

A systematic review of outcome reporting in clinical trials of distal tibia and ankle fractures

THE NEED FOR A CORE OUTCOME SET

**N. A. Pearson,
E. Tutton,
A. Joeris,
S. E. Gwilym,
R. Grant,
D. J. Keene,
K. L. Haywood**

From University of
Warwick, Coventry, UK

Aims

To describe outcome reporting variation and trends in non-pharmacological randomized clinical trials (RCTs) of distal tibia and/or ankle fractures.

Methods

Five electronic databases and three clinical trial registries were searched (January 2000 to February 2022). Trials including patients with distal tibia and/or ankle fractures without concomitant injuries were included. One reviewer conducted all searches, screened titles and abstracts, assessed eligibility, and completed data extraction; a random 10% subset were independently assessed and extracted by a second reviewer at each stage. All extracted outcomes were mapped to a modified version of the International Classification of Functioning, Disability and Health framework. The quality of outcome reporting (reproducibility) was assessed.

Results

Overall, 105 trials (n = 16 to 669 participants) from 27 countries were included. Trials compared surgical interventions (n = 62), post-surgical management options (n = 17), rehabilitative interventions (n = 14), surgical versus non-surgical interventions (n = 6), and pre-surgical management strategies (n = 5). In total, 888 outcome assessments were reported across seven domains: 263 assessed body structure or function (85.7% of trials), 136 activities (68.6% of trials), 34 participation (23.8% of trials), 159 health-related quality of life (61.9% of trials), 247 processes of care (80% of trials), 21 patient experiences (15.2% of trials), and 28 economic impact (8.6% of trials). From these, 337 discrete outcomes were described. Outcome reporting was inconsistent across trials. The quality of reporting varied widely (reproducibility ranged 4.8% patient experience to 100% complications).

Conclusion

Substantial heterogeneity in outcome selection, assessment methods, and reporting quality were described. Despite the large number of outcomes, few are reported across multiple trials. Most outcomes are clinically focused, with little attention to the long-term consequences important to patients. Poor reporting quality reduces confidence in data quality, inhibiting data synthesis by which to inform care decisions. Outcome reporting guidance and standardization, which captures the outcomes that matter to multiple stakeholders, are urgently required.

Cite this article: *Bone Jt Open* 2022;3-10:832–840.

Keywords: Systematic review, Outcome reporting, Distal tibia, Ankle fracture, Randomized control trial

Correspondence should be sent to
Kirstie L Haywood; email:
k.l.haywood@warwick.ac.uk

doi: 10.1302/2633-1462.310.BJO-
2022-0080.R1

Bone Jt Open 2022;3-10:832–840.

Introduction

Ankle fractures, including distal tibial fractures, account for approximately 14% of fractures requiring hospitalization.¹ Irrespective of age, they are associated with

significant morbidity,¹⁻³ pain, and impaired function, with a substantial reduction in ability to perform activities of daily life widely described. Evidence suggests that recovery from ankle fracture can be slow,^{4,5} with

protracted physical and psychological implications,^{2,6} and a variable rate at which patients return to their pre-injury lives.^{4,5} Understanding and accurately assessing outcomes following ankle fracture is therefore important to informing high-quality research.

The results of clinical trials and subsequent evidence reviews are essential for guideline development for the management and care of patients with ankle fractures.⁷⁻⁹ However, where trials include a wide range of different outcomes (e.g. clinical, clinician-reported, patient-reported, economic impact, or resource use), such heterogeneity in outcome assessment can limit evidence synthesis or meta-analysis, detract from the developing evidence base to inform clinical practice, and contribute to research waste.¹⁰

Guidance for outcome reporting in ankle fracture trials do not currently exist, outcome reporting standards have not been defined, and the magnitude of inconsistencies in outcome reporting remains unknown. This review aims to describe the degree of variation and trends in outcome reporting in non-pharmacological randomized clinical trials (RCTs) of ankle fractures, including distal tibial fractures.

Methods

The review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.¹¹

Search strategy. Five online databases were searched (January 2000 to February 2022): Medline (OVID), Embase (OVID), PsycINFO (OVID), CINAHL, and AMED. Three trial registries were screened: ISRCTN, ICTRP, and ClinicalTrials.gov. A comprehensive search strategy, translated for each database, was defined (Supplementary Table i). Hits were imported into EndNote (Clarivate Analytics, USA) and duplicates removed.

Eligibility criteria. Inclusion criteria were: randomized clinical trials examining non-pharmacological interventions for adults with an ankle fracture (AO/OTA type 43 and 44 fractures,¹² and available in English as full-text articles published in peer-reviewed journals. Exclusion criteria were: multiple lower limb fractures that include the ankle (e.g. foot, tibia), or paediatric patients; screening, diagnostic methods, or involving animals or cadavers; conference proceedings, abstracts, or editorials.

