Surgical Versus Nonsurgical Management of Humeral Shaft Fractures

Citation for published version:
Oliver, WM, Bell, KR, Molyneux, S, White, TO, Clement, ND & Duckworth, A 2023, 'Surgical Versus Nonsurgical Management of Humeral Shaft Fractures: a systematic review and meta-analysis of randomised trials', *Journal of the American Academy of Orthopaedic Surgeons*, vol. 31, no. 2, pp. e82-e93.
https://doi.org/10.5435/JAAOS-D-22-00432

Digital Object Identifier (DOI):
10.5435/JAAOS-D-22-00432

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal of the American Academy of Orthopaedic Surgeons

Operative *versus* non-operative management of humeral shaft fractures: a systematic review and meta-analysis of randomised trials

William M. Oliver, LLB(Hons), MBBS(Hons), MRCSEd – william.m.oliver@doctors.org.uk

Katrina R. Bell, MBChB (Hons), MRCSEd – katrinabell@doctors.org.uk

Samuel G. Molyneux, MSc, FRCSEd(Tr&Orth) – sgmolyneux@gmail.com

Timothy O. White, MD, FRCSEd(Tr&Orth) – twhite@rcsed.ac.uk

Nicholas D. Clement, MD, PhD, FRCSEd(Tr&Orth) – nickclement@doctors.org.uk

Andrew D. Duckworth, MSc, FRCSEd(Tr&Orth), PhD – andrew.duckworth@ed.ac.uk

All authors affiliated with:

Edinburgh Orthopaedics, Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh, Midlothian, EH16 4SA, UK

ADD also affiliated with: Centre for Population Health Sciences, Usher Institute, University of Edinburgh, 49 Little France Crescent, EH16 4SB, UK

Corresponding author:

William M. Oliver, Orthopaedic Trauma Research Fellow

Address: Edinburgh Orthopaedics, Royal Infirmary of Edinburgh, 51 Little France Crescent, Edinburgh, Midlothian, EH16 4SA, UK

Email: william.m.oliver@doctors.org.uk

Telephone: +44 131 242 3459

Sources of funding:

The study was supported by the Scottish Orthopaedic Research Trust into Trauma (SORT-iT).

No external funding was received.

Running head (max. 40 characters):

Management of humeral shaft fractures

ABSTRACT

Introduction: The aim was to compare operative and non-operative management for adults with humeral shaft fractures, in terms of patient-reported upper limb function, health-related quality of life (HRQoL), radiographic outcomes and complications.

Methods: MEDLINE, Embase, CINAHL (Cumulative Index to Nursing and Allied Health Literature), PubMed, CENTRAL (Cochrane Central Register of Controlled Trials), ClinicalTrials.gov, ISRCTN (International Clinical Trials Registry) and OpenGrey (Repository for Grey Literature in Europe) were searched in September 2021. All published prospective randomised trials comparing operative and non-operative management of humeral shaft fractures in adults were included. Of 715 studies identified, five were included in the systematic review and four in the meta-analysis. Data were extracted by two independent reviewers according to the PRISMA statement. Methodological quality was assessed using the revised Cochrane risk-of-bias tool for randomised trials. Pooled data were analysed using a random-effects model.

Results: The meta-analysis comprised 292 patients (mean age 41yrs [18-83], 67% male). Surgery was associated with superior DASH and Constant-Murley scores at six months (mean DASH difference 7.6, p=0.01; mean Constant-Murley difference 8.0, p=0.003) but there was no difference at one year (DASH, p=0.30; Constant-Murley, p=0.33). No differences in HRQoL or pain scores were found. Surgery was associated with a lower risk of nonunion (0.7% *versus* 15.7%; odds ratio [OR] 0.13, p=0.004). The number-needed-to-treat (NNT) with surgery to avoid one nonunion was 7. Surgery was associated with a higher risk of transient radial nerve palsy (17.4% *versus* 0.7%; OR 8.23, p=0.01) but not infection (OR 3.57, p=0.13).

Surgery was also associated with a lower risk of re-intervention (1.4% *versus* 19.3%; OR 0.14, p=0.04).

Conclusions: Surgery may confer an early functional advantage to adults with humeral shaft fractures, but this is not sustained beyond six months. The lower risk of nonunion should be balanced against the higher risk of transient radial nerve palsy.

Level of evidence: I

Keywords: Humeral shaft; fracture; operative; fixation; non-operative; outcomes; nonunion; meta-analysis

INTRODUCTION

Humeral shaft fractures are relatively common injuries, with an incidence of around 12 per 100,000 per year[1]. However, the optimal management of patients with a fracture of the humeral shaft remains uncertain. Non-operative management is the default strategy in many centres, although nonunion may complicate approximately 20% of injuries managed this way[2–6]. There has been a recent increase in the proportion of patients managed operatively[7,8], perhaps in recognition of the lower nonunion rate in retrospective comparative studies[3,6,9–12] and the reported unpredictable functional outcomes following humeral bracing[13].

Previous systematic reviews have identified a paucity of level one evidence to inform the treatment of patients with humeral shaft fractures[14–18]. Existing meta-analyses comparing operative with non-operative management are predominantly based on level two and three studies[16–18] that are subject to selection and reporting biases, and are inconclusive regarding which strategy is most effective. However, several randomised trials have recently been published aiming to address this uncertainty[5,19,20]. The authors are not aware of any meta-analyses presenting only level one evidence comparing operative and non-operative treatment of these injuries.

The aim of this systematic review and meta-analysis was to synthesise data from all randomised trials comparing the outcomes of operative and non-operative management for patients with humeral shaft fractures, in terms of patient-reported upper limb function, health-related quality of life (HRQoL), radiographic outcomes and complications.

METHODS

The study was registered with the PROSPERO International Prospective Register of Systematic Reviews on 31 August 2021 (CRD42021276079).

Selection criteria

Inclusion criteria consisted of all published prospective randomised trials comparing the outcomes of operative *versus* non-operative management of adults (aged ≥18 years) with an acute, traumatic humeral shaft fracture. Surgery was defined as any procedure involving operative fixation, including open reduction and internal fixation (ORIF) and minimally invasive plate osteosynthesis (MIPO). Non-operative management was defined as any method of fracture immobilisation in which surgery was not performed, including plaster splinting/casting and functional bracing. Included studies were required to report at least one pre-defined outcome of interest. Studies comparing different techniques of non-operative management alone, or different techniques of surgical fixation alone, were excluded. Non-randomised comparative studies, review articles, conference abstracts, non-clinical studies or those not available in English were also excluded.

