Title: Proximal humeral fracture. A commentary on systematic reviews of surgical versus non-surgical management in older adults.

Authors: Pauline May[1] (Senior Physiotherapist), Joanna Harrison[2] (Research Fellow), Charlotte Williams[1] (Advanced Clinical Practitioner), James Hill [2] (Senior Research Fellow)

[1]Integrated Musculoskeletal Pain and Rheumatology Service, East Lancashire Hospitals NHS Trust, Burnley, UK.

[2]Synthesis, Economic Evaluation and Decision Science Group, University of Central Lancashire, Preston, UK.

Keywords: Proximal Humeral Fracture, management, surgical, non-surgical, older adults

Abstract:

Background

Proximal humeral fractures (PHFs) are common fractures in older adults and their prevalence is on the rise. Recovery following this fracture can be complex and disabling. Treatment varies from non-surgical management such as immobilisation to surgical procedures, with choice dependent on type and severity of fracture and patient health.

Objective

Several systematic reviews have considered the evidence for non-surgical versus surgical management of PHF in older adults. This commentary considers these findings for clinical practice and further research.

Methods

Three systematic reviews exploring non-surgical versus surgical management were selected based on the quality of their included evidence, and individually critically appraised. Findings from the reviews were reported for each outcome, and the implications considered for clinical practice and future research.

Results

Findings from the three reviews suggest that surgical management of PHF in older adults does not result in better functional outcomes or quality of life and non-surgical management should achieve acceptable upper limb function while decreasing the risks of surgery. More complex three-part fractures may also be managed non-surgically with fair to good functional results relative to fracture severity.

Conclusion

The findings align with current guidance to offer non-surgical management to uncomplicated cases of PHF in adults and older adults. More complex three-part PHFs may also be managed well non-surgically. There is however a lack of evidence and guidance on the specifics of rehabilitation for this type of management and further research is needed to evaluate the factors that contribute to the effectiveness of non-surgical interventions.

Introduction

Proximal humeral fractures (PHF, or shoulder fractures) are painful and debilitating injuries and account for approximately 6% of all adult fractures [1]. PHF symptoms include pain, swelling, and loss of movement [2], with functional capacity impaired for an average of two to three months [3]. Recovery from a shoulder fracture can be a long and often incomplete process that can be hindered by complications [4], including long-term consequences of mal union, non-union, avascular necrosis, and traumatic arthritis [5]. PHFs are also associated with a higher risk of hospitalisation or further fracture within the first year, and an increased utilisation of healthcare services and hospital costs [6,7].

The incidence of PHF varies with estimations ranging from 45.7 per 100,000 person years in Australia to 60.1 per 100,000 in Southern Europe and 74.2 per 100,000 in Northern Europe, during the period 2016-2018 [8-10]. The incidence of PHF is also increasing over time [9,10,11], with significant increases in females and older adults [8,9,11]. Shoulder fractures are most common in people over 65 who fall from a standing height, accounting for the third most common fracture in this population [12-15]. The escalating incidence of PHF in the older population is driven by an aging population, a suspected decline in the bone health of older adults and an increase in more severe falls [16,17].

The management of PHF varies from non-surgical management to surgical procedures, with choice of treatment depending on factors such as fracture type, severity and patient health status [18,19]. Non-surgical management of PHF usually involves a period of immobilisation (typically of three-four weeks) providing support and pain relief, followed by physiotherapy to restore function and mobility [20]. Variation exists in the recommended period of immobilisation [21], however evidence suggests that early mobilisation (within one week) may have beneficial effects on function [22]. Current guidelines in England advise that surgical management should be considered for complex PHF in adults, whereas non-surgical management is recommended for uncomplicated injuries (National Institute for Health and Care Excellence [23]. The most common definition for PHF is the Neer classification system with fractures defined by the number of parts involved (one to four part) [24].

The increasing incidence of PHF, together with the uncertainty of treatment options, variations in practice and emerging research, all endorse the need for updated evidence. This commentary aims to critically appraise the methods used in three systematic reviews exploring surgical versus non-surgical management for PHF in older adults; Beks et al. 2018 [25], Handoll et al. 2022 [26], and non-surgical management in more complex three- and four-part fractures (Soler-Peiro et al. 2020) [27]. The findings are subsequently discussed in the context of clinical practice and further research.