Trial selection. Two authors (NAP, KLH) independently screened titles and abstracts for eligibility, and reviewed retained full-text articles: all were screened by NAP; KLH screened a 10% subset. Agreement was checked; a third author (ET) helped to resolve any disagreements. Reference lists of included trials and meta-analyses were reviewed.¹³⁻¹⁵

Data extraction. A pre-defined data extraction form (Excel; Microsoft, USA) was developed.¹⁶ Trial-specific information included: author, publication year,

country, type of fracture, patient population, trial intervention(s), reported outcomes, and assessment method. All outcomes were extracted verbatim based on assessment focus, method, timing, and reproducibility (i.e. adequate citation or detail supporting reproduction). Outcomes were mapped to pre-defined domains informed by the core International Classification of Functioning, Disability and Health (ICF) framework:¹⁷ body structure and function, activities, and participation. Additional (sub)domains were iteratively added to reflect commonly reported outcomes in included trials.¹⁶

Data analysis. Descriptive statistics were used to describe pattern of outcome reporting including numbers of trials and frequency of outcome reporting, assessment focus (what was assessed), and methods (how were outcomes assessed).¹⁶ Assessment timings and the quality of outcome reporting were also analyzed. Reporting quality considered the reproducibility of outcome assessments: i.e. was sufficient detail, or appropriate citation, provided to allow assessment reproduction.¹⁶ The proportion of outcomes judged to be reproducible were categorized as: poor (< 30%); low (31% to 49%); moderate (50% to 69%); or high (> 70%).

Patient and public involvement. Three experienced patient research partners contributed to all stages of the review process.

Results

From 7,623 results, 148 full-text articles were reviewed and 48 excluded (Figure 1). Five additional articles were identified after reference screening. In total, 105 trials were included (Supplementary Table ii).

Trial characteristics. Included trials compared: surgical interventions (n = 62), post-surgical injury management options (n = 17), rehabilitation approaches (n = 14), surgical versus non-surgical interventions (n = 6), and pre-surgery management strategies (n = 5) (Table I). Almost half (n = 47; 45%) were conducted in the past five years. Most were small trials, including < 100 participants (n = 79/105; ranging from 16 to 669; median 64 (interquartile range (IQR) 58 to 605)). Trials were from 27 countries; most were from the UK (n = 12), China (n = 11), Finland, the USA, and South Korea (n = 8).

Which outcomes are assessed in ankle fracture trials? From a total of 888 extracted outcome assessments (i.e. the total number of outcome assessments reported across all trials), 337 clearly defined, discrete outcomes were identified e.g. the Olerud-Molander Ankle Score (OMAS),¹⁸ pain severity with a clearly described visual analogue scale (VAS), or the American Orthopaedic Foot Ankle Score (AOFAS)¹⁹ (Table II). While a third (32.3%; n = 123) of these discrete outcomes were each assessed in two or more trials, most (67.8%; n = 258) were assessed just once in single trials. The number of outcomes reported