Search strategy

Searches of MEDLINE, Embase, PubMed, CINAHL (Cumulative Index to Nursing and Allied Health Literature), CENTRAL (Cochrane Central Register of Controlled Trials), ClinicalTrials.gov, ISRCTN (International Clinical Trials Registry) and OpenGrey (Repository for Grey Literature in Europe) databases were performed according to the updated Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement[21]. Search terms included ((('humerus' [All fields]) OR ('humeral' [All fields])) AND (('shaft' [All fields]) OR (‘diaphysis’ [All fields])) AND (‘fracture’ [All fields]) AND ((‘randomised’ [All fields]) OR (‘randomized’ [All fields]))).

Study selection

Electronic searches were performed in September 2021. A combined search of Ovid MEDLINE and EMBASE yielded 224 abstracts for review. A search of PubMed yielded 111 abstracts for review. A search of CINAHL Plus, limited to studies of human subjects and excluding MEDLINE records, yielded 268 results. A search of clinical trial registries yielded 100 trial registrations (CENTRAL, n=90; ClinicalTrials.gov, n=8; ISRCTN, n=2). Finally, a search of the OpenGrey repository yielded 11 results. One additional study was identified from the reference list of a previous review[18].

Search results were collated using the Covidence online platform (Melbourne, Australia) to facilitate exclusion of duplicate studies and a centralised screening and review process. Of 715 studies initially identified, five were included in the systematic review (**Figure 1**). One study[20] comprised two-year follow-up of a previously reported cohort, and thus only the one-year paper[5] was included in the meta-analysis to avoid duplication.

Data extraction

Included studies were evaluated for title, authorship, year of publication, location (country in which the study was conducted), design, number of participants, population description (age and sex of participants) and inclusion/exclusion criteria. Fracture characteristics (descriptive and AO-OTA classification[22]) and details regarding operative and non-operative management were obtained.

Quality assessment

Study methodology was assessed according to the revised Cochrane risk-of-bias tool for randomised trials (RoB 2)[23]. Each study was assessed by two reviewers with respect to five domains: Randomisation sequence generation and allocation concealment (selection bias); blinding of participants and personnel (performance bias); blinding of outcome assessment (detection bias); potential for incomplete outcome data (attrition bias); and selective reporting/non-reporting (reporting bias). A risk of bias score (low, high or uncertain) was assigned to each domain, with any disagreement resolved through consensus with a third, more senior reviewer.

Outcome measurement

Patient-reported measures of upper limb function included the Disabilities of the Arm, Shoulder and Hand score[24] (DASH; 0 = no disability, 100 = complete disability; minimal clinically-important difference [MCID] 10 points[25]), abbreviated DASH[26] (QuickDASH; 0 = no disability, 100 = complete disability; MCID 16 points[25]) and Constant-Murley shoulder score[27] (0 = worst function, 100 = perfect function; MCID 10-11 points[28,29]). Other patient-reported outcome measures included HRQoL – according to the 36-item Short-Form Health Survey (SF-36; 0 = worst health, 100 = perfect health)[30] or 15-Dimensional instrument (15D; 0 = worst health, 1 = perfect health)[31] – and pain score (0 = no pain, 10 = worst pain imaginable).

Radiographic outcomes included union/nonunion and malunion. Nonunion was generally defined as a failure of the fracture to unite after between three[5] and six[32,33] months. Only one of the included studies specified criteria for malunion, defined as angulation >20° in the sagittal or coronal planes[19]. Other treatment-related complications included radial nerve palsy (RNP) and infection. No studies provided an explicit definition of infection. Several studies reported the broader complication of treatment failure requiring further surgical intervention ('re-intervention'), encompassing failure of initial management due to loss of fracture reduction, intolerance of functional bracing, symptomatic metalwork prominence or nonunion.

Statistical methods

Meta-analysis was performed using Review Manager 5.3 (Cochrane Collaboration, Oxford, UK). Pooled effect estimates for continuous outcome variables were determined using inverse variance, and the mean difference was calculated with 95% confidence intervals (CIs). Where necessary, the standard deviation was determined using the group mean and 95% CIs. Pooled effect estimates for dichotomous outcome variables were determined using a Mantel-Haenszel analysis, and odds ratios (ORs) were calculated with 95% CIs. Cochran's Q and I^2 tests were used to assess heterogeneity. However, due to suspected clinical variability and effect sizes between studies, random models were employed to assess the influence of management on outcome. A p-value of <0.05 was considered statistically significant where studies reported no event in one treatment arm.

RESULTS

Four studies included 292 patients, with 66.8% (n=195/292) male and 33.2% (n=97/292) female. The mean age was 41 years (range 18 to 83). Fracture location was documented for 214 patients, with the middle-third (79.9%, n=171/214) most frequently involved (distal-third 11.2%, n=24/214; proximal-third 8.9%, n=19/214). AO-OTA type A injuries were most common (74.0%, n=214/289; type B 18.3%, n=53/289; type C 7.6%, n=22/289). Half of patients (n=146/292) were managed operatively, either with ORIF (60.3%, n=88/146)[5,19,32] or MIPO (39.7%, n=58/146)[33]. Half were managed non-operatively with an initial period (two-to-four weeks) in a U-slab or hanging cast[19,32,33], after which a functional brace was applied (**Table 1**).

Risk of bias in included studies

Overall, the risk of selection bias in patient allocation was low, although two studies[19,32] did not clearly document the method of allocation concealment. All studies were vulnerable to performance bias, as blinding of participants and personnel was not possible given the interventions being compared. Two groups[5,20,33] detailed strategies by which outcome assessment was successfully blinded (low risk of detection bias). Two studies[19,32] were at high risk of attrition and reporting biases, due to incomplete reporting of exclusions, attritions and outcome data as specified in respective study methods. Other potential sources of bias included lack of central trial registration[19,32], incomplete details regarding management of nonunions[19,32], and loss to follow-up leading to a potential loss of power[33] (**Table 2**).

Upper limb function

DASH and QuickDASH scores

Three studies[5,32,33] (213 patients) reported the DASH at six months, with a statistically significant difference in favour of surgery found (difference 7.6, 95% CI 1.8 to 13.4, p=0.01; **Figure 2A**). Two studies[5,33] (172 patients) reported the DASH at 12 months, with no difference between the groups (difference 1.4, 95% CI -1.3 to 4.1, p=0.30; **Figure 2B**).

Matsunaga *et al.*[33] reported the DASH at two weeks, one and two months, finding no difference at any of these timepoints. Rämö *et al.* reported differences in the DASH in favour of surgery at six weeks (difference 9.9, 95% CI 3.5 to 16.3, p=0.002) and three months (difference 10.1, 95% CI 3.6 to 16.6, p=0.002)[5], but no difference at two years (difference 0.8, 95% CI -6.0 to 7.6, p=0.81)[20].

Khameneh *et al.*[19] reported a trend towards greater improvement in the QuickDASH following surgery (mean change = 29.1) compared with non-operative management (mean change = 26.7, p=0.065). However, it was unclear at which timepoints the initial and final QuickDASH measurements were obtained, and therefore we opted not to incorporate this data into the meta-analysis.

