complications were poorly reported and described mainly respiratory complications.

Sensitivity analyses

In sensitivity analysis for study quality, results did not change significantly except for HLOS which increased in favor of rib fixation in studies with higher quality with a mean difference of -3.53 (95%CI -7.27, -0.21, *p* =0.06) (Table 3). Results from studies published after 2012 did not show a reduced HLOS, ILOS, incidence of pneumonia or need for tracheostomy after rib fixation (Table 3).

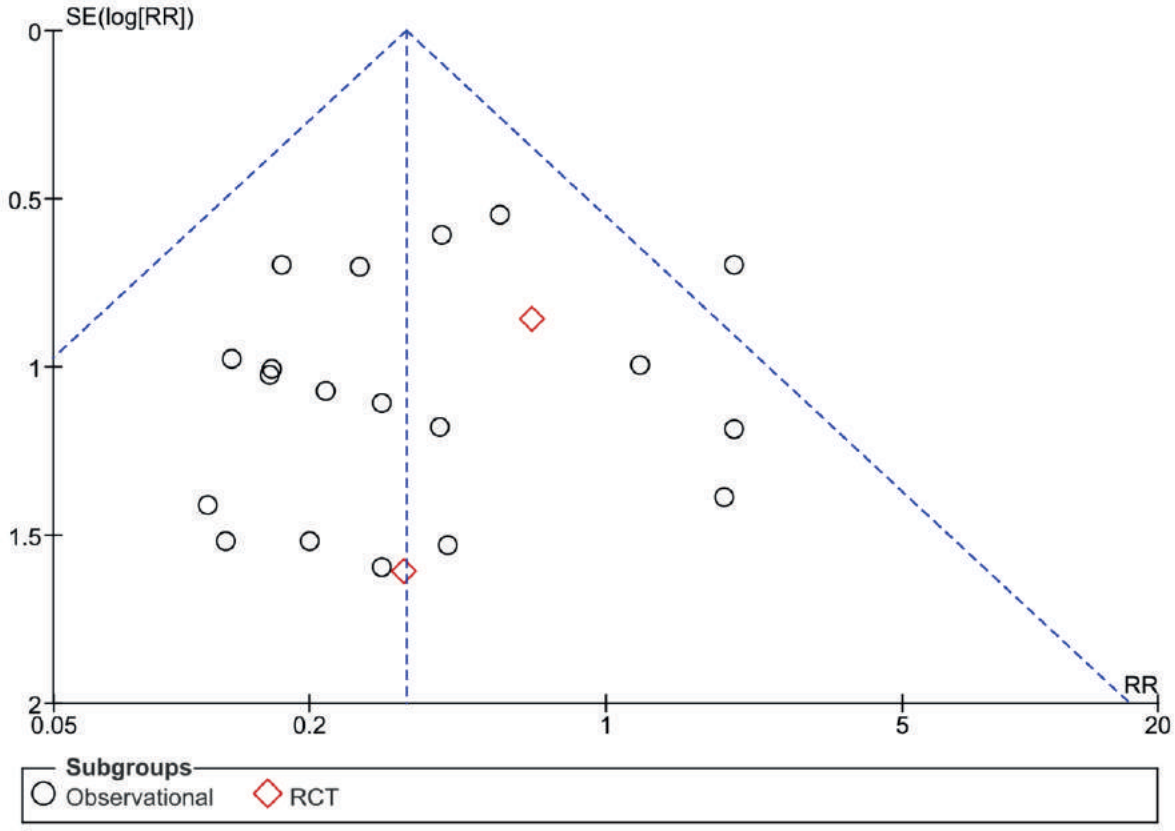


Fig. 3 Funnel plot of studies included in a meta-analysis reporting mortality rates after operative or nonoperative treatment of rib fractures. (RR risk ratio; SE standard error)

Table 1. Baseline characteristics of the included studies comparing rib fixation versus nonoperative treatment of traumatic rib fractures.

Study	Study design	Country		Number of patients	Follow-up (months)	Age (years, range or ± SD)	Male (%)	Number of fractured ribs	ISS-score
Dehghan 2018	RC	Canada	RF: NOM:	77 1631	NR	52 ± 18 58 ± 18	55 (76) 1176 (72)	NR	NR
Ali-Osman 2018	RC	USA	RF: NOM:	64 135	NR	68.5 [63-74] 72 [66-81]	41 (64) 73 (54)	7 [5.25-9] 5 [3-7.25]	17.5 [9-25} 14 [8-24]
Wijffels 2018	CC		RF: NOM:	20 20	NR	60 [41-69] 57 [44-69]	15 (75) 15 (75)	9 [8-11] 10 [9-14]	31 [21-48] 32 [21-41]
Kane 2018	RC		RF: NOM:	116 1000	NR	58.3 ± 14.4 46.9 ± 29.3	NR	NR	21.6 (9.1) 16.1 (11.4)
Fitzgerald 2017	CC	USA	RF: NOM:	23 50	NR	68 (63–89) 75 (65–97)	NR	NR	21 (16-26) 19 (14-23)
Farquhar 2016	CC	Canada	RF: NOM:	19 36	21.9 ± 13.2 16.0 ± 12.1	53 ± 14 57 ± 16	15 (79) 25 (69)	NR	31.4 ± 9.6 29.3 ± 8.1
Pieracci 2016	PC	USA	NOM:	35	16.0 [10.0, 23.0]	50 + 15	24 (69)	9.0 [6.0, 13.0]	22.0 [17.0,38.0]
Defreest 2016	RC	USA	RF: NOM:	41 45	28.3 (9–69) 13.0 (3–43)	51 (19–80) 56 (23–89)	32 (78) 39 (87)	11.2 (6-19) 10.6 (6-23)	27.5 (16-48) 29.3 (16-66)
Uchida 2016	CC	Japan	RF: NOM:	10 10	NR	63 [51,72] 57 [53,75]	7 (70) 7 (70)	5 [4, 8] 5 [2, 7]	NR
Velasquez 2016	CC	USA	RF: NOM:	20 20	6 [4,10] 16 [11,22]	51 [41,63] 45 [36,55]	NR	5 [4, 8] 5 [4.6, 5]	9 [9,16] 13 [9,17]
Qiu a 2016	RC	China	RF: NOM:	21 17	NR	35 ± 13 36 ± 14	15 (48) 12 (71)	6.0 ± 1.3 5.9 ± 1.3	NR
Qiu b 2016	RC	China	RF: NOM:	65 59	NR	38 ± 12 36 ± 12	46 (71) 42 (71)	3.2 ± 1.2 3.5 ± 1.2	NR
Jayle 2015	CC	France	RF: NOM:	10 10	21.7 ± 7.8 32.3 ± 19.3	48 ± 11 51 ± 13	8 (80) 8 (80)	7.7 ± 2.4 6.6 ± 2.9	21.7 ± 7.80 32.3 ± 19.3
Zhang Y 2015	RC	China	RF: NOM:	24 15	38 [33, 54.25] 60 [38, 99.75]	43 [34,50] 47 [35,55]	19 (79) 14 (93)	11.5 [8, 15.3] 11 [7, 16]	38 [34,43] 38 [35,43]
Zhang X 2015	CC	China	RF: NOM:	23 29	419,4 ± 107.1 419,4 ± 107.1	58 ± 12 60 ± 10	21 (72) 16 (70)	7.8 ± 1.5 7.4 ± 1.7	NR
Wada 2015	CC	Japan	RF: NOM:	84 336	33 (24-45) 42 (23-58)	NR	59 (70) 225 (76)	NR	NR
Wu 2015	PC	China	RF: NOM:	75 89	15.3 ± 6.4 26.5 ± 6.9	52 ± 5 51 ± 3	75 (100) 89 (100)	8.1 (6-12) 7.9 (6-11)	NR
Majercik 2015	CC	USA	RF: NOM:	137 274	11.4 + 5.7 12.3 + 9.1	56 ± 16 55 ± 20	110 (80) 56 (80)	6.5 ± 2.0 4.6 ± 2.3	21 ± 10.7 22 ± 11.8
Xu 2015	RC	China	RF: NOM:	17 15	NR	36 ± 14 39 ± 12	12 (71) 12 (80)	6.8 ± 2.1 7.4 ± 1.6	21.8 ± 7.8 24.0 ± 8.0
Granhed 2014	CC	Sweden	RF: NOM:	60 153	NR	NR NR	53 (77) NR	7.5 (2-14) NR	21.7 ± 10.7 30.9 ± 13.3

(Table 1 continued)

Doben 2014	CC	USA	RF: NOM:	10 11	21.6 (8-59) 28.5 (6-50)	47 ± 15 57 ± 17	9 (90) 7 (64)	8.3 (4-20) 9.2 (6-16)	26.3 ± 9.5 35.7 ± 12.7
Marasco 2013	RCT	Australia	RF: NOM:	23 23	90 90	58 ± 17 59 ± 10	20 (87) 20 (87)	11.0 ± 3.1 11.3 ± 4.7	35.0 ± 11.4 30.0 ± 6.3
Khandelwal 2011	PC	India	RF: NOM:	31 29	30 30	47 45	40 (66) = total group	3.1 3.3	NR
Moya 2011	CC	USA	RF: NOM:	16 32	18 ± 12 16 ± 11	45 ± 16 47 ± 14	14 (88) 26 (81)	8 ± 4 8 ± 3	24 ± 7 25 ± 9
Althausen 2011	CC	USA	RF: NOM:	22 28	17.84 ± 4.51 NR	48 51	17 (74) 23 (79)	5.9 7.3	25.1 24.3
Solberg 2009	RC	USA	RF: NOM:	9 7	16.1 ± 6.7 12.0 ± 2.3	39 ± 17 41 ± 13	6 (67) 5 (71)	NR	24.9 ± 6.5 24.8 ± 6.2
Nirula 2006	CC	USA	RF: NOM:	30 30	NR	52 50	NR	NR	25.7 27.5
Granetzny 2006	RCT	Germany	RF: NOM:	20 20	2 2	41 ± 8 36 ± 15	17 (85) 16 (80)	4.4 4.0	16.8 ± 3.5 18.0 ± 5,1
Balci 2004	RC	Turkey	RF: NOM:	27 37	NR	35 ± 8 31 ± 10	20 (74) 28 (76)	NR	21.0 ± 7.4 18.4 ± 8.1
Tanaka 2002	RCT	Japan	RF: NOM:	18 19	360 360	43 ± 12 46 ± 9	12 (67) 14 (74)	8,2 ± 3.3 8.2 ± 2.6	33 ± 11 30 ± 8
Voggenreiter a 1996	RC	Germany	RF: NOM:	10 18	NR	55 ± 8 44 ± 19	NR	NR	31.0 ± 7.0 36.6 ± 12.3
Voggenreiter b 1996	RC	Germany	RF: NOM:	10 4	NR	50 ± 16 48 ± 27	NR	NR	37.0 ± 7.9 37.8 ± 19.5
Ahmed 1995	RC	United Arab Emirates	RF: NOM:	26 38	(3-9) (3-9)	20-60 (range) 10-60 (range)	23 (88) 36 (95)	NR	NR NR
Kim 1981	RC	France	RF: NOM:	18 142	NR	NR	NR	NR	NR
Aubert 1981	RC	France	NOM:	224	NR	NR	NR	NR	NR

CC case control; PC prospective cohort; RC retrospective cohort; RCT randomized controlled trial; RF rib fixation; NOM nonoperative treatment; NR not reported

Table 2. Treatment characteristics of the included studies comparing operative versus nonoperative management of traumatic rib fractures.

Study	Treatment groups	Included fractures	Flail chest in surgery group n (%)	Indication for surgery
Dehghan 2018	NR	FC	77 (100%)	NR
Ali-Osman 2018	RF: Plates + screws NOM: aggressive pain management	FC + MRF	NR	displaced rib fractures, uncontrolled pain, rib crepitus with breathing
Wijffels 2018	RF: Plates + intramedullary nails NOM: supportive management	FC	20 (100%)	flail chest
Kane 2018	RF: NR NOM: aggressive multimodal analgesia protoco	FC + MRF	75 (65%)	3 consecutively displaced rib fractures plus FEV1 and FVC less than 50% predicted
Fitzgerald 2017	RF: Plates + screws NOM: NR	FC + MRF	NR	NR
Farquhar 2016	RF: Plates + screws NOM: Standard conservative treatment	FC	19 (100%)	FC (≥ 3 fractures), displaced, segmental rib fractures with respiratory insufficiency
Pieracci 2016	RF: Titanium plates + screws NOM: Standard conservative treatment	FC + MRF	28 (80%)	FC (≥ 3 fractures), ≥ 3 displaced fractures; ≥ 30% thorax volume loss, failure treatment within first 72h
Defreest 2016	RF: Titanium locking plates + screws NOM: NR	FC	41 (100%)	failure to wean, intractable pain, or respiratory failure
Uchida 2016	RF: Titanium plates + locking screws NOM: Conservative management + chest strap	FC + MRF	NR	flail segment, massive dislocation, >15mm fracture overlapping, or pain
Velasquez 2016	RF: Thoracic Osteosynthesis System (STRATOS) NOM: NR	FC + MRF	NR	FC (≥ 3), ≥ 3 ribs fractured + respiratory failure, intractable pain, thorax deformity, or displacement
Qiu a 2016	RF: AO standard plates + cancellous screws NOM: NR	FC	21 (100%)	NR
Qiu b 2016	RF: AO standard plates + cancellous screws NOM: NR	MRF	0 (0%)	NR
Jayle 2015	RF: Titanium plates + screws NOM: NR	FC	10 (100%)	FC (≥ 3 fractures)
Zhang Y 2015	RF: ORIF NOM: NR	FC with PC	24 (100%)	NR
Zhang X 2015	RF: Claw-type titanium plates NOM: Standard conservative treatment	FC	23 (100%)	FC (≥ 3 fractures)
Wada 2015	RF: ORIF NOM: NR	FC + MRF	84 (100%)	NR
Wu 2015	RF: Nickel titanium alloy devices NOM: Conservative management + chest strap	FC + MRF	31 (41%)	FC (≥ 3 fractures), ≥ 3 rib fractures, dislocation, thorax deformity, or chest cavity active bleeding
Majercik 2015	RF: Plates + locking screws NOM: Standard conservative management	FC + MRF	101 (75%)	FC, severely displaced fractures, intractable pain, failure to wean, or combination of these

(table 2 continued)

Xu 2015	RF: Titanium locking plates NOM: Standard conservative management	FC	17 (100%)	NR
Granhed 2014	RF: Titanium plates + intramedullary splints NOM: NR	FC + MRF	56 (93%)	impaired saturation in spite of oxygen administration; intractable pain
Doben 2014	RF: Plates + intramedullary nails NOM: Standard conservative management	FC	10 (100%)	failure of nonoperative management
Marasco 2013	RF: Inion resorbable plates + bicortical screws NOM: Mechanical ventilator management	FC	23 (100%)	FC (≥ 3 fractures) and ventilator dependent without prospect of weaning within 48h
Khandelwal 2011	RF: Titanium plates + screws NOM: NR	FC + MRF	2 (5.3%)	NRS score > 7 on 10 days after trauma
Moya 2011	RF: Titanium or steel plates NOM: NR	FC + MRF	9 (56%)	intractable pain, ≥ 2 severely displaced rib fractures with pain, and respiratory failure
Althausen 2011	RF: Locking plates + locking screws NOM: NR	FC	22 (100%)	FC with displacement, failure to wean, respiratory failure, or need of thoracotomy
Solberg 2009	RF: Titanium plates NOM: Ventilatory pneumatic stabilization	FC	9 (100%)	superolateral chest wall deformity
Nirula 2006	RF: Adkin struts NOM: NR	FC + MRF	15 (50%)	FC, intractable pain, bleeding, and inability to wean
Granetzny 2006	RF: K-wires and/or stainless steel wire NOM: Strapping and packing	FC	20 (100%)	FC (≥ 3 rib fractures) with paradoxical chest wall movement
Balci 2004	RF: Suture and traction NOM: Endotracheal intubation	FC	27 (100%)	FC with paradoxical chest wall movement, respiratory failure, dyspnea, and insufficient blood gas
Tanaka 2002	RF: Judet struts NOM: Internal pneumatic stabilization	FC	18 (100%)	FC (≥ 6 fractures) with respiratory failure requiring mechanical ventilation and failure to wean
Voggenreiter a '96	RF: ASIF reconstruction plates NOM: Standard conservative management	FC without PC	10 (100%)	FC and thoracotomy for other injury, respiratory failure, paradoxical chest wall movement, or deformity
Voggenreiter b '96	RF: ASIF reconstruction plates NOM: Standard conservative management	FC with PC	10 (100%)	FC and thoracotomy for other injury, respiratory failure, paradoxical chest wall movement, severe deformity
Ahmed 1995	RF: K-wires NOM: Endotracheal intubation	FC	26 (100%)	NR
Kim 1981	RF: Judet struts NOM: Internal pneumatic stabilization	FC	18 (100%)	NR
Aubert 1981	RF: Osteosynthesis NOM: Ventilator assistance, physiotherapy	FC	22 (100%)	NR

RF rib fixation; NOM nonoperative management; NR not reported; FC flail chest; MRF multiple rib fractures; PC pulmonary contusion

Table 3. Subgroup & sensitivity analyses of studies included in a meta-analysis of rib fractures comparing rib fixation versus nonoperative treatment for patients with a flail chest

Analysis description	Mortality			HLOS			ILOS	
	n	RR (95% CI)	P value	n	MD (95% CI)	P value	n	MD (95% CI)	P value
All studies	25	0.41 (0.27, 0.61)	p<0.001	21	-1.46 (-4.31, 1.39)	0.32	26	-2.00 (-3.61, -0.38)	0.02
Subgroup analysis									
RCT	3	0.57 (0.13, 2.52)	0.46	2	-8.33 (-14.60, -2.07)	0.009	3	-6.37 (-9.72, -3.03)	p<0.001
Observational studies	22	0.40 (0.26, 0.60)	p<0.001	19	-0.77 (-3.72, 2.18)	0.61	23	-1.53 (-3.21, 0.15)	0.07
Sensitivity analysis									
High-quality studies	13	0.71 (0.35, 1.44)	0.34	15	-3.53 (-7.27, 0.21)	0.06	17	-2.83 (-4.75, -0.91)	0.004
Studies after 2012	17	0.43 (0.25, 0.77)	0.004	16	-0.64 (-3.98, 2.69)	0.71	19	-1.51 (-3.40, 0.37)	0.12

(table 3 continued)

Analysis description	DMV			Pneumonia			Tracheostomy		
	n	MD (95% CI)	P value	n	RR (95% CI)	P value	n	RR (95% CI)	P value
All studies	27	-4.01 (-5.58, -2.45)	p<0.001	25	0.59 (0.42, 0.83)	p<0.001	16	0.59 (0.39, 0.90)	0.01
Subgroup analysis									
RCT	3	-5.88 (-11.32, -0.44)	0.03	3	0.36 (0.15, 0.85)	0.02	2	0.38 (0.14, 1.02)	0.05
Observational studies	23	-3.79 (-5.46, -2.11)	p<0.001	22	0.63 (0.44, 0.92)	0.02	14	0.63 (0.40, 1.01)	0.05
Sensitivity analysis									
High-quality studies	17	-3.87 (-6.06, -1.68)	0.000	16	0.55 (0.37, 0.82)	0.004	10	0.57 (0.41, 0.80)	0.001
Studies after 2012	18	-3.27 (-5.11, -1.43)	0.000	16	0.73 (0.50, 1.06)	0.10	12	0.73 (0.47, 1.14)	0.16

RCT randomized controlled trial; RR risk ratio; MD mean difference; CI confidence interval; n no. of studies; RR risk ratio; MD mean difference;

Discussion

In this systematic review and meta-analysis of RCTs and observational studies, rib fixation for patients with flail chest resulted in lower mortality, shorter ILOS and DMV, lower pneumonia rate, and lower need for tracheostomy. Pooled results from RCTs and observational studies were similar for all studied outcome measures although results from RCTs showed a larger treatment effect for HLOS, ILOS, and DMV. Results from recent studies showed lower mortality and shorter DMV after rib fixation but there were no significant differences for the other outcome measures. The implant removal rate ranged from 1.5 – 4.9%. There were not enough studies of only patients with multiple rib fractures to perform meta-analyses on rib fixation for this patient population.

This meta-analysis included a large number of studies demonstrating the potential short-term benefit of rib fixation over nonoperative treatment for flail chest. Most often the indication for rib fixation was the presence of flail chest and to a lesser extent respiratory failure or intractable pain. Even though almost all studies included patients with flail chest, in many cases it was unclear whether it was a radiological or clinical flail chest making results harder to interpret. It is important to distinguish between these subgroups as respiratory compromise as well as injury severity are thought to mark important differences and influence outcome. The heterogeneous indication and patient populations reported on in the literature mask the exact indication and patient subgroup that would benefit most from rib fixation and consequently the adaptation of rib fixation in current practice.