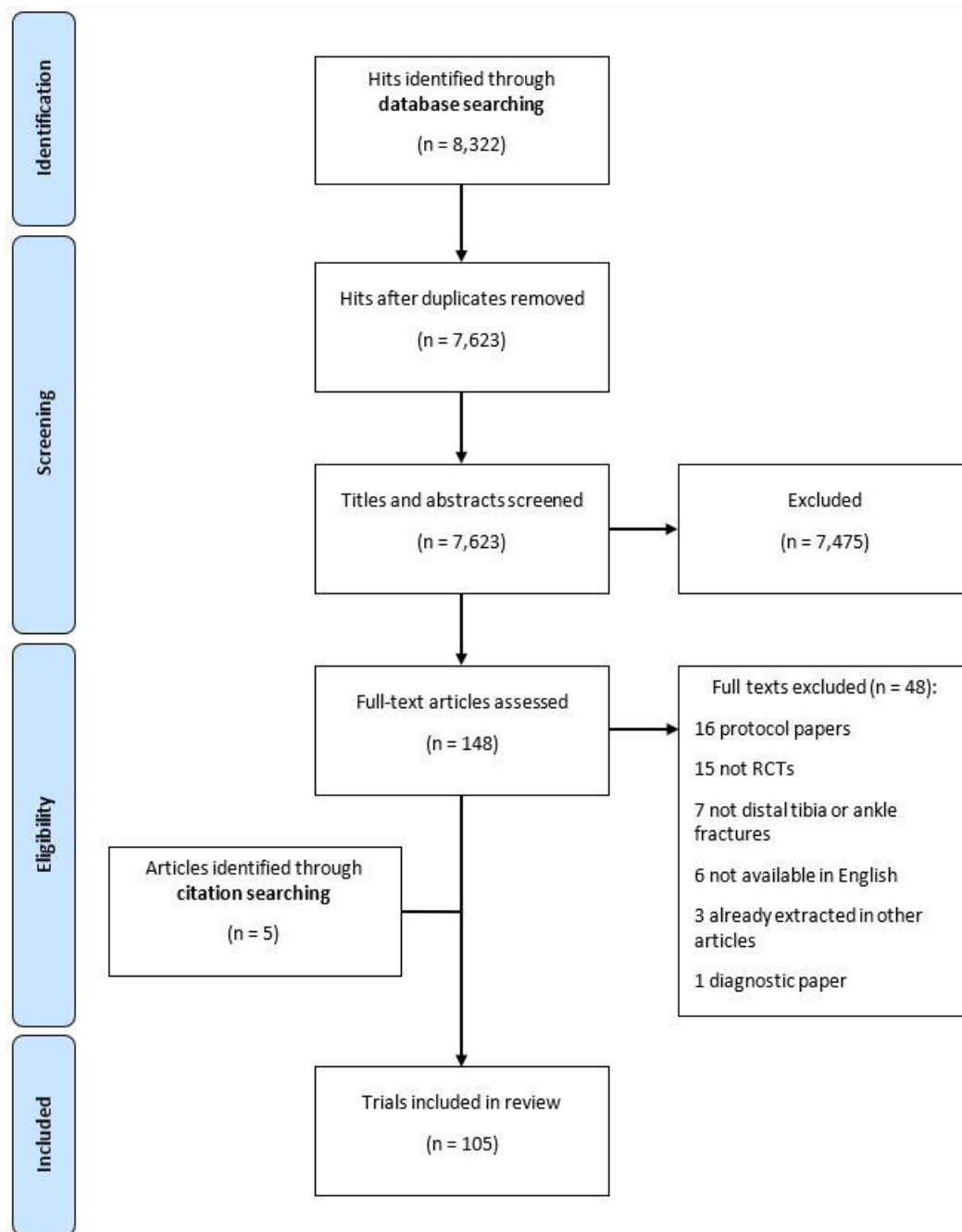


Fig. 1

Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of trial selection process.

per trial ranged from 1 to 25 (median 8 (IQR 6 to 17)) (Table III).

All outcome assessments were assigned to one of seven domains, with 24 sub-domains further defined (Table II). When assessed by frequency of outcome reporting, most assessments focused on body structure or function ($n = 263/880$ outcomes, 29.6%), processes of care ($n = 247/880$, 27.8%), health-related quality of life (HRQoL) ($n = 159/880$, 17.9%), and activities ($n = 136/880$, 15.3%). Participation ($n = 34/880$, 3.8%), economic assessments

($n = 28/880$, 3.2%), and patient experiences ($n = 21/880$, 2.4%) were less frequently assessed.

Body structure and function. Most trials reported at least one assessment of body structure or function ($n = 90/105$; 85.7%): 99 discrete outcomes contributed to 263 outcome assessments across the five sub-domains. Radiological assessments ($n = 104$ outcome assessments; $n = 32$ discrete outcomes; 55.2% of trials) and joint range of motion ($n = 88$ outcome assessments; $n = 32$ discrete outcomes; 42.9% of trials) were the most

frequently assessed outcomes. Assessment reproducibility ranged from 42% for range of motion to 100% for blood chemistry.

Processes of care. Processes of care were assessed in most trials ($n = 84/105$; 80%): 95 discrete outcomes contributed to 247 outcome assessments across three sub-domains. Complications were most frequently reported ($n = 77/105$ trials (73.3%)), with 74 discrete outcomes contributing to 186 outcome assessments. Reproducibility was high (82.8%). Hospital-specific outcomes (e.g. length of hospital stay, and re-admissions) were assessed in 21% of trials ($n = 22/105$). The 25 outcome assessments (of which seven were discrete) were highly reproducible (100%). Surgery-specific outcomes were assessed in 23.8% ($n = 25/105$) trials. The 36 outcome assessments (of which 14 were discrete) were highly reproducible (100%), and included the duration of surgery or assessments of blood loss.

Health-related quality of life. More than 60% of trials included at least one health-related quality of life (HRQoL) assessment ($n = 65/105$; 61.9%): 61 discrete outcomes contributed to 159 outcome assessments across five sub-domains. A total of 11 discrete, multi-domain assessments, e.g. the 36-Item Short-Form Health Survey (SF-36)²⁰ and EuroQol five-dimension questionnaire (EQ-5D),²¹ contributed to 29 outcome assessments across 29 trials (27.6%). Seven discrete assessments of general health, often with single-item VAS such as the EQ-VAS, contributed to 13 outcome assessments across 13 trials (12.4%). The most widely reported discrete outcome assessments of HRQoL were the EQ-5D, SF-36, and SF-12. Overall, reproducibility was high (72.3%).

