

## **Surgical versus conservative management of ankle fractures in adults: a systematic review and meta-analysis**

Omar A. Javed<sup>1\*</sup>, Qasim A. Javed<sup>2\*</sup>, Obioha Ukoumunne<sup>3</sup>, Livio Di Mascio<sup>4</sup>

\* These two authors contributed equally to this work as joint first authors

<sup>1</sup> Specialist Registrar in Trauma & Orthopaedic Surgery, University Hospitals Plymouth NHS Trust, Plymouth, PL6 8DH, UK

<sup>2</sup> Medical Student, University of Birmingham, Birmingham, B15 2TT, UK

<sup>3</sup> Associate Professor in Medical Statistics, NIHR CLAHRC South West Peninsula, University of Exeter Medical School, Exeter, EX1 2LU, UK

<sup>4</sup> Consultant Orthopaedic Surgeon, Barts and the Royal London Hospitals, London, E1 1BB, UK

### **Corresponding author:**

Omar Ali Javed

Trauma & Orthopaedic Surgery Department

University Hospitals Plymouth NHS Trust

Derriford Road

Plymouth, PL6 8DH, UK

Tel: 01752 202082

Email: [omar.javed@nhs.net](mailto:omar.javed@nhs.net)

### Acknowledgements

The authors would like to express their gratitude to the Royal Cornwall Hospital librarian, Katy Oak, for searching the literature. No funding was received for this project to influence on the study and the views expressed are those of the authors and not of any associated organisations.

Obioha Ukoumunne is funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) for the South West Peninsula at the Royal Devon and Exeter NHS Foundation Trust. The views expressed in this publication are those of the authors and not necessarily those of the NHS and the NIHR or the Department of Health and Social Care in England.

Word count: 4429 words

## **Aims**

This study compared outcomes of surgical versus conservative management of ankle fractures in adults through a systematic review and meta-analysis.

## **Methods**

We searched CINAHL, EMBASE, MEDLINE and CENTRAL databases (1946 to June 2019) for randomised and quasi-randomised controlled trials comparing surgical versus conservative management of closed adult ankle fractures of any type. Estimates of effect were pooled using random effects meta-analysis.

## **Results**

1153 patients from 7 trials were included. Our primary outcome, ankle function score, was not statistically significantly different at 6-months (pooled mean difference (surgical minus conservative) = 1.0; 95% CI: -2.3 to 4.3; p=0.55) or 12-months or more (pooled mean difference = 4.6; 95% CI: -1.0 to 10.2; p=0.11) between surgical and conservative groups in three trials assessing displaced or unstable fractures, and two trials using non-validated questionnaires. One trial assessing AO-type-B1 fractures without talar shift had a statistically significant difference favouring conservative management, which was not clinically meaningful. Surgery had lower rates of early treatment failure and malunion/non-union, but higher rates of further surgery and infection.

## **Conclusions**

Surgical and conservative management of displaced or unstable ankle fractures produce similar short-term functional outcomes. The higher risk of early treatment failure and malunion/non-union in the conservative group versus higher rates of further surgery and infection in the surgical group should be considered. Trials are needed to assess longer-term results and inform management of select patient groups.

## **Keywords**

ankle; fracture; systematic review; surgical; conservative; management; meta-analysis

## **1. Introduction**

Ankle fractures represent a rising healthcare burden in ageing populations and an ongoing debate exists regarding their optimal management [1,2]. It has been shown that only 1mm of lateral displacement of the talus can cause a 42% reduction in contact area between the talus and tibia, resulting in increased loads across the joint [3]. Restoration of anatomical congruency of the ankle joint following injury is therefore thought to be important to improve function and reduce post-traumatic osteoarthritis. Surgical or conservative methods are employed to allow the joint to heal in this way.

The incidence of ankle fractures has been reported as 168.7 per 100,000 people per year and they make up 10.2% of all fractures [2]. Many surgeons feel surgical management of unstable ankle fractures is more likely to maintain anatomical alignment, facilitate a quicker return to function and attain better results than conservative methods. However, the risks of surgery and potential complications lead to others favouring conservative management.

There is a large variability in practice and a study in the United States showed the rate of surgery for ankle fractures ranged from 14% to 72% depending on the hospital region [4]. A previous Cochrane review found insufficient evidence to conclude whether surgical or conservative management of ankle fractures in adults provided better outcomes [1]. Large, multicentre studies have been performed since then and provide further information. A more recent review by Larsen et al. performed searches up to 2017 and found equal short-term ankle function results between surgical and conservative groups [5]. However, they did not analyse outcomes such as time to weight-bearing, patient satisfaction and cost-effectiveness. Long-term follow up results of a large study were also published after the review by Larsen et al. An updated review with more comprehensive results was needed.

This review outlined results of randomised and quasi-randomised controlled trials comparing surgical and conservative management of adult ankle fractures. Ankle fractures are a heterogeneous group and we did not restrict studies based on displacement or stability, in order to summarise the best available evidence on this topic. The primary outcome was ankle function score.

## **2. Methods**

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist was used to report the study and the protocol was registered on PROSPERO [6,7].

### 2.1 Participants, Intervention and Comparison

Studies that reported on the management of acute ankle fractures in adults were included. Those with more extensive injuries, open fractures or studies evaluating paediatric or revision surgery were excluded.

Surgical interventions such as operative procedures with plates, screws, tension bands, internal or external fixation were compared with non-operative treatments such as plaster cast, walking cast or orthosis.

### 2.2 Outcomes

The primary outcome assessed was ankle function score. Examples include Olerud-Molander Ankle Score (OMAS) [8] and Foot and Ankle Outcomes Questionnaire (FAOQ) [9].

Other outcomes assessed were quality of life measures (e.g. Short Form 36 [SF-36] [10], Short Form 12 [SF-12] [11] and EuroQol 5D [EQ-5D] [12]), pain scores, ankle motion and patient mobility, complications, radiological outcomes, health resource use and cost-effectiveness, patient satisfaction, time to weight-bearing and time to return to work.

We considered outcomes measured before 12 months as short-term, and outcomes at 12 months or over as long-term.

### 2.3 Search methods

Our initial search was performed in March 2018 and then searches were updated up to June 2019 before publication to ensure no studies were missed. We searched MEDLINE (1946 to 3rd June 2019), EMBASE (1974 to 3rd June 2019), CINAHL (1946 to 3rd June 2019) and The Cochrane Central Register of Controlled Trials [CENTRAL] (inception to 3<sup>rd</sup> June 2019). The search strategies were an adaptation of Cochrane, Scottish Intercollegiate Guidelines Network [SIGN] filters and a previous Cochrane review (Donken et al. 2012) [1,13,14]. They

were developed with a research librarian and only texts available in English were in the final inclusion.

#### 2.4 Study selection

All studies were screened and duplicates removed. Two independent reviewers (O.A.J. and Q.A.J.) assessed titles and abstracts of all identified studies. Full texts of potentially relevant studies were assessed by the two independent reviewers and inclusion/exclusion criteria applied.

#### 2.5 Data collection and bias

The two independent reviewers completed data extraction using a piloted form. Study characteristics, characteristics of participants, intervention type, follow-up and outcome measures were extracted. Study authors were contacted for data where necessary.

Risk of bias was assessed independently by the two reviewers using the Cochrane Collaboration's Risk of Bias tool [15].

#### 2.6 Data synthesis

Where studies analysed the same outcome measure, we pooled the estimates using random effects meta-analysis using the DerSimonian-Laird method. Estimates were pooled at short-term follow-up and at long-term follow-up. The intervention effect was quantified using the mean difference for continuous outcomes and the risk ratio for binary outcomes. The pooled analyses were performed using means, standard deviations and sample sizes for continuous outcomes and numerator and denominator data for binary outcomes. Heterogeneity across studies was quantified using the I-square statistic. Analyses were carried out using Stata software (Version 15.1) and Review Manager (Version 5.3).

#### 2.7 Assessing quality of evidence

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) tool was used to assess quality of evidence of outcomes independently by the two reviewers (Table 1) [16].