Methods

The selection of reviews for this commentary was based on surgical versus non-surgical management of PHF in older adults, where evidence from the included studies' outcomes of interest was deemed to be of moderate to good quality. The reviews that matched these criteria explored randomised controlled trials (RCTs) or quasi RCTs pertinent to the treatment and rehabilitation of PHF in adults [26] or focused on surgical versus non-surgical management through RCTs and observational studies [25]. One review also reported outcomes for comparisons such as early mobilization versus delayed [26]. As our focus was on the comparison of surgical versus non-surgical treatment, only the outcomes related to this comparison were reported here. Despite the inclusion of similar trials across the two systematic reviews, we incorporated the less recent review [25], as the authors argued that the addition of observational studies provided a broader study population. Furthermore, they included an analysis of function by Constant-Murley score which was reported in the more recent review [26], but with limited data. This commentary also reports on a review of non-surgical management for more complex fractures (three and four-part fractures) that explored RCTs and observational studies for three- or four-part PHFs [27].

Using the PICO (Population, Intervention, Control, Outcome) variables, methodological components of clinical evidence were compared for each review (Table 1). Inclusion criteria were not specific to older adults, however all three reviews subsequently included older populations (mostly over 60). Exclusion criteria from the studies within the reviews consisted of fracture dislocations, open fractures, multiple trauma, clear indication for surgery and comorbidities precluding surgery. Outcomes for the three reviews included upper-limb function, quality of life, additional surgery, and adverse events. In one review, secondary outcomes for constant score, pain and power were reported for a limited number of studies and downgraded to mostly low or very low certainty evidence [26], and are therefore not reported here. Two reviews reported outcome follow-up periods of at least one year [25, 27], and one review reported at six months, one and two years [26].

(Insert Table 1 here)

Using the Joanna Briggs Institute critical appraisal checklist for systematic reviews and research syntheses [28], all three systematic reviews were judged to be methodologically robust (Table 2) with some areas of concern. These were: 1) lack of publication bias assessment in [26, 27], explained as being due to an insufficient number of trials, and 2) an unclear description of the number of reviewers for critical appraisal [27]. The use of an arbitrary score for study quality [25] was also questioned due to the difficulties this poses for valuing the importance of individual items. However, the subsequent analysis included studies of all quality and good quality which allowed for comparison. Thus, despite some concerns, the three systematic reviews were overall deemed to provide an accurate and comprehensive summary of the evidence available.

(Insert Table 2 here)

Unions of effect

Effect sizes are reported as mean difference (MD), standardised mean difference (SMD) or Risk Ratio (RR). SMD effect sizes are interpreted as small (0.2), moderate (0.5) or large (0.80 with a significance level of p=0.05 [29]. Heterogeneity is reported using the I^2 statistic and interpreted as 0-40% (might not be important), 30-60% (may represent moderate), 50-90% (may represent substantial), 75-100% (may represent considerable) [30].

Results

Study characteristics (including reported primary outcome measures) are described for the three systematic reviews in Table 3.

(Insert Table 3 here)

Estimates of effectiveness from the meta-analyses reported in Beks et al. 2018 [25] and Handoll et al. 2022 [26] can be found in Table 4. These include the reported outcomes of function, quality of life,

mortality, major reinterventions, adverse events, and include as assessment of quality. The conservative treatment of more complex fractures including the review by Solar-Peiro 2020 [27] is reported as a narrative only.

(Insert Table 4 here)

Function

The most recent review [26] reported no important clinical difference in patient reported functional outcomes (physical function or shoulder and upper limb function) at six months, one- and two-years follow-up comparing surgical and non-surgical treatments of PHF, based on high certainty evidence [the authors have confidence that the true effect is similar to the estimated effect]. The earlier review [25] found similar findings in that there was no functional difference between the two groups at least one year post follow-up, based on mostly good quality evidence but with substantial heterogeneity. A sub-analysis of studies interpreted as good quality, showed no difference in surgical versus non-surgical treatment [25](MD=0.55, 95% CI: -2.93 to 4.03, p=0.76).

Quality of life

One review [26] reported no clinically important difference in quality of life (EQ-5D score >0.12) between surgical and non-surgical treatment at one and two years follow up, based on high-and moderate certainty evidence respectively.

Mortality

One review [26] reported no or little difference in mortality up to two years follow-up between surgical and non-surgical treatment, based on low certainty evidence [the true effect might be markedly different from the estimated effect] and no reported heterogeneity.

Major reinterventions

Major re-interventions (additional and unplanned surgery for implant removal) occurred statistically more often with surgical treatment compared to non-surgical based on mostly good quality evidence and no reported heterogeneity [25]. A sub-analysis of studies interpreted as good quality showed a similar result (RR=2.52, 95% CI: 1.55 to 4.11). One review [26] reported a statistically higher risk of additional or secondary surgery in the surgery treatment group at two-year follow-up based on low certainty evidence.