Constant-Murley score

Two studies[5,33] reported the Constant-Murley score at six (174 patients) and 12 months (170 patients). At both timepoints a difference in favour of surgery was observed, although statistical significance was only reached at six months (six-month difference = 8.0, 95% CI 2.7 to 13.4, p=0.003; 12-month difference = 1.8, 95% CI -1.8 to 5.3, p=0.33; **Figures 2C and 2D**). Matsunaga *et al.*[33] found a trend towards superior Constant-Murley scores at two weeks in the surgery group (difference 5.5, p=0.076), but not at one or two months. Similarly, Rämö *et al.* observed superior Constant-Murley scores at six weeks (difference 30.7, 95% CI 22.8 to 38.7, p<0.001) and three months (difference 14.9, 95% CI 6.9 to 22.9, p<0.001)[5], but not at two years (difference -3.3, 95% CI -11.7 to 5.1, p-value not reported)[20].

Other indicators of upper limb function

Rämö *et al.*[5,20] found a superior elbow motion arc (in degrees) at six weeks (difference 29, 95% CI 21 to 37, p<0.001) and three months (difference 14, 95% CI 6 to 21, p<0.001) following surgery, which was maintained (but non-significant) at six months (difference 6, 95% CI -2 to 14, p=0.14), one year (difference 7, 95% CI -1 to 15, p=0.10) and two years (difference 4, 95% CI -4 to 12, p-value not reported).

Rämö *et al.*[5,20] found the surgery group had superior DASH 'work module' scores up to six weeks (difference 21.7, 95% CI 5.8 to 37.6, p=0.008) and superior DASH 'sports/performing arts module' scores up to one year (difference 21.2, 95% CI 3.4 to 38.9, p=0.02). There was no difference in either module scores beyond these timepoints.

Health-related quality of life

No two studies reported the same HRQoL measure and thus meta-analysis was not performed. Matsunaga *et al.*[33] found a difference in the 'physical functioning' domain of the SF-36 at one month (surgery score 67.6, non-operative score 76.0, p=0.025), but no difference in any other domains up to one year. Rämö *et al.*[5,20] compared 15D scores between groups, but found no difference at any timepoint up to two years.

Pain

Matsunaga *et al.*[33] reported overall pain level, while Rämö *et al.*[5] documented pain at rest and pain on activity. There was no difference in pain scores between groups at six (difference 0.1, 95% CI -0.3 to 0.6, p=0.57) or 12 months (difference 0.1, 95% CI -0.4 to 0.6, p=0.70; **Figure**

3). Matsunaga *et al.*[33] also reported pain scores at two weeks, one and two months; although pain was marginally greater among the surgery group, this did not reach significance (two-week difference 0.7, p=0.076; one-month difference 0.6, p=0.152; two-month difference 0.7, p=0.074). Ramo *et al.*[5,20] found no difference in pain at rest up to two years of follow-up. At six weeks, pain on activity was marginally lower in the surgery group (difference 1.2, 95% CI 0.1 to 2.3, p=0.04), but there were no differences thereafter.

Radiographic outcomes

All studies (278 patients) documented union and nonunion. The pooled nonunion rate was 8.3% (n=23/278). Surgery (0.7%, n=1/138) was associated with a lower risk of nonunion (OR 0.13, 95% CI 0.03 to 0.52, p=0.004) compared with non-operative management (15.7%, n=22/140; **Figure 4A**). The number-needed-to-treat (NNT) with surgery to avoid one additional nonunion was seven. The pooled malunion rate was 1.8% (n=5/278). There was no significant difference in the risk of malunion between the groups (operative rate 0%, n=0/138; non-operative rate 3.6%, n=5/140; OR 0.23, 95% CI 0.04 to 1.44, p=0.12; **Figure 4B**).

Other complications

All studies (278 patients) reported details of other complications. The pooled rate of transient radial nerve palsy (TRNP) was 9.0% (n=25/278). No instances of permanent RNP were documented. Surgery (17.4%, n=24/138) was associated with an increased risk of TRNP (OR 8.23, 95% CI 1.62 to 41.77, p=0.01) compared with non-operative management (rate 0.7%, n=1/140; **Figure 5A**). However, this finding was predominantly based on a single study documenting a 60% rate of transient postoperative radial neurapraxia (n=18/30)[19]. With this study omitted, the pooled TRNP rate decreased to 3.2% (n=7/218) and the differential risk of

TRNP was non-significant (operative rate 5.6%, n=6/108; non-operative rate 0.9%, n=1/110; OR 3.81, 95% CI 0.77 to 18.97, p=0.10).

The pooled infection rate was 2.5% (n=7/278). All infections were superficial and successfully managed with antibiotics and dressings, with no requirement for surgical re-intervention. There was no significant difference in the infection risk between the groups (operative rate 4.3%, n=6/138; non-operative rate 0.7%, n=1/140; OR 3.57, 95% CI 0.69 to 18.38, p=0.13; **Figure 5B**).

The pooled re-intervention rate was 10.4% (n=29/278). Surgery (1.4%, n=2/138) was associated with a lower risk of re-intervention (OR 0.14, 95% CI 0.02 to 0.92, p=0.04) compared with non-operative management (19.3%, n=27/140; **Figure 5C**). The NNT with surgery to avoid one additional re-intervention was six.

DISCUSSION

This systematic review and meta-analysis of randomised trials found that surgical fixation of humeral shaft fractures was associated with superior early functional outcome scores compared with non-operative management, although this was not sustained beyond six months. Surgery was not associated with any clear benefit in HRQoL or pain scores. Surgery was, however, associated with a lower rate of nonunion and re-intervention. The increased likelihood of nonunion and re-intervention should be considered when balancing the benefits and risks of non-operative management. These data are useful for surgeons when counselling patients during shared decision-making. A pragmatic approach, whereby surgery is selectively offered to those at increased risk of nonunion, would appear a cogent strategy based on current evidence.

To the authors' knowledge, this is the first meta-analysis of humeral shaft fracture management to include only randomised trials. Although including observational studies increases the sample size (and thus the ability to examine smaller sub-groups or less common outcomes), the risk of bias is a concern. Several previous reviews, which combined data from one randomised study[33] with data from several observational studies, are limited in this way[16,17]. Meta-analyses of randomised trials represent the highest level of evidence in the evaluation of treatment effects, and thus this study offers the best available evidence in the management of these injuries.

One previous review[18] attempted a meta-analysis of patient-reported outcomes following a humeral shaft fracture, finding a non-significant trend towards superior function (according to the DASH) among patients managed operatively. However, this analysis included only 150 patients (around half the number in the present study) and the authors acknowledged the lack of statistical significance was likely due to a lack of data[18]. Our analysis demonstrated a statistical difference in favour of surgery according to the DASH and Constant-

Murley scores. However, these differences were not sustained beyond six months and fell below the MCID in these scores (10 points[25] and 10-11 points[28,29], respectively). This also reflects provisional data from a recently-completed multicentre prospective randomised trial comparing operative and non-operative treatment[34], which found a significant difference in the DASH score in favour of surgery at six weeks and four months but not at one year[35]. Like previous reviews we were unable to undertake meta-analysis of patient-reported HRQoL, although no studies in our analysis found any difference in these measures at any timepoint[5,20,33]. Our pooled analysis of pain scores failed to demonstrate any significant differences based on treatment strategy.