The most widely assessed sub-component of HRQoL was pain, reported by almost 50% of trials (46.7%): 31 discrete 'pain' outcomes contributed to 80 outcome assessments. Overall, 18 of the discrete outcomes were single-item assessments of pain intensity, and 13 were sub-domains or components of existing measures (e.g. from the American Academy of Orthopaedic Surgeons Foot and Ankle Questionnaire (AAOS),²² American College of Foot and Ankle Surgeons (ACFAS),²³ EQ-5D, or Mazur Ankle Score²⁴ and McGuire Scoring Criteria).²⁵ Pain experienced under specific conditions was also assessed (e.g. during walking, at rest, at night). Reproducibility was moderate (66.3%).

Few trials specifically assessed the impact of ankle fracture on mental health (17.1%). From 24 outcome assessments, eight discrete outcomes were described, taken from existing HRQoL assessments, most commonly the SF-36 ($n = 9$) and SF-12 ($n = 4$) mental health scores. Reproducibility was high (75%).

Activities. Many trials reported at least one measure of activity ($n = 72/105$; 68.6%): 61 discrete outcomes contributed to 136 outcome assessments across the four sub-domains. Most frequently assessed was 'Ankle function'

Table 1. Characteristics of included trials ($n = 105$).

Characteristics	Trials, n
Number of participants	
1 to 50	37
51 to 100	42
101 to 150	12
> 150	14
Age, yrs	
Mean (range)	45.9 (15 to 98)
Median (IQR)	43.5 (7.22 to 45.50)
Not stated	5
Interventions	
Pre-surgical management	5
Surgical	62
Surgical vs non-surgical	6
Post-surgical management	18
Rehabilitation	14
Continent	
Africa	3
Asia	34
Australasia	7
Europe	47
North America	14

IQR, interquartile range.

($n = 59$ outcome assessments); the nine discrete outcomes included the Ankle Osteoarthritis Scale,²⁶ Foot Function Index,²⁷ and the OMAS (most widely used in 49 (46.7%) trials). Basic activities of daily living were also widely assessed with 25 discrete outcomes contributing to 36 outcome assessments. Except for 'basic activities of daily living' (reproducibility 61.1%), reproducibility was high (84.0% to 93.2%).

Participation. Fewer than 25% of trials assessed participation; 13 discrete outcomes contributed to 34 outcome assessments across the three sub-domains. Multi-domain assessments were largely reproducible (75%), but this reduced substantially when single-item VAS or numeric rating scales were used (5% and 0%, respectively). Content predominantly reflected returning to work, sport, or leisure activity.

Economic assessments. Across nine trials, 25 discrete outcomes contributed to 28 economic outcome assessments e.g. absenteeism, costs of treating complications, days away from paid work or unpaid activities, and direct healthcare costs. Broadly, outcomes were well described and reproducible (82.1%).

Patient experiences. Few trials assessed patient experience (15.2%): 16 discrete outcomes contributed to a total of 21 outcome assessments. Assessments were typically related to patient satisfaction with care or treatment ($n = 12$), outcomes e.g. scar appearance or comfort ($n = 6$), or recovery ($n = 3$). Reproducibility was poor (range 0% to 16.7%). Simple VASs were widely used, with little detail supporting reproduction.

Table II. Trial outcome reporting across health domains. Expanded definitions can be found in Supplementary Table iii.

Domain	Outcome reporting across trials			
	Outcome assessments, n	Discrete outcomes, n	Trials, n (%)	Reproducibility of outcomes, %
Body structure and function	263	99	90 (85.7)	60.1
Body structure				
Radiological fracture assessments	104	32	58 (56.2)	66.0
Body function				
Ankle swelling and alignment assessments	53	20	40 (38.0)	71.7
Blood chemistry	11	11	2 (1.9)	100
Calf muscle assessments	7	4	7 (46.7)	85.7
Joint range of motion	88	32	45 (42.8)	42.0
Activities	136	56	72 (68.6)	82.4
Ankle function	59	9	56 (53.3)	93.2
Basic activities of daily living	36	25	21 (20.0)	61.1
Instrumental activities of daily living	25	14	20 (19.0)	84.0
Physical activities	16	8	11 (10.5)	87.5
Participation	34	13	25 (23.8)	20.6
Participation in social and leisure activities	8	4	8 (7.6)	75.0
Return to work	20	6	20 (19.0)	5.0
Return to sport or leisure activity	6	3	6 (5.7)	0
Health-related quality of life	159	61	65 (61.9)	72.3
General health	13	7	13 (12.4)	53.8
Mental health	24	8	18 (17.1)	75.0
Multi-domain assessment	29	11	29 (27.6)	89.7
Pain	80	31	49 (46.7)	66.3
Physical health	13	4	12 (11.4)	84.6
Processes of care	247	95	84 (80.0)	87.0
Complications	186	74	77 (73.3)	82.8
Hospital-specific outcomes	25	7	22 (20.9)	100
Surgery-specific outcomes	36	14	25 (20.0)	100.0
Patient experiences	21	13	16 (6.7)	4.8
Satisfaction with care or treatment	12	6	11 (10.5)	0
Satisfaction with outcomes	6	4	5	16.7
Perceptions of recovery	3	3	2	0
Economic assessment	28	25	9	82.1
Total	888	337	105	58.5