Adverse events

One review [26] reported a non-significant, higher risk of complications with surgery at two-year follow-up based on low certainty evidence (RR=1.46, 0.92 to 2.31, p=0.11). Looking at complications individually, one review [26] reported that nonunion and avascular necrosis were more common in the non-surgical group but stated that the clinical implications of these radiological findings were unclear as many cases were asymptomatic. One review [25] also reported that nonunion was statistically more common in the non-surgical group and there was no difference in the rate of avascular necrosis based on mostly good quality evidence and low reported heterogeneity. A sensitivity analysis of good quality studies maintained these findings.

Conservative treatment of more complex fractures

Treatment of three-part fractures with conservative management resulted in fair to good functional outcomes (mean constant score, 64.5) at a minimum of 12 month follow up, based on evidence considered by the study authors to be mostly good quality [27]. For four-part fractures, lower functional outcomes were achieved (mean constant score 54.9). There were some complications reported for three and four-part fractures treated conservatively (21% malunion, 9% avascular necrosis) with less avascular necrosis reported in three-part, compared to four-part fractures (7 and 10% respectively). Malunion however was higher in the three-part fractures (27%) compared to four-part fractures (17%). Consolidation was achieved in 96% of three-part fractures and 90% of four-part fractures.

A sub-group analysis in Beks et al. 2018 [25] reported that in studies where patients with a three- or four-part fracture underwent treatment, there was no difference in functional outcome between operative and non-operative treatment (SMD 0.02, 95% CI: –0.20 to 0.24, *p*=0.86).

Commentary

Using the JBI checklist [28], the three reviews overall can be considered to provide an adequate and comprehensive summary of evidence that address the question of interest. The findings suggest that for older adults, surgical management of PHF does not typically lead to better functional outcomes or quality of life compared to non-surgical approaches. Non-surgical management is likely to provide acceptable upper limb function while also reducing the risks associated with surgery. It is worth noting that in one review the functional outcome reported is based on studies of substantial heterogeneity [25]. However, the other review [26] reported high certainty GRADE evidence for functional outcomes.

These results align with NICE recommendations to offer non-surgical management as a definitive treatment for uncomplicated PHF in adults [23], and the review findings show that this is also relevant for older adults. Based on the review of more complex fractures [27], most three-part PHFs can also be managed non-surgically with fair to good functional results (in accordance with the severity of the fracture), a high rate of consolidation and few complications. Four-part PHFs also achieved a high rate of consolidation from non-surgical management and few complications but with poorer functional results than three-part PHFs. It is worth noting that in Handoll et al. 2022 [26], 66% of the fractures in the study population were also three- or four-part fractures and in Beks et al. 2018 [25], a sub-group analysis of three- and four-part fractures showed no

difference in functional outcome between surgical and non-surgical treatment. Current NICE guidance however recommends that surgical management is considered for those with complicated fractures such as fractural dislocation or a split of the humeral head [23].

Despite the data supporting the use of non-surgical management for PHF, there is a lack of current evidence and guidance on the specifics of rehabilitation for this type of management. The effectiveness of early versus delayed mobilisation after injury was explored, but the available data for this comparison were limited and uncertain [26]. Similarly, another systematic review found that early mobilisation may have a beneficial effect on function, but quality of evidence was low [31]. A more recent systematic review comparing early mobilisation (one week) to three-week immobilisation suggested early mobilisation may be beneficial for improving function at 6 month follow-up with long-term results less certain [22]. Exercise programmes for PHF, supervised or non-supervised have not been shown to reduce impairment or improve activity [32]. The consequences of immobilising older people however, should be considered due to the potential impact of physical inactivity on both physical and mental health [33]. Where prescription of exercise is appropriate, evidence has suggested that starting exercise early combined with a shorter immobilisation period may be more effective than a longer immobilisation period [31,32,34]. When considering intensity of supervised exercise, one trial reported no advantages to a more intensive rehabilitation regime over a conventional programme [35]. Exercise programmes can also be managed at home [34] with high satisfaction levels reported by patients due to good functional outcome, the availability and ease of being at home and maintaining independence [36].