This study has provided convincing evidence that surgery is associated with a lower risk of nonunion (operative rate 0.7%, non-operative rate 15.7%), consistent with previous reviews involving level two and three studies (operative rate 6.0-8.1%, non-operative rate 15.3-17.7%)[16–18]. Most observational studies support the concept that surgery confers a lower risk of nonunion than functional bracing[3,6,9–12]. It is also apparent that nonunion may result in inferior functional outcomes and HRQoL for years after union has subsequently been achieved following reintervention with surgery[20,36]. The differential nonunion rate between operative and non-operative management, and strategies by which surgery might be targeted to patients at risk of nonunion at an early stage in their treatment[37,38], warrant careful consideration.

Previous reviews are somewhat contradictory regarding the association between humeral shaft fracture management and RNP. One review reported an increased risk of TRNP with surgery (operative rate 3.4%, non-operative rate 0%)[16], but others have found no difference[17,18]. The rate of permanent RNP is reported to be similarly low after both operative (2.6%) and non-operative management (1.3%)[18]. Our analysis suggested that surgery carried a higher risk of TRNP (operative rate 17.4%, non-operative rate 0.7%), but this finding was disproportionately affected by one study[19]. No instances of permanent RNP were documented in any of the included studies. Intuitively one might anticipate an increased risk of TRNP with surgical fixation, but this is not clearly borne out in existing randomised trials. It may be concluded that any difference in RNP risk is modest, and that such injuries are invariably transient and self-resolving. Similarly, two previous reviews[16,17] reported an increased infection rate with surgery (operative rate 3.3-3.7%, non-operative rate 0-0.6%), while another[18] found no difference. We did not find a significant difference in the rate of superficial infection (operative rate 4.3%, non-operative rate 0.7%) between the groups based on included level one studies.

Despite being the largest meta-analysis of level one evidence to date, we acknowledge the number of patients is still relatively small which increases the risk of bias. Moreover, the methodological quality of the included studies was variable, with two[19,32] demonstrating a high risk of bias across all domains. Further high-quality randomised trials are needed[34,39,40]. Existing randomised trials involve predominantly younger male patients sustaining higher-energy injuries, which may not reflect the modern epidemiology of humeral shaft fractures[1] and thus may limit generalisability. Other limitations in existing trials include a lack of American Society of Anaesthesiologists (ASA) grading and frailty data, and limited data on the outcomes of MIPO (one study[33]) or intramedullary nailing (no studies). Although the DASH is frequently documented, there is inconsistency in other patient-reported outcome measures especially relating to HRQoL. Furthermore, both the DASH and Constant-Murley scores may involve ceiling effects, potentially limiting the ability to detect subtle differences in outcome[41], particularly among younger, higher-functioning patients[42]. Health economic analyses are of clear importance, and though this data is currently lacking we note some future randomised trials plan to incorporate such analyses[40,43].

SUMMARY

In this meta-analysis of randomised trials comparing surgical fixation with functional bracing for adults with humeral shaft fractures (292 patients), surgery resulted in superior patient-reported function at six months compared with functional bracing but there was no difference thereafter. Surgery conferred no clear benefit in terms of health-related quality of life or pain scores at any timepoint. Surgery was associated with a lower rate of nonunion and treatment failure requiring re-intervention, but a higher rate of transient radial nerve palsy.

TABLES

Table 1: Baseline details for included studies

Study	Location	Recruitment period	Participants (n)	Mean age (yrs)	M:F	Injury energy (low:high)	Fracture location	AO-OTA classification	Non-op (n, tech)	Operative (n, tech)	Follow-up (months)
Kumar 2017[32]	India	2012-2014	40	35	29:11 (73% M)	17:23 (58% high)	Prox N/S Mid 63% Dist N/S	A 97.5% B 2.5% C 0%	20, functional brace	20, ORIF	6
Matsunaga 2017[33]	Brazil	2012-2015	110	38	73:37 (66% M)	N/S	Prox 11% Mid 69% Dist 20%	A 62% B 30% C 8%	52, functional brace	58, MIPO	12
Khameneh 2019[19]	Iran	2016	60	43	49:11 (82% M)	32:28 (47% high)	N/S	A 65% B 15% C 20%	30, functional brace	30, ORIF	'Until complete union'
Rämö 2020[5] & 2021[20]	Finland	2012-2018	82	48	44:38 (54% M)	72:10 (12% high)	Prox 9% Mid 88% Dist 4%	A 85% B 13% C 1%	44, functional brace	38, ORIF	12 & 24

AO-OTA, Arbeitsgemeinschaft für Osteosynthesefragen-Orthopaedic Trauma Association; Dist, distal-third; F, female; M, male; Mid, middle-third; N/S, not specified; Prox, proximal-third; tech, technique

Table 2: Risk of bias assessment for included studies (using RoB 2 tool)

Study	Selection bias	Performance bias	Detection bias	Attrition bias	Reporting bias	Other bias	Notes
Kumar 2017[32]	Unsure	High	High	High	High	High	• Patients randomised using a random number table • Method of allocation concealment not described • No blinding of outcome assessment • Attritions and exclusions not reported • Outcome data incompletely reported • No central trial registration • Nonunion occurred in n=1/20 pts in ORIF group and n=2/20 pts in functional brace group; no details regarding secondary treatment
Matsunaga 2017[33]	Low	High	Low	Low	Low	High	• 'Before the outcome assessments, the participants were instructed to not reveal the treatment that they had undergone, and an identical opaque gown was used to cover the injured arm in both groups' • 'Blinded assessment of the self-reported questionnaires (DASH, SF-36, and pain VAS)' • Study under-powered (*a priori* power calculation specified 50 pts in each treatment arm, but only 44 in final analysis of functional brace group)
Khameneh 2019[19]	Low/unsure	High	High	High	High	High	• Method of allocation concealment not described • No blinding of outcome assessment • Attritions and exclusions not reported • Outcome data incompletely reported • No central trial registration • Nonunion occurred in n=2/30 pts in functional brace group; underwent surgical management but timing unclear
Rämö 2020[5] & 2021[20]	Low	High	Low	Low	Low	Low	• 'Outcome assessors carrying out objective measures were blinded to the treatment group during the follow-up visits by having the patients wear a long-sleeved shirt and not verbally reveal their study group'

NB. All included studies involved a high risk of performance bias, as blinding of participants/personnel was not possible in the context of a prospective randomised trial of non-operative *versus* operative management