### **3. Results**

#### 3.1 Study selection

The search yielded 9628 potentially eligible studies (Figure 1). After removal of duplicates, 5916 titles and abstracts were screened followed by 15 full-text articles, 7 of which met our inclusion criteria [17-23]. Six of the included studies were randomised controlled trials and 1 was a quasi-randomised controlled trial (Rowley et al. [21]) where the allocation was based on the last digit of the accident and emergency number. Overall, 1153 patients were recruited across the trials.

#### 3.2 Study characteristics

A full description of studies is given in Table 2. A summary of characteristics is provided below.

#### 3.3 Study size

The number of participants randomised in each trial were: 111 (Bauer et al. [17]), 43 (Makwana et al. [18]), 160 (Mittal et al. [19]), 96 (Phillips et al. [20]), 42 (Rowley et al. [21]), 81 (Sanders et al. [22]) and 620 (Willett et al. [23]).

### 3.4 Setting

The trials were conducted in hospitals in Sweden [17], the UK [18,21,23] Australia/New Zealand [19] and North America [20,22]. Three of the studies were large, multicentre trials [19,22,23].

### 3.5 Participants

Gender was reported by 6 studies [20,22,23] with a total of 376 males and 732 females. Age ranged from 15 to 81 years across the studies. Three studies looked at skeletally mature patients of all ages [17,20,21], two only included patients up to the age of 65 years [19,22], and two focussed on older patients above 55 years [18] and 60 years [23]. Three studies noted patients with diabetes in their study, which were in similar numbers between conservative and surgical groups [19, 22, 23].

Six of the seven studies examined displaced or unstable ankle fractures [17,18,20-23]. Displacement was either on initial radiograph or following external rotation stress test. One study (Mittal et al. [19]) investigated AO-type-B1 distal fibula fractures without significant talar shift [19,24].

### 3.6 Interventions

All studies compared Open Reduction Internal Fixation (ORIF) with closed reduction and cast immobilisation. Surgical techniques were according to AO/ASIF principles [25,26]. Of note, the largest trial in our review used a close contact casting technique under anaesthetic in operating theatre for the conservative group [23].

### 3.7 Outcomes

Time-points for data collection at follow-up ranged from 6 weeks to 8 years across the studies. The short-term and long-term results of one study were published by different authors and have been reported as such (Willett 2016 and Keene 2018) [23, 27]. The studies used different functional outcome tools such as OMAS [18,22,23,27], FAOQ [19] and non-validated questionnaires [17,20]. Validated quality-of-life measures were used by three studies [19,22,23,27]. All 7 studies reported on complications. Other outcomes that were assessed included radiological outcomes, ankle motion, health resource use or cost, patient satisfaction, time to weight-bearing and time to return to work.

### 3.8 Risk of bias

The assessment of bias is shown in Table 3 and detailed reasoning is found in Table 2.

### 3.9 Study results

#### *3.9.1 Primary outcomes*

Data from the three studies reporting OMAS scores were pooled (Figures 2 and 3) [18,22,23,27]. There was no statistically significant difference between surgical and conservative groups at 6-months (pooled mean difference [MD] = 1.0; 95% CI: -2.3 to 4.3; p

= 0.55;  $I^2 = 0\%$ ) and long-term follow-up (pooled MD = 4.6; 95% CI: -1.0 to 10.2;  $p = 0.11$ ;  $I^2 = 23\%$ ).

Mittal et al. [19] found a statistically significant difference in FAOQ score favouring conservative management at 6-month (MD = -2.7; 95% CI: -5.1 to -0.4;  $p = 0.025$ ) and 12-month follow-up (MD = -3.2; 95% CI: -5.9 to -0.4;  $p = 0.028$ ), but this was not clinically meaningful.

Bauer et al. [17] and Phillips et al. [20] used non-validated questionnaires. Bauer et al. [17] found no statistically significant difference at a mean follow-up of 7 years, between surgical and conservative management in significant ankle “troubles” (26% versus 29%; risk ratio (RR) = 0.90; 95% CI 0.46 to 1.76). Phillips et al. [20] found no significant difference in clinical score between the two groups at a mean follow-up of 3.5 years (88.8/100 versus 84.3/100, MD = 4.5).

### 3.9.2 Secondary outcomes

Data were pooled for three studies that reported Physical Component Summary (PCS) scores assessing quality-of-life [19,22,23,27] (Figures 4 and 5). Analysis showed no statistically significant difference between surgical and conservative groups at short-term (pooled MD = -0.05; 95% CI: -2.6 to 2.6;  $p = 0.97$ ;  $I^2 = 52\%$ ) and long-term follow-up (pooled MD = 1.5; 95% CI: -0.1 to 3.1;  $p = 0.06$ ;  $I^2 = 0\%$ ). The study by Willett et al. and Keene et al. [22,27] also did not find a statistically significant difference in EQ-5D at 6 months (MD = 0.001; 95% CI: -0.04 to 0.04) and 3 years (MD = -0.04; 95% CI: -0.09 to 0.01).

Data were pooled from six studies that reported on early treatment failure (Figure 6); surgery had a lower rate of failure (pooled RR = 0.12; 95% CI: 0.06 to 0.22;  $p < 0.001$ ;  $I^2 = 0\%$ ) [17,18,20-23]. Pooled analysis of six studies on infection (Figure 7) found a higher risk in surgical patients (pooled RR = 3.42; 95% CI: 0.99 to 11.76;  $p = 0.05$ ;  $I^2 = 35\%$ ) [17-20,22,23]. Rates of further surgery were pooled from four studies (Figure 8) and favoured conservative treatment (pooled RR = 3.38; 95% CI: 1.59 to 7.17;  $p = 0.002$ ;  $I^2 = 0\%$ ) [18,19,22,23]. There was a significant difference favouring surgery in three studies on malunion (pooled RR 0.23, 95% CI 0.13 to 0.43,  $p < 0.001$ ,  $I^2 = 0\%$ ) and delayed or non-union (pooled RR = 0.10; 95% CI: 0.04 to 0.28;  $p < 0.001$ ;  $I^2 = 0\%$ ) [18,22,23] (Figures 9 and 10). Other secondary outcomes are shown in Table 4.

## 4. Discussion

### 4.1 Main findings

This study reviewed 7 trials with a total of 1153 participants. Of these, 6 trials looked at displaced or unstable ankle fractures, whereas Mittal et al. [19] focussed on AO-type-B1 distal fibula fractures without significant talar shift. Overall, 4 of the 6 studies that reported functional outcomes found no difference between surgical and conservative management [17,20,22,23,27]. Of the remaining two, one study favoured surgery but was a small study with bias [18], and the other had ankle function scores favouring conservative management but was not clinically meaningful [19]. The two groups were similar for quality-of-life, pain, range of movement, venous thromboembolism (VTE), patient satisfaction, time to weight-bearing and return to work.

Surgery outperformed conservative management in achieving fewer incidences of early treatment failure, achieving more anatomic reduction, fewer occurrences of malunion/non-union and less readmissions.

Conservative management was more favourable with respect to infection rates, rates of further surgery after the acute period, and cost-effectiveness at 1-year [28,29].

Willett et al. [23] was the largest trial included in this review and contributed 54% (620/1153) of the total participants. There was a high rate of early treatment failure (26%) in the close contact casting group, primarily due to an inability to achieve or maintain reduction. At the initial intervention, 18 of the casting patients underwent surgery instead. After the initial intervention, 52 patients were converted to internal fixation and were kept in the per-protocol analysis as it was “allowable and expected as part of the close contact casting intervention pathway.” The other studies also reported higher levels of early treatment failure in the conservative group. However, this could be deemed to be acceptable in order to attempt to avoid the risks of surgery initially. The conservative group also had higher levels of malunion (42/275 [15%]) and non-union (28/274 [10%]). The 3-year follow-up by Keene et al. [27] stated that patients with malunion or non-union were found to have statistically significantly lower OMAS scores. This highlights the importance of achieving a satisfactory reduction and aiming for a well-united fracture.