Given the findings supporting a non-surgical approach to PHF management, it may be useful to provide further guidance on what this means to older patients, notably that non-surgical treatment should achieve acceptable upper limb function without the risks of surgery. Providing information to patients following a fracture is recommended within NICE guidelines [23] and should include expected outcomes of treatment, activities to work on independently, homecare options if needed and information on rehabilitation, mobilisation and weight bearing. For older patients, a booklet may be preferable to other formats [36]. In addition to information provision, positive relationships with healthcare professionals following PHF in the older population contributes to increased levels of patient trust, perceptions of recovery and improvement in emotional state [37]. Communication of treatment options and consideration of other risk factors for poor function could therefore be explored by healthcare professionals when treating patients post PHF. For example, social deprivation is associated with an increased incidence of adult fractures [38], and in those over 60, longer hospital stays, hospital readmission and higher mortality [39]. Another factor to consider for patients with PHF is psychological health and its impact on recovery. The reviews did not specifically address psychological outcomes for non-surgical vs surgical treatment, yet in recovery from a fracture, high fear avoidance beliefs and levels of catastrophising have been shown to substantially increase the risk of future pain and less than full recovery of strength respectively [40]. Self-efficacy interventions such as goal focused rehabilitation may help to improve coping abilities, reduce anxiety and depression and improve quality of life in people with post-traumatic fractures [41]. At present, there is no clear guidance to provide direction for these psychological factors when considering PHF management and NICE guidelines would benefit from an update.

Questions remain around the most effective rehabilitation protocol following non-surgical management of PHF. Further research is needed to evaluate the factors that contribute to the effectiveness of non-surgical interventions for PHF which may include sling use, exercise programmes, psychological support, and provision of patient information.

Acknowledgement/Disclaimer:

This research was partly-funded by the National Institute for Health and Care Research Applied Research Collaboration North-West Coast (NIHR ARC NWC) and the NIHR Pre-Doctoral Clinical Academic Fellowship (NIHR PCAF). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

Conflicts of interest: The authors have no conflict of interest to report.

References:

[1] Court-Brown CM, & Caesar B. Epidemiology of adult fractures: A review. Inj. 2006;37:691–697.

[2] John Hopkins Medicine. [internet]. Humerus fracture (upper arm fracture). John Hopkins Medicine; 2023 [cited 2024 Feb 14). Available from: https://www.hopkinsmedicine.org/health/conditions-and-diseases/humerus-fracture-upper-arm-fracture

[3] Lee SH, Dargent-Molina P, & Bréart G. Risk Factors for Fractures of the Proximal Humerus: Results From the EPIDOS Prospective Study. J. Bone Miner. Res. 2002;17:817–825.

[4] Handoll HH, Keding A, Corbacho B, Brealey SD, Hewitt C, & Rangan A. 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. Bone Joint J. 2017;99B:383–392.

[5] Pandey R, Raval P, Manibanakar N, Nanjayan S, McDonald C, Singh H. Proximal humerus fractures: A review of current practice. J Clin Orthop Trauma. 2023;10:102233.

[6] Maravic M, Briot K, & Roux C. Burden of proximal humerus fractures in the French National Hospital Database. Orthop Traumatol-Sur. 2014;100:931–934.

[7] Clinton J, Franta A, Polissar NL, Neradilek B, Mounce D, Fink HA, Schousboe JT, & Matsen FA. Proximal Humeral Fracture as a Risk Factor for Subsequent Hip Fractures. J Bone Joint Surg AM. 2009;91:503–511.

[8] Iglesias-Rodríguez S, Domínguez-Prado DM, García-Reza A, et al. Epidemiology of proximal humerus fractures. *J Orthop Surg*. 2021(16):402.

[9] Klug A, Gramlich Y, Wincheringer D, et al. Trends in surgical management of proximal humeral fractures in adults: a nationwide study of records in Germany from 2007 to 2016. Arch Orthop Trauma Surg. 2019;139:1713–1721.

[10] McLean AS, Price N, Graves S, Hatton A, Taylor FJ. Nationwide trends in management of proximal humeral fractures: an analysis of 77,966 cases from 2008 to 2017. *J Shoulder Elbow Surg*. 2019;28:2072–2078.

[11] Dimai HP, Svedbom A, Fahrleitner-Pammer A, Pieber T, Resch H, Zwettler E, Thaler H, Szivak M, Amrein K, Borgström F. Epidemiology of proximal humeral fractures in Austria between 1989 and 2008. Osteoporos Int. 2013;24:2413-21.

[12] Kelsey JL, Browner WS, Seeley DG, Nevitt MC, Cummings SR. Risk factors for fractures of the distal forearm and proximal humerus. Am J Epidemiol. 1992;135:477-489.