*Total number (frequency) of outcome assessments reported across trials.

†Clearly defined and reproducible discrete outcome.

Assessment timing. The 888 reported outcomes were captured across 1,955 timepoints. Most were collected between months two and three (38.2%; n = 744) post-injury. Around 20% were collected between months four to six (n = 387), and similarly between months seven and 12. Only 132 assessments (6.8%) were taken after one year (Figure 2). The number of outcome assessments by domain and timepoint is presented in Figure 3.

Quality of outcome reporting. Across all domains, assessment reproducibility was just 58.5%. Assessments of blood chemistry (100%), ankle function (93.2%), and processes of care outcomes (87%) were among the most clearly reported and hence reproducible assessments. Assessments of patient experience (4.8%), participation (20.6%), joint range of motion (42.0%), general health (53.8%), and pain (66.3%) were less clearly reported and hence less reproducible (Table II). Overall, just over 50%

(n = 53) of reviewed trials included two or more non-reproducible outcomes.

Discussion

This review describes the substantial heterogeneity and inconsistency in outcome reporting across published trials of distal tibial and ankle fracture with regard to which outcomes are assessed, and how and when assessments are undertaken. Such inconsistency of outcome reporting is described across the recovery journey, expanding on an earlier review by McKeown et al,⁶ which highlighted the wide range of primary outcome measures reported in ankle fracture trials.

The lack of outcome reporting guidance following ankle fracture is evident. Despite the large number of outcomes reported, few are reported across multiple trials. The short-term assessment of clinically focused

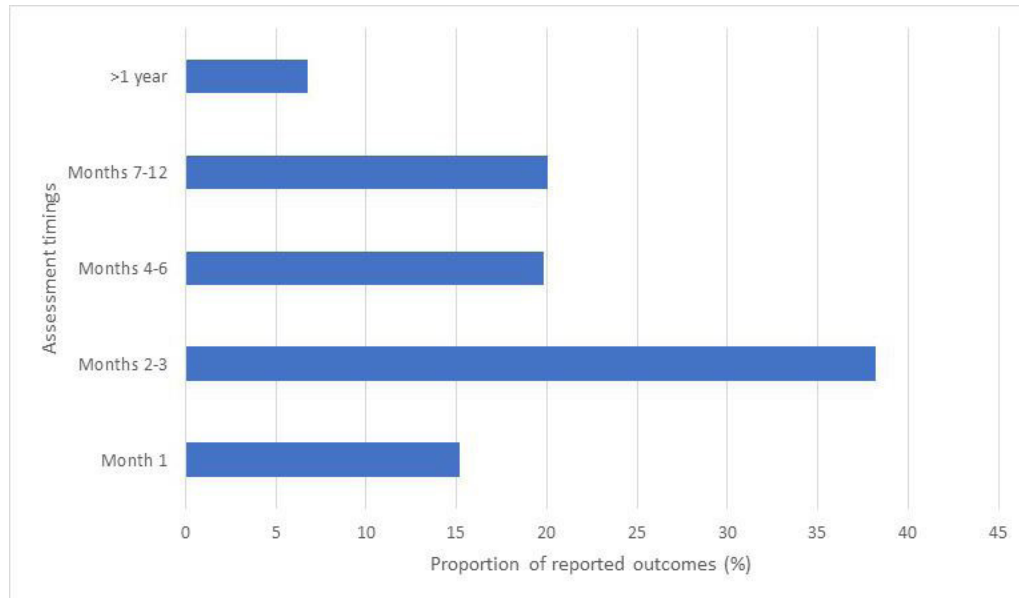


Fig. 2

Proportion of outcomes assessed at each timepoint.

domains of body structure and body function dominate outcome reporting, with little attention to the long-term consequences from the perspective of patients. While such assessments are important to understanding the clinical impact of ankle fracture and its management, they tell us little about what really matters to patients, how they understand their recovery journey, and their healthcare needs. Apart from pain, outcomes that may be important to patients, such as emotional impact and return to normal activities, are rarely assessed. Moreover, the quality of outcome reporting (assessment reproducibility) varied widely; when poor, this reduces confidence in data quality and inhibits data synthesis with which to inform care decisions. High-quality, standardized assessment that captures the outcomes that matter to key stakeholders is urgently required.