Very few studies are available investigating patients with multiple rib fractures without flail chest. In a retrospective study, Qiu et al. performed separate analysis on patients with multiple rib fractures without flail segment and showed good short term results and an earlier return to 'normal activity' after rib fixation.[18] Another notable study on multiple rib fractures was from Khandelwal et al. who described a prospective cohort of patients with multiple rib fractures where most patients had two or three rib fractures and only two (5.3%) had a flail chest.[29] They reported a significant reduction of pain and earlier return to work after rib fixation. No other studies have reported on rib fixation compared to nonoperative treatment focused on multiple rib fractures even though this is the largest subgroup of patients seen in daily practice.

In this review we have included both RCTs and observational studies and show similar results for all outcome measures between both designs. Concato et al., Benson et al., and Ioannides et al. have provided an empirical basis for the comparison of RCTs and observational studies and showed results from these different designs can be remarkably similar, but can be rather different as well.[52–54] Although, treatment effects can be similar across studies regardless of design, genuine differences in treatment effects between different patient populations may be masked by biases in observational studies. Pooling results across different design could

then lead to incorrect inferences. The judgement about validity of pooling results from different designs should be made on a case-by-case basis, since for instance the potential for confounding bias is context- and research-specific. Still, within the field of (orthopedic) trauma surgery there is growing evidence showing the potential of observational studies in meta-analyses leading to more robust conclusions without decreasing quality of the results.[7–9]

Interestingly, RCTs in this study showed a larger treatment effect for some of the outcome measures as compared to observational studies. It is thought that observational studies tend to overestimate treatment effect which is possibly the result of the surgeon introducing a selection bias by choosing the optimal patient or publication bias.[55,56] The three RCTs available on this subject all had very strict in- and exclusion criteria resulting in specific patient groups where treatment effects could be demonstrated yet with limited generalizability.[22,23,50] In observational studies, usually with less strict in- and exclusion criteria, an unclear indication together with other serious concomitant injuries can result in a selection of patients including patients who would benefit more from nonoperative treatment. A wrong patient selection can reduce measured treatment effects after rib fixation which could explain differences found between RCTs and observational studies in this specific topic. Additionally, differences in timing of the surgical procedure between studies might have introduced bias in comparability as early surgical stabilization is associated with favorable outcomes.[57] However, data regarding timing of surgery was not sufficiently reported in the included studies to further explore these effects. Finally, improvement of intensive care management over time could have attributed to differences in treatment effects as shown by our sensitivity analysis. In more recent studies only mortality and DMV improved after rib fixation but there was no difference for the other outcome measures.

This study had some limitations. First, the results may be altered by missed studies in the literature search or by publication bias. However, we performed an extensive search using multiple databases with citation and reference checking of included studies. A funnel plot of the primary outcome measure did not suggest bias due to selective publication. Therefore we are confident that we have a representative overview of the current literature. Second, we did not distinguish between studies with both flail chest and multiple rib fractures and studies including only flail chest patients. Very few patients with multiple rib fractures were included in these studies. Therefore, we think results from these studies translate to flail chest patients and should not be excluded from analyses. Still, cautious interpretation of study results is necessary as the variety of definitions used in the included studies might have resulted in a high in-between study variability of patient samples.

More research is needed to further identify the right indication and right patient for rib fixation. As previously mentioned, RCTs in this heterogenic population are very difficult to perform and for adequate subgroup analyses sufficiently large sample sizes are needed. In the rapidly developing area of surgery, RCTs can be expensive, time consuming, and often

have limitations in terms of generalizability and small sample sizes due to strict in- and exclusion criteria.[58,59] Observational studies show similar results as compared to RCTs and might be an achievable first step in gathering high quality evidence. Currently a large prospective multicenter database is created in the Netherlands including both patients with flail chest and multiple rib fractures from multiple level-1 trauma centers, aiming to answer the above questions with the use of large sample sizes and long-term follow-up.[60]

Conclusion

Rib fixation significantly improves short-term outcome for patients with flail chest, although the indication and patient subgroup who would benefit most from this treatment remains unclear. There is not enough data regarding patients with multiple rib fractures without flail segment. Observational studies show similar results as compared to RCTs and might be an achievable first step in gathering high quality evidence. Larger prospective studies are required to investigate proper indications and relevant outcome after rib fixation.

Appendices

Appendix 1. Search Syntax

Date searched: June 16th 2018

Search string Pubmed (n=698)

("Rib Fractures"[Mesh] OR rib fracture* OR "flail chest"[Mesh]) AND (surgical management OR fixation OR plating OR orif)

Search string Embase (n=847)
('rib fracture'/exp OR (rib NEAR/1 fracture*):ab,ti OR 'flail chest':ab,ti) AND ('fracture treatment'/exp OR orif:ab,ti OR fixation:ab,ti OR plating:ab,ti)

Search string CENTRAL (n=195)
("rib fracture*" OR "flail chest")

Search string CINAHL (n=612)
(''rib fracture*'')

Appendix 2. MINORS assessment criteria

Criteria	2	1	0
A clearly stated aim	Aim or hypothesis including outcomes have been reported	Aim or hypothesis have been reported without a clear outcome	Not reported
Inclusion of consecutive patients	Explicit inclusion and exclusion criteria have been reported	Unclear or poor description inclusion and exclusion criteria have been reported	Not reported
Prospective collection of data	Prospective	Retrospective	Not reported
Endpoints appropriate to the aim of the study	Outcomes are appropriate to the aim of the study	Outcomes are not appropriate to the aim of the study	Not reported
Unbiased assessment of the study endpoint	Blind evaluation of objective outcomes and double-blind evaluation of subjective outcomes	One or more outcomes have been blinded	No blinding / not reported
Follow-up period appropriate to the aim of the study	≥ 1 year	< 1 year	Not reported
Loss to follow-up less than 5%	≤ 5%	> 5% and ≤ 20%	Not reported / >20%
Prospective calculation of the study size	Power analysis has been performed	Explanation for the number of included patients without a power analysis	Not reported / not performed
An adequate control group	Plate or intramedullary fixation compared with a conservative treatment	Not applicable	Not reported
Contemporary groups	Study group and controls have been managed during the same time period	Study group and controls have not been managed during the same time period	Not reported / unclear discription
Baseline equivalence of groups	Baseline characteristics have been described for both groups and are comparable	Baseline characteristics have not been described thoroughly or are not comparable	Not reported
Adequate statistical analyses	Statistical analysis has been described including the type of test	Inadequate statistical analysis	Not reported

Appendix 3. Quality assessment of all included studies in a systematic review of proximal humerus fractures comparing operative to nonoperative treatment.

Criteria	Aubert 1981	Kim 1981	Ahmed 1995	Voggenreiter 1996	Tanaka 2002	Balci 2004	Granetzny 2005	Nirula 2006	Solberg 2009	Althausen 2011	Moya 2011	Khandelwal 2011	Marasco 2013	Doben 2014	Granhed 2014	Xu 2015	Majercik a + b 2015	Wu 2015	Wada 2015	Zhang X 2015	Zhang Y 2015	Jayle 2015	Qiu 2016	Velasquez 2016	Uchida 2016	DeFreest 2016	Pieracci 2016	Farquhar 2016	Fitzgerald 2017	Dehghan 2018	Ali-Osman 2018	Wijffels 2018	Kane 2018
A Clearly stated aim	2	2	2	2	2	2	2	2	1	2	2	1	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Inclusion of consecutive patients	2	1	2	2	2	2	2	0	2	2	1	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	0
Prospective collection of data	1	1	1	1	2	1	2	1	1	1	1	2	2	1	1	1	1	2	1	1	1	1	1	1	1	1	2	1	0	0	0	0	0
Endpoints appropriate to the aim of the study	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	1
Unbiased assessment of the study endpoint	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Follow-up period appropriate to the aim of the study	2	1	1	1	2	0	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2
Loss to follow-up less than 5%	1	1	2	0	2	0	2	0	0	0	2	2	1	2	0	0	2	0	0	0	1	0	0	2	0	2	1	2	1	0	1	1	0
Prospective calculation of the study size	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
An adequate control group	0	0	0	0	2	1	2	2	1	1	1	1	2	1	1	0	2	2	2	2	0	2	0	1	2	1	2	1	2	1	1	2	1
Contemporary groups	1	1	1	2	2	1	2	2	2	2	1	2	2	1	0	2	2	2	2	2	2	1	1	1	2	1	1	1	1	1	1	1	1
Baseline equivalence of groups	0	0	1	2	2	1	1	1	2	2	2	0	2	2	0	2	1	2	2	2	2	2	2	2	2	2	1	1	2	1	1	2	1
Adequate statistical analyses	2	2	0	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	1	2	2	1
Total quality score MINORS	**13**	**11**	**12**	**14**	**20**	**12**	**19**	**13**	**15**	**18**	**16**	**17**	**21**	**17**	**12**	**15**	**18**	**17**	**17**	**16**	**16**	**16**	**14**	**17**	**17**	**17**	**17**	**16**	**15**	**11**	**14**	**16**	**9**

Appendix 4. Results of the included studies comparing operative versus non-operative management of traumatic rib fractures.

Study	Treatment groups	Mortality	Hospital LOS (days)	ICU LOS (days)	Duration of mechanical ventilation (days)	Pneumonia	Tracheostomy
Dehghan 2018	Operative	2 (2.6%)	21 ± 20	15 ± 13	NR	45 (48%)	17 (22%)
	Non-operative	160 (9.8%)	17 ± 26	13 ± 15		614 (38%)	182 (11%)
Ali-Osman 2018	Operative	1 (1.6%)	12 [9-16]	6 [3-10]	3 [1-15]	5 (7.8%)	NR
	Non-operative	13 (9.6%)	4.8 [2.9-8.4]	4 [3-7]	4 [1-10]	16 (12%)	
Wijffels 2018	Operative	2 (10%)	21 [12-33]	5 [3-13]	4 [2-10]	7 (35%)	NR
	Non-operative	1 (5%)	23 [17-42]	12 [3-29]	18 [12-26]	16 (80%)	NR
Kane 2018	Operative	1 (0.9%)	12 [10-14]	3 [0-6]	NR	7 (6%)	10 (8.6%)
	Non-operative	13 (1.3%)	5 [3-9]	0 [0-3]		59 (6%)	45 (4.5%)
Fitzgerald 2017	Operative	0 (0%)	18 (14-23)	12 (7-17)	NR	0 (0%)	NR
	Non-operative	2 (4%)	17 (10-23)	8 (5-11)		7 (14%)	
Farquhar 2016	Operative	1 (5,3%)	21.9 ± 13.2	7.4 ± 6.7	6.1 ± 5.9	12 (63%)	NR
	Non-operative	1 (2,8%)	16.0 ± 12.1	3.7 ± 6.0	3.1 ± 5.5	8 (22%)	
Pieracci 2016	Operative	0 (0%)	13.0 [9.0, 21.0]	6.0 [3.0, 10.0]	0 [0.0, 8.0]	7 (20%)	5 (14%)
	Non-operative	0 (0%)	16.0 [10.0, 23.0]	9.0 [4.0, 15.0]	5.0 [0, 18]	11 (31%)	16 (46%)
Defreest 2016	Operative	1 (2,4%)	28.3 (9-69)	14.0 (0-43)	9.3 (0-39)	11 (27%)	10 (24%)
	Non-operative	5 (11,1%)	13.0 (3-43)	8.0 (0-43)	5.8 (0-39)	10 (22%)	8 (18%)
Uchida 2016	Operative	0 (0%)	NR	6.5 [3, 9]	5.5 [1, 8]	2 (20%)	1 (10%)
	Non-operative	0 (0%)		12 [8, 14]	9 [7, 12]	9 (90%)	3 (30%)
Velasquez 2016	Operative	0 (0%)	6 [4, 10]	4.5 [1, 8]	2 [1, 3]	3 (15%)	NR
	Non-operative	2 (10%)	16 [11, 22]	8 [6, 10.5]	10 [6, 16]	13 (65%)	
Qiu a 2016	Operative	1 (4,8%)	NR	7.2 ± 1.7	5.7 ± 1.4	NR	2 (9,5%)
	Non-operative	2 (11,8%)		10.3 ± 2.3	9.1 ± 3.6		8 (47%)
Qiu b 2016	Operative	0 (0%)	11.1 ± 1.9	NR	NR	3 (4,6%)	NR
	Non-operative	0 (0%)	15.9 ± 2.8			10 (17%)	
Jayle 2015	Operative	NR	21.7 ± 7.8	9.0 ± 4.3	3.1 ± 5.2	4 (40%)	NR
	Non-operative	NR	32.3 ± 19.3	12.3 ± 8.5	5.9 ± 9.4	3 (30%)	

(table continued)

Zhang Y 2015	Operative	0 (0%)	38 [33, 54.25]	4.5 [21.3, 30.7]	12 [7.5, 17.8]	16 (67%)	12 (50%)
	Non-operative	2 (13,3%)	60 [38, 99.75]	21.5 [18, 33.5]	7 [4, 14]	7 (47%)	7 (9,7%)
Zhang X 2015	Operative	0 (0%)	NR	5.5 ± 6.4	4.1 ± 6.1	NR	NR
	Non-operative	0 (0%)		14.2 ± 6.5	14 ± 7.6		
Wada 2015	Operative	3 (3,6%)	33 [22, 45]	NR	NR	NR	10 (12%)
	Non-operative	6 (1,8%)	42 [23, 58]				68 (20%)
Wu 2015	Operative	1 (1,3%)	15.3 ± 6.4	8.2 ± 4.3	3.7 ± 1.4	5 (6,7%)	4 (5,3%)
	Non-operative	4 (4,5%)	26.5 ± 6.9	14.6 ± 3.2	9.5 ± 4.3	17 (19%)	7 (7,9%)
Majercik 2015	Operative	NR	11.4 ± 5.7	4.6 ± 5.6	0 [0, 3]	12 (8,8%)	8 (5,8%)
	Non-operative		12.3 ± 9.1	5.9 ± 7.7	0 [0, 4]	55 (20%)	30 (11%)
Xu 2015	Operative	0 (0%)	NR	15.9 ± 5.0	10.5 ± 3.7	10 (59%)	2 (12%)
	Non-operative	1 (6,7%)		19.6 ± 5.0	13.7 ± 4.4	12 (93%)	6 (40%)
Granhed 2014	Operative	2 (3,3%)	NR	NR	2.7 (0-21)	0 (0%)	NR
	Non-operative	NR			9.0 (1-76)	NR	
Doben 2014	Operative	N/A	21.6 (8-59)	12.5 (5-21)	8.2 (0-30)	NR	NR
	Non-operative	0 (0%)	28.5 (6-50)	15.3 (5-22)	18.0 (4-40)		
Marasco 2013	Operative	0 (0%)	20 [18, 28]	13.5 [9.9, 15.8]	6.3 ± 3.4	11 (48%)	9 (3,9%)
	Non-operative	1 (4,3%)	25 [18, 38]	18.7 [13.4, 26.9]	7.5 ± 5.4	17 (74%)	16 (7,0%)
Khandelwal 2011	Operative	NR	NR	NR	NR	NR	NR
	Non-operative						
Moya 2011	Operative	NR	18 ± 12	9 ± 8	7 ± 8	5 (31%)	NR
	Non-operative		16 ± 11	7 ± 10	6 ± 10	12 (38%)	
Althausen 2011	Operative	NR	11.9 ± 7.8	7.6 ± 7.4	4.2 ± 6.6	1 (4,5%)	3 (3,9%)
	Non-operative		19.0 ± 12.6	9.7 ± 9.2	9.7 ± 9.2	7 (25%)	11 (3,9%)
Solberg 2009	Operative	NR	NR	5.4 ± 1.5	1.9 ± 1.1	0 (0%)	NR
	Non-operative			21 ± 13.6	13.3 ± 5.3	3 (43%)	
Nirula 2006	Operative	NR	18.8 ± 1.8	12.1 ± 1.2	6.5 ± 1.3	NR	NR
	Non-operative		21.1 ± 3.9	14.1 ± 2.7	11.2 ± 2.6		
Granetzny 2006	Operative	2 (10%)	11.7 ± 10.1	9.6 ± 12.0	2 ± 8.9	NR	NR
	Non-operative	3 (15%)	23.1 ± 10.1	14.6 ± 12.0	12 ± 8.9		

(table continued)

Balci 2004	Operative	3 (1,11%)	18.3 ± 7.6	NR	3.1 ± 1.8	NR	0 (0%)
	Non-operative	10 (27,0%)	19.2 ± 7.2	NR	7.2 ± 5.8		7 (19%)
Tanaka 2002	Operative	0 (0%)	NR	16.5 ± 7.4	10.8 ± 3.4	4 (22%)	3 (17%)
	Non-operative	0 (0%)		26.8 ± 13.2	18.3 ± 7.4	17 (90%)	15 (79%)
Voggenreiter a 1996	Operative	0 (0%)	NR	NR	6.5 ± 7.0	1 (10%)	NR
	Non-operative	7 (38,9%)			26.7 ± 29.0	5 (28%)	
Voggenreiter b 1996	Operative	3 (30%)	NR	NR	30.8 ± 33.7	4 (40%)	NR
	Non-operative	1 (25%)			29.3 ± 22.5	2 (50%)	
Ahmed 1995	Operative	2 (10%)	NR	9	3.9	NR	3 (15%)
	Non-operative	11 (57,9%)		21	15		14 (74%)
Kim 1981	Operative	1 (5,9%)	NR	NR	24 ± 15	NR	NR
	Non-operative	60 (42,2%)			22.1 ± 13.5	7 (4,9%)	
Aubert 1981	Operative	3 (13,6%)	NR	NR	NR	NR	NR
	Non-operative	54 (24,1%)					135 (60%)

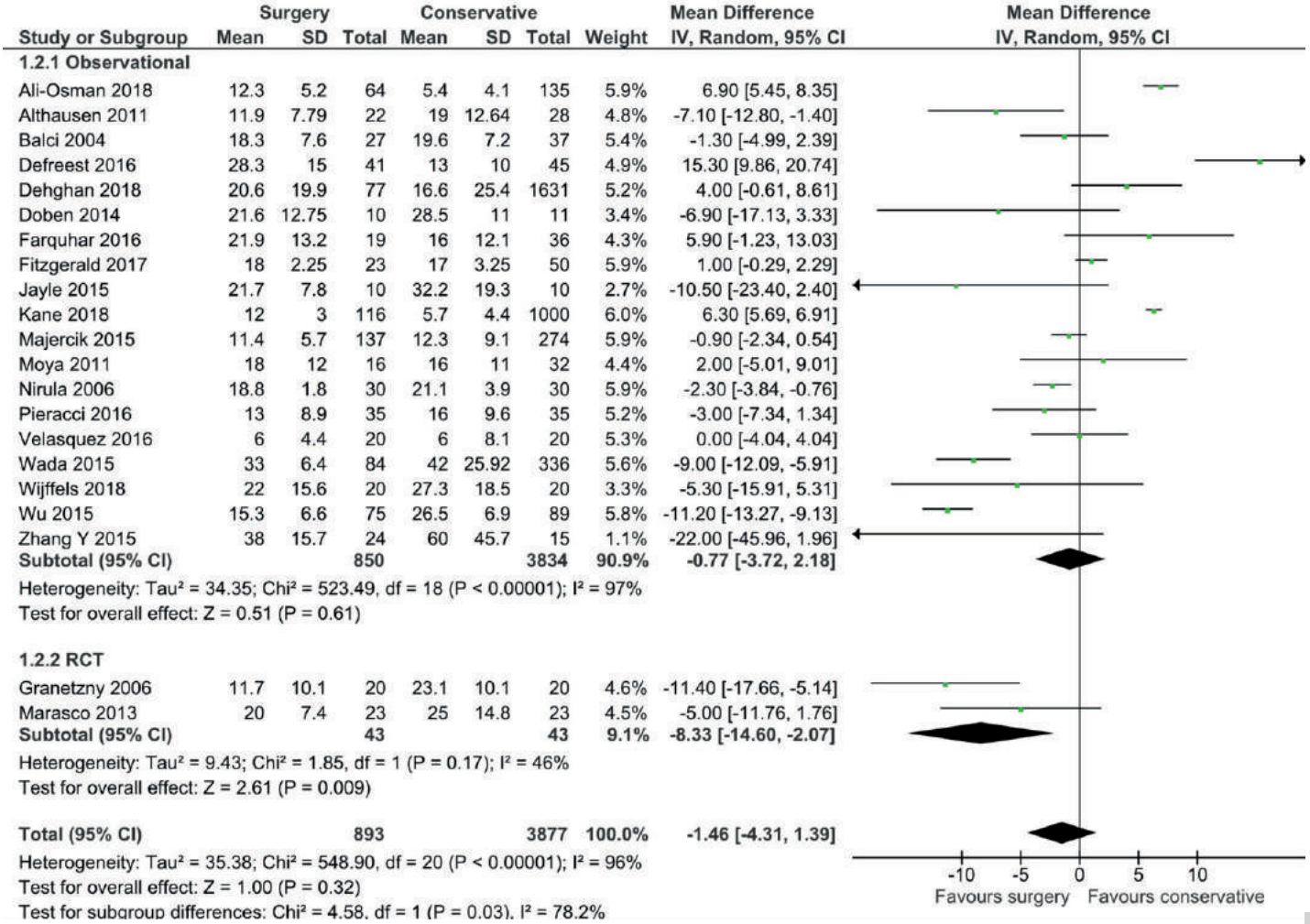
Appendix 5. Impact of different methods to handle zero-event data in a meta-analysis of operative versus nonoperative treatment of rib fractures and mortality