REFERENCES

1. Oliver WM, Searle HKC, Ng ZH, et al. Fractures of the proximal- and middle-thirds of the humeral shaft should be considered as fragility fractures: An epidemiological study of 900 consecutive injuries. *Bone Joint J*. 2020;102-B(11):1475-1483. doi:10.1302/0301-620X.102B11.BJJ-2020-0993.R1
2. Ali E, Griffiths D, Obi N, Tytherleigh-Strong G, Van Rensburg L. Nonoperative treatment of humeral shaft fractures revisited. *J Shoulder Elbow Surg*. 2015;24(2):210-214. doi:10.1016/j.jse.2014.05.009
3. Harkin FE, Large RJ. Humeral shaft fractures: union outcomes in a large cohort. *J Shoulder Elbow Surg*. 2017;26(11):1881-1888. doi:10.1016/j.jse.2017.07.001
4. Serrano R, Mir HR, Sagi HC, et al. Modern results of functional bracing of humeral shaft fractures: A multicenter retrospective analysis. *J Orthop Trauma*. 2020;34(4):206-209. doi:10.1097/BOT.0000000000001666
5. Rämö L, Sumrein BO, Lepola V, et al. Effect of surgery vs functional bracing on functional outcome among patients with closed displaced humeral shaft fractures: The FISH Randomized Clinical Trial. *JAMA*. 2020;323(18):1792-1801. doi:10.1001/jama.2020.3182
6. Oliver WM, Searle HKC, Ng ZH, et al. Factors associated with humeral shaft nonunion. *J Shoulder Elbow Surg*. 2021;30(10):2283-2295. doi:10.1016/J.JSE.2021.01.029
7. Huttunen TT, Kannus P, Lepola V, Pihlajamäki H, Mattila VM. Surgical treatment of humeral-shaft fractures: A register-based study in Finland between 1987 and 2009. *Injury*. 2012;43(10):1704-1708. doi:10.1016/j.injury.2012.06.011
8. Schoch BS, Padegimas EM, Maltenfort M, Krieg J, Namdari S. Humeral shaft fractures: national trends in management. *J Orthop Traumatol*. 2017;18(3):259-263.

doi:10.1007/s10195-017-0459-6

9. Wallny T, Sagebiel C, Westerman K, Wagner UA, Reimer M. Comparative results of bracing and interlocking nailing in the treatment of humeral shaft fractures. *Int Orthop*. 1997;21(6):374-379. doi:10.1007/s002640050189

10. Denard AJ, Richards JE, Obremskey WT, Tucker MC, Floyd M, Herzog GA. Outcome of nonoperative vs operative treatment of humeral shaft fractures: a retrospective study of 213 patients. *Orthopedics*. 2010;33(8). doi:10.3928/01477447-20100625-16

11. Westrick E, Hamilton B, Toogood P, Henley B, Firoozabadi R. Humeral shaft fractures: results of operative and non-operative treatment. *Int Orthop*. 2017;41(2):385-395. doi:10.1007/s00264-016-3210-7

12. Olson JJ, Entezari V, Vallier HA. Risk factors for nonunion after traumatic humeral shaft fractures in adults. *JSES Int*. 2020;4(4):734-738. doi:10.1016/j.jseint.2020.06.009

13. Papasoulis E, Drosos GI, Ververidis AN, Verettas DA. Functional bracing of humeral shaft fractures. A review of clinical studies. *Injury*. 2010;41(7):e21-e27. doi:10.1016/j.injury.2009.05.004

14. Gosler MW, Testroote M, Morrenhof JW, Janzing HMJ. Surgical versus non-surgical interventions for treating humeral shaft fractures in adults. *Cochrane Database Syst Rev*. 2012;1:CD008832. doi:10.1002/14651858.CD008832.pub2

15. Clement ND. Management of humeral shaft fractures; non-operative versus operative. *Arch Trauma Res*. 2015;4(2):e28013. doi:10.5812/atr.28013v2

16. Sargeant HW, Farrow L, Barker S, Kumar K. Operative versus non-operative treatment of humeral shaft fractures: A systematic review. *Shoulder Elb*. 2020;12(4):229-242. doi:10.1177/1758573218825477

17. Lode I, Nordviste V, Erichsen JL, Schmal H, Viberg B. Operative versus nonoperative

treatment of humeral shaft fractures: a systematic review and meta-analysis. *J Shoulder Elbow Surg*. 2020;29(12):2495-2504. doi:10.1016/J.JSE.2020.05.030

18. van de Wall BJM, Ochen Y, Beeres FJP, et al. Conservative vs. operative treatment for humeral shaft fractures: a meta-analysis and systematic review of randomized clinical trials and observational studies. *J Shoulder Elbow Surg*. 2020;29(7):1493-1504. doi:10.1016/j.jse.2020.01.072

19. Khameneh SMH, Abbasian M, Abrishamkarzadeh H, et al. Humeral shaft fracture: a randomized controlled trial of nonoperative versus operative management (plate fixation). *Orthop Res Rev*. 2019;11:141-147. doi:10.2147/ORR.S212998

20. Rämö L, Paavola M, Sumrein BO, et al. Outcomes With Surgery vs Functional Bracing for Patients With Closed, Displaced Humeral Shaft Fractures and the Need for Secondary Surgery: A Prespecified Secondary Analysis of the FISH Randomized Clinical Trial. *JAMA Surg*. 2021;156(6):1-9. doi:10.1001/jamasurg.2021.0906

21. Page M, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/BMJ.N71

22. Meinberg EG, Agel J, Roberts CS, Karam MD, Kellam JF. Fracture and Dislocation Classification Compendium–2018. *J Orthop Trauma*. 2018;32(Suppl 1):S1-S170. doi:10.1097/BOT.0000000000001063

23. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. doi:10.1136/BMJ.L4898

24. Hudak PL, Amadio PC, Bombardier C, et al. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med*. 1996;29(6):602-608. doi:10.1002/(SICI)1097-0274(199606)29:6<602::AID-AJIM4>3.0.CO;2-L

25. Franchignoni F, Vercelli S, Giordano A, Sartorio F, Bravini E, Ferriero G. Minimal

clinically important difference of the disabilities of the arm, shoulder and hand outcome measure (DASH) and its shortened version (QuickDASH). *J Orthop Sports Phys Ther*. 2014;44(1):30-39. doi:10.2519/jospt.2014.4893

26. Beaton DE, Wright JG, Katz JN, Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am*. 2005;87(5):1038-1046. doi:10.2106/JBJS.D.02060

27. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res*. 1987;(214):160-164.