The review benefited from a transparent data extraction process which captured the many outcome reporting challenges.¹⁶ An extensive literature search of five major medical and three clinical trial databases was supplemented by citation searching of included trials. The inclusion of all eligible international publications aimed to reduce any publication bias that may have occurred by including only English-language publications. With a focus towards improving the standardization of outcome assessment in clinical trials, the review did not include alternative studies, such as observational or cohort studies. The review only included trials of adults with fractures; a review of outcome reporting in trials of childhood fractures has similarly highlighted outcome reporting heterogeneity.²⁸ Although of potential relevance to the quality of outcome reporting,

trial quality was not evaluated. Moreover, although important to informing outcome selection, this review did not extend to assessing the relative quality and acceptability of identified assessment methods.

Well-developed patient-reported outcome measures (PROMs) are single or multi-item questionnaires, intended to measure aspects of health that are important to patients.^{29,30} Providing important complementary evidence to traditional clinically based outcomes, their inclusion in clinical trials is strongly advocated.^{29,30} Moreover, where treatment goals include symptom reduction and enhancement of function and quality of life, PROMs have a central role in treatment optimization. However, this review describes the infrequent use of PROMs in trials of ankle fracture: while around 50% of trials included a patient-reported assessment of ankle function (most often the OMAS) and/or pain, fewer than 30% included multi-domain assessments of HRQoL, and just 12% assessed general health. Moreover, recent reviews highlight the inadequacies of several ankle-specific patient-reported outcomes including the OMAS, the AOFAS, and the Ankle Fracture Outcome of Rehabilitation Measure,³¹ suggesting insufficient evidence of essential measurement properties to recommend their use following ankle fracture.^{32,33} Major concerns related to inadequate content validity, measurement focus, evidence of measurement validity, responsiveness to change, and interpretative guidance. The widespread use of measures that lack fundamental measurement properties undermines the quality of trial data, reducing confidence in results.

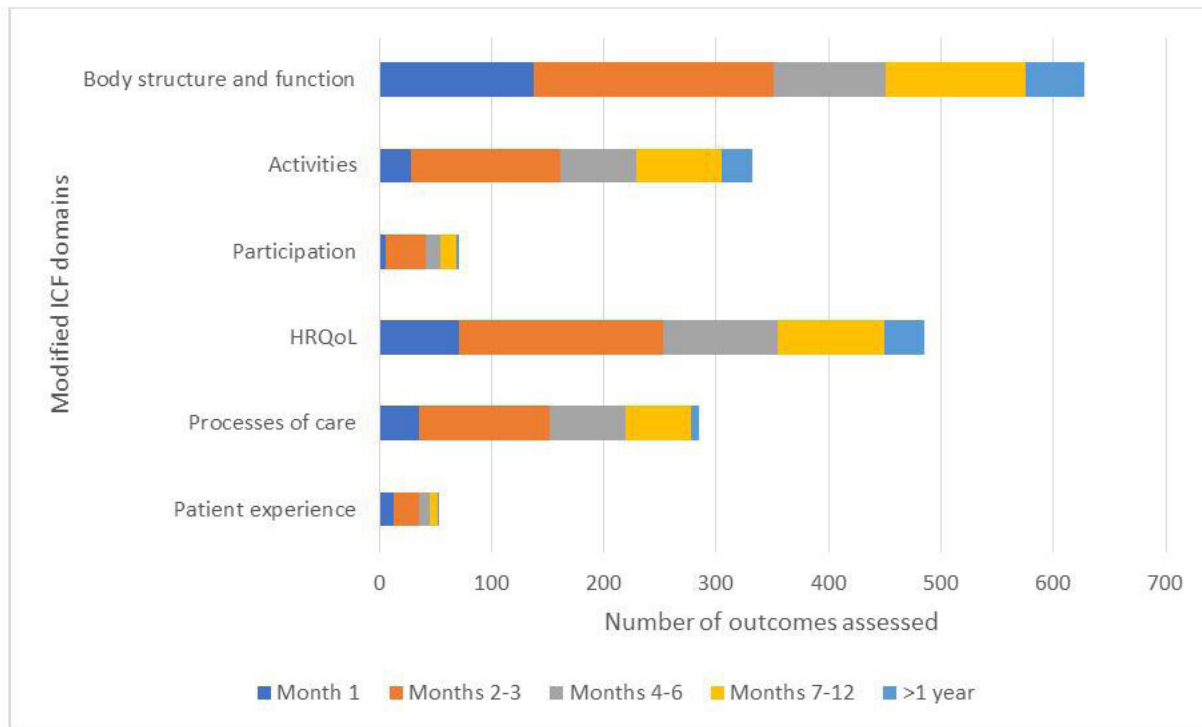


Fig. 3

Timing of outcome assessment by domain. HRQoL, health-related quality of life; ICF, International Classification of Functioning, Disability and Health framework.