Mittal et al. [19] was a high-quality study that investigated AO-type-B1 fractures with minimal talar shift, which showed better ankle function scores in the conservative group with fewer adverse events. They commented on a study showing a large variability in management of this injury in Australia, with around 40% of surgeons choosing operative management [30]. The difficulties of achieving and maintaining anatomical reduction with conservative methods do not apply in stable fracture patterns as they do with displaced and unstable fractures. Radiological outcomes and early treatment failure are therefore less problematic in this group. Although they did not elaborate on the stability of fractures and radiological outcomes, higher ankle function scores and fewer complications in the conservative group supports the view that undisplaced, stable B1 fractures can be managed conservatively.

Our review shows that surgical and conservative management of displaced or unstable ankle fractures in adults produce similar functional outcomes in the short-term and are both acceptable treatment modalities. An informed discussion can be held with the patient to balance the risks of early treatment failure, poor reduction and malunion/non-union in conservative treatment versus the risks of infection and further operations with the surgical route. Stable lateral malleolar fractures can be managed conservatively due to less risk of poor anatomical reduction and malunion/non-union. The long-term effects of having a poorly reduced ankle fracture needs more study, but the 3-year follow-up by Keene et al. [27] has raised concerns of lower ankle function scores in poorly united fractures. Development of arthritis is a concern too, although longer-term follow-up would be needed to investigate this.

#### 4.2 Strengths and limitations

Two previous reviews on this topic (Donken et al. [1] and Petrisor et al. [31]) could not provide meaningful conclusions due to insufficient evidence. More studies have been performed since then and have been included in this review. Our review was thorough and systematic with two independent reviewers screening studies, assessing risk of bias, and performing data extraction. The previous Cochrane review (Donken et al. [1]) only had one



author to screen titles and abstracts. We only included randomised and quasi-randomised controlled trials to find the best quality evidence and found recent multicentre trials with a low risk of bias. Overall, the quality of evidence was high according to the GRADE assessment. The studies all compared open reduction and internal fixation to cast or walking boot, with most studies focussing on displaced or unstable fractures. They took place in various countries and adopted different weight-bearing protocols. This allowed a review of surgical and conservative management of ankle fractures consistent with clinical practice across the world and we were able to draw clinically meaningful conclusions from the studies.

However, the large variability in practice across the trials could make it difficult to compare findings. Rowley et al. [21], a quasi-randomised controlled trial, was at high risk of bias for sequence generation and allocation concealment. Makwana et al. [18] did not blind outcome assessors and excluded patients after randomisation, leading to incomplete outcome data. All 7 studies were at high risk of bias for blinding of participants and personnel, which was not possible due to the nature of interventions. However, the more recent studies blinded outcome assessors for primary outcomes [19,22,23]. We were unable to perform subgroup analysis based on ankle fracture classification or patient age due to limited data and variability in selection criteria of the studies, but the majority of patients in the studies were Weber B fractures (Table 5).

Larsen et al. [5] produced a thorough systematic review and meta-analysis on this topic and found equal short-term functional results between surgical and conservative groups. Our study has more up-to-date searches to June 2019 and has included long-term results of the Ankle Injury Management (AIM) trial reported by Keene et al [27]. This study contributed large amounts of data to the meta-analysis of our primary outcome, ankle function score, which was not included in the meta-analysis by Larsen et al. [5] due to it being published after their searches were performed. The 3-year results by Keene et al. [27] has confirmed equivalent ankle function scores between the two groups at this time-point, although it raised an important finding that poorly united fractures at 6-months went on to have significantly lower OMAS scores at 3-years.

Our review also reported many secondary outcomes that were not analysed by Larsen et al. [5], such as patient satisfaction, time to weight-bearing, time to return to work, health resource use and cost-effectiveness. Overall however, the fact that we used different search strategies to Larsen et al. [5] and arrived at similar conclusions strengthens the findings of both studies.

#### 4.3 Implications for practice and research

Overall, our study has shown that surgical and conservative management of displaced or unstable ankle fractures in adults both produce similar functional outcomes in the short-term and are acceptable treatment modalities. It is important to note that our study did not show that unstable ankle fractures do not require surgery. Clinicians will be able to have a more informed discussion with patients regarding the benefits and risks of both management strategies. Surgery remains the primary treatment strategy for achieving better reduction, union and treatment success. Conservative methods provide an alternative management option, particularly where patients have a strong preference for avoiding surgery or there are concerns about infection.



The National Institute for Health and Care Excellence (NICE) guidelines offer limited recommendations for managing ankle fractures [32]. They state that undisplaced unimalleolar fractures can often be managed conservatively. However, they offer no recommendations on surgical versus conservative management of displaced or unstable ankle fractures. The British Orthopaedic Association Audit Standards for Trauma (BOAST) guidelines make a similar recommendation for stable ankle fractures based on the NICE guideline [33]. In addition, they state “in patients over 60 years close contact casts are an option if reduction can be maintained.” Following our review, we believe more meaningful recommendations should be implemented in guidelines. The superiority of surgery in obtaining and maintaining reduction and fracture union in unstable ankle fractures should be emphasised. It should be advised that patients should be informed of the risks of early treatment failure, malunion and non-union with conservative methods versus infection and further surgery for metalwork or wound complications with surgical management. The conservative management of stable lateral malleolar fractures should be supported by the evidence we have reported in this review.

More research is needed into the longer-term effects of surgical and conservative methods. A review by Larsen et al. [5] noted a lack of trials assessing long-term effects of these injuries and made similar recommendations. In particular, the rate of development of osteoarthritis in these two groups is largely unknown. Studies could also investigate the subgroups of patients that are more likely to benefit from one treatment choice over another, based on fracture pattern and patient characteristics.