Method	Observational studies OR (95% CI)	RCT OR (95% CI)	Total OR (95% CI)
Mantel-Haenzel*	0.43 (0.27 – 0.69)	0.57 (0.13 – 2.52)	0.44 (0.28 – 0.69)
Crude	0.21 (0.13 – 0.35)	0.49 (0.09 – 2.79)	0.22 (0.14 – 0.35)
Inverse variance - no correction	0.41 (0.23 – 0.73)	0.63 (0.09 – 4.24)	0.43 (0.25 – 0.74)
Inverse variance - with correction	0.39 (0.23 – 0.65)	0.59 (0.13 – 2.68)	0.41 (0.25 – 0.66)
DerSimonian Laird with correction	0.37 (0.17 – 0.79)	0.58 (0.11 – 3.23)	0.39 (0.20 – 0.78)
Peto	0.28 (0.16 – 0.49)	0.50 (0.07 – 3.47)	0.29 (0.17 – 0.50)

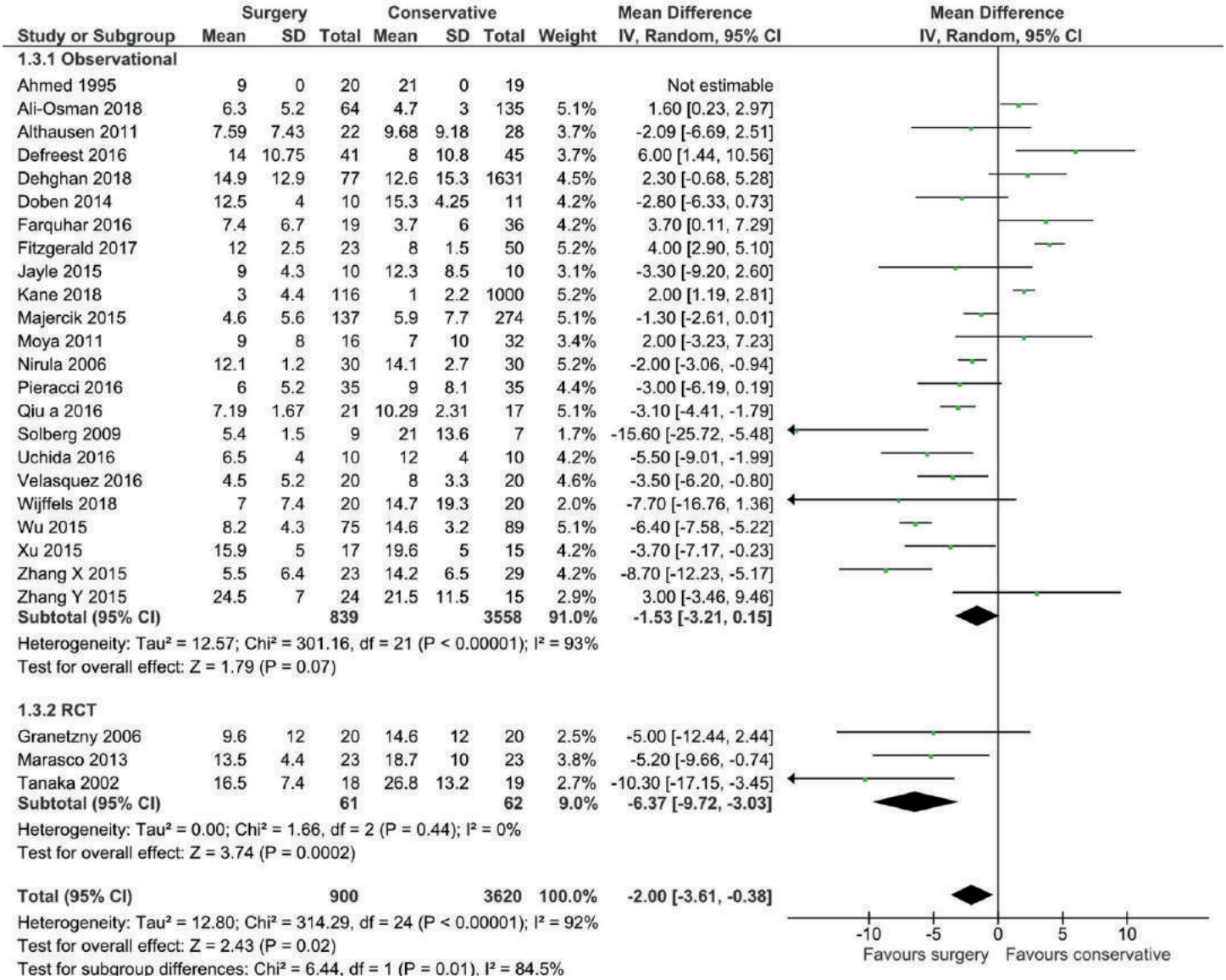
* Method used in meta-analysis; OR odds-ratio; CI confidence interval

In a model with correction 0.5 is added to every table of the 2x2 table

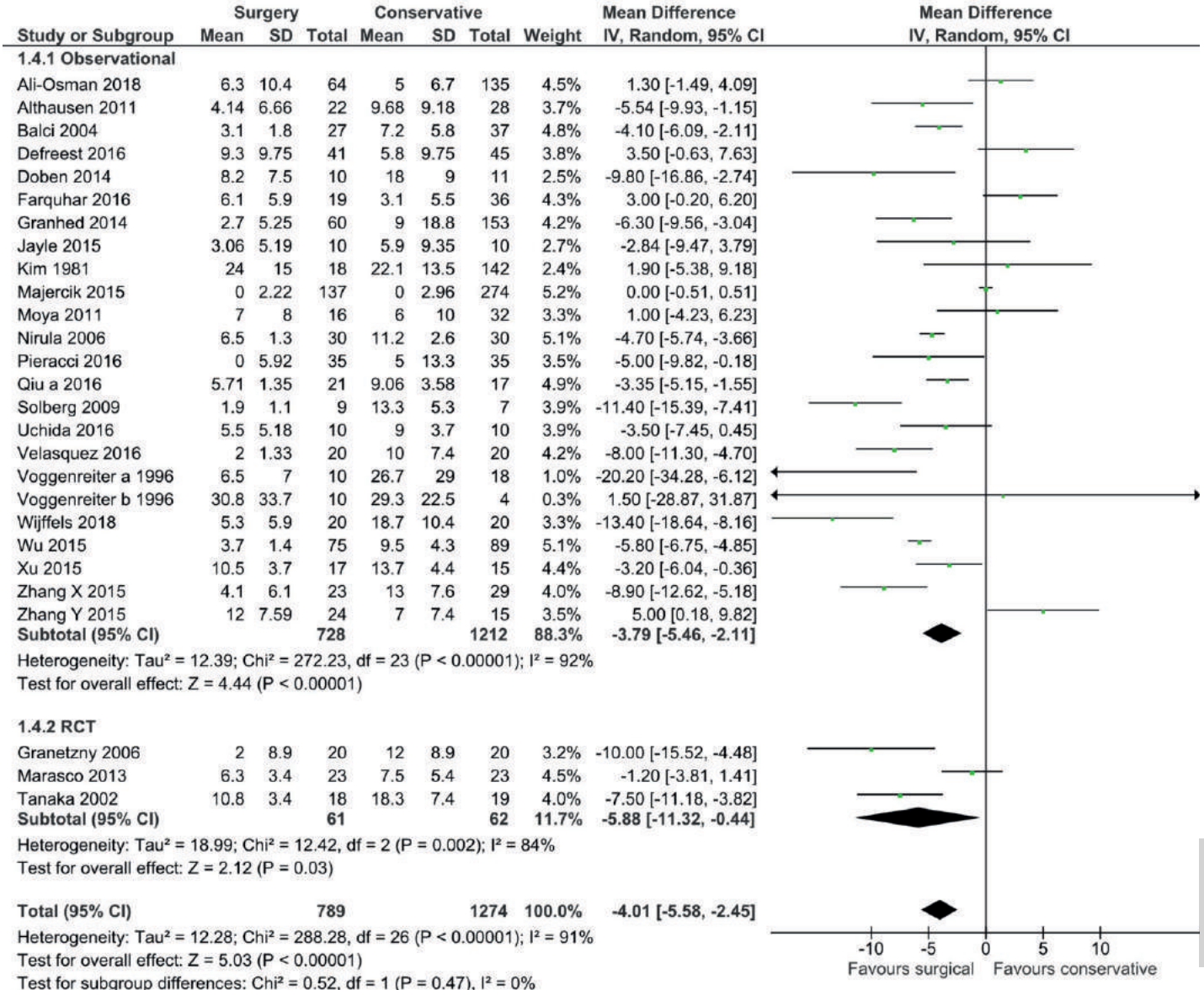
Appendix 6. Hospital length of stay in a systematic review of rib fractures comparing operative to nonoperative treatment



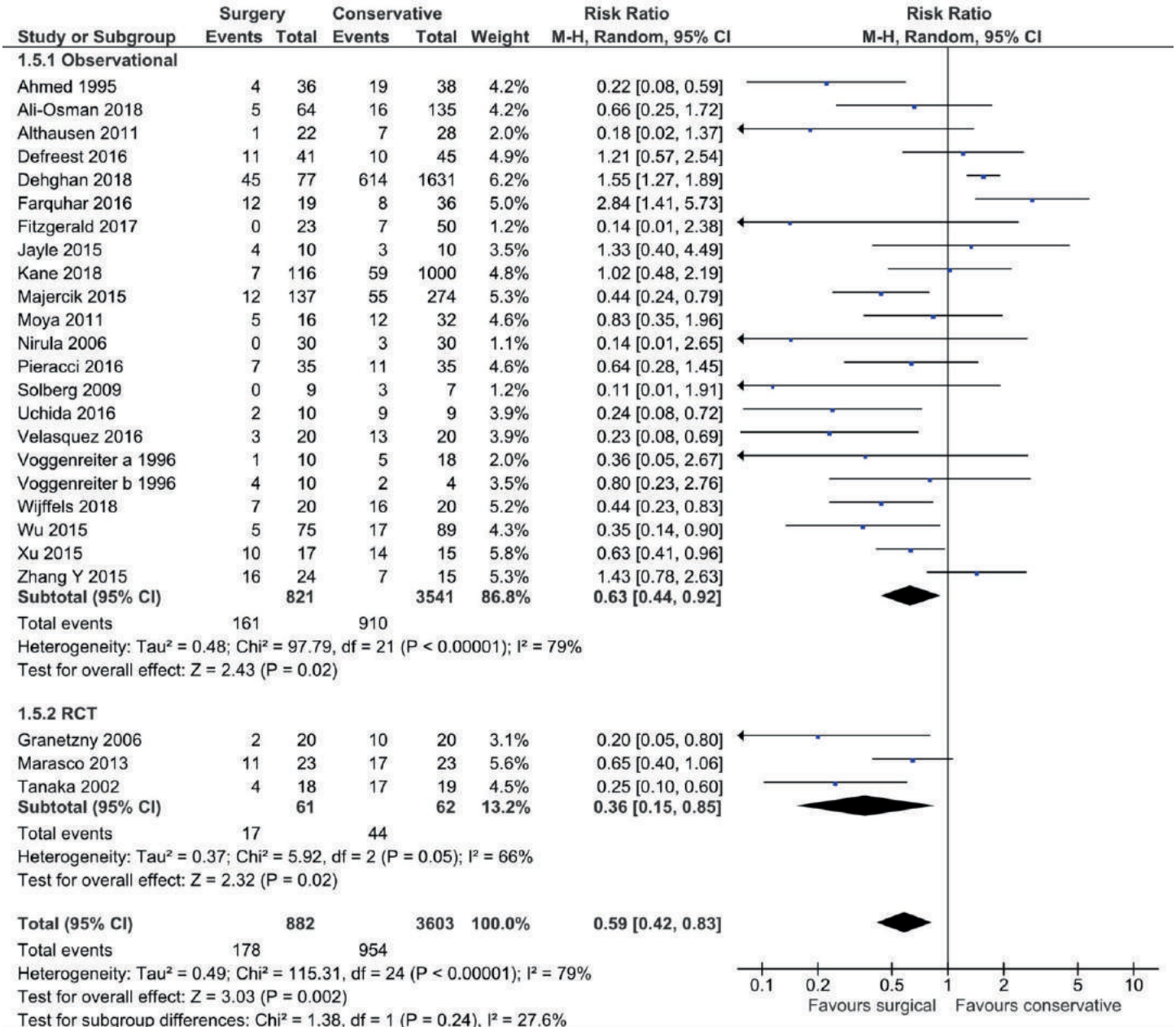
Appendix 7. Intensive care length of stay in a systematic review of rib fractures comparing operative to nonoperative treatment



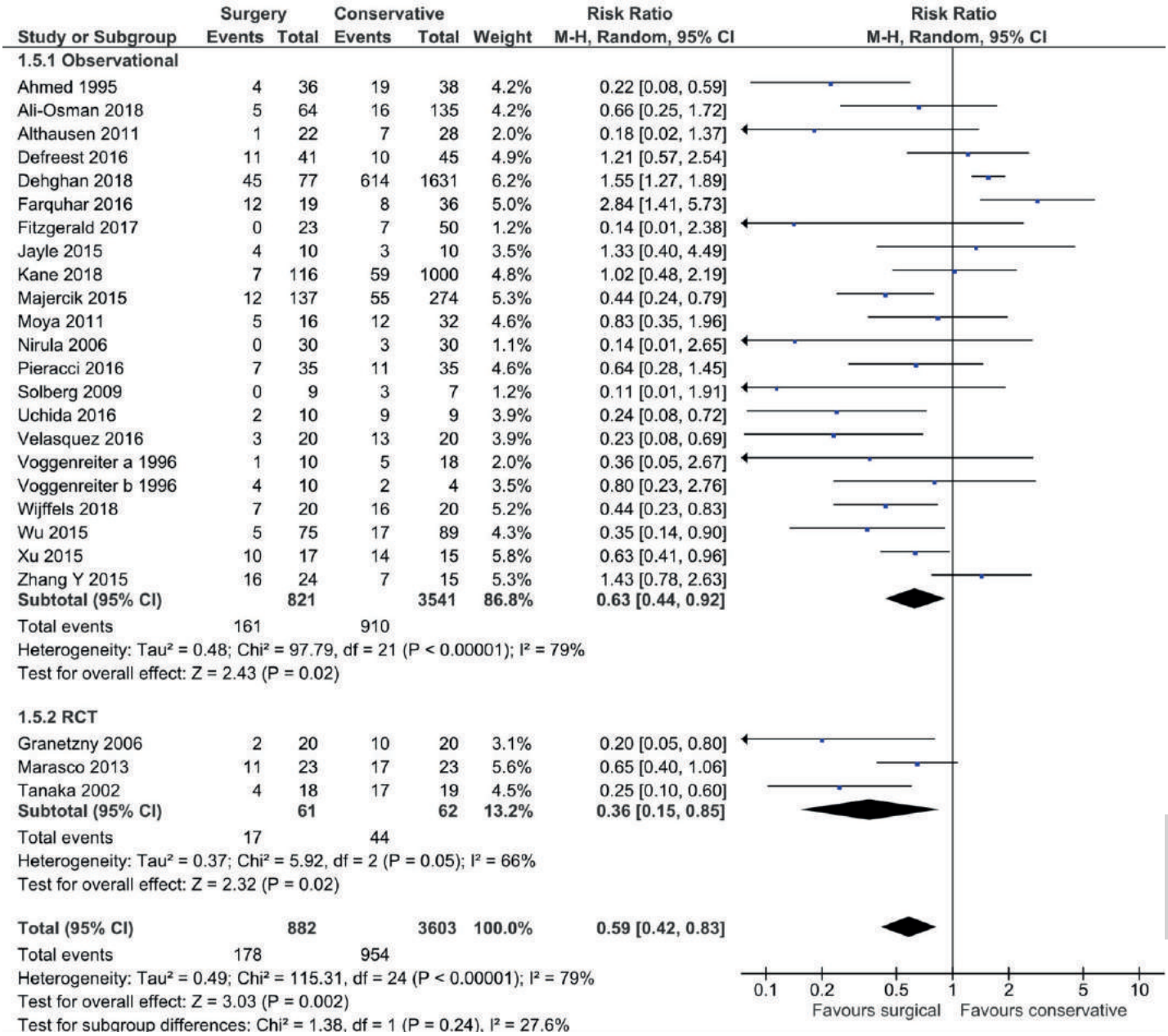
Appendix 8. Duration of mechanical ventilation in a systematic review of rib fractures comparing operative to nonoperative treatment



Appendix 9. Pneumonia in a systematic review of rib fractures comparing operative to nonoperative treatment



Appendix 10. Tracheostomy in a systematic review of rib fractures comparing operative to nonoperative treatment



References

1. Bulger EM, Arneson MA, Mock CN, Jurkovich GJ. Rib fractures in the elderly. *J Trauma*. 2000;48(6):1040-1047.

2. Ziegler DW, Agarwal NN. The morbidity and mortality of rib fractures. *The Journal of trauma*. 1994;37(6):975-979.

3. Vana PG, Neubauer DC, Luchette FA. Contemporary management of flail chest. *The American surgeon*. 2014;80(6):527-535.

4. Lin FC-F, Li R-Y, Tung Y-W, Jeng K-C, Tsai SC-S. Morbidity, mortality, associated injuries, and management of traumatic rib fractures. *Journal of the Chinese Medical Association*. 2016;79(6):329-334. doi:10.1016/j.jcma.2016.01.006.

5. Cannon RM, Smith JW, Franklin GA, Harbrecht BG, Miller FB, Richardson JD. Flail chest injury: are we making any progress? *Am Surg*. 2012;78(4):398-402.

6. Dehghan N, De Mestral C, McKee MD, Schemitsch EH, Nathens A. Flail chest injuries: A review of outcomes and treatment practices from the national trauma data bank. *Journal of Trauma and Acute Care Surgery*. 2014;76(2):462-468. doi:10.1097/TA.0000000000000086.

7. 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. doi:10.1177/0363546516673615.

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. doi:10.1016/j.jse.2016.01.018.

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. doi:10.1016/j.jclinepi.2009.04.005.

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;0(0). doi:10.1016/j.jse.2018.03.009.

11. 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(1):598. doi:10.1186/s12913-016-1816-5.

12. 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. doi:10.1371/journal.pmed.1000100.

13. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD RD. MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies. *Jama*. 2000;283:2008-2012.

14. 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.

15. Collaboration TC. *Cochrane Handbook for Systematic Reviews of Interventions 4.2.6.*; 2006.

16. Majercik S, Wilson E, Gardner S, Granger S, VanBoerum DH, White TW. In-hospital outcomes and costs of surgical stabilization versus nonoperative management of severe rib fractures. *Journal of Trauma and Acute Care Surgery*. 2015;79(4):533-539. doi:10.1097/TA.0000000000000820.

17. Majercik S, Vijayakumar S, Olsen G, et al. Surgical stabilization of severe rib fractures decreases incidence of retained hemothorax and empyema. *American Journal of Surgery*. 2015;210(6):1112-1117. doi:10.1016/j.amjsurg.2015.08.008.

18. Qiu M, Shi Z, Xiao J, Zhang X, Ling S, Ling H. Potential Benefits of Rib Fracture Fixation in Patients with Flail Chest and Multiple Non-flail Rib Fractures. *The Indian journal of surgery*. 2016;78(6):458-463. doi:10.1007/s12262-015-1409-2.

19. Voggenreiter G, Neudeck F, Aufmkolk M, et al. Outcome of operative chest wall stabilization in fail chest with or without pulmonary contusion. *Unfallchirurg*. 1996;99(6):425-434.

20. 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. doi:10.1177/0363546516673615.

21. 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. doi:10.1002/sim.2528.

22. Tanaka H, Yukioka T, Yamaguti Y, et al. Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients. *Journal of Trauma*. 2002;52(4):727-732. doi:10.1097/00005373-200204000-00020.

23. Granetzny A, Abd El-Aal M, Emam E, Shalaby A, Boseila A. Surgical versus conservative

treatment of flail chest. Evaluation of the pulmonary status. *Interactive Cardiovascular and Thoracic Surgery*. 2005;4(6):583-587. doi:10.1510/icvts.2005.111807.