[13] Launonen AP, Lepola V, Saranko A, Flinkkila T, Laitinen M, Mattila VM. Epidemiology of proximal humerus fractures. Arch Osteoporos. 2015;10:2.

[14] Lauritzen JB, Schwarz P, Lund B, McNair P, Transbol I. Changing incidence and residual lifetime risk of common osteoporosis-related fractures. Osteoporos Int. 1993;3:127-132.

[15] Seeley DJ, Browner WS, Nevitt MC, Genant HK, Scott JC, Cummings SR. Which fractures are associated with low appendicular bone mass in elderly women? Ann Intern Med. 1991;115:837-842.

[16] Palvanen M, Kannus P, Niemi S, & Parkkari J. Update in the epidemiology of proximal humeral fractures. Clin. Orthop. Relat. Res. 2006; 442:87–92.

[17] Vachtsevanos L, Hayden L, Desai AS, Dramis A. Management of proximal humerus fractures in adults. World J Orthop. 2014;18:5:685-93.

[18] Davey MS, Hurley ET, Anil U, Condren S, Kearney J, O'Tuile C, Gaafar M, Mullett H, & Pauzenberger L. Management options for proximal humerus fractures – A systematic review & network meta-analysis of randomized control trials. Inj. 2022;53:244–249.

[19] Kancherla VK, Singh A, & Anakwenze OA. Management of Acute Proximal Humeral Fractures. J M Acad Orthop Surg. 2017;25:42–52.

[20] Martinez-Catalan N. Conservative Treatment of Proximal Humerus Fractures: When, How, and What to Expect. Curr Rev Musculoskelet Med. 2023;16:75-84.

[21] Bhambra A, Souroullas P, Wright AP, Gandhi M. Evidence-based management of proximal humerus fractures. Orthopaedics and Trauma. 2023; 37:296-302.

[22] Ataei M, Moradi A, Ebrahimzadeh MH, Rastaghi S, Daliri M. Immobilization Period for the Non-Operative Treatment of Proximal Humerus Fractures: Systematic Review and Meta-Analysis. Arch Bone Jt Surg. 2024;12:223-233.

[23] National Institute for Health and Care Excellence (NICE). Fractures (non-complex): Assessment and management Guideline NG38 [Internet]. NICE 2016; [cited 2024, Feb 14]. Available from: www.nice.org.uk/guidance/ng38

[24] Neer CS. Displaced proximal humeral fractures. I. Classification and evaluation. J Bone Joint Surg Am. 1970;52:1077–89.

[25] Beks RB, Ochen Y, Frima H, Smeeing PJ, Van Der Meijden O, Timmers TK, Van Der Velde D, Van Heijl M, Leenen LPH, Groenwold RHH, & Marijn Houwert R. Operative versus nonoperative treatment of proximal humeral fractures: a systematic review, meta-analysis, and comparison of observational studies and randomized controlled trials. J Shoulder Elbow Surg. 2018;27:1526-1534.

[26] Handoll HHG, Elliott J, Thillemann TM, Aluko P, & Brorson S. Interventions for treating proximal humeral fractures in adults. Cochrane Database Syst. Rev. 2022;Issue 6:Art. No.: CD000434.

[27] Soler-Peiro M, García-Martínez L, Aguilella L, & Perez-Bermejo M. Conservative treatment of 3-part and 4-part proximal humeral fractures: A systematic review. J. Orthop. Surg Res. 2020;15:347.

[28] Joanna Briggs Institute (JBI). Checklist for Systematic Reviews and Research Syntheses [internet]. JBI 2017; [cited 2024 Feb 14). Available from https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Systematic_Reviews2017_0.pdf

[29] Cohen J. Statistical Power Analysis in the Behavioral Sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates, Inc. 1988.

[30] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.1 [Internet]. Cochrane 2020; [cited 2024 Feb 14]. Available from https://training.cochrane.org/handbook/archive/v6.1/chapter-10

[31] Østergaard HK, Mechlenburg I, Launonen AP, Vestermark MT, Mattila VM, & Ponkilainen VT. The Benefits and Harms of Early Mobilization and Supervised Exercise Therapy after Non-surgically Treated Proximal Humerus or Distal Radius fracture: A systematic Review and Meta-analysis. Curr. Rev. Musculoskelet. Med. 2021;14:107–129.

[32] Bruder AM, Shields N, Dodd KJ, & Taylor NF. Prescribed exercise programs may not be effective in reducing impairments and improving activity during upper limb fracture rehabilitation: a systematic review. J Physiother. 2017;63:205–220.