# References

- [1] Donken CC, Al-Khateeb H, Verhofstad MH, van Laarhoven CJ. Surgical versus conservative interventions for treating ankle fractures in adults. *Cochrane Database Syst Rev.* 2012;8
- [2] Elsoe R, Ostgaard SE, Larsen P. Population-based epidemiology of 9767 ankle fractures. *Foot Ankle Surg.* 2018;24(1):34–9.
- [3] Ramsey PL, Hamilton WI. Changes in tibiotalar area of contact caused by lateral talar shift. *J Bone Jt Surg.* 1976 Apr 1;58(3):356-7.
- [4] Koval KJ, Lurie J, Zhou W, Sparks MB, Cantu RV, Sporer SM et al. Ankle fractures in the elderly: what you get depends on where you live and who you see. *J Orthop Trauma.* 2005 Oct 1;19(9):635
- [5] Larsen P, Rathleff MS, Elsoe R. Surgical Versus Conservative Treatment For Ankle Fractures In Adults: A systematic review and meta-analysis of the benefits and harms. *Foot Ankle Surg.* 2019 Aug; 25(4).
- [6] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 2009 Aug 18;151(4):264-9.
- [7] Javed O. Surgical versus conservative management of ankle fractures in adults: a systematic review [Internet]. PROSPERO 2018 CRD42018100174 [cited 14 November 2018] Available from: [http://www.crd.york.ac.uk/PROSPERO/display\\_record.php?ID=CRD42018100174](http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018100174)
- [8] Olerud C, Molander H. A scoring scale for symptom evaluation after ankle fracture. *Arch Orthop Trauma Surg.* 1984 Sep 1;103(3):190-4.
- [9] Johanson NA, Liang MH, Daltroy L, Rudicel S, Richmond J. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments: reliability, validity, and sensitivity to change. *J Bone Jt Surg.* 2004 May 1;86(5):902-9.
- [10] Stewart AL, Ware JE, Ware Jr JE, editors. *Measuring functioning and well-being: the medical outcomes study approach.* Durham and London: Duke University Press; 1992.
- [11] Ware Jr JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996 Mar 1;34(3):220-33.
- [12] Brooks R. EuroQol: the current state of play. *Health Policy.* 1996 Jul 1;37(1):53-72.
- [13] RCT filters for different databases, *Cochrane Work* [Internet]. [work.cochrane.org](http://work.cochrane.org) 2018 [cited 20 December 2017]. Available from: <http://work.cochrane.org/rct-filters-different-databases>
- [14] Search filters [Internet]. [sign.ac.uk](http://www.sign.ac.uk). 2018 [cited 14 November 2018]. Available from: <http://www.sign.ac.uk/search-filters.html>
- [15] Higgins JP, Green S. *Cochrane handbook for systematic reviews of interventions 5.1.0.* The Cochrane Collaboration. 2011 Mar:33-49. [cited 14 November 2018] Available from <http://www.handbook.cochrane.org>
- [16] Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ.* 2008 Apr;336(7650):924-6.
- [17] Bauer M, Bergström B, Hemborg A, Sandegård J. Malleolar fractures: nonoperative versus operative treatment. A controlled study. *Clinic Orthop Relat Res.* 1985 Oct(199):17-27.
- [18] Makwana NK, Bhowal B, Harper WM, Hui AW. Conservative versus operative treatment for displaced ankle fractures in patients over 55 years of age: a prospective, randomised study. *Bone Jt J.* 2001 May 1;83(4):525-9.
- [19] Mittal R, Harris IA, Adie S, Naylor JM. Surgery for type B ankle fracture treatment: a Combined Randomised and Observational Study (CROSSBAT). *BMJ Open.* 2017 Mar 1;7(3):e013298.
- [20] Phillips WA, Schwartz HS, Keller CS, Woodward HR, Rudd WS, Spiegel PG et al. A prospective, randomized study of the management of severe ankle fractures. *J Bone Jt Surg. American volume.* 1985 Jan;67(1):67-78.
- [21] Rowley DI, Norris SH, Duckworth T. A prospective trial comparing operative and manipulative treatment of ankle fractures. *Bone Jt J.* 1986 Aug 1;68(4):610-3.
- [22] Sanders DW, Tieszer C, Corbett B. Operative versus nonoperative treatment of unstable lateral malleolar fractures: a randomized multicenter trial. *J Orthop Trauma.* 2012 Mar 1;26(3):129-34.
- [23] Willett K, Keene DJ, Mistry D, Nam J, Tutton E, Handley R et al. Close contact casting vs surgery for initial treatment of unstable ankle fractures in older adults: a randomized clinical trial. *J Am Med Assoc.* 2016 Oct 11;316(14):1455-63.
- [24] Müller ME, Nazarian S, Koch P, Schatzker J. *The Comprehensive Classification of Fractures of Long Bones.* New York: Springer-Verlag; 1990.
- [25] Müller ME, Allgöwer M, Allgower M, Schneider R, Willenegger H. *Manual of internal fixation: techniques recommended by the AO-ASIF group.* Berlin: Springer Science & Business Media; 1991.

- [26] Rüedi TP, Murphy WM, Colton CL, Fackelman GE, Harder Y. AO principles of fracture management. New York: Thieme; 2000.
- [27] Keene DJ, Lamb SE, Mistry D, Tutton E, Lall R, Handley R, et al. Three-year follow-up of a trial of close contact casting vs surgery for initial treatment of unstable ankle fractures in older adults. *J Am Med Assoc*. 2018 Mar 27;319(12):1274-6.
- [28] Keene D, Mistry D, Nam J, Tutton E, Handley R, Morgan L, et al. The ankle injury management (AIM) trial: a pragmatic, multicentre, equivalence randomised controlled trial and economic evaluation comparing close contact casting with open surgical reduction and internal fixation in the treatment of unstable ankle fractures in patients aged over 60 years. *Health Technol Assess*. 2016 Oct 5;20(75).
- [29] Slobogean GP, Marra CA, Sadatsafavi M, Sanders DW. Is surgical fixation for stress-positive unstable ankle fractures cost effective? Results of a multicenter randomized control trial. *J Orthop Trauma*. 2012 Nov 1;26(11):652-8.
- [30] Ansari U, Adie S, Harris IA, Naylor JM. Practice variation in common fracture presentations: a survey of orthopaedic surgeons. *Inj*. 2011 Apr 1;42(4):403-7.
- [31] Petrisor BA, Poolman R, Koval K, Tornetta III P, Bhandari M, Evidence-Based Orthopaedic Trauma Working Group. Management of displaced ankle fractures. *J Orthop Trauma*. 2006 Aug 1;20(7):515-8.
- [32] Fractures (non-complex): assessment and management Guidance and guidelines [Internet]. NICE; 2018 [cited 26 November 2018]. Available from: <https://www.nice.org.uk/guidance/ng38>
- [33] BOAST 12: The Management of Ankle Fractures [Internet]. British Orthopaedic Association Audit Standards for Trauma; 2016 [cited 26 November 2018]. Available from: <https://www.boa.ac.uk/wp-content/uploads/2016/09/BOAST-12-Ankle-Fractures.pdf>

<b>Table 1. GRADE Quality Assessment of Trials for surgical versus conservative management of ankle fractures in adults</b>						
<b>No of studies</b>	<b>Study design</b>	<b>Risk of bias</b>	<b>Inconsistency</b>	<b>Indirectness</b>	<b>Imprecision</b>	<b>Other considerations</b>
Ankle function score (follow up: range 6 months to 7 years; assessed with: OMAS, FAOQ, self-validated scores)						
6	randomised trials	not serious	not serious	not serious	not serious	none
Quality-of-life score (follow up: range 6 months to 3 years; assessed with: SF-12, SF-36, EQ-5D)						
3	randomised trials	not serious	not serious	not serious	not serious	none
Pain (follow up: range 6 months to 7 years; assessed with: Self-made questionnaires, EQ-5D, VAS)						
3	randomised trials	serious <sup>a</sup>	not serious	not serious	not serious	none
Range of motion (follow up: range 6 months to 7 years; assessed with: Geniometer, Timed Up and Go, range of dorsiflexion)						
4	3 randomised trials & 1 quasi-randomised trial	serious <sup>b</sup>	not serious	not serious	not serious	none
Early treatment failure						
6	5 randomised trials & 1 quasi-randomised trial	serious <sup>c</sup>	not serious	not serious	not serious	none
Infection						
6	randomised trials	not serious	not serious	not serious	serious <sup>d</sup>	none
VTE						
3	randomised trials	not serious	serious <sup>e</sup>	not serious	not serious	none
Further surgery (follow up: range 6 months to 27 months)						
4	randomised trials	not serious	not serious	not serious	not serious	none
Malunion/non-union (follow up: range 6 months to 27 months)						
3	randomised trials	not serious	not serious	not serious	not serious	none
Health resource use						
3	randomised trials	not serious	not serious	not serious	not serious	none
Patient satisfaction						
2	randomised trials	not serious	not serious	not serious	not serious	none
Time to weight-bearing						
2	1 randomised trial & 1 quasi-randomised trial	not serious	not serious	not serious	not serious	none
Time to return to work						
2	randomised trials	serious <sup>f</sup>	not serious	not serious	serious <sup>g</sup>	none