24. Solberg BD, Moon CN, Nissim AA, Wilson MT, Margulies DR. Treatment of chest wall implosion injuries without thoracotomy: technique and clinical outcomes. *Journal of Trauma*. 2009;67(1):8-13. doi:10.1097/TA.0b013e3181a8b3be.

25. Zhang Y, Tang X, Xie H, Wang RL. Comparison of surgical fixation and nonsurgical management of flail chest and pulmonary contusion. *American Journal of Emergency Medicine*. 2015;33(7):937-940. doi:10.1016/j.ajem.2015.04.005.

26. Althausen PL, Shannon S, Watts C, et al. Early stabilization of flail chest with locked plate fixation. *Journal of orthopaedic trauma*. 2011;25(11):648. doi:10.1097/BOT.0b013e31822a542d.

27. Granhed HP, Pazooki D. A feasibility study of 60 consecutive patients operated for unstable thoracic cage. *Journal of trauma management & outcomes*. 2014;8(1):20. doi:10.1186/s13032-014-0020-z.

28. Xu J-QQ, Qiu P-LL, Yu R-GG, et al. Better short-term efficacy of treating severe flail chest with internal fixation surgery compared with conservative treatments. *European Journal of Medical Research*. 2015;20(1):55. doi:10.1186/s40001-015-0146-0.

29. Khandelwal G, Mathur RK, Shukla S, Maheshwari A. A prospective single center study to assess the impact of surgical stabilization in patients with rib fracture. *International journal of surgery (London, England)*. 2011;9(6):478-481. doi:10.1016/j.ijsu.2011.06.003.

30. Uchida K, Nishimura T, Takesada H, et al. Evaluation of efficacy and indications of surgical fixation for multiple rib fractures: a propensity-score matched analysis. *European journal of trauma and emergency surgery : official publication of the European Trauma Society*. 2016;43(4):541-547. doi:10.1007/s00068-016-0687-0.

31. de Moya M, Bramos T, Agarwal S, et al. Pain as an indication for rib fixation: a bi-institutional pilot study. *Journal of Trauma*. 2011;71(6):1750-1754. doi:10.1097/TA.0b013e31823c85e9.

32. Wu WM, Yang Y, Gao ZL, Zhao TC, He WW. Which is better to multiple rib fractures, surgical treatment or conservative treatment? *Int J Clin Exp Med*. 2015;8(5):7930-7936.

33. Fitzgerald MT, Ashley DW, Abukhdeir H, Christie DB. Rib fracture fixation in the 65 years and older population. *Journal of Trauma and Acute Care Surgery*. 2017;82(3):524-527. doi:10.1097/TA.0000000000001330.

34. Wada T, Yasunaga H, Inokuchi R, et al. Effectiveness of surgical rib fixation on prolonged mechanical ventilation in patients with traumatic rib fractures: A propensity

score-matched analysis. *Journal of Critical Care*. 2015;30(6):1227-1231. doi:10.1016/j.jcrc.2015.07.027.

35. Nirula R, Allen B, Layman R, Falimirski ME, Somberg LB. Rib fracture stabilization in patients sustaining blunt chest injury. *American Surgeon*. 2006;72(4):307-309.

36. Ahmed Z, Mohyuddin Z. Management of flail chest injury: Internal fixation versus endotracheal intubation and ventilation. *Journal of Thoracic and Cardiovascular Surgery*. 1995;110(6):1676-1680. doi:10.1016/S0022-5223(95)70030-7.

37. Pieracci FM, Lin Y, Rodil M, et al. A prospective, controlled clinical evaluation of surgical stabilization of severe rib fractures. *Journal of Trauma and Acute Care Surgery*. 2016;80(2):187-194. doi:10.1097/TA.0000000000000925.

38. DeFreest L, Tafen M, Bhakta A, et al. Open reduction and internal fixation of rib fractures in polytrauma patients with flail chest. *American Journal of Surgery*. 2016;211(4):761-767. doi:10.1016/j.amjsurg.2015.11.014.

39. Farquhar J, Almahrabi Y, Slobogean G, et al. No benefit to surgical fixation of flail chest injuries compared with modern comprehensive management: results of a retrospective cohort study. *Canadian journal of surgery Journal canadien de chirurgie*. 2016;59(5):299-303.

40. Doben AR, Eriksson EA, Denlinger CE, et al. Surgical rib fixation for flail chest deformity improves liberation from mechanical ventilation. *Journal of Critical Care*. 2014;29(1):139-143. doi:10.1016/j.jcrc.2013.08.003.

41. Wijffels MME, Hagenaars T, Latifi D, Lieshout EMM Van, Verhofstad MHJ. Early results after operatively versus non-operatively treated flail chest : a retrospective study focusing on outcome and complications. *European Journal of Trauma and Emergency Surgery*. 2018;0(0):0. doi:10.1007/s00068-018-0961-4.

42. Ali-Osman F, Mangram A, Sucher J, et al. Geriatric (G60) trauma patients with severe rib fractures: Is muscle sparing minimally invasive thoracotomy rib fixation safe and does it improve post-operative pulmonary function? *American Journal of Surgery*. 2018:1-6. doi:10.1016/j.amjsurg.2018.02.022.

43. Dehghan N, Mah JM, Schemitsch EH, Nauth A, Vicente M, McKee MD. Operative stabilization of flail chest injuries reduces mortality to that of stable chest wall injuries. *Journal of Orthopaedic Trauma*. 2018;32(1):15-21. doi:10.1097/BOT.0000000000000992.

44. Kane ED, Jeremitsky E, Bittner KR, Kartiko S, Doben AR. Surgical Stabilization of Rib Fractures: A Single Institution Experience. *Journal of the American College of Surgeons*. 2018;226(6):961-966. doi:10.1016/j.jamcollsurg.2017.11.008.

45. Balci AE, Eren S, Cakir O, Eren MN. Open fixation in flail chest: Review of 64 patients.

Asian Cardiovascular and Thoracic Annals. 2004;12(1):11-15. doi:10.1177/021849230401200104.

46. Zhang X, Guo Z, Zhao C, Xu C, Wang Z. Management of patients with flail chest by surgical fixation using claw-type titanium plate. *Journal of cardiothoracic surgery*. 2015;10:145. doi:10.1186/s13019-015-0363-1.

47. Velasquez M, Ordonez CA, Parra MW, Dominguez A, Puyana JC. Operative versus Nonoperative Management of Multiple Rib Fractures. *The American surgeon*. 2016;82(5):103-105.

48. Aubert M, Antoine P, Pilichowski P. Flail chests. Study of 224 cases. *Annales de Chirurgie*. 1981;35(1):33-39.

49. Kim M, Brutus P, Christides C, et al. Compared results of flail chests treatments: Standard internal pneumatic stabilization, new technics of assisted ventilation, oseosynthesis. *Journal de Chirurgie*. 1981;118(8-9):499-503.

50. Marasco SF, Davies AR, Cooper J, et al. Prospective randomized controlled trial of operative rib fixation in traumatic flail chest. *Journal of the American College of Surgeons*. 2013;216(5):924-932. doi:10.1016/j.jamcollsurg.2012.12.024.

51. Jayle CPM, Allain G, Ingrand P, et al. Flail chest in polytraumatized patients: Surgical fixation using stracos reduces ventilator time and hospital stay. *BioMed Research International*. 2015;2015:624723. doi:10.1155/2015/624723.

52. Concato J, Shah N, Horwitz RI. Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs. *New England Journal of Medicine*. 2000;342(25):1887-1892. doi:10.1056/NEJM200006223422507.

53. Benson K, Hartz AJ. A Comparison of Observational Studies and Randomized, Controlled Trials. *New England Journal of Medicine*. 2000;342(25):1878-1886. doi:10.1056/NEJM200006223422506.

54. Ioannidis JPA, Haidich A-B, Pappa M, et al. Comparison of Evidence of Treatment Effects in Randomized and Nonrandomized Studies. *Jama*. 2001;286(7):821-830. doi:10.1001/jama.286.7.821.

55. 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. doi:10.1001/jama.297.11.1233.

56. 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.

57. Pieracci FM, Coleman J, Ali-Osman F, et al. A multicenter evaluation of the optimal

timing of surgical stabilization of rib fractures. *Journal of Trauma and Acute Care Surgery*. 2018;84(1):1-10. doi:10.1097/TA.0000000000001729.

58. 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. doi:10.1056/NEJMra1614394.

59. Jacobs WCH, Kruyt MC, Verbout AJ, Oner FC. Spine surgery research: on and beyond current strategies. *The Spine Journal*. 2012;12(8):706-713. doi:10.1016/j.spinee.2012.08.424.

60. Beks RB, Houwert RM. Nonoperative versus operative treatment for flail chest and multiple rib fractures after blunt thoracic trauma. A multicenter prospective cohort study. NTR6833. http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=6833.


PART 3
WHERE TO GO
PG.164

CHAPTER 9

WHEN OBSERVATIONAL STUDIES ARE AS HELPFUL AS RANDOMIZED TRIALS: EXAMPLES FROM ORTHOPAEDIC TRAUMA

R.B. Beks, A. R Bhashyam, R.M. Houwert,
D. van der Velde, M. van Heijl, D.P.J. Smeeing,
F. Hietbrink, L.P.H. Leenen, R.H.H. Groenwold

Submitted

Abstract

Randomized controlled trials (RCTs) are often challenging to design in surgical fields and can be misleading when poorly executed. While the prevailing belief is that observational studies on therapeutic efficacy are credible only in exceptional circumstances due to unrecognized confounding, we identify three types of intervention in orthopaedic trauma and illustrate relevant features that allow observational studies in orthopaedic surgery to be as helpful as RCTs.

Introduction

Randomized controlled trials (RCTs) were initially developed to evaluate the efficacy of streptomycin, subsequently expanded to other interventions within medicine,[1,2] and are now increasingly used in surgical disciplines.[3,4] RCTs in surgery have been fraught with challenges including recruitment difficulty, inadequate power associated with small sample size, failed randomization, and inadequate blinding.[3–6] Surgical RCTs are difficult to design and execute, and can lead to misleading results when poorly conducted. Obstacles to high-quality surgical RCTs include surgeon-related factors (culture, surgeon-preference, equipoise/consent, procedural/technique variety), patient-related factors (unwillingness to participate, patient preferences), and methodological factors (follow-up, strict inclusion/exclusion criteria, sample size estimation, funding).[3,4] These challenges are especially pronounced in trauma surgery where urgent life- or limb-threatening situations and surgical variability can lead to practical and methodological difficulties with recruitment, consent, randomization, and execution of RCTs.[3,4,7] While some authors have suggested increasing sample sizes and improving blinding to improve RCT quality, these changes are particularly difficult to implement in trauma surgery,[5,8] and other study types may be more appropriate.[3]

Observational studies and RCTs in orthopaedic trauma surgery

Historically, observational studies have been used to increase our understanding of etiologic risk factors and of prognosis and diagnosis in clinical medicine, especially in settings where RCTs would be logistically or ethically impossible.[9] Observational studies have several advantages over RCTs including lower cost and often more participants are more representative of patients in clinical practice. In trauma research, what further favors an observational design over RCTs is the acuity of trauma, while in observational studies patients enter a cohort at a well-defined moment in time with immediate subsequent hospitalization, and many important outcomes during admission are rapidly available.[10] Still, due to the potential for unmeasured bias (notably confounding), many insist on RCTs as the gold standard design for testing the effectiveness of therapeutic treatment.[1,9,11] However, well-designed observational studies accurately estimate the treatment effects of medical[1,9] and surgical[12–14] interventions when compared to RCTs on the same topic, and may represent an alternate study type that can improve our understanding of the care of orthopaedic trauma patients when applied within the appropriate context. This is especially important since traumatic injuries continue to be a leading cause of morbidity and mortality throughout the world, and merit further study.[15] In this article, we highlight and provide examples of scenarios in orthopaedic trauma surgery where observational studies of treatment effectiveness are as credible as RCTs.

When are observational studies as credible as RCTs?

Many sources of bias in either RCTs or observational intervention studies have been previously identified.[16–18] The Cochrane Handbook lists possible sources of bias in a randomized trial: inappropriate random sequence generation, no concealment of allocation, no blinding of participants and personnel, no blinding of endpoint assessment, incomplete endpoint data, and selective reporting. While it is clear that the first three sources of bias are inherently present in observational studies, the extent to which they impact the validity of a

study differs based on the research topic and setting. In a general context, Vandenbroucke proposed a three-pronged restriction to make observational research as methodologically credible as RCTs: (1) Restriction of research topics to those where exposure allocation is reasonably unrelated to the outcome, (2) Restriction in design to studies where allocation of the exposure is nearly random, and (3) Restriction to topics where confounding variables can be identified, measured, and adjusted for in the analysis.[19] In orthopaedic trauma surgery, we identify three types of intervention and the differential risk of confounding using this criteria (Table 1).[20]

Table 1. Overview of different research types in trauma surgery

	Type 1 comparison	Type 2 comparison	Type 3 comparison
Comparison	Pharmacological intervention versus placebo / other drug	Operative intervention versus alternative operative intervention	Operative intervention versus non-operative intervention
Treatment choice	Mostly dependent on clinical indication and patient characteristics	Mostly dependent on surgeon preference	Mostly dependent on surgeon preference and patient characteristics
Comparability of intervention groups	Most likely not comparable	Usually comparable	Possibly comparable
Risk of confounding	+++	+/-	+

Research designs in orthopaedic trauma surgery: Using cases to illustrate when trauma investigations are amenable to observational studies

A type 1 comparison focuses on pharmacological interventions in surgical patients, comparing a particular medication with a placebo or other drug. For this type of study, observational research is well-suited to study adverse or rare events as the surgeon is aware that (s)he is prescribing a drug but does not know the risk of an unintended, particular effect.[20,21] For example, recent literature demonstrating an increased risk of wound drainage and post-operative inflammation when using rhBMP-2 as an osteoinductive agent meets all 3 restrictive criteria and is likely to have an effect size similar to a dedicated RCT.[22,23] Nevertheless, observational studies of pharmacological therapeutic effect in surgical patients remain prone to confounding by indication and are best studied by RCT. For instance, an observational study exploring the efficacy of vancomycin powder in preventing wound complications would likely be prone to bias by unmeasured factors (e.g., patient frailty) that influence a surgeon to choose vancomycin powder treatment in a non-random fashion. As a result, we would expect different study designs to differ in terms of their ability to control for

confounding. This is nicely illustrated in a recent meta-analysis where the pooled therapeutic effect size of vancomycin powder was directly related to observational study quality.[24]

A type 2 comparison evaluates one particular operative procedure against a control group that undergoes a different intervention (including a sham operation). Differences between the two procedures can involve changes in techniques or implants. Observational studies are particularly well suited for type 2 comparisons when there is therapeutic uncertainty and near-random variation in treatment selection due to surgeon preference. This scenario is quite common in trauma surgery because patients are typically randomized to an "on-call" surgeon, and their treatment selection largely varies due to surgeon preference in the absence of strong patient-related factors.[25,26] For example, in a recent meta-analysis of RCTs and observational studies comparing plate fixation with intramedullary fixation for mid-shaft clavicle fractures, effect estimates for the primary and secondary outcome measures were comparable between observational studies and randomized trials.[14] This is not surprising because fixation choice for mid-shaft clavicle fractures is an area of therapeutic uncertainty, largely dependent on surgeon preference, and allocated to patients in a quasi-random fashion. In this setting, Vandenbroucke's restriction criteria are met and we expect observational results to be of similar credibility as an RCT.

Finally, a type 3 comparison evaluates the effectiveness of operative versus non-operative treatment. As in a type 2 study, surgeon preference is an important factor in decision making, but, in this setting, patient factors may play a stronger role. Even in cases of therapeutic uncertainty, patient factors can strongly bias the choice of treatment (e.g. preference for operative versus non-operative management, medical comorbidity, socioeconomic factors). In this setting, execution of a credible observational study requires great care in topic selection and study design. For instance, a meta-analysis comparing operative to non-operative treatment of displaced proximal humerus fractures found no difference in functional results with either treatment, and the effect sizes of observational studies were similar to those of RCTs.[13] Critical analysis demonstrates why the observational results were credible in this type 3 setting (i.e. fulfilling Vandenbroucke's restriction criteria): (1) There was therapeutic uncertainty in selection of treatment, (2) Treatment assignment was largely based on quasi-random allocation to a surgeon whose treatment selection was based on personal preference, and (3) Patient-related factors were balanced between the two groups. Similar conclusions could be drawn from a meta-analysis comparing non-operative and operative treatment for multiple rib fractures or flail chest and a recently published meta-analysis on treatment for Achilles tendon ruptures[27,28]

Another study example of a type 3 comparison is the comparison of intracranial pressure monitoring and targeted intensive care on functional outcome after severe head injury compared to supportive intensive care.[29] This oft cited (retrospective) observational study, with a clear uncertainty of superiority of one treatment over the other, made use of geographic randomization and patient related factors were similar across both treatment groups (the three restrictions of Vandenbroucke). In this study, no benefit was demonstrated of one treatment option over the other. This was confirmed in an RCT that was subsequently published, further supporting the validity of an observational design in this setting.[30] Using these criteria, we suggest that multiple recent observational studies are likely to be as credible as RCTs based on their research design and study topic.[31,32] Still, great care must be exercised when designing observational experiments for type 3 studies. For example, the

treatment of calcaneal fractures is highly influenced by patient characteristics and an observational study comparing operative versus non-operative treatment of calcaneal fractures would be more prone to bias because patient risk factors for wound complications after operative intervention (obesity, smoking) are the most significant factors in selection of non-operative treatment.[33] In this example, the observational study's findings are unlikely to be as credible as an RCT's results.

Summary

While the prevailing belief is that observational studies of therapeutic efficacy are credible only in exceptional circumstances due to unmeasured confounding, we highlight several instances where observational studies in orthopaedic trauma surgery are likely to be as credible as RCTs.[3,9] Surgical RCTs are challenging to design and can be misleading when poorly executed. Not all research topics need to be investigated by an RCT, and in many cases, observational studies in trauma surgery can provide credible results at lower cost with a more representative set of patients.[21] In general, RCTs tend to have better compliance rates and (potentially) allow for blinding. Both issues are relevant for a type 1 comparisons, but, for type 2 comparisons, their role is much smaller because compliance with an operation is often quite good and blinding is not present in practice either. For type 3 comparisons, these issues are dependent on the research topic and study design so general statements are more difficult. We acknowledge that the greatest threat to the validity of observational trials of treatment effectiveness is the potential for unmeasured bias, but this potential is context- and comparison-dependent and should not be presumed to be universal. In orthopaedic trauma surgery, with careful planning and execution, an observational study can, to the least, add valuable information to that obtained from an RCT and sometimes even be as informative as an RCT.

References

1. 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. doi:10.1056/NEJM200006223422507.

2. Chalmers I. Why transition from alternation to randomisation in clinical trials was made. *BMJ*. 1999;319(7221):1372.

3. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *BMJ*. 2002;324(7351):1448-1451.

4. Sibai T, Carlisle H, Tornetta P. The darker side of randomized trials: recruitment challenges. *J Bone Joint Surg Am*. 2012;94 Suppl 1(Suppl 1):49-55. doi:10.2106/JBJS.L.00240.

5. Chan S, Bhandari M. The Quality of Reporting of Orthopaedic Randomized Trials with Use of a Checklist for Nonpharmacological Therapies. *J Bone Jt Surg*. 2007;89(9):1970. doi:10.2106/JBJS.F.01591.

6. James MA. Insufficient Post Hoc Statistical Power: A Potential Pitfall of a Well-Designed Randomized Controlled Surgical Trial: Commentary on an article by Geert A. Buijze, MD, PhD, et al.: "Three-Dimensional Compared with Two-Dimensional Preoperative Planning of Corrective Osteotomy for Extra-Articular Distal Radial Malunion. A Multicenter Randomized Controlled Trial". *J Bone Joint Surg Am*. 2018;100(14):e98. doi:10.2106/JBJS.18.00256.

7. Buijze GA, Leong NL, Stockmans F, et al. Three-Dimensional Compared with Two-Dimensional Preoperative Planning of Corrective Osteotomy for Extra-Articular Distal Radial Malunion. *J Bone Jt Surg*. 2018;100(14):1191-1202. doi:10.2106/JBJS.17.00544.

8. Lurati Buse GAL, Botto F, Devereaux PJ. Revisiting sample size: are big trials the answer? *J Bone Joint Surg Am*. 2012;94 Suppl 1(Suppl 1):75-79. doi:10.2106/JBJS.K.01270.

9. Benson K, Hartz AJ. A Comparison of Observational Studies and Randomized, Controlled Trials. *N Engl J Med*. 2000;342(25):1878-1886. doi:10.1056/NEJM200006223422506.

10. Balogh ZJ, Martin AB. Prospective cohorts and risk adjusted outcomes for trauma. *Injury*. 2010;41:S24-S26. doi:10.1016/j.injury.2010.03.034.