[33] Cunningham C, O' Sullivan R, Caserotti P, & Tully MA. Consequences of physical inactivity in older adults: A systematic review of reviews and meta-analyses. Scand J Med Sci Sports. 2020;30:816–827.

[34] Aguado HJ, Ariño B, Moreno-Mateo F, Bustinza EY, Simón-Pérez C, Martínez-Zarzuela M, García-Virto V, Ventura PS, & Martín-Ferrero MÁ. Does an early mobilization and immediate home-based self-therapy exercise program displace proximal humeral fractures in conservative treatment? Observational study. J Shoulder Elbow Surg. 2018;27:2021–2029.

[35] Carbone S, Razzano C, Albino P, & Mezzoprete R. Immediate intensive mobilization compared with immediate conventional mobilization for the impacted osteoporotic conservatively treated proximal humeral fracture: a randomized controlled trial. Musculoskelet. Surg. 2017;101(Suppl 2):137–143.

[36] Aguado HJ, Ventura-Wichner, PS, Perez-Hickman L, Polo-Pérez I, Alonso-Olmo JA, Bragado M, Pereda-Manso A, Martínez-Zarzuela M, García-Virto V, Simón-Pérez C, Barajas EJ, & Martín-Ferrero MA. Patient Satisfaction Using a Home-Based Rehabilitation Protocol for the Non-Surgical Treatment of Proximal Humeral Fractures: A Prospective Longitudinal Cohort Study. Geriatr Orthop Surg Rehabil. 2021;12:1–8.

[37] Sabharwal S, Archer S, Cadoux-Hudson D, Griffiths D, Gupte CM, & Reilly P. Exploring elderly patients' experiences of recovery following complex proximal humerus fracture: A qualitative study. J. Health Psychol. 2021;26:880–891.

[38] Court-Brown CM, Aitken SA, Duckworth AD, Clement ND, & McQueen MM. The relationship between social deprivation and the incidence of adult fractures. J Bone JT Surg. 2013;95:e32.

[39] Patel R, Bhimjiyani A, Ben-Shlomo Y, & Gregson CL. Social deprivation predicts adverse health outcomes after hospital admission with hip fracture in England. Osteoporos. Int. 2021;32:1129–1141.

[40] Steven JL, Buer N, Samuelsson L, Harms-Ringdahl K. Pain-related fear, catastrophizing and pain in the recovery from a fracture. Scand J Pain. 2010;1:38-42.

[41] Ma Y, Wu H, Wang Y, Guo L. A feasibility study of modified self-efficacy for the improvement of adverse emotions and quality of life in traumatic fracture patients. Am J Transl Res. 2021;15:13:6507-6515.

[42] Furlan AD, Pennick V, Bombardier C, & Van Tulder M. 2009 Updated method guidelines for systematic reviews in the cochrane back review group. Spin. 2009;34:1929–1941.

Table 1: PICO variables for the three included systematic reviews

PICO	Beks et al. (2018) [25]	Handoll et al. (2022) [26]	Soler-Peiro et al. (2020) [27]
Population	Adults, proximal humeral fracture.	Adults, proximal humeral fracture.	Adults, three and four-part proximal humeral fracture.
Intervention	Surgical management. External osteosynthesis as an operative treatment was excluded.	Non-surgical and surgical management. Pharmacological, biological and acupuncture trials were excluded.	Conservative management. Surgical treatment was excluded.
Comparison	Non-surgical management.	Two or more treatments for management of PHF (our focus on surgical versus non-surgical).	None.

Outcome	Functional outcomes and complications including major reinterventions (additional and unplanned surgery), and adverse events.	For comparison of surgical versus non-surgical: functional outcomes, health related quality of life, mortality, additional surgery and adverse events.	Functional outcomes, complications and consolidation.