Table 1. GRADE continued					
№ of patients		Effect		Certainty	Importance
Surgical	Conservative	Relative (95% CI)	Absolute (95% CI)		
Ankle function score (follow up: range 6 months to 7 years; assessed with: OMAS, FAOQ, self-validated scores)					
496	490	Most studies found no significance in ankle function scores between the 2 groups.		⊕⊕⊕⊕ HIGH	CRITICAL
Quality-of-life score (follow up: range 6 months to 3 years; assessed with: SF-12, SF-36, EQ-5D)					
411	403	The studies found no statistical significance between the 2 groups.		⊕⊕⊕⊕ HIGH	IMPORTANT
Pain (follow up: range 6 months to 7 years; assessed with: Self-made questionnaires, EQ-5D, VAS)					
335	342	There was no statistical significance in any of the 3 studies.		⊕⊕⊕○ MODERATE	IMPORTANT
Range of motion (follow up: range 6 months to 7 years; assessed with: Geniometer, Timed Up and Go, range of dorsiflexion)					
375	377	Three of the four studies found no statistical significance between the 2 groups.		⊕⊕⊕○ MODERATE	IMPORTANT
Early treatment failure					
9/463 (1.9%)	100/468 (21.4%)	<b>RR ranged from</b> 0.06 to 0.38 across the studies	not estimable	⊕⊕⊕○ MODERATE	IMPORTANT
Infection					
29/516 (5.6%)	7/518 (1.4%)	<b>RR ranged from</b> 0.12 to 12.82 across the studies	not estimable	⊕⊕⊕○ MODERATE	IMPORTANT
VTE					
11/433 (2.5%)	18/440 (4.1%)	<b>RR ranged from</b> 0.34 to 1.63 across the studies	not estimable	⊕⊕⊕○ MODERATE	IMPORTANT
Further surgery (follow up: range 6 months to 27 months)					
30/441 (6.8%)	8/432 (1.9%)	<b>RR ranged from</b> 2.47 to 4.88 across the studies	not estimable	⊕⊕⊕⊕ HIGH	IMPORTANT
Malunion/non-union (follow up: range 6 months to 27 months)					
15/341 (4.4%)	90/327 (27.5%)	<b>RR ranged from</b> 0.06 to 0.26 across the studies	not estimable	⊕⊕⊕⊕ HIGH	IMPORTANT
Health resource use					
Follow-up of Willett et al. <sup>22</sup> found significantly more theatre time for the surgical group, similar length of hospital stay between the 2 groups, and a higher readmission rate in the conservative group.				⊕⊕⊕⊕ HIGH	IMPORTANT
Patient satisfaction					

Table 1. GRADE continued					
№ of patients		Effect		Certainty	Importance
Surgical	Conservative	Relative (95% CI)	Absolute (95% CI)		
271	262	Satisfaction was similar between the 2 groups.		⊕⊕⊕⊕ HIGH	IMPORTANT
Time to weight-bearing					
312	295	Keene et al. <sup>27</sup> reported return to partial weight-bearing was similar between the 2 groups. Rowley et al. <sup>20</sup> was a smaller trial and found a delay of 4 weeks in the surgical group compared to conservative.		⊕⊕⊕⊕ HIGH	IMPORTANT
Time to return to work					
115	118	Both studies found similar time to return to work between the 2 groups.		⊕⊕○○ LOW	IMPORTANT

**CI:** Confidence interval; **EQ-5D:** EuroQol Five-dimensions **FAOQ:** Foot and Ankle Outcomes Questionnaire; **SF:** Short-Form Health Survey; **OMAS:** Olerud-Molander Ankle Score; **RR:** Risk ratio; **VAS:** Visual Analogue Score

#### Explanations

- Two of the three studies were at a high risk of bias (Bauer et al.<sup>16</sup> and Makwana et al.<sup>17</sup>).
- Three of the four studies were at a high risk of bias (Bauer et al.<sup>16</sup>, Makwana et al.<sup>17</sup> and Rowley et al.<sup>20</sup>).
- Four of the six studies were at a high risk of bias (Bauer et al.<sup>16</sup>, Makwana et al.<sup>17</sup>, Phillips et al.<sup>19</sup>, Rowley et al.<sup>20</sup>).
- The studies displayed wide confidence intervals, particularly Sanders et al.<sup>21</sup> (0.74 to 218.11)
- Differing results between the three studies.
- One of the two studies was at a high risk of bias (Bauer et al.<sup>16</sup>).
- The sample size was not large enough for the result to be precise.

Table 2. Characteristics of studies				
Study	Design/setting	Participants	Interventions	Outcomes
Bauer 1985	Randomised controlled trial (RCT), Sweden	111 randomised, baseline data for 108, follow-up for 92.  Mean age 44 years, range 16 to 77 years. 44 male, 64 female  Displaced malleolar fractures included. Weber C fractures excluded.	Surgery: Use of tourniquet, suturing of ligaments and a 24 hour suction drain. 6-weeks of non-weight-bearing (NWB), then partial-weight-bearing (PWB) for 3-weeks then full-weight-bearing (FWB).  Conservative: Closed reduction and plaster cast. Weight-bearing as in surgical group.	Mean follow-up time was 7 years. Range was 6 to 8 years.  A self-made non-validated questionnaire was used to obtain functional outcomes such as significant troubles, pain and stiffness. Secondary outcomes were range of motion with a goniometer, complications, radiological outcomes according to Cedell and Magnusson, time to discharge, time in cast and time of sick-leave.
Makwana 2001	RCT, United Kingdom	43 randomised, baseline data for 43, follow-up data for 31.  Mean age 66 years, range 55 to 81 years.  12 male, 31 female  Patients aged 55 years or over with a displaced ankle fracture included.	Surgery. AO principles and below-knee cast. PWB from 48 hours, change of cast at 2 weeks. FWB under supervision of physiotherapist on removal of cast.  Conservative: Below-knee cast. 48 hours limb elevation. PWB for 6 weeks. Weekly cast inspection and check xray for first 3-weeks. Cast removed and FWB with physiotherapist at 6 weeks.	Mean follow-up 27 months, range 15 to 42 months.  OMAS score, visual analogue score for pain, ankle girth, range of motion, inpatient stay, radiological outcomes and complications.
Mittal 2017	A pragmatic, multicentre, single-blinded, combined RCT and observational study in 22 hospitals in Australia and New Zealand.	436 patients recruited; 160 randomised and all 276 who declined randomisation were included in an observational cohort (not included in this review). Of the randomised cohort, follow-up data was obtained for 139 patients.  Mean age of 39 years (inclusion criteria allowed age between 18 and 65 years)  77 male, 83 female  AO type 44-B1 distal fibula fracture without significant talar shift included.	Surgery: AO principles. NWB in cast or boot post-operatively. Change of cast or boot at 10-14 days and allow FWB. Review at 6 weeks with x-ray and removal of cast or boot.  Conservative: FWB in walking boot. X-ray at 10-14 days. X-ray at 6 weeks and removal of cast or boot.	Length of follow-up was 12 months.  Primary outcomes were FAOQ and PCS using SF-12v2. Secondary outcomes were adverse events and return to work.
Phillips 1985	RCT in the United States of America.	96 patients randomised, follow-up data obtained for 49 patients.  Mean age 41.6 years, range 15 to 78 years.  42 male, 54 female.	Surgery: ORIF and below-knee cast. Walking on crutches without weight-bearing a few days after surgery until tenth week. Cast removed after 7 days with early active motion exercises. Trans-syndesmotic screw (if	Follow-up: Mean follow-up time was 3.5 years. Range was 1.7 to 6 years.  A self-made non-validated questionnaire was used to score clinical (up to 100 points for pain, range of



			<p>present) removed after 10 weeks.</p> <p>Conservative: Above-knee cast for 6 weeks, without weight-bearing. Then below-knee cast for 4 weeks.</p>	<p>motion and function), anatomical (up to 35 points for talocrural angle medial clear space, integrity of tibiofibular syndesmosis, medial malleolar displacement, size of posterior malleolar fragment, lateral malleolar displacement and shortening, talar tilt, talar subluxation, aspect of anteromedial corner) and arthritis (up to 15 points for non-union, synostosis, osteoporosis, irregular articular surface, osteophytes, periarticular cysts and joint-space narrowing). They also assessed complications.</p>
Rowley 1986	Quasi-RCT in United Kingdom	<p>42 patients randomised, follow-up data obtained for 40 patients.</p> <p>Age range 16 to 70 years.</p> <p>Gender not provided.</p> <p>Patients with a displaced ankle fracture included.</p>	<p>Surgery: AO principles and below-knee cast with early weight-bearing encouraged. Plaster removed at 6 weeks and weight-bearing encouraged.</p> <p>Conservative: Above-knee plaster for 6 weeks and early weight-bearing. X-ray at 1-week and 2-weeks. Removal of cast at 6 weeks and weight-bearing encouraged.</p>	<p>Length of follow-up was 20 weeks.</p> <p>Time to weight-bearing, time for dorsiflexion and foot angle to return to normal (a footprint analysis was used) and complications.</p>
Sanders 2012 and Slobogean 2012	RCT in 6 Level 1 trauma centres in North America	<p>81 patients randomised, follow-up data obtained for 81 patients.</p> <p>Mean age of 41 years, range 18 to 65 years.</p> <p>41 male, 40 female.</p> <p>Patients under 65 years with an isolated lateral malleolar Weber B fracture with a positive stress examination included - defined as an increase in medial clear space of 1mm or greater and an absolute MCS value of 5 mm or greater.</p>	<p>Surgery: ORIF and splinting in neutral dorsiflexion for 2 weeks. Then PWB in a removable cast brace for 4 weeks.</p> <p>Conservative: Plaster or fibreglass cast or brace and PWB for 6 weeks. X-ray at 1 or 2 weeks.</p>	<p>Length of follow-up was 12 months.</p> <p>Primary outcomes were PCS score on the SF-36 and OMAS score. Secondary outcomes included radiological outcomes, complications and cost-effectiveness</p>