11. Ioannidis JPA, Haidich A-B, Pappa M, et al. Comparison of Evidence of Treatment Effects in Randomized and Nonrandomized Studies. *Jama*. 2001;286(7):821-830. doi:10.1001/jama.286.7.821.

12. 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 Sport Med*. 2016:363546516673615. doi:10.1177/0363546516673615.

13. 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;0(0). doi:10.1016/j.jse.2018.03.009.

14. 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 Elb Surg*. 2016;25(7):1195-1203. doi:10.1016/j.jse.2016.01.018.

15. GBD 2013 DALYs and HALE Collaborators CJL, Murray CJL, Barber RM, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: quantifying the epidemiological transition. *Lancet (London, England)*. 2015;386(10009):2145-2191. doi:10.1016/S0140-6736(15)61340-X.

16. 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.

17. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355(7040):i4919. doi:10.1136/bmj.i4919.

18. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley; 2011.

19. Vandenbroucke JP. When are observational studies as credible as randomised trials? *Lancet*. 2004;363(9422):1728-1731. doi:10.1016/S0140-6736(04)16261-2.

20. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med*. 2008;5(3):e67. doi:10.1371/journal.pmed.0050067.

21. Concato J, Horwitz RI. Beyond randomised versus observational studies. *Lancet (London, England)*. 2004;363(9422):1660-1661. doi:10.1016/S0140-6736(04)16285-5.

22. Chan DS, Garland J, Infante A, Sanders RW, Sagi HC. Wound Complications Associated With Bone Morphogenetic Protein-2 in Orthopaedic Trauma Surgery. *J Orthop Trauma*. 2014;28(10):599-604. doi:10.1097/BOT.0000000000000117.

23. James AW, LaChaud G, Shen J, et al. A Review of the Clinical Side Effects of Bone Morphogenetic Protein-2. *Tissue Eng Part B Rev*. 2016;22(4):284-297. doi:10.1089/ten.TEB.2015.0357.

24. Lemans JVC, Wijdicks SPJ, Boot W, et al. Intrawound Treatment for Prevention of Surgical Site Infections in Instrumented Spinal Surgery: A Systematic Comparative Effectiveness Review and Meta-Analysis. *Glob Spine J*. July 2018:219256821878625. doi:10.1177/2192568218786252.

25. Nowak LL, Vicente MR, McKee MD, Hall JA, Nauth A, Schemitsch EH. Orthopaedic surgeons' opinions surrounding the management of proximal humerus fractures: an international survey. *Int Orthop*. 2017;41(9):1749-1755. doi:10.1007/s00264-017-3569-0.

26. Little Z, Smith TO, McMahon SE, et al. The treatment of segmental tibial fractures: does patient preference differ from surgeon choice? *Injury*. 2017;48(10):2306-2310. doi:10.1016/j.injury.2017.08.014.

27. 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*. 2018;0(0):0. doi:10.1007/s00068-018-1020-x.

28. 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.

29. Cremer OL, van Dijk GW, van Wensen E, et al. Effect of intracranial pressure monitoring and targeted intensive care on functional outcome after severe head injury. *Crit Care Med*. 2005;33(10):2207-2213.

30. Chesnut RM, Temkin N, Carney N, et al. A Trial of Intracranial-Pressure Monitoring in Traumatic Brain Injury. *N Engl J Med*. 2012;367(26):2471-2481. doi:10.1056/NEJMoa1207363.

31. Wlodarczyk JR, Thomas AS, Schroll R, et al. To shunt or not to shunt in combined orthopedic and vascular extremity trauma. *J Trauma Acute Care Surg*. 2018;85(6):1038-1042. doi:10.1097/TA.0000000000002065.

32. Pieracci FM, Lin Y, Rodil M, et al. A prospective, controlled clinical evaluation of surgical stabilization of severe rib fractures. *J Trauma Acute Care Surg*. 2016;80(2):187-194. doi:10.1097/TA.0000000000000925.

33. Abidi NA, Dhawan S, Gruen GS, Vogt MT, Conti SF. Wound-Healing Risk Factors After Open Reduction and Internal Fixation of Calcaneal Fractures. *Foot Ankle Int*. 1998;19(12):856-861. doi:10.1177/107110079801901211.

CHAPTER 10

STUDY PROTOCOL FOR A MULTICENTER PROSPECTIVE COHORT STUDY OF NONOPERATIVE VERSUS OPERATIVE TREATMENT FOR FLAIL CHEST AND MULTIPLE RIB FRACTURES AFTER BLUNT THORACIC TRAUMA.

R.B. Beks, M.B. de Jong, A.A.R Sweet,
J. Peek, B. van Wageningen, T.N. Tromp,
F. IJpma, R. Wouters, K.W. Lansink, M. Bemelman,
M.C. van Baal, J. Hoogendoorn, T.P. Saltzherr,
R.H.H. Groenwold, L.P.H. Leenen, R.M. Houwert

Submitted

Abstract

Introduction

A trend has evolved towards rib fixation for flail chest although evidence is limited. Little is known about rib fixation for multiple rib fractures without flail chest. The aim of this study is to compare rib fixation with nonoperative treatment for both patients with flail chest or multiple rib fractures.

Methods and analysis

In this multicenter prospective cohort study all patients with ≥ 3 rib fractures admitted to one of the five participating centers will be included. In two centers rib fixation is performed and in three centers nonoperative treatment is the standard-of-care for flail chest or multiple rib fractures. The primary outcome measures are intensive care unit length of stay and hospital length of stay for patients with a flail chest and patients with multiple rib fractures, respectively. Propensity score matching will be used to control for potential confounding of the relation between treatment modality and length of stay. All analyses will be performed separately for patients with flail chest and patients with multiple rib fractures without flail chest.

Ethics and Dissemination

The regional Medical Research Ethics Committee UMC Utrecht approved a waiver of consent (reference number WAG/mb/17/024787) and confirmed that the Medical Research Involving Human Subjects Act (WMO) does not apply to this study and therefore an official approval of this study is not required under the WMO (METC protocol number 17-544/C). Before discharge, patients will be asked to participate in this follow-up study by the supervising doctor of the surgical ward and written informed consent will be obtained in order to follow-up on the patient and use individual patient data. Study results will be submitted for peer review publication.

Introduction

Thoracic injury is currently the second leading cause of trauma-related death and rib fractures are the most common of these injuries.[1] Mortality rates after rib fractures are around 10% with higher rates observed in the elderly trauma patient.[2–4] Flail chest, which is defined as fractures of three or more adjacent ribs in at least two places, is associated with an even higher mortality rate and significant morbidity.[5,6] Rib fractures after chest wall trauma are usually accompanied by internal thoracic injuries including pulmonary pathology.[4] Pain associated with rib fractures can lead to inadequate ventilation and ineffective clearance of secretions resulting in atelectasis. Consequently, there is a high risk of superinfection leading to pneumonia and prolonged mechanical ventilation.[4]

Nonoperative treatment for multiple rib fractures and flail chest has been the gold standard for the past decades and consists of (non) invasive ventilation and pain management. Recently, a trend towards operative treatment of flail chest has evolved in part due to commercially available rib fixation systems and consensus statements.[7,8] Several randomized controlled trials (RCT) on patients with flail chest showed reduced duration of hospital length of stay (HLOS), intensive care unit length of stay (ILOS), days on mechanical ventilation (DMV), mortality rate, pneumonia rate, and treatment costs. Although, a recent meta-analysis showed that only 3 RCTs were available on this subject including a total of just 123 patients.[9] For multiple rib fractures, very few studies exist investigating the effects of rib fixation and the indication for surgery remains unclear.[9]

More research is needed on different treatment modalities since evidence remains inconclusive, which is illustrated by the large variation in treatment choices between different trauma centers. Recent meta-analyses in orthopaedic trauma surgery suggested that evidence from high-quality observational studies provide valuable insight, which may complement knowledge obtained through RCTs.[10,11] This has contributed to the growing evidence of the potential of observational studies in orthopaedic trauma surgery. Results from thoroughly designed observational studies in representative cohorts, in which participation rates among patients as well as surgeons is relatively high, may provide detailed information about infrequent complications and outcomes.

The aim of this large multicenter prospective cohort study is to compare rib fixation with nonoperative treatment for flail chest or multiple rib fractures by evaluating treatment effects in a representative patient population treated in different level 1 trauma centers.

Methods and analysis

This will be a large multicenter prospective non-randomized cohort study with the following participating centers: University Medical Center Utrecht (UMCU), Radboud University Medical Center (RUMC), University Medical Center Groningen (UMCG), Elisabeth-Tweesteden Ziekenhuis (ETZ), and Haaglanden Medical Center (HMC). All these hospitals are level 1 trauma centers and have a similar volume of trauma patients admitted to the emergency department with comparable injury severity.[12]

Research questions

1) What is the short-term outcome after rib fixation for flail chest or multiple rib fractures compared to nonoperative treatment?
2) What is the long-term outcome after rib fixation for flail chest or multiple rib fractures compared to nonoperative treatment?
3) What are surgery related complications after rib fixation?

Participants selection

Inclusion criteria

All admitted adult (18 years and older) patients presenting at the Emergency Department (ED) of the participating hospitals with a CT scan confirmed flail chest or multiple rib fractures after blunt thoracic trauma will be enrolled in this cohort study. Flail chest is defined as three or more adjacent ribs fractured in at least two places leading to paradoxical breathing. Multiple rib fractures are defined as three or more ipsilateral rib fractures without paradoxical breathing. Patients will be enrolled from January 1st, 2018, onwards.

Exclusion criteria

Patients with non-traumatic rib fractures or rib fractures as a result of cardiopulmonary resuscitation will be excluded.

Intervention

In all participating centers, a multidisciplinary team following the ATLS guidelines will provide trauma care. Primary and secondary surveys are performed during initial trauma care in the Emergency Department. All patients will be treated according to the standard of care of the hospital of admission which will be determined by the treating surgeon. Rib fixation is current practice in the UMCU and ETZ. The UMCG, RUMC and HMC will perform nonoperative treatment as standard care.

Nonoperative treatment

Nonoperative treatment consists of adequate pain management, supportive mechanical ventilation when indicated, and physiotherapy for breathing exercises according to standard national guidelines.

Rib fixation

In the participating centers where rib fixation is performed, treatment will be according to the algorithm presented in Figure 1. In those centers, flail chest or a severe thorax deformity are strict indications for rib fixation. Intractable pain at the fracture site(s) insufficiently manageable with epidural or intravenous analgesia is an indication for rib fixation as well. Rib fixation will be performed or supervised by senior orthopaedic trauma surgeons experienced with surgical treatment of rib fractures. Preoperative planning of the procedure will be done using chest computed tomography (CT) with 3D reconstructions. Preoperative antibiotic prophylaxis (2 grams of Cefazolin) will be administered intravenously in all patients. Depending on the site of the fractures, patients are positioned in the supine, lateral or prone position. The surgical approach is performed as described by Taylor.[5] After reduction, internal fixation using the MatrixRIB™ system (Depuy Synthes®, Amersfoort, The Netherlands) will be performed. Fixation is preferably done with 3 bicortical screws on each side of the fracture. If plate fixation is not possible due to anatomical boundaries and rib fixation is deemed necessary, splints will be used. The number of fixed ribs will be at the discretion of the surgeon, and depended on anatomical boundaries and the possibility to regain stability of the chest wall during respiration. Tube thoracostomy will be performed in case of pneumothorax and/or hemothorax at initial presentation or clinical suspicion of pneumothorax during surgery at the discretion of the surgeon. Postoperative chest radiography will be performed in all patients to document surgical result and to rule out early complications. Patients will be encouraged to mobilize as soon as possible with the help of physiotherapy and adequate pain management.

Outcome measures

Primary Objectives

- Number of days in the ICU for patients with flail chest
- Hospital length of stay for patients with multiple rib fractures

Secondary Objectives:

Short-term health related outcome measures

- Pneumonia rate
- Number of days in the ICU for patients with multiple rib fractures
- Number of days on mechanical ventilation
- Need for tracheostomy
- In-hospital complication rate
- Hospital length of stay for patients with a flail chest
- In-hospital mortality
- Pain with breathing on day 3, 5, and 7 after admission
- Pain with coughing on day 3, 5, and 7 after admission
- General pain on day 3, 5, and 7 after admission

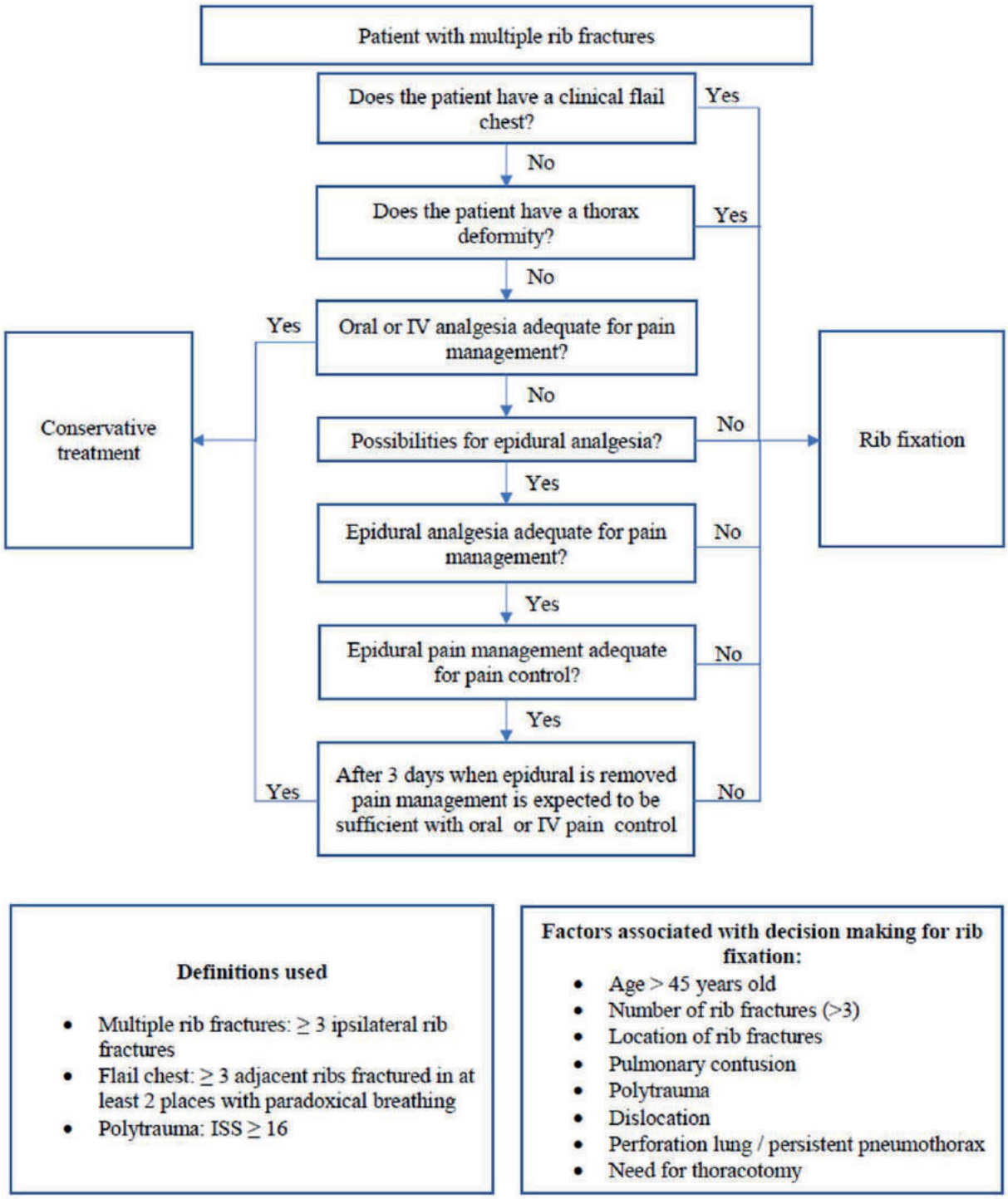


Fig. 1 Treatment algorithm for patient with flail chest or multiple rib fractures

Long-term outcome measures
- Pain with breathing/coughing
- Quality of life
- Dyspnea burden
- Health costs
- Return to work

Surgery related outcome measures
- Incidence of surgical complications
- Implant removal rate

Conservative related outcome measures
- Incidence of pseudoarthrosis
- Late rib fixation (more than 10 days after trauma)

Data collection & patient follow-up

Measurements of explanatory and outcome variables will be obtained from the electronic medical patient files, from each of the participating centers. No extra information will be collected during hospital stay. We will appoint a project leader in every participating institution who will have access to Research Online for Researchers to enable centralized tracking of potentially eligible patients in compliance with the data management policy of the UMCU and the GCP guidelines for electronic data collection. All data will be stored in a research folder that can only be accessed by the principal investigator and the project leaders. Data will be pseudonymized and a file to decode this data will be stored in a research folder only accessible by the principal investigator and project leaders.

Outpatient measurements

A standard outpatient department visit will be scheduled six weeks after discharge. During this visit the EQ5D-5L, modified Medical Research Council Dyspnea Scale (mMRC), the general numerical rating scale (NRS), the NRS with breathing and coughing, and the vital capacity will be assessed. Visits after 12 weeks and 6 months will be scheduled on indication. Patients will be asked for a one year follow-up telephone interview to reassess EQ5D-5L, mMRC, general NRS, NRS during breathing and coughing, and additionally to evaluate implant removal (appendix). The EQ5D-5L is a standardized instrument for generic health status measurements to assess quality of life.[13] The mMRC is a five-category scale that characterizes the level of dyspnoea with physical activity.[14]

Cost analysis

Cost analysis will be performed from a provider perspective and includes only costs made during the index hospital admission, defined as admission for multiple rib fractures or flail chest after blunt thoracic trauma. Additional costs prior to admission or after discharge will

not be taken into account. Reference prices will be derived from a national ledger for standardization of healthcare costs and calculated using activity based costing to measure costs for surgery and hospital day unit costs.[15] Costs of ICU stay and general ward costs will be defined separately. Hospital day unit costs include costs for physician care, nursing, materials, medication, writing-off equipment, housing and other overhead costs. Surgery costs include specialist' fee, costs of personnel, equipment, materials and other overhead costs. Volumes of blood products, radiology, laboratory tests, physiotherapy and consultation of other specialties will be extracted from the medical file. Total costs will be calculated as the summed product of volumes and resources used and their corresponding unit costs. Because costs between the different participating hospitals can differ due to contracts with different suppliers of materials and equipment, and the fact that healthcare reimbursements in the Netherlands are based upon agreements between individual hospitals and insurance companies, costs will be calculated as if all patients are treated in the same hospital (UMCU).

Sample size considerations

Previous studies show a mean duration of ICU stay of 14 days for patients who had nonoperative treatment for a flail chest with a standard deviation (SD) of 4.5.16,17 To detect three days difference in ICU stay with a power of 80% and two-sided alpha of 0.05, accounting for 15% loss after propensity score matching, the total number of patients needed is approximately 82.

Based on previous studies and our own unpublished retrospective data the mean duration of hospital stay for conservatively treated patients with multiple rib fractures is 15 days with a SD of 6.3.18–20 To detect a difference of three days hospital stay difference with a power of 80% and an alfa of 0.05, accounting for 15% loss after propensity score matching, a total of approximately 160 patients are needed.

This study is expected to take four years to reach the required sample size based on retrospective data on rib fracture patients in the participating centers. An additional year is needed to collect sufficient follow-up data on all patients.

Statistical analysis

All analyses will be performed separately for the group of patients with a flail chest and the group of patients with multiple rib fractures.

Baseline characteristics will be presented as means (standard deviation) for continuous variables and proportions for categorical variables. Differences in distributions of baseline characteristics between the study groups will be quantified by means of standardized differences.21

We will apply multiple imputation to impute missing values for baseline characteristics using the mice() algorithm in R.22

To control for potential confounding, we will conduct propensity score (PS) matching. The PS will be estimated using binary logistic regression analysis, with rib fracture fixation as the dependent variable and age, gender, ASA-score, trauma mechanism, AIS head, ISS, number of rib fractures, number of rib fractures in flail chest, and concomitant injuries will be included as covariates in the model. A 1:1 nearest neighbor matching will be performed, with a maximum caliper of 0.15 of the standard deviation of the logit of the PS using the Matchit() algorithm in R.23 After PS matching, the distributions of baseline characteristics will be compared between the study groups and quantified using standardized differences.

The primary analysis will be conducted within the dataset of PS matched subjects. For the primary analysis, the relation between rib fracture fixation and the number of days on mechanical ventilation will be assessed by means of Poisson regression, while pain scores will be compared by means of linear regression.
Sensitivity and subgroups analysis

As a sensitivity analysis, the above analyses will be repeated, but in subgroups of (i) patients without a CT confirmed pulmonary contusion and (ii) patients with epidural anesthesia. This study is expected to require four years of patient enrollment, therefore an additional sensitivity analysis will be performed dividing patients in year of treatment to study outcome measures over time.