346

347

348

349

350

351

352

353

354 **Table 2: JBI critical appraisal checklist for systematic reviews [28]**

Criteria	Beks et al. (2018) [25]	Handoll et al. (2022) [26]	Soler-Peiro et al. (2020) [27]
Is the review question clearly and explicitly stated?	Yes: 'To compare operative versus nonoperative treatment of displaced PHF'	Yes: 'To assess the effects (benefits and harms) of treatment and rehabilitation interventions for proximal humeral fractures in adults'.	Yes: 'To assess criteria for indications, treatment protocols, and outcomes obtained with conservative treatment of three--part and four-part PHFs'.
Were the inclusion criteria appropriate for	Yes: PICO structure was followed according to question.	Yes: PICO structure was followed according to question.	Yes: PICO structure was followed according to question.

the review question?			
Was the search strategy appropriate?	Yes: A clear search strategy addressing each of the identifiable PICO components of the review question was conducted up to September 5th, 2017. Studies in a language other than English, Dutch or German were excluded.	Yes: A clear search strategy addressing each of the identifiable PICO components of the review question was conducted up to September 2020. No language or publication restrictions.	Yes: A clear search strategy addressing each of the identifiable PICO components of the review question was conducted from 2000-January 20th, 2020. Restricted to English publications.
Were the sources and resources used to search for studies appropriate?	Yes: MEDLINE, Embase, CENTRAL and CINAHL. Reference and citation tracking was performed.	Yes: CENTRAL, MEDLINE, Embase, CINAHL, AMED and PEDro. Trial databases, reference lists and conference proceedings were also searched.	Yes: PubMed and the Cochrane Library.
Were the criteria for appraising studies appropriate?	Yes: Methodological quality was assessed using the Methodological Index for Non-Randomised Studies (MINORS). Scores ranged from 0-24 with an author interpreted score of 16+ representing good methodological quality.	Yes: Risk of bias was assessed using the Cochrane handbook, plus four other aspects of trial quality. The GRADE approach was used to rate the certainty of evidence: very low, low, moderate or high.	Yes: Risk of bias was evaluated [33] and considered to be low risk (good quality) when 6/12 criteria were met.
Was critical appraisal conducted by two or more	Yes: Critical appraisal was carried out by two reviewers independently and disagreements resolved by a third reviewer.	Yes: Critical appraisal was carried out by two reviewers independently and differences resolved through discussion.	Unclear: No indication as to how many reviewers evaluated risk of bias.

reviewers independently?			
Were there methods to minimise errors in data extraction?	Yes: Data extraction was completed independently by two reviewers with a data extraction file.	Yes: two reviewers independently completed a data extraction tool. Differences were discussed.	Yes: A piloted data extraction form was completed independently by two reviewers.
Were the methods used to combine studies appropriate?	Yes: Outcomes reported by two or more studies were pooled in a meta-analysis. When heterogeneity was present, a random-effects model was used.	Yes: Where possible, data were pooled using both fixed-effect and random-effects models (depending on clinical heterogeneity).	Yes: A descriptive synthesis of outcomes was reported.
Was the likelihood of publication bias assessed?	Yes: Inspection of a funnel plot of the primary outcome measure. Publication bias not detected.	No: not assessed due to - insufficient number of trials to merit production of funnel plots.	No: not assessed.
Were recommendations for policy and/or practice supported by the reported data?	Yes: Appropriate recommendations were made based on the findings of the review.	Yes: Appropriate implications for practice were made based on the findings of the review.	Yes: Appropriate implications for practice were made based on the findings of the review.
Were the specific directives for new research appropriate?	N/A: no recommendations for new research were made	Yes: a need for similar trials to address key treatment uncertainties and optimisation of non-surgical treatments, plus decisions	Yes: Future research of conservative treatments conservative treatment of PHFs, including subgroups of fractures

		on priority topics identified from the ongoing trial data.	and comparing diverse treatment protocols
Total criteria	11/11	10/11	9/11

355

356

357 **Table 3. Study characteristics of Beks et al. 2018 [25], Handoll et al. 2022 [26] and Soler-Peiro et al.**

358 **2020 [27]**

Systematic Review	Number of included studies	Participants	Primary outcome	Follow-up period
Beks et al. (2018) [25]	22 studies (7 RCTs, 15 observational studies)	Total of 1743 patients of which the average age was 68 years, and 75% were women. Included patients with two-, three- or four--part fractures (Neer classification).	The primary outcome measure for function was the Constant-Murley Score.	Follow-up ranged from 12 to 86 months. Reported as at least one year.
Handoll et al. (2022) [26]	For the treatment comparison of surgical versus non-surgical, there were 10 RCTs included.	For the treatment comparison of surgical versus non-surgical, there were 717 participants of which 66% were three or four-part fractures (Neer classification). Most participants were over 60 and over two-thirds were women.	The primary outcome for function was measured using four different scores: The American Shoulder and Elbow Surgeons (ASES), the Disability of the Arm, Shoulder, and Hand questionnaire (DASH), Oxford Shoulder Score (OSS) and Simple Shoulder Test (SST). Quality of life was evaluated using the EQ-5D.	For the pooled results, the follow-up period was up to two years. Reported as 6 months, 1 and 2 years.