28. Kukkonen J, Kauko T, Vahlberg T, Joukainen A, Äärimaa V. Investigating minimal clinically important difference for Constant score in patients undergoing rotator cuff surgery. *J Shoulder Elbow Surg*. 2013;22(12):1650-1655. doi:10.1016/J.JSE.2013.05.002

29. Christiansen DH, Frost P, Falla D, Haahr JP, Frich LH, Svendsen SW. Responsiveness and Minimal Clinically Important Change: A Comparison Between 2 Shoulder Outcome Measures. *J Orthop Sport Phys Ther*. 2015;45(8):620-625. doi:10.2519/JOSPT.2015.5760

30. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.

31. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med*. 2001;33(5):328-336. doi:10.3109/07853890109002086

32. Kumar S, Shanmugam N, Kumar S, Ramanusan K. Comparison between operative and non operative treatment of fracture shaft of humerus: an outcome analysis. *Int J Res Orthop*. 2017;3(3):445-450. doi:10.18203/issn.2455-4510.IntJResOrthop20171537

33. Matsunaga FT, Tamaoki MJS, Matsumoto MH, Netto NA, Faloppa F, Belloti JC. Minimally invasive osteosynthesis with a bridge plate versus a functional brace for

humeral shaft fractures: A randomized controlled trial. *J Bone Joint Surg Am*. 2017;99(7):583-592. doi:10.2106/JBJS.16.00628

34. A Randomized Controlled Trial Comparing Operative and Nonoperative Treatment of Humeral Diaphyseal Fractures. https://clinicaltrials.gov/ct2/show/NCT00878319.

35. Bergeron S, Schneider P, Liew A, Kreder HJ, Berry G. A Randomized Controlled Trial Comparing Operative and Nonoperative Treatment of Humeral Diaphyseal Fractures (Abstract). OTA 2021 Annual Meeting: Program Highlights. https://ota.org/education/meetings-and-courses/abstracts/randomized-controlled-trial-comparing-operative-and.

36. Oliver WM, Searle HKC, Molyneux SG, White TO, Clement ND, Duckworth AD. Factors Associated with Patient-Reported Outcomes Following a Humeral Shaft Fracture: Nonunion Results in a Poorer Outcome Despite Union After Surgical Fixation. *J Orthop Trauma*. 2021:In press.

37. Driesman AS, Fisher N, Karia R, Konda S, Egol KA. Fracture site mobility at 6 weeks after humeral shaft fracture predicts nonunion without surgery. *J Orthop Trauma*. 2017;31(12):657-662. doi:10.1097/BOT.0000000000000960

38. Oliver WM, Smith TJ, Nicholson JA, et al. The Radiographic Union Score for HUmeral fractures (RUSHU) predicts humeral shaft nonunion. *Bone Joint J*. 2019;101-B(10). doi:10.1302/0301-620X.101B10.BJJ-2019-0304.R1

39. Oliver WM, Smith TJ, Nicholson JA, et al. The Radiographic Union Score for HUmeral fractures (RUSHU) predicts humeral shaft nonunion. *Bone Joint J*. 2019;101-B(10):1300-1306. doi:10.1302/0301-620x.101b10.bjj-2019-0304.r1

40. The humeral shaft fracture trial (HUSH): surgical versus non-surgical interventions for humeral shaft fractures in patients aged 18 years or older. https://www.isrctn.com/ISRCTN17108318.

41. Mahabier KC, Den Hartog D, Theyskens N, Verhofstad MHJ, Van Lieshout EMM, HUMMER Trial Investigators. Reliability, validity, responsiveness, and minimal important change of the Disablities of the Arm, Shoulder and Hand and Constant-Murley scores in patients with a humeral shaft fracture. *J Shoulder Elbow Surg*. 2017;26(1):e1-e12. doi:10.1016/j.jse.2016.07.072

42. Hsu JE, Nacke E, Park MJ, Sennett BJ, Huffman GR. The Disabilities of the Arm, Shoulder, and Hand questionnaire in intercollegiate athletes: validity limited by ceiling effect. *J Shoulder Elbow Surg*. 2010;19(3):349-354. doi:10.1016/J.JSE.2009.11.006

43. Oliver WM, Carter TH, Graham C, et al. A prospective randomised controlled trial of operative versus non-operative management of fractures of the humeral diaphysis: the HUmeral Shaft Fracture FIXation (HU-FIX) Study protocol. *Trials*. 2019;20(1):475. doi:10.1186/s13063-019-3576-0

FIGURES

Figure 1: Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) flow diagram

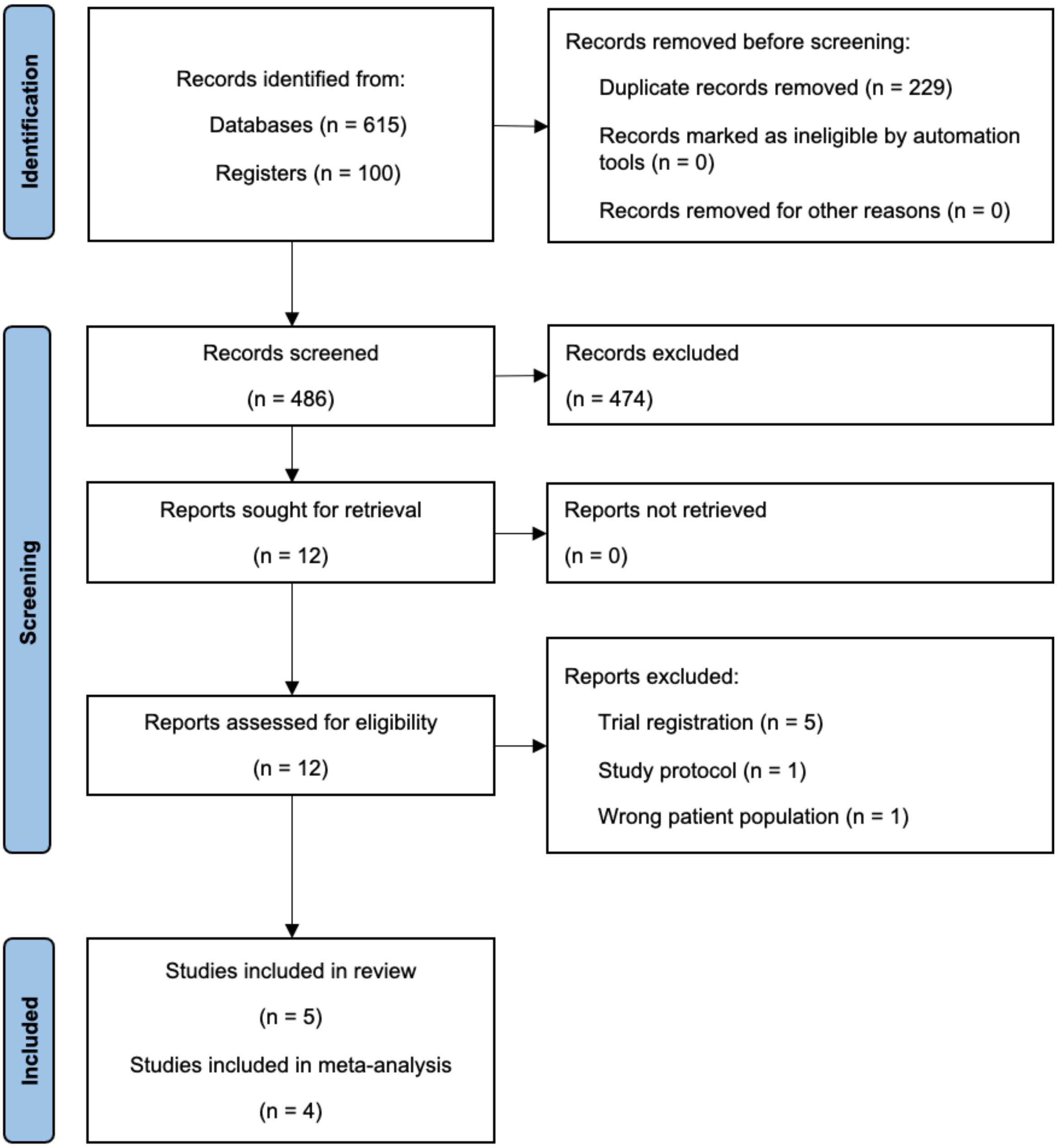
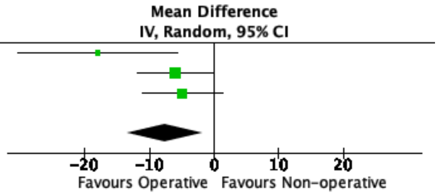