Furthermore, to explore potential differences in outcome measures because of local differences in treatment regimens between hospitals, a subgroup analysis will be performed for each hospital. Additional analyses will be performed to adjust for clustering by hospital. Finally, to investigate the effect of timing of rib fixation on the outcome measures, subgroup analysis will be performed dividing patients between early rib fixation (within 48 hours) and late rib fixation (after 48 hours).

Ethics and dissemination

The study protocol was evaluated by The Medical Research Ethics Committee (METC). The METC approved a waiver of consent (reference number WAG/mb/17/024787) and confirmed that the Medical Research Involving Human Subjects Act (WMO) does not apply to this study and therefore an official approval of this study by the METC UMC Utrecht is not required under the WMO (METC protocol number 17-544/C). Before discharge, patients will be asked to participate in this follow-up study by the supervising doctor of the surgical ward. All patients will be given six weeks to consider their decision as after six weeks the outpatient department visit is planned. Subjects can leave the study at any time for any reason if they

wish to do so without any consequences. Study results will be submitted for peer review publication.

Patient and public involvement

There was no patient or public involvement in the development of this study.

Trial registration number

This study has been registered with the Netherlands Trial Register (NTR number: NTR6833) and can be viewed via: http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=6833

Appendices

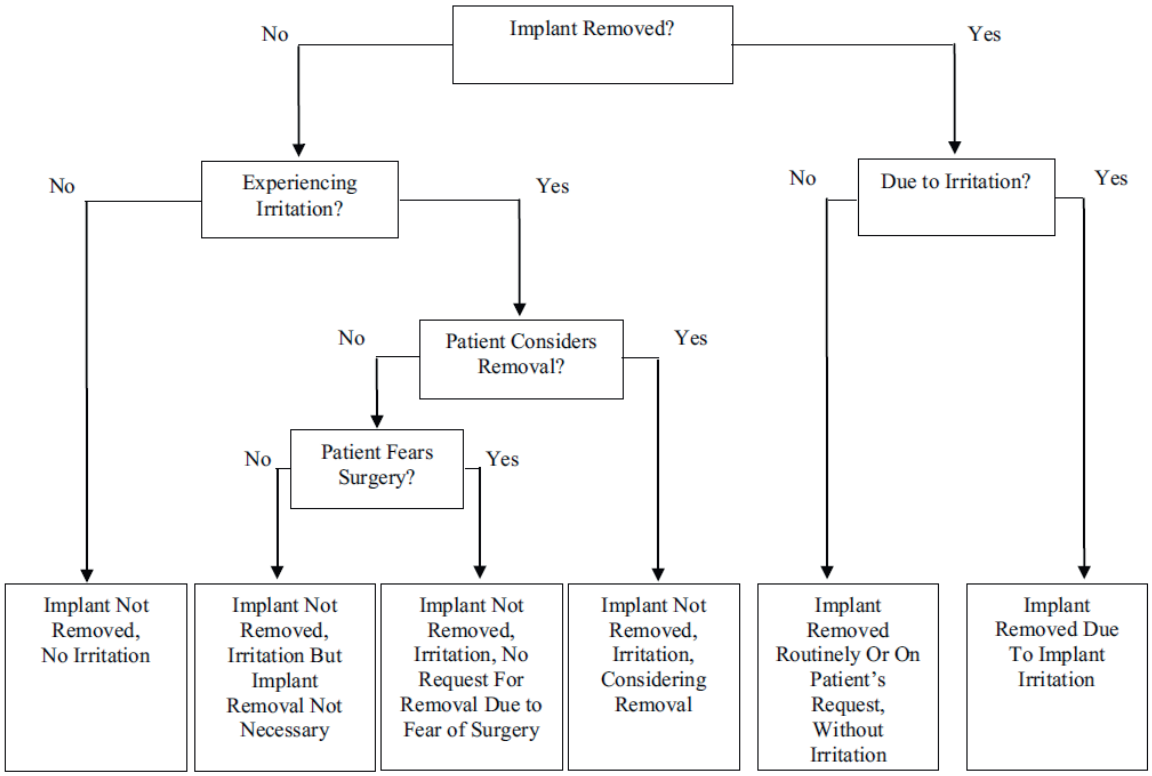
Validated Dutch versions of these questionnaires will be used as this research will be conducted in the Netherlands

Appendix 1. Modified Medical Research Council Dyspnoea Scale

Modified Medical Research Council Dyspnoea Scale (mMRC)

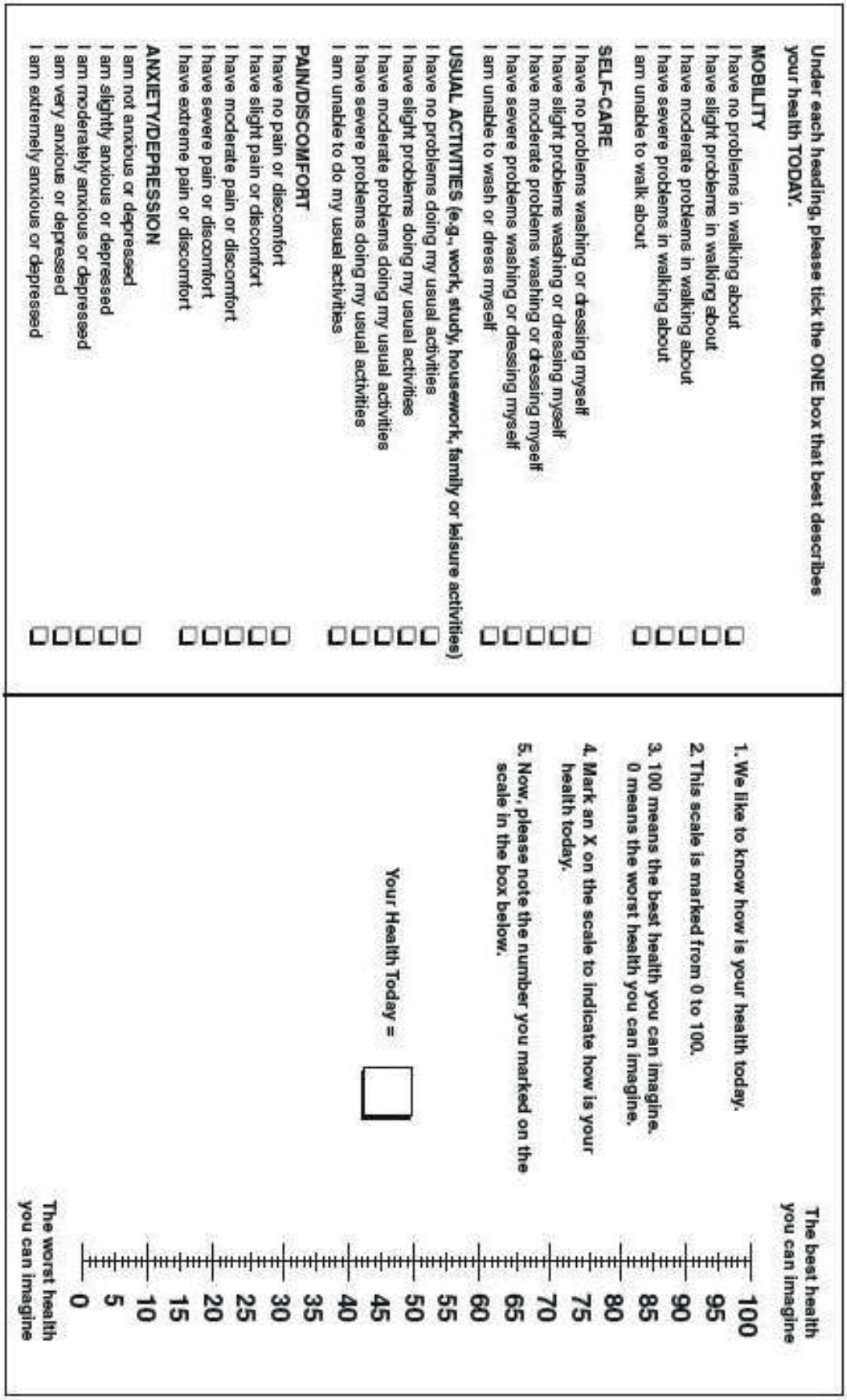
Grade	Patients description of breathlessness
mMRC 0	I only get breathless with strenuous exercise
mMRC 1	I get short of breath when hurrying on the level or walking up a slight hill
mMRC 2	I walk slower than people of the same age on the level because of breathlessness or have to stop for breath when walking on my own pace on the level
mMRC 3	I stop for breath after walking about 100 meters or after a few minutes on the level
mMRC 4	I am too breathless to leave the house or I am breathless when dressing

Appendix 2. Implant-related irritation and removal questionnaire.



A schematic representation of the questions regarding implant-related irritation and removal is shown. Patients who had other causes of irritation such as infection or nonunion should be disregarded.[14]

Appendix 3. EQ-5D 5L – Quality of Life questionnaire



References

1. Vana PG, Neubauer DC, Luchette FA. Contemporary management of flail chest. *The American surgeon*. 2014;80(6):527-535.
2. Bulger EM, Arneson MA, Mock CN, Jurkovich GJ. Rib fractures in the elderly. *J Trauma*. 2000;48(6):1040-1047.
3. Ziegler DW, Agarwal NN. The morbidity and mortality of rib fractures. *The Journal of trauma*. 1994;37(6):975-979.
4. Lin FC-F, Li R-Y, Tung Y-W, Jeng K-C, Tsai SC-S. Morbidity, mortality, associated injuries, and management of traumatic rib fractures. *Journal of the Chinese Medical Association*. 2016;79(6):329-334. doi:10.1016/j.jcma.2016.01.006.
5. Cannon RM, Smith JW, Franklin GA, Harbrecht BG, Miller FB, Richardson JD. Flail chest injury: are we making any progress? *Am Surg*. 2012;78(4):398-402.
6. Dehghan N, De Mestral C, McKee MD, Schemitsch EH, Nathens A. Flail chest injuries: A review of outcomes and treatment practices from the national trauma data bank. *Journal of Trauma and Acute Care Surgery*. 2014;76(2):462-468. doi:10.1097/TA.0000000000000086.
7. Pieracci FM, Majercik S, Ali-Osman F, et al. Consensus statement: Surgical stabilization of rib fractures rib fracture colloquium clinical practice guidelines. *Injury*. 2017;48(2):307-321. doi:10.1016/j.injury.2016.11.026.
8. Kane ED, Jeremitsky E, Pieracci FM, Majercik S, Doben AR. Quantifying and exploring the recent national increase in surgical stabilization of rib fractures. *The journal of trauma and acute care surgery*. 2017;83(6):1047-1052. doi:10.1097/TA.0000000000001648.
9. Kasotakis G, Hasenboehler EA, Streib EW, et al. Operative fixation of rib fractures after blunt trauma: A practice management guideline from the Eastern Assocaiation for the Surgery of Trauma. *Journal of Trauma and Acute Care Surgery*. 2017;82(3):618-626. doi:10.1097/TA.0000000000001350.
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. *Am J Sports Med*. 2016. doi:10.1177/0363546516673615.
11. 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. doi:10.1016/j.jse.2016.01.018.
12. Berden HJJ, Leenen LPH. Landelijke Traumaregistratie. 2015.
13. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x.
14. Mahler DA, Wells CK. Evaluation of clinical methods for rating dyspnea. *Chest*.

1988;93(3):580-586.
15. Hakkaart-van Roijen L, van der Linden N, Bouwmans C, Kanters T, Swan Tan S. Kostenhandleiding: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. *Zorginstituut Nederland*. 2016:1-73.
16. Qiu M, Shi Z, Xiao J, Zhang X, Ling S, Ling H. Potential Benefits of Rib Fracture Fixation in Patients with Flail Chest and Multiple Non-flail Rib Fractures. *The Indian journal of surgery*. 2016;78(6):458-463. doi:10.1007/s12262-015-1409-2.
17. Zhang X, Guo Z, Zhao C, Xu C, Wang Z. Management of patients with flail chest by surgical fixation using claw-type titanium plate. *Journal of cardiothoracic surgery*. 2015;10:145. doi:10.1186/s13019-015-0363-1.
18. Wu WM, Yang Y, Gao ZL, Zhao TC, He WW. Which is better to multiple rib fractures, surgical treatment or conservative treatment? *Int J Clin Exp Med*. 2015;8(5):7930-7936.
19. Velasquez M, Ordonez CA, Parra MW, Dominguez A, Puyana JC. Operative versus Nonoperative Management of Multiple Rib Fractures. *The American surgeon*. 2016;82(5):103-105.
20. Fitzgerald MT, Ashley DW, Abukhdeir H, Christie DB. Rib fracture fixation in the 65 years and older population. *Journal of Trauma and Acute Care Surgery*. 2017;82(3):524-527. doi:10.1097/TA.0000000000001330.
21. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behav Res*. 2011;46(3):399-424. doi:10.1080/00273171.2011.568786.
22. Buuren S van, Groothuis-Oudshoorn K. mice : Multivariate Imputation by Chained Equations in *R*. *Journal of Statistical Software*. 2011;45(3):1-67. doi:10.18637/jss.v045.i03.
23. Ho D, Imai K, King G, Stuart E. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis*. 2007.

CHAPTER 11

GENERAL DISCUSSION

This chapter aims to aggregate different important aspects of the previous chapters into one overview. A brief summary is given of different study designs for intervention studies in trauma patients and their potential for confounding and bias. Methodological differences between observational studies and RCTs are further explored and future perspectives are provided using a new study protocol.

Study designs for intervention studies

Different study designs are being used to assess effectiveness of medical interventions, yet the randomized controlled trial (RCT) is often considered the most valuable.[1] Concato et al., Benson et al., and Ioannidis et al. provide an empirical basis for a comparison between the different designs.[2–4] The majority of the topics they considered included pharmacological treatments, but surgical intervention studies were included as well. Based on their reviews, one could get the impression that results from the different designs are rather similar (Concato and Benson), but also that results can be quite different (Ioannidis). The overall conclusion, based on theoretical rather than empirical considerations, seems to be that RCTs have superior validity (notably due to the randomization), yet in certain situations the different designs may yield actually quite similar results. This is an interesting observation as the use of randomization can be challenging and the RCT does have its downsides.[5] An obvious question that emerges is what are drivers of difference between the different designs and in which situation might an observational study design be considered appropriate as well? Here we address these questions and focus on intervention studies in traumatology patients.

Types of intervention and potential for confounding

A key difference between RCTs and observational studies is the potential for confounding in the latter due to the absence of randomization. In observational research of a medical treatment effect, allocation of treatment follows clinical practice and therefore treated and untreated patients often differ on prognostic characteristics.[1] When comparing groups of treated and untreated patients, such incomparability almost always leads to bias, known as confounding. One may adjust for observed (and validly measured) confounders, however, although several methods have been described to adjust for so-called unobserved confounding (due to unknown or invalidly measured confounders) it can be difficult to adequately adjust for this type of confounding.[6] Particularly the potential for unobserved confounding is considered an important reason for differences in the effects obtained from RCTs and effects obtained from observational studies. The potential for confounding likely depends on the type of intervention that is studied. In intervention studies in trauma surgery patients, three types of comparisons can be distinguished[7,8], which are discussed below and summarized in Table 1.

Type 1 – Comparison of pharmacological intervention

A type 1 comparison focuses on pharmacological interventions in surgical patients, comparing a particular drug with a placebo or active comparator (other drug). The focus of this kind of research is to quantify the effects of a particular compound (efficacy) or a particular treatment strategy (effectiveness). Efficacy is the effect of an intervention under ideal controlled circumstances, whereas effectiveness is the effect of an intervention in daily clinical practice.[9] In daily practice, prescriptions of pharmacological interventions often are based on a clinical indication for that drug. Patient characteristics and severity of disease are important factors in decision making (e.g. about initiating treatment) and consequently who ends up in which treatment group. Observational studies of pharmacological interventions are therefore generally highly susceptible to confounding, which in this context is referred to as confounding by indication. Consequently, it is very challenging to study a type 1 comparison using an observational study design and randomization is usually considered indispensable.

Type 2 – Comparison of different operative intervention

Type 2 comparisons aim at comparing two operative interventions. Hence, patients in both research groups undergo an operative procedure, be it that the actual procedure differs. A particular operative procedure is compared to a control group that undergoes an alternative operative intervention or a sham operation. This type of research aims to study the effectiveness of an operative procedure as it encompasses a course of pre-, peri-, and postoperative treatment and the treatment effects are subject to skill and expertise of the surgical team; treatment strategies, from admission to discharge, are optimized to the hospital and can differ from other hospitals. Surveys amongst surgeons suggest that for common fracture types different treatment strategies exist due to surgeon preference.[10,11] Consequently, when two surgical treatment options are considered, skill and expertise of the surgical team will be the most important factor used in decision making. Patient reference to the surgeon, especially in the trauma setting, is to a large extent a random process, therefore, compared to type 1 research, a comparison of operative interventions is less sensitive to confounding if there are no indications that the surgeon's treatment preference is strongly subject to patient characteristics.[12,13] In other words, a surgeon will treat patients according to his own preference and since ‘allocation of surgeon’ is a rather random process, one can speculate that patient groups that underwent different surgical interventions may actually be rather similar. Indeed, in a recent meta-analysis of RCTs and observational studies comparing plate fixation with intramedullary fixation for midshaft clavicle fractures, treatment groups in observational studies were surprisingly similar and consequently effect estimates for the primary and secondary outcome measures were comparable between observational studies and randomized trials.[14] Although a type 2 comparison in trauma surgery will generally involve a surgical procedure, type 2 comparisons might focus on an single aspect of the treatment strategy such as post-operative treatment after the surgical

procedure; e.g. cast immobilization compared to other post-operative treatment protocols after surgical treatment of ankle fractures[15].

Type 3 – Comparison of operative and nonoperative intervention

In type 3 comparisons, an operative intervention is compared to a nonoperative intervention, for example surgery versus conservative treatment in patients with a particular fracture (see also **chapters 7 & 8)**. Similar to a type 2 comparison, the focus is on the effectiveness of the treatment. Again a surgeon's preference is an important factor in decision making, however, type 3 comparison may have a greater risk of confounding because patient characteristics may have a greater influence on the decision between an operation or conservative treatment than on the decision between two operative treatments. Still, for many fracture types there is discussion on whether to operate or not and one could argue decision making is based on surgeon preferences or hospital guidelines rather than patient characteristics; depending on hospital expertise and the surgical team, treatment can be different between different hospitals. For certain type 3 comparisons an observational study design may therefore be suitable. Similar to type 2 research, multiple recent meta-analyses where surgery was compared to conservative treatment, demonstrated treatment groups that were similar which resulted in comparable effect estimates from observational studies and randomized trials.[16,17] Selected type 3 studies can produce patient cohorts that are comparable if the surgeon's preference for one treatment over the other is of greater magnitude than influences of unmeasured patient characteristics.

Table 1. Overview of different research types in trauma surgery

	Type 1 comparison	Type 2 comparison	Type 3 comparison
Comparison	Pharmacological intervention versus placebo / other drug	Operative intervention versus alternative operative intervention	Operative intervention versus nonoperative intervention
Treatment choice	Mostly dependent on clinical indication and patient characteristics	Mostly dependent on surgeon preference	Mostly dependent on surgeon preference and patient characteristics
Focus	Efficacy or effectiveness	Effectiveness	Effectiveness
Comparability of intervention groups	Most likely not comparable	Possibly comparable	Possibly comparable
Risk of confounding	+++	+/-	+
Observational design	Not recommended	To be considered	To be considered

Representativeness

To what extent results from a study are representative to a larger group of patients (now and in the future) depends at least on three factors: 1) presence of effect modifiers, 2) representativeness of study patients and 3) representativeness of study setting.

An effect modifier is a certain (patient) characteristic that influences the magnitude of a treatment effect. When an effect modifier is not accounted for in the statistical analysis, the (average) effect obtained from a particular study will depend on the distribution of the effect modifier.[18] For example, if the effect of an operative intervention depends on the age of the patient, such that the intervention is more effective in young adults, the overall average effect size obtained from a study of that intervention depends on the age distribution of study participants; a study with predominantly young patients probably will show a larger effect than a study of mostly elderly patients. This does not necessarily cause differences between observational studies and randomized trials, because it is not possible to 'control' for effect modification (e.g. by means of randomization) as often the effect modifier (e.g. age) cannot be allocated randomly.[19] However, compared to observational studies, RCTs tend to have more restrictive in- and exclusion criteria, thereby possibly leading to different distributions of effect modifiers and hence different (average) effects may be observed. Note that in the absence of effect modification, restrictive in- and exclusion criteria will not affect the (average) treatment effect observed.