<p>Willett 2016 and Keene 2018</p>	<p>Pragmatic, equivalence RCT in 24 hospitals in the United Kingdom.</p>	<p>620 patients randomised, follow-up data obtained for 593 patients at 6 months and 450 patients at 3 years.</p> <p>Mean age of 71 years. Adults over 60 years with an unstable ankle fracture were included.</p> <p>160 male, 460 female.</p>	<p>Surgery: ORIF. Local practice and surgeon decided selection of implants, splinting, weight bearing, and clinical follow-up.</p> <p>Conservative: Close contact cast under general or spinal anaesthetic in operating room. Follow up with x-rays. Touch or NWB for 4 weeks then FWB by 6 to 8 weeks at surgeon's discretion and patient volition.</p>	<p>Length of follow-up was 3 years.</p> <p>Primary outcome was OMAS score. Secondary outcomes were SF-12 (version 1), EQ-5D, pain (using subscales of OMAS and EQ-5D), patient satisfaction (rated 1-5, where 5 is "very satisfied"), health resource use, cost-effectiveness, time to weight-bearing, range of motion using goniometer (plantar and dorsiflexion), mobility with the Timed Up and Go test, radiological outcomes (non-union and malunion) and complications.</p>
------------------------------------	--	---	--	---

**Table 3. Assessment of risk of bias using the Cochrane Collaboration's Risk of Bias tool<sup>14</sup>**

	Sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessors	Incomplete outcome data	Selective outcome reporting	Other sources of bias
	Low	Unclear	High	Unclear	Unclear	High	Unclear
Bauer 1985	Random allocation performed with a list of random numbers	Not commented upon.	Not possible due to the nature of interventions	Not commented upon.	Results of 8 patients with Type A fractures not included, whereas baseline data was shown. It is unclear why.	Excluded Weber A fracture outcome reporting.	Insufficient information to assess.
	Low	Unclear	High	High	High	Unclear	Unclear
Makwana 2001	Randomisation using computer-generated numbers.	Not commented upon.	Not possible due to the nature of interventions	All patients were assessed by the same observer (the first-named author of the trial).	8 patients in the closed group with early treatment failure were excluded from the group and analysed separately with no intention-to-treat analysis.	Insufficient detail on those lost to follow-up.	Unclear details of the characteristics of the 2 final groups that were analysed.
	Low	Low	High	Low	Low	Low	Unclear
Mittal 2017	The National Health and Medical Research Council Clinical Trials Centre (not otherwise involved in the study) generated the randomisation schedule using a permuted block approach with variable block size and stratified by site.	Automated telephone system provided allocation concealment	Not possible due to the nature of interventions	Outcome assessors were independent of the treating teams, and collected data using a standardised telephone interview.	Intention-to-treat analysis was performed. Attrition and exclusions were reported, as were the numbers in each intervention group (compared with total randomised participants) and reasons for attrition/exclusions.	Good correlation with protocol.	Lack of reporting with regards to variability between sites.
	Low	Low	High	Unclear	High	Unclear	Unclear
Phillips 1985	Sealed envelope chosen from a "sequential pool maintained by the statisticians, based on table of randomized numbers."	Attending surgeon provided with a sealed envelope.	Not possible due to the nature of interventions	Not commented upon.	Only 51% of patients were followed-up. In 3, baseline characteristics were missing due to lost charts.	Protocol unavailable. Data reported for all scores set out in abstract and all 49 patients of Group A where outcome data was collected. Self-made score is a potential source of bias.	Insufficient information to assess.
	High	High	High	Unclear	Low	Unclear	Unclear

Rowley 1986	Quasi-randomised using last digit of Accident & Emergency department number.	Not commented upon and unlikely due to nature of randomisation	Not possible due to the nature of interventions	Not commented upon.	Loss of follow-up was mentioned as 2 patients.	No protocol available to compare with.	Insufficient information to assess.
Sanders 2012 and Slobogean 2012	Low	Low	High	Primary outcomes - low Secondary outcomes - high	Unclear	Low	Unclear
	Sealed envelope technique. Randomisation was stratified according to the site using a block size of 5.	Sealed envelope technique.	Not possible due to the nature of interventions	Data analysis for the primary outcome was conducted independently by blinded assessors. Blinding of assessors for primary outcomes was specified but not for secondary outcomes, which was unlikely.	Insufficient detail provided regarding outcome data.	Good correlation with protocol.	Insufficient detail. Baseline characteristics not shown.
Willett 2016 and Keene 2018	Low	Low	High	6-week and 3-year follow up – high  6-month follow-up - low	Low	Low	Unclear
	Randomised in a 1:1 allocation using a 24-hour telephone service at an independent organisation (Aberdeen University).	Participants were registered before computer generation of the allocation.	Not possible due to the nature of interventions	Outcome assessments at the primary end point were performed by a blinded health professional. "The assessments at 6 weeks were not blinded because the assessor needed knowledge of postoperative instructions for weight bearing and movement." Blinding is not commented upon at 3-year follow-up and is unlikely.	The study had a high follow-up rate and well-explained reasons for loss of follow-up.	All outcomes were well-reported.	Unclear if the overall results were reflective of all the hospitals or if there were disparities between centres. It is unclear if the decision to keep the conservative patients with early conversion to surgery in the per-protocol analysis affected interpretation of results.