Few trials assessed emotional wellbeing, participation or return to 'usual' activities, or patients' perceptions of recovery. Moreover, most assessments were short-term (two to three months post-fracture); few assessments extended beyond one year. The life and wellbeing of people who sustain an ankle fracture is substantially disrupted, bringing both short- and longer-term challenges.^{2,4,5} Awareness of the broad-ranging injury impact, and the time taken to return to 'normality', is important to ensuring appropriate, long-term assessment of meaningful outcomes.

This review describes an important lack of data clarity and integrity in outcome reporting, which undermined the quality of outcome reporting and assessment reproducibility: numerous potential outcome assessments were excluded from the review due to insufficient detail. This was particularly troublesome for the assessment of ankle range of motion where patient position and direction of ankle movement (e.g. inversion, eversion, dorsi- or plantar-flexion) was poorly conveyed. The utilization of trial data to inform evidence syntheses, decision-making, and guideline development is dependent upon good-quality data, and requires urgent attention in this field. The use of established trial outcome reporting guidelines such as the CONSORT-PRO extension³⁴ is recommended.

An iterative approach to mapping outcomes to domains, informed by the ICF framework,^{16,17} evidenced

Table III. Number of outcomes reported per trial.

Number of outcomes	Trials (n = 105)
1	3
2 to 5	29
6 to 10	47
11 to 20	24
21+	2

the diversity of outcome assessment in ankle fracture trials. Most commonly, trials included outcomes within the body structure and function (85.7%), processes of care (80%), and activities (68.5%) domains.

Complications, one aspect of the processes of care domain, were the most frequently reported outcomes. However, 74 discrete outcomes reported across 77 trials (73%) suggests that most trials included an assessment of different complications. Radiological outcomes were the second most frequently assessed (56% of trials). Substantial variation is evident, with 32 discrete outcomes reported across 58 trials. However, the limited clinical relevance of follow-up radiological assessment in ankle fracture management, in the absence of patient-reported symptoms, has been described.³⁵ Moreover, there is little evidence of the association between radiological and patient-reported outcomes.⁶ Of note, almost 45% of trials did not include

radiological assessment, perhaps reflecting the different focus of reviewed trials, e.g. rehabilitation.

Pain was assessed in almost 50% of trials. However, further outcome reporting heterogeneity is described, with 31 discrete outcomes described across the 49 trials. Moreover, assessment largely used poorly defined, single-item scales, such as VAS. While pain is clearly an important outcome following an acute injury, the relative importance afforded to pain by patients in comparison to other patient-important outcomes, particularly over the longer term, is not clear. Understanding how patients prioritize outcomes following an ankle fracture would ensure that future trials capture patient-important outcomes.

This review has described the substantial inconsistencies in outcome reporting, over both the short and longer term, the greater weight afforded to clinically based assessment, and the limited engagement with patients to better understand their perceptions of recovery and what a 'good outcome' following ankle fracture looks like. Moreover, the general quality and reproducibility of outcome reporting was often poor. Standardization of outcome reporting, in the form of a core outcome set, for clinical trials of ankle fracture is highly recommended.³⁶ A core outcome set seeks to define agreed and evidence-based guidance for a minimum number of outcomes which should be reported in all trials for a defined health condition.³⁷ Such guidance is currently available for several adult fractures, including proximal humerus,³⁸ distal radius,³⁹ hip,⁴⁰ and open lower limb fracture.⁴¹

Development of a Core Outcome Set for distal Tibia and Ankle fractures (COSTA) would support the standardization of outcome reporting in clinical trials and future evidence syntheses upon which clinical, policy, and healthcare decisions are based. Adopting a multi-stakeholder perspective including clinicians and patients will ensure that outcomes of relevance to key stakeholders are considered alongside methods of assessment that are both robust and relevant, supporting future adoption.



Take home message

- Substantial variation in outcome selection, assessment methods, and reporting quality inhibit efforts to synthesize research findings to inform care decisions.

- Poor outcome reporting quality reduces confidence in the quality of trial data and the potential contribution to evidence-based healthcare.

Twitter

Follow A. Joeris @AJJoeris

Follow S. E. Gwilym @stevegwilym

Follow D. J. Keene @davidkeenePT

Follow K. L. Haywood @haywood_kirstie

Follow Oxford Trauma and Emergency Care @Oxford_Trauma

Follow Warwick Research in Nursing @TheWRinN

Follow Warwick Medical School @warwickmed

Supplementary material



Tables showing example search strategy, list of included trials, and expanded definitions of terms used in Table II.