A representative patient sample reflects the target patient population in a particular setting and can be obtained via, e.g., a random selection of patients called random sampling. As effect modifiers are distributed in the study sample as they are in the population of interest, effect modification will affect both in the same way. Treatment effects found in different studies that are all representative of the same population are therefore equally affected by effect modification (i.e., effect modification is not an explanation for any observed differences in effects obtained from different studies).

Representativeness of study setting is of extra importance in studies of surgical interventions, because of the influence on treatment effects of pre-, peri- and postoperative care and expertise of the surgical team. Results from a highly specialized center with a high patient volume are probably less representative for general care. Furthermore, when small patient numbers are studied or when many different surgical teams contribute to the study data, treatment effect estimates might not reflect the real potential of a procedure. One study that illustrates possible limitations in representativeness of surgical care is the FAITH trial that investigated the effect of a sliding hip screw versus cancellous screws.[20] In this trial an impressive number of 1108 patients were randomized. However, with a study period of six years and 81 contributing centers, the average number of included patients per year per center was just over two. Consequently, high treatment variability can be expected as a very

limited number of cases were performed by many different surgeons in different regions with different nuances in local care regimens. Another example is the PROFHER trial that compared nonoperative and operative treatment for proximal humoral fractures: 66 surgeons treated a total 109 patients in 30 different centers.[21] This trial is the largest trial investigating treatment of this difficult fracture and consequently impacted clinical practice.[22] However, one could argue that representativeness of the studied patient sample and representativeness of care due to the low number of patients treated per surgeon and center could affect the generalizability of the results of the study. Observational studies, on the other hand, are possibly more representative of everyday patients and clinical practice, compared to RCTs.[23]

Study Quality

Many sources of bias in either RCTs or observational intervention studies have been identified.[24–26] The Cochrane handbook lists possible sources of bias in a randomized trial: inappropriate random sequence generation, no concealment of allocation, no blinding of participants and personnel, no blinding of endpoint assessment, incomplete endpoint data, and selective reporting. It is obvious that the first three sources of bias are present in observational studies, but to what extent they will impact the validity of the study may differ between different types of interventions (as outlined above).

An empirical basis for the effect of blinding of outcome assessment was made by Pildal et al. who compared RCTs with and without double blinding in a meta-analysis and found an average 9% overestimation of treatment effects in the absence of blinding.[27] Over the years it has been argued that observational studies systematically overestimate treatment effects, compared to RCTs, which could in part be the consequence of the absence of blinding in this study design. However, several meta-analyses concluded that results from observational studies and RCTs are similar and observational studies do not seem to consistently report larger treatment effects.[2,3,28] Karanicolas et al. found that in orthopedic trauma surgery studies blinding of outcome assessment is rarely performed, and is likely the result of the unique situation of surgical trials with a setting of pre-, peri- and postoperative care which makes blinding often not possible.[29] They also found that the majority of the studied outcomes in orthopedic trauma surgery studies have a directly related and objective nature such as mortality, infection, length of stay, fracture union, and adverse events. In another meta-analysis, Wood et al. found that lack of blinding mostly impacted subjective outcome measures, but no difference in treatment effect was seen for objective outcome measures.[30] Incomplete endpoint data are the result of excluded patients, but more importantly from patient drop-out (attrition) and can have important impact on study results.[31]. While observational studies will more often experience missing exposure and covariate data, compared to RCTs, both study designs need to deal with missing outcome data. Different

approaches, such as imputation, are available to deal with missing data and apply to both study designs.[32]

Selective reporting is, for example, a tendency for non-significant results to be withheld from publication by researchers. A subset of the originally recorded variables is presented tailored to the message the authors want to deliver, consequently introducing bias. Several studies published on selective reporting in RCTs and its impact on study results. [33,34] Similarly to incomplete endpoint data, selective reporting could impact both RCTs and observational studies to the same extent and will noticeably impact the results of meta-analyses. However, over recent years an increasing number of journals have made publication of a study protocol mandatory for RCTs. Consequently, deviation from study protocol will need to be justified, partly preventing selective reporting in RCTs. Publication of protocols for observational studies, be it either hypothesis-driven or hypothesis generating studies, is not common practice (yet). To prevent selective reporting, study protocols of observational studies should be published similar to those of RCTs, in order to prevent selective reporting and improve study quality.

Combining information from different study designs

Meta-analyses have shown to be a valuable tool to assess differences in treatment effects. An increasing number of meta-analyses in traumatology topics are published including both observational studies and RCTs (see also **chapters 7 & 8)**. Provided observational studies are of sufficient quality, adding information from observational studies to meta-analyses will increase the number of patients available for analysis and can lead to more precise effects estimates, possibilities for subgroup analysis, and provide more insight in rare outcomes. Obviously, only high quality observational studies should be included for meta-analysis, or at least sensitivity analysis, stratified by study quality, needs to be conducted.

The role of observational studies in trauma surgery

Vandenbroucke proposed a three-pronged restriction to make observational research as methodologically credible as RCTs: (1) Restriction of research topics to those where the intervention allocation is reasonably unrelated to the outcome, (2) Restriction in design to studies where allocation of the intervention is nearly random, and (3) Restriction in analysis to topics where confounding variables can be identified and measured. With the criteria of Vandenbroucke, observational studies in traumatology have the potential to also produce credible results in intervention studies. For example, observational studies have an important role in reporting of adverse events (**chapter 2)**, long term outcomes (**chapters 3 & 4)**, and the effects of new surgical techniques (**chapter 5)**.[35,36]

The example of the OPVENT2 study

An example of an area of research particularly useful for an observational design is rib fracture treatment. Rib fractures are very common in daily clinical practice and nonoperative treatment is the standard of care in many hospitals throughout the world. Over the years surgical rib fixation has gained popularity and three small randomized trials available on this subject show benefit of rib fixation over nonoperative treatment in patients with a flail chest.[37–39] However, more research is needed to optimize surgical indication for patients with multiple rib fractures without flail chest. The OPVENT2 study (presented in chapter 10) aims to answer these questions in a multicenter prospective observational study. We will further exemplify benefits of an observational design using this study protocol.[BMJ OPEN, *under review*]

The three RCTs published on this subject used very strict in and exclusion criteria creating a specific patient subgroup with limited generalizability. Consequently the inclusion period stretched multiple years and patient numbers in each treatment arm were small. In the OPVENT2 every patient with three or more rib fractures is included, except patients transferred to another hospital. This will likely result in a large representative and consecutive patient cohort. In the OPVENT2 patients receive treatment according to the standard of care of the admission hospital. Therefore, the previously discussed surgeon and patient participation rate is expected to be high as compared to randomized studies. Since the OPVENT2 lacks randomization, treatment groups are not automatically comparable in terms of patient characteristics. However, treatment groups are based on geographical location; two centers perform rib fixation following a treatment algorithm and the three other centers perform nonoperative treatment as standard of care. Patient differences between the different geographic locations and therefore differences in treatment groups are expected to be small, yet differences will be addressed in data analysis using propensity score matching. A patient with rib fixation is matched to a patient from the nonoperative patient pool with similar objective baseline characteristics, such as location of rib fractures and concomitant injuries. This study will produce the large patient numbers needed to investigate rib fixation and provide treatment effects. Large patient numbers allow for subgroup analysis and reporting incidence of rare complications, which is important in this very heterogeneous patient population.

Conclusion

Three types of comparisons of interventions can be identified in trauma surgery, of which type 2 and 3 (comparison of surgical intervention and comparison of surgical intervention with nonoperative treatment, respectively) can be studied by using an observational study design. Observational studies in surgery should be assessed based on the type of comparison that is made and its associated potential sources of confounding. Meta-analysis of

observational studies and RCTs in trauma surgery showed that high-quality observational studies indeed yield similar effects as those obtained from RCTs, supporting the idea to consider these observational studies as complementary to RCTs. This has the additional benefit of a better representativeness and increased samples size, which allows for subgroups analysis. In the future, well designed and thoroughly conducted prospective observational studies in trauma surgery, like the OPVENT2 study, can provide valuable results that are relevant to daily clinical trauma surgery practice.

References

1. Vandenbroucke JP. Observational research, randomised trials, and two views of medical science. *PLoS Med*. 2008;5(3):e67. doi:10.1371/journal.pmed.0050067.
2. Concato J, Shah N, Horwitz RI. Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs. *New England Journal of Medicine*. 2000;342(25):1887-1892. doi:10.1056/NEJM200006223422507.
3. Benson K, Hartz AJ. A Comparison of Observational Studies and Randomized, Controlled Trials. *New England Journal of Medicine*. 2000;342(25):1878-1886. doi:10.1056/NEJM200006223422506.
4. Ioannidis JPA, Haidich A-B, Pappa M, et al. Comparison of Evidence of Treatment Effects in Randomized and Nonrandomized Studies. *Jama*. 2001;286(7):821-830. doi:10.1001/jama.286.7.821.
5. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the Gold Standard — Lessons from the History of RCTs. Malina D, ed. *New England Journal of Medicine*. 2016;374(22):2175-2181. doi:10.1056/NEJMms1604593.
6. Klungel OH, Martens EP, Psaty BM, et al. Methods to assess intended effects of drug treatment in observational studies are reviewed. *Journal of Clinical Epidemiology*. 2004;57(12):1223-1231. doi:10.1016/j.jclinepi.2004.03.011.
7. Beks RB, Houwert RM, Groenwold RHH. [Added value of observational studies in surgery: the hierarchical structure of study designs requires a more refined approach]. *Nederlands tijdschrift voor geneeskunde*. 2017;161(0):D1493.
8. McCulloch P, Taylor I, Sasako M, Lovett B, Griffin D. Randomised trials in surgery: problems and possible solutions. *BMJ*. 2002;324(7351):1448-1451.
9. Zuidgeest MG, Goetz I, Grobbee DE. PRECIS-2 in perspective: what is next for pragmatic trials? *Journal of Clinical Epidemiology*. 2017;84:22-24. doi:10.1016/j.jclinepi.2016.02.027.
10. Nowak LL, Vicente MR, McKee MD, Hall JA, Nauth A, Schemitsch EH. Orthopaedic surgeons' opinions surrounding the management of proximal humerus fractures: an international survey. *International Orthopaedics*. 2017;41(9):1749-1755. doi:10.1007/s00264-017-3569-0.
11. Little Z, Smith TO, McMahon SE, et al. The treatment of segmental tibial fractures: does patient preference differ from surgeon choice? *Injury*. 2017;48(10):2306-2310. doi:10.1016/j.injury.2017.08.014.
12. Wijdicks FJ, Houwert M, Dijkgraaf M, et al. Complications after plate fixation and elastic stable intramedullary nailing of dislocated midshaft clavicle fractures: a retrospective comparison. *Int Orthop*. 2012;36(10):2139-2145. doi:10.1007/s00264-012-1615-5.
13. Bhandari M, Guyatt GH, Swiontkowski MF, et al. Surgeons ' Preferences for the. 2011.
14. 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. doi:10.1016/j.jse.2016.01.018.

15. Briet JP, Houwert RM, Smeeing DPJ, et al. Weight bearing or non-weight bearing after surgically fixed ankle fractures, the WOW! Study: study protocol for a randomized controlled trial. *Trials*. 2015;16(1):175. doi:10.1186/s13063-015-0714-1.
16. 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. doi:10.1177/0363546516673615.
17. 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;0(0). doi:10.1016/j.jse.2018.03.009.
18. Lunt M, Solomon D, Rothman K, et al. Different methods of balancing covariates leading to different effect estimates in the presence of effect modification. *American Journal of Epidemiology*. 2009;169(7):909-917. doi:10.1093/aje/kwn391.
19. Groenwold RHH, Donders ART, van der Heijden GJMG, Hoes AW, Rovers MM. Confounding of Subgroup Analyses in Randomized Data. *Archives of Internal Medicine*. 2009;165(16):1917. doi:10.1001/archinte.165.16.1917.
20. Nauth A, Creek AT, Zellar A, et al. Fracture fixation in the operative management of hip fractures (FAITH): an international, multicentre, randomised controlled trial. *The Lancet*. 2017;389(10078):1519-1527. doi:10.1016/S0140-6736(17)30066-1.
21. 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. doi:10.1001/jama.2015.1629.
22. Jefferson L, Brealey S, Handoll H, et al. Impact of the PROFHER trial findings on surgeons ' clinical practice. 2017;6(10):590-599. doi:10.1302/2046-3758.610.BJR-2017-0170.
23. Concato J. Observational Versus Experimental Studies : What ' s the Evidence for a Hierarchy ? 2004;1(July):341-347.
24. 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.
25. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ (Clinical research ed)*. 2016;355(7040):i4919. doi:10.1136/bmj.i4919.
26. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Wiley; 2011.
27. Pildal J, Hróbjartsson A, Jörgensen KJ, Hilden J, Altman DG, Gøtzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *International Journal of Epidemiology*. 2007;36(4):847-857. doi:10.1093/ije/dym087.

28. 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. doi:10.1002/14651858.MR000034.pub2.
29. Karanicolas PJ, Bhandari M, Taromi B, et al. Blinding of Outcomes in Trials of Orthopaedic Trauma: An Opportunity to Enhance the Validity of Clinical Trials. *The Journal of Bone and Joint Surgery-American Volume*. 2008;90(5):1026-1033. doi:10.2106/JBJS.G.00963.
30. Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in. 2001;(November). doi:10.1136/bmj.39465.451748.AD.
31. Groenwold RHH, Moons KGM, Vandenbroucke JP. Randomized trials with missing outcome data: How to analyze and what to report. *Cmaj*. 2014;186(15):1153-1157. doi:10.1503/cmaj.131353.
32. Groenwold RHH, Donders ART, Roes KCB, Harrell FE, Moons KGM. Practice of Epidemiology Dealing With Missing Outcome Data in Randomized Trials and Observational Studies. 2012;175(3):210-217. doi:10.1093/aje/kwr302.
33. Chan A-W, Hrobjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials. 2004;291(20):2457-2465.
34. Chan A, Altman DG. Papers Identifying outcome reporting bias in randomised trials on PubMed : review of publications and survey of authors. 2005;(January):1-6. doi:10.1136/bmj.38356.424606.8F.
35. Hughes RE, Batra A, Hallstrom BR. Arthroplasty registries around the world: valuable sources of hip implant revision risk data. *Current reviews in musculoskeletal medicine*. 2017;10(2):240-252. doi:10.1007/s12178-017-9408-5.
36. Post MWM, Nachtegaal J, van Langeveld SA, et al. Progress of the Dutch Spinal Cord Injury Database: Completeness of Database and Profile of Patients Admitted for Inpatient Rehabilitation in 2015. *Topics in Spinal Cord Injury Rehabilitation*. 2018;24(2):141-150. doi:10.1310/sci2402-141.
37. Marasco SF, Davies AR, Cooper J, et al. Prospective randomized controlled trial of operative rib fixation in traumatic flail chest. *Journal of the American College of Surgeons*. 2013;216(5):924-932. doi:10.1016/j.jamcollsurg.2012.12.024.
38. Granetzny A, Abd El-Aal M, Emam E, Shalaby A, Boseila A. Surgical versus conservative treatment of flail chest. Evaluation of the pulmonary status. *Interactive Cardiovascular and Thoracic Surgery*. 2005;4(6):583-587. doi:10.1510/icvts.2005.111807.
39. Tanaka H, Yukioka T, Yamaguti Y, et al. Surgical stabilization of internal pneumatic stabilization? A prospective randomized study of management of severe flail chest patients. *Journal of Trauma*. 2002;52(4):727-732. doi:10.1097/00005373-200204000-00020.

CHAPTER 12

NEDERLANDSE SAMENVATTING

Introductie

De medische besluitvorming was tot halverwege de vorige eeuw het resultaat van de subjectieve interpretatie en de 'klinische blik' van de dokter. De toenemende kritiek op deze benadering leidde tot de opkomst van *Evidence Based Medicine* (EBM). EBM streeft naar medische besluitvorming op basis van het best beschikbare bewijs van het moment en heeft geresulteerd in een duidelijke hiërarchie binnen de onderzoeksmethoden. Daarin wordt de gerandomiseerde placebo-gecontroleerde trial (RCT) gezien als de gouden standaard.

In een RCT worden patiënten gerandomiseerd tussen een behandel- en controlegroep waarbij zowel de patiënt als de behandelaar wordt geblindeerd voor de behandelkeuze. Zo ontstaan twee vergelijkbare groepen en kunnen waargenomen verschillen in gezondheidsuitkomsten tussen deze onderzoeksgroepen worden toegeschreven aan de behandeling. In observationeel onderzoek naar de effecten van medische behandeling vindt geen randomisatie plaats en is blindering niet mogelijk. Hierdoor is deze onderzoeksopzet gevoeliger voor bias en krijgt observationeel onderzoek doorgaans minder waardering.

De laatste jaren heeft EBM zich in toenemende mate toegespitst op het gebruik van RCTs en worden nauwelijks meer aanbevelingen gedaan op basis van observationeel onderzoek. RCTs hebben een vaste plek gekregen in het geneesmiddelenonderzoek. Ook in chirurgisch onderzoek wordt in toenemende mate gebruik gemaakt van RCTs. Maar juist bij chirurgische onderwerpen zijn er uitdagingen die het uitvoeren van een RCT bemoeilijken. De keuze tussen wel of niet opereren is voor de patiënt veel tastbaarder en inzichtelijker dan de meer abstractere keuze tussen twee geneesmiddelen. De behandelingsvoorkeur van de patiënt en de welwillendheid om de behandeling te laten bepalen door loting, resulteren bij chirurgische RCTs daarmee vaak tot lage aantallen deelnemers. Ook de voorkeur van een chirurg voor een bepaalde behandeling speelt een belangrijke rol. Deze kan afhankelijk zijn van cultuur, persoonlijke ervaring en technische expertise. Als gevolg hiervan is de noodzakelijke equipoise (daadwerkelijke onzekerheid over superioriteit van behandeling A of B en daarmee de verantwoording voor het uitvoeren van randomisatie) in chirurgische RCTs vaak lastiger. Ook zijn er de methodologische uitdagingen die interpretatie en implementatie van de resultaten belemmeren, zoals de vaak lage aantallen deelnemers in chirurgische RCTs als gevolg van de strikte in-/exclusie criteria. Het oplossen van dergelijke methodologische uitdagingen wordt verder bemoeilijkt door de zeer beperkte subsidiemogelijkheden voor (trauma)chirurgisch onderzoek.
Tot slot omvat een operatie, in tegenstelling tot een farmacologische behandeling, een traject van pre-, peri- en postoperatieve handelingen onder invloed van de vaardigheid en expertise van het chirurgisch team. Men kan zich afvragen in hoeverre een dergelijk complex proces zich goed laat onderzoeken met behulp van een RCT dat juist streeft naar maximale controle over alle variabelen die van invloed zijn op de uitkomstmaat.

Observationeel onderzoek is belangrijk bij het evalueren van mogelijk gunstige effecten van operatieve behandelingen en het in kaart brengen van (zeldzame) complicaties. In vergelijking met een RCT zijn resultaten van observationeel onderzoek naar behandelingseffect vaak beter te generaliseren door de minder strikte inclusiecriteria. Ook zijn er meer mogelijkheden tot subgroep analyse en is het onderzoek makkelijker, sneller en met minder geld uit te voeren. Deze voordelen maken observationeel onderzoek een aantrekkelijk alternatief voor een RCT in een medische wereld met snelle technologische vooruitgang, implementatie van nieuwe technieken en beperkte financiële mogelijkheden.

Doelstellingen van het proefschrift

Het doel van dit proefschrift is de waarde van observationeel onderzoek te vergelijken met RCTs in de huidige hiërarchie van onderzoeksmethoden. De volgende vragen komen daarbij specifiek aan bod.

1. Wat zijn de complicaties na een chirurgische techniek? In dit proefschrift is dit onderzocht voor de operatieve behandeling van distale bicepspeesrupturen.
2. Wat zijn de lange termijn uitkomsten na evaluatie van een bestaande chirurgische techniek? In dit proefschrift is dit onderzocht voor patiënten die geopereerd zijn aan een proximale humerusfractuur.
3. Wat zijn de lange termijn uitkomsten na implementatie van een nieuwe chirurgische techniek? In dit proefschrift is dit onderzocht voor ribfixatie.
4. Hoe verhouden de resultaten van een nieuwe operatieve techniek zich ten opzichte van standaard uitgevoerde niet operatieve behandeling? In dit proefschrift is dit onderzocht voor patiënten met een fladderthorax en multipele ribfracturen.
5. Hoe moeten we de waarde van observationele studies interpreteren voor traumapatiënten?
6. Hoe verschillen de uitkomsten van observationele studies en RCTs in traumapatiënten? In dit proefschrift is dat onderzocht voor patiënten met een proximale humerus fractuur, fladderthorax en multipele ribfracturen.
7. Wanneer kan observationeel onderzoek valide informatie geven over effecten van behandeling in traumapatiënten?
8. Hoe ziet toekomstig onderzoek in traumapatiënten eruit?