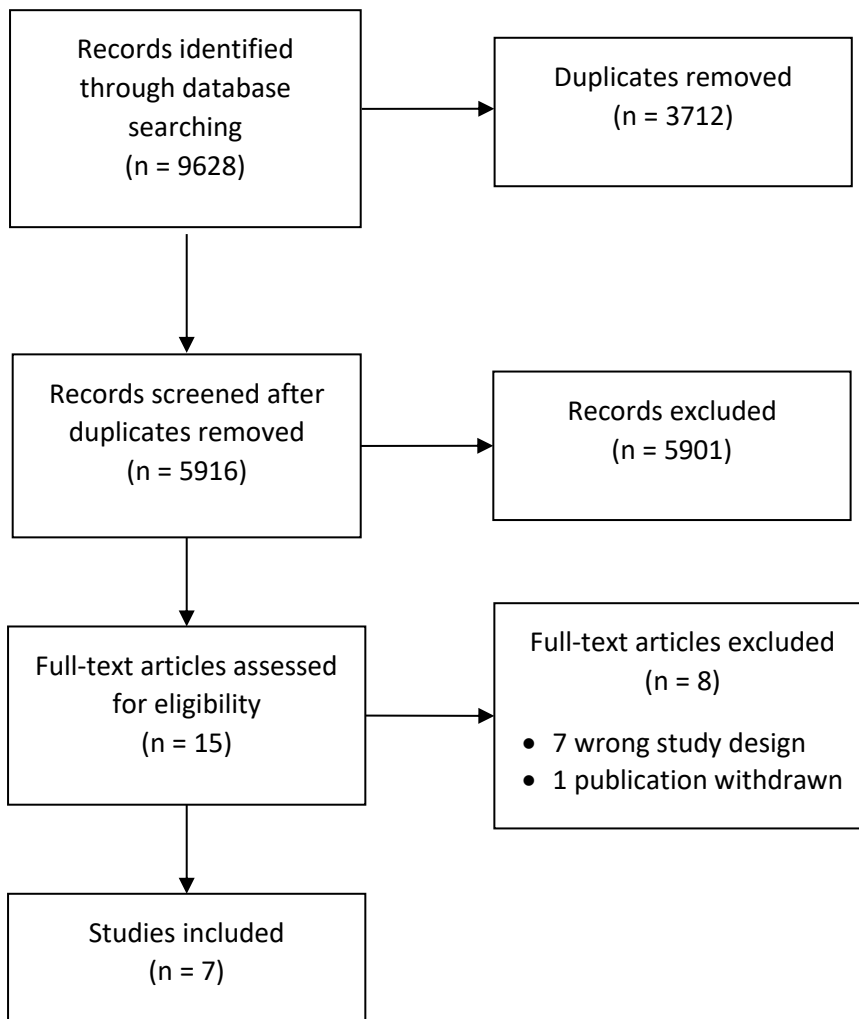
<b>Table 4. Other secondary outcomes</b>			
Study	Outcome	Result	Favours
Bauer 1985	Pain	RR 1.06 (95% CI 0.56 to 2.00)	No significant difference
	Degrees of dorsiflexion	MD -0.70 (95% CI -3.49 to 2.09)	No significant difference
	VTE	RR 0.75 (95% CI 0.13 to 4.28)	No significant difference
	Poor reduction	RR 0.36 (95% CI 0.21 to 0.64)	Surgery
	Moderate/severe arthritis	RR 0.81 (95% CI 0.28 to 2.38)	No significant difference
	Inpatient stay	MD -4.5, reported significant as p<0.05	Conservative
	Sick leave	Median time of 14 weeks for the surgical and conservative groups	No significant difference
Makwana 2001	Pain	MD -0.30 (95% CI -17.50 to 16.90)	No significant difference
	Range of motion	MD 7 degrees, reported as p = 0.044	Surgery
	Poor reduction	RR 0.32 (95% CI 0.10 to 1.02)	No significant difference
	Patient satisfaction	RR 0.97 (95% CI 0.54 to 1.76)	No significant difference
Mittal 2017	VTE	RR 1.63 (95% CI 0.41 to 6.63)	No significant difference
	Return to work at 3 months	RR 0.92 (95% CI 0.82 to 1.04)	No significant difference
Phillips 1985	Poor reduction	MD 4.6, reported significant as p<0.05	Surgery
	Moderate/severe arthritis	RR 0.09 (95% CI 0.01 to 1.46)	No significant difference
Rowley 1986	Restricted dorsiflexion	RR 1.33 (95% CI 0.34 to 5.21)	No significant difference
Sanders 2012 and Slobogean 2012	Cost-effectiveness	Base case 1-year incremental cost-effectiveness ratio (ICER) of the surgical group was \$205,090 per quality-adjusted life years (QALY) gained.  The lifetime horizon found surgical treatment to be preferred with an ICER of \$16,404 per QALY gained	Conservative at 1-year  Surgery at lifetime horizon
Willett 2016 and Keene 2018	Pain	MD -0.01 (95% CI -0.10 to 0.08) at 6 months  MD -0.05 (95% CI -0.2 to 0.1) at 3 years	No significant difference
	Range of dorsiflexion	MD -0.3 (95% CI -1.9 to 1.3)	No significant difference
	VTE	RR 0.34 (95% CI 0.11 to 1.03)	No significant difference
	Inpatient stay	MD -0.26 (95% CI -3.38 to 2.86)	No significant difference
	Readmissions	RR 0.38 (95% CI 0.24 to 0.60)	Surgery
	QALY	MD -0.01 (95% CI -0.04 to 0.02)	No significant difference
	National Health Service (NHS) costs at 6 months	MD -692 (95% CI -1438 to 24)	No significant difference
	Patient satisfaction	Adjusted MD 0.03 (95% CI -0.1 to 0.2)	No significant difference
Return to partial weightbearing	RR 0.89 (95% CI 0.72 to 1.10)	No significant difference	

<b>Table 5. Types of fractures</b>	
<u>Study</u>	<u>Fracture type</u>
Bauer 1985	<p>Initial data:</p> <ul style="list-style-type: none"> <li>Weber A – 8/108 patients</li> <li>4 surgical</li> <li>4 conservative</li> <li>Weber B – 100/108 patients</li> <li>47 surgical</li> <li>53 conservative</li> <li>Unimalleolar – 62/108 patients</li> <li>24 surgical</li> <li>38 conservative</li> <li>Bimalleolar – 16/108 patients</li> <li>8 surgical</li> <li>8 conservative</li> <li>Trimalleolar – 30/108 patients</li> <li>19 surgical</li> <li>11 conservative</li> </ul> <p>Follow-up data:</p> <ul style="list-style-type: none"> <li>Weber B – 92/92 patients</li> <li>43 surgical</li> <li>49 conservative</li> <li>Unimalleolar – 51/92 patients</li> <li>20 surgical</li> <li>31 conservative</li> <li>Bimalleolar – 14/92 patients</li> <li>7 surgical</li> <li>7 conservative</li> <li>Trimalleolar – 27/92 patients</li> <li>16 surgical</li> <li>11 conservative</li> </ul>
Makwana 2001	Unknown
Mittal 2017	<p>Initial data:</p> <ul style="list-style-type: none"> <li>AO/OTA-type 44-B1 (undisplaced) – 160/160 patients</li> <li>80 surgical</li> <li>80 conservative</li> </ul> <p>Follow-up data:</p> <ul style="list-style-type: none"> <li>AO/OTA-type 44-B1 – 141/141 patients</li> <li>72 surgical</li> <li>69 conservative</li> </ul>
Phillips 1985	<p>Initial data:</p> <ul style="list-style-type: none"> <li>Supination-external rotation type 4 – 84/96 patients</li> <li>Pronation-external rotation type 4 – 12/96 patients</li> </ul> <p>Follow-up data:</p> <ul style="list-style-type: none"> <li>Supination-external rotation type 4 – 44/49 patients</li> <li>22 surgical</li> <li>22 conservative</li> <li>Pronation-external rotation type 4 – 5/49 patients</li> <li>1 surgical</li> <li>4 conservative</li> </ul>
Rowley 1986	<ul style="list-style-type: none"> <li>Weber A – 0/40 patients</li> <li>Weber B – 32/40 patients</li> <li>15 surgical</li> <li>17 conservative</li> <li>Weber C – 8/40 patients</li> <li>5 surgical</li> <li>3 conservative</li> </ul>

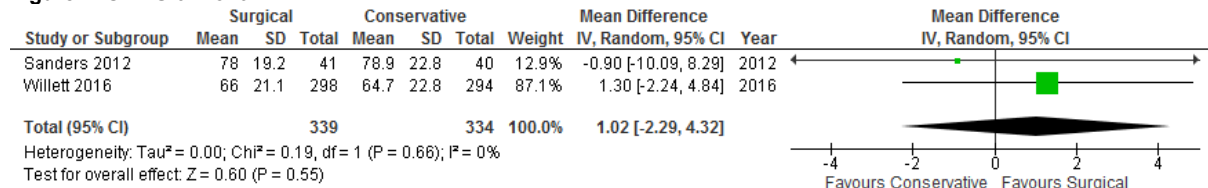
Sanders 2012	Weber B lateral malleolus (undisplaced but unstable on external rotation stress test) – 81/81 patients 41 surgical 40 conservative
Willett 2016 and Keene 2018	<p>Initial data:</p> <p>Weber A - 35/620 patients 21 surgical 14 conservative</p> <p>Weber B – 507/620 patients 251 surgical 256 conservative</p> <p>Weber C – 78/620 patients 37 surgical 41 conservative</p> <p>3-year follow-up data:</p> <p>Weber A / Weber B – 389/450 patients 197 surgical 192 conservative</p> <p>Weber C – 61/450 patients 29 surgical 32 conservative</p>