Figure 2: Forest plots for studies reporting upper limb function following a humeral shaft fracture: (A) the Disabilities of the Arm, Shoulder and Hand score (DASH) at six months; (B) the DASH at 12 months; (C) the Constant-Murley score at six months; (D) the Constant-Murley score at 12 months; CI, confidence interval; IV, inverse variance

(A) DASH at six months

	Operative			Non-operative				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
Kumar 2017	23.4	22.87	20	41.35	16.32	20	17.2%	−17.95 [−30.26, −5.64]
Matsunaga 2017	10.9	10.5	50	16.9	18	46	42.3%	−6.00 [−11.96, −0.04]
Rämö 2020	13.5	13.8	34	18.4	14	43	40.5%	−4.90 [−11.15, 1.35]
Total (95% CI)			104			109	100.0%	−7.61 [−13.42, −1.81]

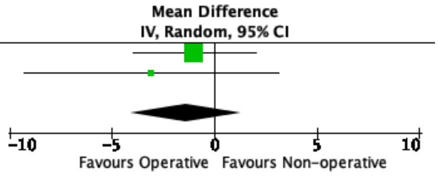
Heterogeneity: Tau² = 11.50; Chi² = 3.57, df = 2 (P = 0.17); I² = 44%
Test for overall effect: Z = 2.57 (P = 0.01)



(B) DASH at 12 months

	Operative			Non-operative				Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	Year
Matsunaga 2017	5.5	5.8	50	6.5	8.6	44	81.0%	−1.00 [−4.01, 2.01]	2017
Rämö 2020	8.9	13.9	36	12	14	42	19.0%	−3.10 [−9.31, 3.11]	2020
Total (95% CI)			86			86	100.0%	−1.40 [−4.11, 1.31]	

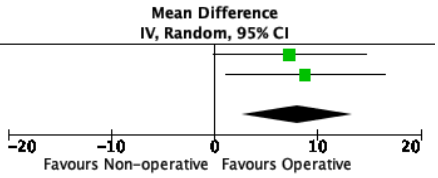
Heterogeneity: Tau² = 0.00; Chi² = 0.36, df = 1 (P = 0.55); I² = 0%
Test for overall effect: Z = 1.01 (P = 0.31)



(C) Constant-Murley at six months

	Operative			Non-operative				Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	Year
Matsunaga 2017	87.3	13.8	50	80	22.1	46	51.8%	7.30 [−0.14, 14.74]	2017
Rämö 2020	73.1	17.1	35	64.3	17.5	43	48.2%	8.80 [1.09, 16.51]	2020
Total (95% CI)			85			89	100.0%	8.02 [2.67, 13.38]	

Heterogeneity: Tau² = 0.00; Chi² = 0.08, df = 1 (P = 0.78); I² = 0%
Test for overall effect: Z = 2.94 (P = 0.003)



(D) Constant-Murley at 12 months

	Operative			Non-operative				Mean Difference	
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	Year
Matsunaga 2017	92.2	8.6	50	90.4	10.8	44	79.8%	1.80 [−2.18, 5.78]	2017
Rämö 2020	78.1	17.6	35	76.4	17.5	41	20.2%	1.70 [−6.22, 9.62]	2020
Total (95% CI)			85			85	100.0%	1.78 [−1.78, 5.34]	

Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.98); I² = 0%
Test for overall effect: Z = 0.98 (P = 0.33)

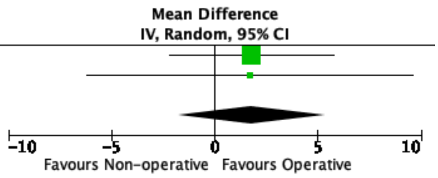
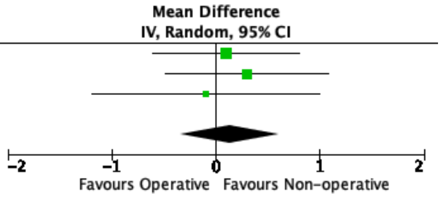


Figure 3: Forest plots for studies reporting pain score (using a visual analogue scale) following a humeral shaft fracture: (A) at six months; (B) at 12 months; CI, confidence interval; IV, inverse variance

(A) Pain score at six months

Study or Subgroup	Operative Mean	Operative SD	Operative Total	Non-operative Mean	Non-operative SD	Non-operative Total	Weight	Mean Difference IV, Random, 95% CI	Year
Matsunaga 2017	1.5	1.6	50	1.4	1.9	46	45.0%	0.10 [−0.61, 0.81]	2017
Rämö 2020	1	1.7	34	0.7	1.8	43	36.4%	0.30 [−0.48, 1.08]	2020
Rämö 2021	2.4	2.3	34	2.5	2.6	43	18.6%	−0.10 [−1.20, 1.00]	2021
Total (95% CI)			118			132	100.0%	0.14 [−0.34, 0.61]	

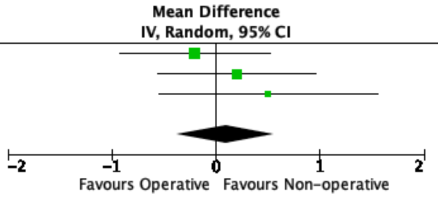
Heterogeneity: Tau² = 0.00; Chi² = 0.36, df = 2 (P = 0.84); I² = 0%
Test for overall effect: Z = 0.56 (P = 0.57)