Soler-Peiro et al. (2020) [27]	6 studies (3 RCTs, 3 observational)	133 patients, of which the average age was 74, and 79% were women. Using the Neer classification, there were 41% three-part fractures and 59% four-part fractures.	The primary outcome for function was the Constant-Murley Score.	Follow-up was between 12 to 68 months. Reported as a minimum follow-up of one year

359

360

361

Table 4. Estimates of effectiveness for surgical versus non-surgical treatment on outcomes of function, quality of life, adverse events, and mortality outcomes (Beks et al. 2018 [25]; Handoll et al.2022[26])

Systematic Review	**Number and type of trial**	**Follow-up period**	**Estimate of effect** MD, SMD, RR (95% CI), *p* value, I^2	**Interpretation of effect and heterogeneity**	**Quality Assessment of included studies** (summary)
Functional outcome					
Beks et al. (2018) [25]	14 studies (5 RCTs, 9 observational)	At least 1 year	MD= -0.87 (-5.13 to 3.38), *p*=0.69, I^2=69%	No difference in functional outcome between groups, substantial heterogeneity.	Mostly good quality studies (11/14)
Handoll et al. (2022) [26]	3 RCTs	6 months	SMD = 0.17, (-0.04 to 0.38)	No clinically important difference in patient reported functional scores between groups, no reported heterogeneity	GRADE: Moderate Certainty

Handoll et al. (2022) [26]	7 RCTs	1 year	SMD= 0.10 (-0.07 to 0.27), p =0.24, I^2=0%	No clinically important difference in patient reported functional scores between groups, no reported heterogeneity.	GRADE: High certainty
Handoll et al. (2022) [26]	5 RCTS	2 years	SMD= 0.06, (-0.13 to 0.25), p=0.54, I^2=0%	No clinically important difference in patient reported functional scores between groups, no reported heterogeneity.	GRADE: High certainty
Major Reintervention					
Beks et al. (2018) [25]	15 studies (6 RCTs, 9 observational)	At least 1 year	RR= 2.72 (1.71 to 4.34), p= <.0001, I^2 =0%	Major reinterventions occurred more often in the surgical treatment than in non-surgical, no heterogeneity reported.	Mostly good quality studies (13/15)
Handoll et al. (2022) [26]	9 RCTs	Up to 2 years	RR 2.06 (1.21 to 3.51), p=0.007, I^2= 23%	A higher risk of additional surgery in the surgery group, low heterogeneity.	GRADE: low certainty
Nonunion					
Beks et al. (2018) [25]	13 studies (6 RCTs, 7 observational)	At least 1 year	RR =0.45 (0.23 to 0.89), p=.02, I^2 =0%	Surgical treatment resulted in fewer nonunions than non-surgical treatment, no heterogeneity reported.	Mostly good (11/13)
Handoll et al. (2022) [26]	8 RCTs	Up to 2 years	RR =0.42 (0.19 to 0.94), p=0.04, I^2=0%	Nonunion was more common in the non-surgical treatment group, no heterogeneity reported.	Unclear
Avascular Necrosis					

Beks et al. (2018) [25]	13 studies (6 RCTs, 7 observational)	At least 1 year	RR 1.24 (0.87 to 1.77), p=0.24, I^2 =24%	No difference in the rate of avascular necrosis between groups, low heterogeneity.	Mostly good quality studies (10/13)
Handoll et al. (2022) [26]	8 RCTs	Up to 2 years	RR 0.52 (0.33 to 0.81), p=0.004, I^2=50%	Avascular Necrosis was more common in the non-surgical treatment group, moderate heterogeneity.	Unclear
Quality of Life					
Handoll et al. (2022) [26]	6 RCTs	1 year	MD =0.01 (-0.02 to 0.04), p=0.51, I^2 =0%	No clinically important difference in quality of life between groups, no reported heterogeneity.	GRADE: high certainty evidence
Handoll et al. (2022) [26]	5 RCTS	2 years	MD=0.01 (-0.02 to 0.05), p=0.42), I^2=56%	No clinically important difference in quality of life between groups, moderate heterogeneity.	GRADE: moderate certainty evidence
Mortality					
Handoll et al. (2022) [26]	8 RCTs	2 years	RR 1.35 (0.70 to 2.62), p=0.37, I^2=0%	Little difference between groups, no reported heterogeneity.	GRADE: low certainty evidence

362