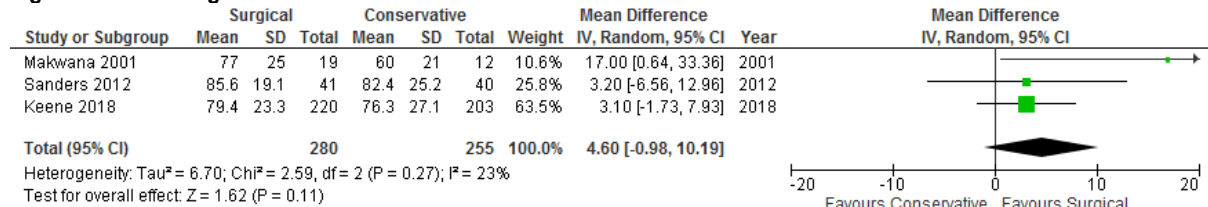
Figure 1. PRISMA Flow Diagram



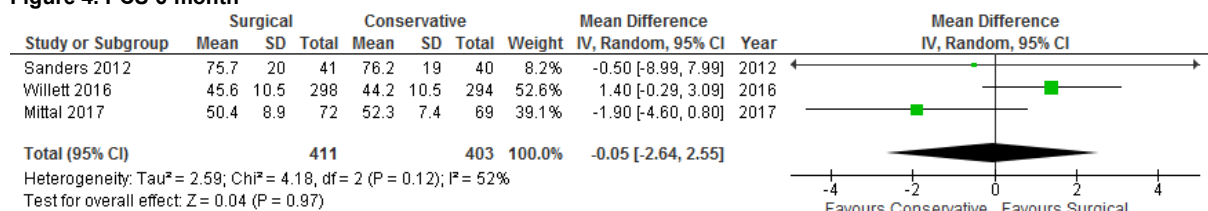
**Figure 2. OMAS 6-month**



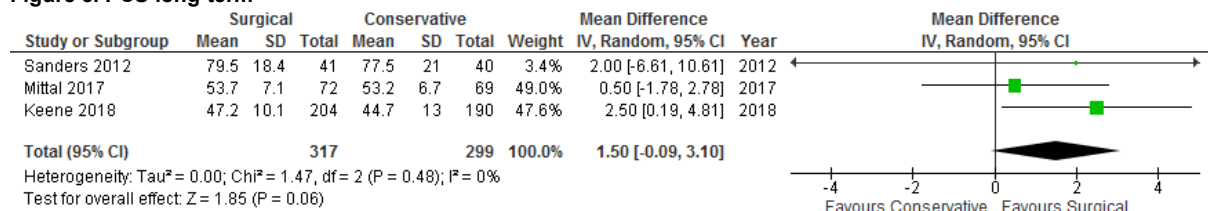
**Figure 3. OMAS long-term**



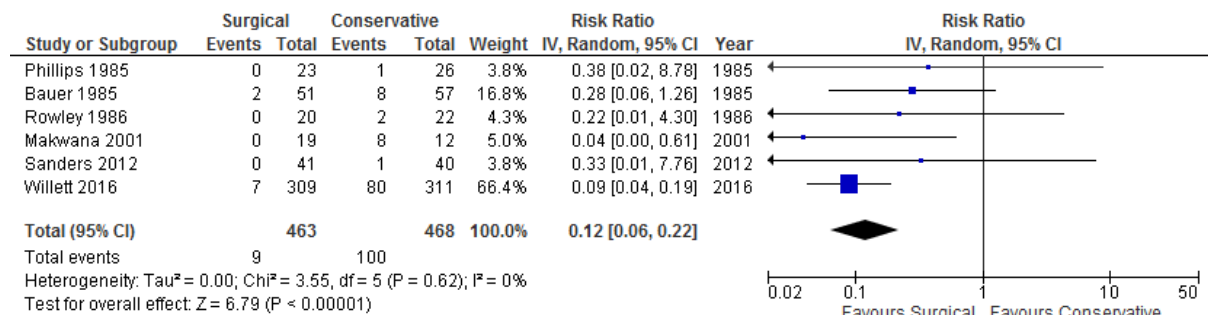
**Figure 4. PCS 6-month**



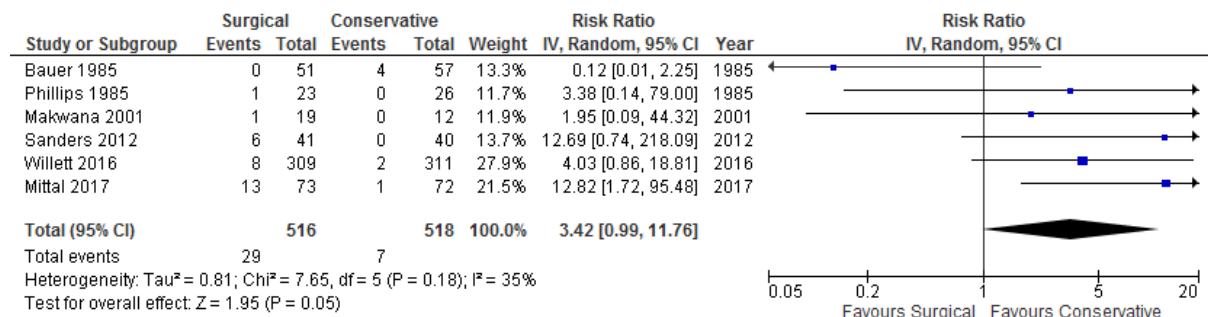
**Figure 5. PCS long-term**



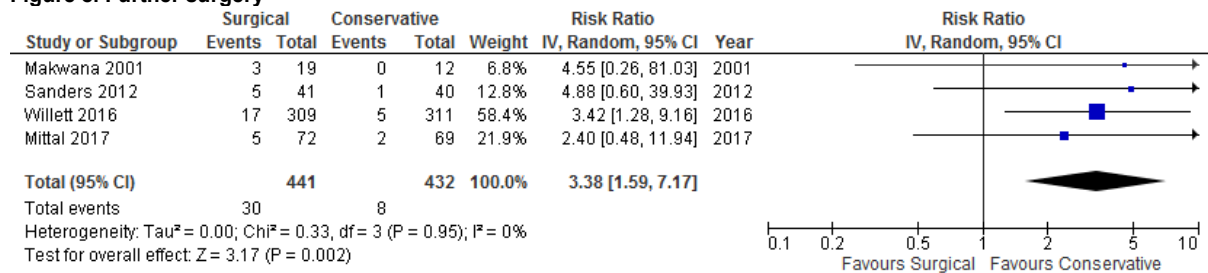
**Figure 6. Early treatment failure**



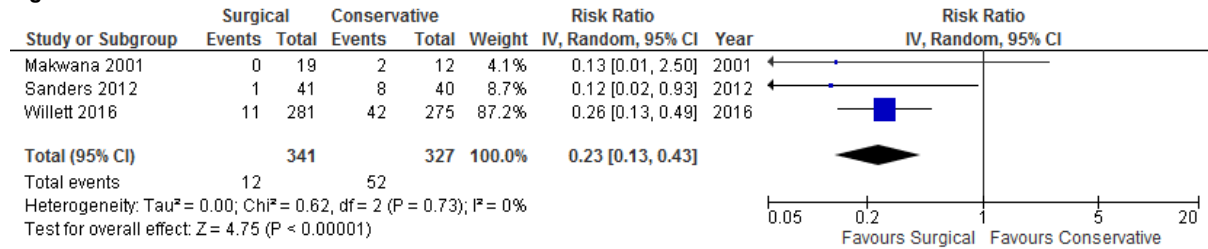
**Figure 7. Infection**



**Figure 8. Further surgery**



**Figure 9. Malunion**



**Figure 10. Delayed or non-union**