(A) Pain score at 12 months

Study or Subgroup	Operative Mean	Operative SD	Operative Total	Non-operative Mean	Non-operative SD	Non-operative Total	Weight	Mean Difference IV, Random, 95% CI	Year
Matsunaga 2017	1.2	1.7	50	1.4	1.9	46	42.0%	−0.20 [−0.92, 0.52]	2017
Rämö 2020	0.9	1.6	35	0.7	1.8	42	38.1%	0.20 [−0.56, 0.96]	2020
Rämö 2021	2.2	2.3	35	1.7	2.4	42	19.9%	0.50 [−0.55, 1.55]	2021
Total (95% CI)			120			130	100.0%	0.09 [−0.38, 0.56]	

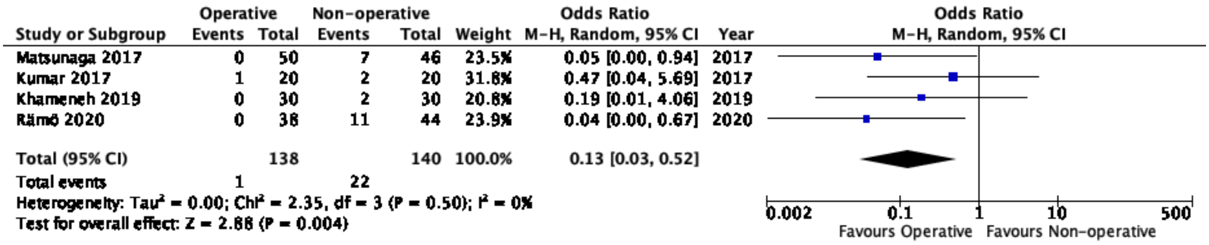
Heterogeneity: Tau² = 0.00; Chi² = 1.28, df = 2 (P = 0.53); I² = 0%
Test for overall effect: Z = 0.38 (P = 0.70)



NB. For the purposes of the above forest plots, Rämö 2020 denotes pain at rest and Rämö 2021 denotes pain on activity

Figure 4: Forest plots for studies reporting radiographic outcomes following a humeral shaft fracture; CI, confidence interval; M-H, Mantel-Haenszel

(A) Nonunion following a humeral shaft fracture



(B) Malunion following a humeral shaft fracture

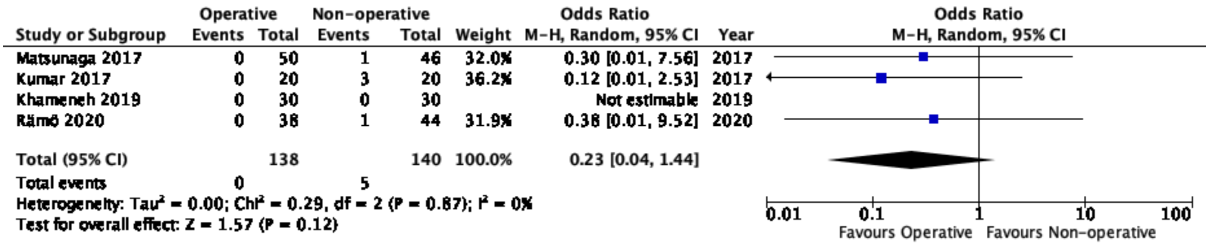
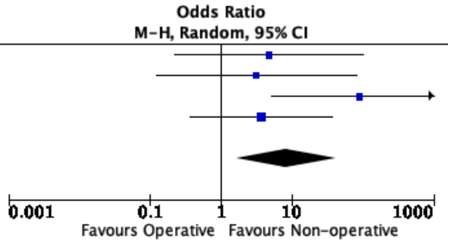


Figure 5: Forest plots for studies reporting complications following a humeral shaft fracture; CI, confidence interval; M-H, Mantel-Haenszel

(A) Transient radial nerve palsy

Study or Subgroup	Operative Events	Operative Total	Non-operative Events	Non-operative Total	Weight	Odds Ratio M-H, Random, 95% CI	Year
Matsunaga 2017	2	50	0	46	22.1%	4.79 [0.22, 102.54]	2017
Kumar 2017	1	20	0	20	20.0%	3.15 [0.12, 82.16]	2017
Khameneh 2019	18	30	0	30	24.3%	90.28 [5.04, 1616.50]	2019
Rämö 2020	3	38	1	44	33.6%	3.69 [0.37, 37.01]	2020
Total (95% CI)		138		140	100.0%	8.23 [1.62, 41.77]	
Total events	24		1				

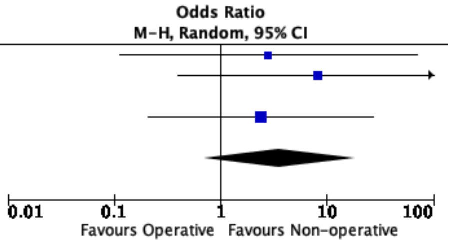
Heterogeneity: $Tau^2 = 0.66$; $Chi^2 = 3.94$, df = 3 (P = 0.27); $I^2 = 24\%$
Test for overall effect: Z = 2.54 (P = 0.01)



(B) Superficial infection

Study or Subgroup	Operative Events	Operative Total	Non-operative Events	Non-operative Total	Weight	Odds Ratio M-H, Random, 95% CI	Year
Matsunaga 2017	1	50	0	46	25.8%	2.82 [0.11, 70.93]	2017
Kumar 2017	3	20	0	20	29.2%	8.20 [0.40, 169.90]	2017
Khameneh 2019	0	30	0	30		Not estimable	2019
Rämö 2020	2	38	1	44	45.0%	2.39 [0.21, 27.43]	2020
Total (95% CI)		138		140	100.0%	3.57 [0.69, 18.38]	
Total events	6		1				

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.42$, df = 2 (P = 0.81); $I^2 = 0\%$
Test for overall effect: Z = 1.52 (P = 0.13)



(C) Re-intervention

Study or Subgroup	Operative Events	Operative Total	Non-operative Events	Non-operative Total	Weight	Odds Ratio M-H, Random, 95% CI	Year
Kumar 2017	2	20	2	20	31.6%	1.00 [0.13, 7.89]	2017
Matsunaga 2017	0	50	10	46	23.4%	0.03 [0.00, 0.61]	2017
Khameneh 2019	0	30	2	30	21.6%	0.19 [0.01, 4.06]	2019
Rämö 2020	0	38	13	44	23.4%	0.03 [0.00, 0.53]	2020
Total (95% CI)		138		140	100.0%	0.14 [0.02, 0.92]	
Total events	2		27				

Heterogeneity: $Tau^2 = 1.81$; $Chi^2 = 5.92$, df = 3 (P = 0.12); $I^2 = 49\%$
Test for overall effect: Z = 2.05 (P = 0.04)