Resultaten

Observationele klinische studies kunnen op verschillende manieren een bijdrage leveren aan de optimale behandelingsstrategie bij patiënten die een trauma hebben doorgemaakt. Zo kan gemakkelijk inzicht worden verkregen in mogelijke (zeldzame) complicaties en factoren die het optreden van complicaties voorspellen (**hoofdstuk 2**). Observationele studies kunnen verder een belangrijk rol vervullen bij het in kaart brengen van lange-termijn follow-up en kwaliteit van leven van de patiënt na een behandeling (**hoofdstuk 3 & 4**). Als laatste kan met

observationele studies een evaluatie worden gedaan van het effect van een chirurgische behandeling ten opzichte van standaard (conservatief) beleid (hoofdstuk 5).

Klinische studies

In hoofdstuk 2 wordt gekeken naar complicaties na een chirurgische behandeling van distale bicepspeesrupturen. De huidige literatuur levert nog maar beperkt inzicht in factoren die bijdragen aan het optreden van een complicatie. Na een onbehandelde bicepspeesruptuur verliest de patiënt supinatiekracht en uithoudingsvermogen van de arm wat deels kan worden verholpen met chirurgische re-insertie van de pees. Voor de patiënt is het relevant om te weten wat mogelijke complicaties zijn bij deze behandeling en voor de chirurg is het belangrijk om te weten welke patiënt een grotere kans heeft op bepaalde complicaties. In totaal werden 373 patiënten onderzocht na een bicepspees reparatie of reconstructie waarna 22% van de patiënten een matige dan wel ernstige complicatie ontwikkelden. De belangrijkste complicatie is uitval van een zenuwtak met als gevolg sensibiliteitsverlies van een deel van de onderarm. Verder blijkt dat patiënten met obesitas een groter risico lopen op het krijgen van een complicatie.

Een operatie kan op korte termijn goede resultaten geven maar voor een traumapatiënt is een goede lange-termijn uitkomst wellicht nog belangrijker. Hierbij is kwaliteit van leven misschien wel de belangrijkste graadmeter van het behandelingseffect. In hoofdstuk 3 wordt de lange termijn follow-up gepresenteerd na minimaal invasieve plaat osteosynthese van proximale humerus fracturen. De proximale humerus fractuur is een zeer veel voorkomende fractuur en vaak kan worden volstaan met een conservatieve behandeling. Voor bepaalde type proximale humerus fracturen kan beter voor een operatie worden gekozen en hiervoor zijn veel verschillende chirurgische opties. In dit hoofdstuk wordt een minimaal invasieve methode (MIPO met een Philos plaat via een deltoid split benadering) onderzocht voor de behandeling van verschillende proximale humerus fracturen (AO classificatie type A, B en C). Analyse van 97 patiënten gemiddeld 8,3 jaar na de operatie laat een zeer goede functionele uitkomst zien. Een verder belangrijk resultaat is dat ongeveer 1 op de 3 van de patiënten opteert voor het verwijderen van het osteosynthese materiaal vanwege klachten hiervan in het dagelijks leven.

Hoofdstuk 4 presenteert de lange termijn follow-up na ribfixatie voor multipele ribfracturen of een fladderthorax. Ribfixatie is een relatief nieuwe techniek en wordt in de dagelijkse praktijk nog maar weinig toegepast. Hierdoor is er nog maar weinig bekend over de lange termijn uitkomsten van deze behandeling. Wel is er enig bewijs dat ribfixatie bij patiënten met een fladderthorax, ten opzichte van conservatief beleid, een positief effect heeft op de korte termijn uitkomsten. In de bestaande literatuur is vooral onderzoek gedaan naar de behandeling van een fladderthorax maar er nauwelijks onderzoek naar de behandeling van multipele ribfracturen terwijl dit letsel veel wordt gezien in de dagelijkse praktijk. Er werden

166 patiënten onderzocht die ribfixatie hadden ondergaan; 66 met een fladderthorax (3 of meer ribfracturen op 2 of meer plaatsen gebroken) en 99 met multipele ribfracturen (3 of meer ribfracturen). Dit, voor de bestaande literatuur, grote cohort patiënten met ribfixatie laat zien dat zowel patiënten met multipele ribfracturen als patiënten met een fladderthorax een goede kwaliteit van leven hebben minimaal 1 jaar na het doorgemaakt trauma. Daarnaast laat dit onderzoek zien dat ongeveer de helft van de patiënten implantaat gerelateerde irritatie ondervindt en dat in ongeveer 1 op de 10 gevallen de patiënt opteert om het geïmplanteerde materiaal te laten verwijderen vanwege deze klachten.

Met name de zeldzame complicaties en de lange termijn follow-up van patiënten na chirurgische behandelingen, worden goed inzichtelijk middels observationele cohort studies. Observationeel onderzoek leent zich ook voor evaluatie van medische behandelingen, echter zal hier moeten worden gecorrigeerd voor verschillende potentiële vormen van bias.

Hoofdstuk 5 laat het effect van ribfixatie zien in vergelijking met een niet-operatieve behandeling bij patiënten met een fladderthorax of multipele ribfracturen. Zoals eerder al werd genoemd is er weinig literatuur over de effecten van ribfixatie. Echter, er lijkt een trend zichtbaar naar operatieve behandeling van ribfracturen, getuige ook de toename in publicaties over dit onderwerp. Toch blijft de bewijslast beperkt en zijn er bijna alleen studies die zich richten op patiënten met een fladderthorax. Dit terwijl patiënten met multipele ribfracturen veel frequenter voorkomen in de dagelijkste praktijk. In dit hoofdstuk worden 332 patiënten van twee level 1 traumacentra beschreven waarvan 92 patiënten met een fladderthorax en 240 met multipele ribfracturen. Het ene centrum behandelde alle patiënten conservatief en in het andere centrum werden alle patiënten behandeld met een ribfixatie volgens een specifiek behandelingsalgoritme. Alle patiënten met ribfixatie werden *gematched* met de conservatief behandelde patiënten op basis van baseline karakteristieken (zoals traumamechanisme, begeleidend letsel en ernst van het letsel). Zo ontstonden vergelijkbare groepen en konden de opname specifieke karakteristieken tegen elkaar afgezet worden. Voor beide patiëntengroepen (fladderthorax en multipele ribfracturen) kon geen voordeel aangetoond worden van ribfixatie op intensive care opnameduur, ziekenhuis opnameduur, mortaliteit en de overige uitkomstmaten. De resultaten onderstrepen het belang van verder onderzoek met voldoende grote patiënten aantallen om de implementatie van ribfixatie in de dagelijkse praktijk te verantwoorden.

Methodologische overwegingen

Zoals genoemd geldt de RCT als de gouden standaard voor het evalueren van medische behandelingen en krijgt observationeel onderzoek doorgaans minder waardering.
Hoofdstuk 6 behandelt verschillende aspecten van de beide onderzoeksmethoden en de verschillende vormen van bias. Er worden verschillende chirurgische onderzoeks-typen geïdentificeerd:

- type 1 geneesmiddel vs. controlemiddel in chirurgische patiënten;
- type 2 operatie vs. controle operatie;
- type 3 operatie vs. geen operatie.

Voor type 2 en type 3 onderzoek lijkt nuancering nodig in de veelgebruikte hiërarchische structuur van onderzoeksopzetten. In plaats van een strikte scheiding aan te brengen tussen de resultaten van RCTs en die van andere studieopzetten zouden we voor het evalueren van operatieve behandelingen de resultaten van verschillende onderzoeksopzetten veel meer als complementair aan elkaar moeten zien.

In hoofdstuk 7 & 8 worden twee uitgebreide systematische literatuurstudies gepresenteerd. Dit betreft studies van de niet-operatieve behandeling versus de operatieve behandeling van respectievelijk proximale humerus fracturen en ribfracturen.
Zowel observationele studies als gerandomiseerde studies werden geïncludeerd en bekeken met behulp van meta-analyses. Op basis van de beschikbare literatuur wordt een conservatieve behandeling geadviseerd voor de oudere patiënt met een gedisloceerde proximale humerusfractuur. Voor patiënten met een fladderthorax geldt dat ribfixatie lijkt te resulteren in verbetering van de korte termijn uitkomsten. Het blijft onduidelijk welke patiënten-subgroep hier het meeste baat van heeft.
Een tweede en belangrijk resultaat uit beide studies is dat er geen verschil wordt gezien in de primaire en secundaire uitkomsten wanneer de resultaten van observationele studies worden vergeleken met RCTs. Ofwel, gemiddeld laten observationele studies eenzelfde effect zien van medische interventie als RCTs. Observationeel onderzoek van chirurgische interventies lijkt minder last te hebben van bias dan tot nu werd gedacht.

Hoe verder

Hoofdstuk 9 heeft als doel om de waarde van observationeel onderzoek naar trauma chirurgische behandelingen af te zetten tegen RCTs. Er wordt verder ingegaan op de drie eerder- genoemde verschillende chirurgische onderzoek types en de daarbij horende vormen van bias. Ook worden de uitdagingen en valkuilen besproken van RCTs naar chirurgische behandelingen. Er wordt een onderbouwing gegeven van de relevante eigenschappen van observationele studies in trauma patiënten en specifieke situaties waarbij resultaten van observationeel onderzoek net zo bijdragend kunnen zijn als de resultaten van een RCT.

Hoofdstuk 10 presenteert het OPVENT2 studieprotocol. Deze grote multicenter observationele studie onderzoekt of ribfixatie voor fladderthorax en multipele ribfracturen zinvol is en welk type patiënt hier het meest baat bij heeft. Alle participerende centra behandelen de patiënt zoals zij dit gewend zijn; ofwel de dagelijkse praktijk wordt geobserveerd. Twee van de participerende centra voeren ribfixatie uit volgens een behandelalgoritme en de overige drie centra behandelen alle patiënten niet-operatief. Alle centra zijn grote level 1 traumacentra met een vergelijkbaar aanbod van patiënten. Doordat

er een vorm van geografische randomisatie plaatsvindt (de locatie van het ongeval bepaalt het centrum de behandeling (en daarmee niet operatieve behandeling of ribfixatie). Doormiddel van een statistisch matchingsmodel kan elke patiënt met ribfixatie gekoppeld worden aan een patiënt met vergelijkbare karakteristieken die niet-operatief is behandeld. Zo worden goede vergelijkbare groepen verkregen en kan het gevonden behandelingseffect toegeschreven worden aan ribfixatie. We verwachten dat de bereidheid tot deelname van patiënten groter zal zijn dan in een gerandomiseerde trial aangezien een patiënt zijn behandeling niet hoeft laten te bepalen door loting. Daarnaast zal de chirurg de behandeling uitvoeren zoals hij/zij dit gewend is. Dit vergroot de generaliseerbaarheid van de resultaten en we verwachten dat er gemakkelijker grote patiënten aantallen kunnen worden ingesloten waarmee ook subgroep analyses mogelijk worden. Dit is bij uitstek van belang in de zeer heterogene populatie van patiënten met ribfracturen. De OPVENT2 beoogt daarmee een raamwerk te zijn voor goed observationeel onderzoek in trauma patiënten.

Conclusie

Dit proefschrift laat zien dat observationeel onderzoek van belang is voor het inzichtelijk maken van (zeldzame) complicaties en lange termijn follow-up. Er worden 3 typen chirurgische onderzoeken geïdentificeerd. Daaraan gekoppeld zijn de verschillende vormen van bias, die van invloed kunnen zijn op het gevonden behandeleffect. Er wordt bewijs gepresenteerd en onderbouwd, waaruit blijkt dat observationeel onderzoek een vergelijkbaar behandelingseffect kan geven als RCTs. Tot slot wordt een prospectief observationeel studieprotocol gepresenteerd. Dit beoogt een raamwerk te zijn voor toekomstig onderzoek in de traumachirurgie.

APPENDICES

Review committee

Prof. Dr. M.R. Vriens

Prof. Dr. F.C. Öner

Prof. Dr. C.J. Kalkman

Prof. Dr. Y. van der Graaf

Prof. Dr. M.G.W. Dijkgraaf

Acknowledgements / Dankwoord

Prof. dr. Leenen, beste professor, ik wil u bedanken voor de mogelijkheid om mijn promotieonderzoek te komen doen in het traumacentrum. Het was erg leerzaam om als krullenbol met u mee te lopen in de trauma chirurgische champions league tijdens de verschillende bijeenkomsten met andere centra, collega's en de industrie.

Prof. dr. Groenwold, beste Rolf, jouw vermogen om complexe materie met dodelijke eenvoud uit te dragen blijft mij verbazen. Het was dan ook fijn hoe ik altijd laagdrempelig bij je terecht kon tijdens de verschillende Micaffé sessies of, na je vertrek naar Leiden, tijdens de station meetings in Bar Beton.

Dr. Houwert, beste Marijn / beste Houwie, ik ben jou enorm veel dank verschuldigd voor alle mogelijkheden en kansen die jij mij hebt geboden de afgelopen jaren. Van de plek in het traumacentrum tot tripjes naar Vancouver, Chur, Valencia en binnenkort Luzern. Altijd weet jij situaties te creëren waarin kansen ontstaan en waarmee je het beste uit mensen naar boven haalt. De vernietigende snelheid waarmee jij mijn manuscripten, en talloze versies daarvan, van feedback verzorgde is een belangrijke factor geweest voor het vlotte ontstaan van dit boekje. Naast je rol op professioneel gebied sta jij ook op persoonlijk vlak voor mij klaar en je hebt daarmee veel betekent voor mijn huidige loopbaan. Heel veel dank en ik kijk er naar uit om onze samenwerking in de toekomst voort te zetten.

Dr. Van der Velde, beste Detlef, ik herinner mij nog goed toen ik je voor het eerst op de OK zag terwijl je een enkel aan het opereren was onder de muzikale begeleiding van Rammstein. Ik bewonder jouw nuchtere kijk op de zaak wat ook voor dit proefschrift zijn toegevoegde waarde had. Dank voor de samenwerking en ik hoop in de toekomst een keer samen te opereren met *Du Hast* op de achtergrond.

Hooggeleerde leden van de leescommissie, **prof. dr. M.R. Vriens, prof. dr. F.C. Öner, prof. dr. C.J. Kalkman, prof. dr. Y. van der Graaf, prof. dr. M.G.W. Dijkgraaf,** dank dat u plaats heeft willen nemen in de leescommissie en dat u dit proefschrift heeft beoordeeld.

Dr. Hietbrink, beste Falco, samen met Marijn vorm jij hét wetenschappelijk duo van de traumachirurgie. Jouw scherpte en inzicht hebben mij veel verder geholpen in mijn onderzoek.

Ook wil ik de overige medeauteurs bedanken: Jesse Peek, Arthur Sweet, Mirjam de Jong, Abhi Bhashyam, Karlijn Wessem, Geertje Govaert, Simmermacher, Ivar de Bruin, Olivier vd Meijden, Tim Timmers, Mark van Heijl, Diederik Smeeing, David Reetz, Edwards, Frölke, Öner, Sommer, Acklin, Michelitsch, Femke Claessen, Luke Oh, Neal Chen, and David Ring

L. Spengler, beste Lukas, zonder jouw goedkeuring had ik nooit mijn kantoor mét view gehad in het traumacentrum. Je zorgde altijd voor een fijne werksfeer op de afdeling. Getuige ook de aanschaf van speciaal bier voor onze koelkast op kantoor voor de vrijmibo. Ook wil ik de andere medewerkers van het traumacentrum: **Marlous, Daniëlle, Bob, Stasja en Natasja**, bedanken voor het fijne jaar.

Beste **Ingrid en Gioya**, dank voor jullie eeuwige enthousiasme en de hulp bij de vele verzoeknummertjes waarmee ik, vaak te last minute, aan kwam zetten.

Beste **Q**, na het wetenschappelijk fundament in Boston te hebben gelegd had ik het geluk om jou als kamergenoot te krijgen op 'de zaak'. Het was fijn om met jou te hebben als sparringpartner voor problemen met onderzoek en het bespreken van carrière plannen.
Ook dank aan de overige **traumaonderzoekers**, in het bijzonder Roy, Lilian, Tom, Robin, Eef en Rogier met als hoogtepunt de gegeven zalfjes in Valencia.

Dr. Frima, beste Herman, dank voor je gastvrijheid tijdens mijn bezoek aan Chur. Naast een mooie wetenschappelijke publicatie leverde het mij helaas ook gezichtsverlies op bij het bedwingen van de Calanda.

Dr. Verleisdonk, beste Egbert Jan, dank voor de mogelijkheid om nog enkele maanden in het Diak als onderzoeker te kunnen werken om zo mijn proefschrift verder te kunnen afronden.

Paranimfen **Pieter en Pierre**, beste Piet & Piet, het is eigenlijk gênant dat ik een snip + en een kaalkop vraag voor deze ceremoniële functie. Maar wát mooi dat jullie naast mij staan op deze dag.

Beste **pap en mam**, het zou een slappe poging zijn om in één alinea proberen te zeggen hoe oneindig dankbaar ik ben met jullie onuitputtelijke liefde, steun, geduld en advies waardoor ik mij vanaf nu doctor mag noemen. Ik hou van jullie.

List of publications

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;0(0). doi:10.1016/j.jse.2018.03.009.

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.

Paulino Pereira NR, **Beks RB**, Janssen SJ, et al. Are allogeneic blood transfusions associated with decreased survival after surgical treatment for spinal metastases? The spine journal : official journal of the North American Spine Society. 2016;16(8):951-961. doi:10.1016/j.spinee.2016.03.043.

Vranceanu AM, **Beks RB**, Guitton TG, Janssen SJ, Ring D. How do Orthopaedic Surgeons Address Psychological Aspects of Illness? The archives of bone and joint surgery. 2017;5(1):2-9. doi:10.22038/abjs.2016.7916.

Beks RB, Houwert RM, Groenwold RHH. [Added value of observational studies in surgery: the hierarchical structure of study designs requires a more refined approach]. Nederlands tijdschrift voor geneeskunde. 2017;161(0):D1493.

Beks RB, Mellema JJ, Menendez ME, Chen NC, Ring D, Vranceanu AM. Does Mindfulness Correlate With Physical Function and Pain Intensity in Patients With Upper Extremity Illness? Hand. 2017:155894471769742. doi:10.1177/1558944717697429.

Beks RB, Drijkoningen T, Claessen F, Guitton T, Ring D. Interobserver Variability of the Diagnosis of Scaphoid Proximal Pole Fractures. Journal of Wrist Surgery. 2018;1(212). doi:10.1055/s-0038-1641716.

Beks RB, de Jong MB, Houwert RM, et al. Long-term follow-up after rib fixation for flail chest and multiple rib fractures. European Journal of Trauma and Emergency Surgery. 2018;0(0):0. doi:10.1007/s00068-018-1009-5.

Frima H, Houwert RM, **Beks RB**, van Heijl M, van der Velde D, Beeres FJP. [Proximal humerus fractures; conservative or surgical treatment?]. Nederlands tijdschrift voor geneeskunde. 2019;163.

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. 2018;0(0):0. doi:10.1007/s00402-018-3063-1.

Beks RB, Claessen FMAPAP, Oh LS, Ring D, Chen NC. Factors associated with adverse events after distal biceps tendon repair or reconstruction. Journal of Shoulder and Elbow Surgery. 2016;25(8):1229-1234. doi:10.1016/j.jse.2016.02.032.

Beks RB, Reetz D, de Jong MB, et al. Rib fixation versus non-operative treatment for flail chest and multiple rib fractures after blunt thoracic trauma: a multicenter cohort study. European Journal of Trauma and Emergency Surgery. 2018;0(0):0. doi:10.1007/s00068-018-1037-1.

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. 2018;0(0):0. doi:10.1007/s00068-018-1020-x.

Curriculum Vitae

Reinier Bart Beks was born on the 24th of August 1989 in Haren, the Netherlands. He attended high school at Maartenscollege in Haren and graduated in 2008. After finishing his bachelor's degree in Biomedial Science at the University of Groningen he was accepted for the pre-Master's program for medical school at the same university. During this program he followed internships abroad in Uganda (Kumi Hospital) and South Africa (Tygerberg, Capetown) focusing on (orthopedic) trauma surgey. He developed his interest for research during his stay in Boston, USA, as a graduate research assistant at the Massachusetts General Hospital's hand and upper extremity service where he worked under supervision of Professor David Ring. After finishing medical school in 2016, he started at the St. Antonius Hospital as a resident not in training in general surgery and worked together with dr. R.M. Houwert and dr. D. van der Velde. In 2017 started as a PhD candidate at the University Medical Center Utrecht in 2017 under the supervision of prof. dr. L.P.H. Leenen, prof. dr. R.H.H. Groenwold, dr. R.M. Houwert, and dr. D. van der Velde which has resulted in this thesis. Currently, to gain further experience in clinical work, he works as a resident not in training in general surgery in the Diakonessenhuis Utrecht.

DEFINING THE VALUE OF OBSERVATIONAL STUDIES IN TRAUMA

REINIER BART BEKS