References

- Jennison T, Brinsden M.** Fracture admission trends in England over a ten-year period. *Ann R Coll Surg Engl.* 2019;101(3):208–214.
- McPhail SM, Dunstan J, Canning J, Haines TP.** Life impact of ankle fractures: qualitative analysis of patient and clinician experiences. *BMC Musculoskelet Disord.* 2012;13(1):224.
- Kyriacou H, Mostafa A, Davies BM, Khan WS.** Principles and guidelines in the management of ankle fractures in adults. *J Perioper Pract.* 2021;31(11):427–434.
- Trickett RW, Mudge E, Price P, Pallister I.** A qualitative approach to recovery after open tibial fracture: the road to A novel, patient-derived recovery scale. *Injury.* 2012;43(7):1071–1078.
- Rees S, Tutton E, Achten J, Bruce J, Costa ML.** Patient experience of long-term recovery after open fracture of the lower limb: a qualitative study using interviews in a community setting. *BMJ Open.* 2019;9(10):e031261.
- McKeown R, Rabiou A-R, Ellard DR, Kearney RS.** Primary outcome measures used in interventional trials for ankle fractures: a systematic review. *BMC Musculoskelet Disord.* 2019;20(1):388.
- No authors listed.** The management of ankle fractures. British Orthopaedic Association. 2016. <https://www.boa.ac.uk/resources/boast-12-pdf.html> (date last accessed 24 August 2022).
- No authors listed.** Fractures (non-complex): assessment and management. NICE guideline [NG38]. National Institute for Health and Care Excellence. 2016. <https://www.nice.org.uk/guidance/ng38> (date last accessed 24 August 2022).
- No authors listed.** Fractures (complex): assessment and management. NICE guideline [NG37]. National Institute for Health and Care Excellence. 2016. <https://www.nice.org.uk/guidance/ng37> (date last accessed 24 August 2022).
- Kirkham JJ, Altman DG, Chan A-W, Gamble C, Dwan KM, Williamson PR.** Outcome reporting bias in trials: a methodological approach for assessment and adjustment in systematic reviews. *BMJ.* 2018;362:k3802.
- Page MJ, McKenzie JE, Bossuyt PM, et al.** The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:71.
- Meinberg EG, Agel J, Roberts CS, Karam MD, Kellam JF.** Fracture and dislocation classification compendium-2018. *J Orthop Trauma.* 2018;32 Suppl 1:S1–S170.
- Wang B, Zhao Y, Wang Q, et al.** Minimally invasive percutaneous plate osteosynthesis versus intramedullary nail fixation for distal tibial fractures: a systematic review and meta-analysis. *J Orthop Surg Res.* 2019;14(1):MC6925456.
- Li Z-H, Yu A-X, Guo X-P, Qi B-W, Zhou M, Wang W-Y.** Absorbable implants versus metal implants for the treatment of ankle fractures: A meta-analysis. *Exp Ther Med.* 2013;5(5):1531–1537.
- Ekman E, Lehtimäki K, Syvänen J, Saltychev M.** Comparison between nailing and plating in the treatment of distal tibial fractures: a meta-analysis. *Scand J Surg.* 2021;110(2):115–122.
- Whitehead L, Perkins GD, Clarey A, Haywood KL.** A systematic review of the outcomes reported in cardiac arrest clinical trials: the need for a core outcome set. *Resuscitation.* 2015;88:150–157.
- No authors listed.** ICF Checklist. World Health Organization. <https://www.who.int/publications/m/item/icf-checklist> (date last accessed 24 August 2022).
- Olerud C, Molander H.** A scoring scale for symptom evaluation after ankle fracture. *Arch Orthop Trauma Surg (1978).* 1984;103(3):190–194.
- Kitaoka HB, Alexander IJ, Adelaar RS, Nunley JA, Myerson MS, Sanders M.** Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. *Foot Ankle Int.* 1994;15(7):349–353.
- Ware JE, Sherbourne CD.** The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30(6):473–483.
- Rabin R, CF.** EQ-5D: a measure of health status from the EuroQol Group. *Ann Med.* 2001;33(5):337–343.
- Johanson NA, Liang MH, Daltroy L, Rudicel SA, Richmond J.** American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. *J Bone Joint Surg Am.* 2004;86-A(5):902–909.
- Cook JJ, Cook EA, Rosenblum BI, Landsman AS, Roukis TS.** Validation of the American College of Foot and Ankle Surgeons Scoring Scales. *J Foot Ankle Surg.* 2011;50(4):420–429.

24. Mazur JM, Schwartz E, Simon SR. Ankle arthrodesis. Long-term follow-up with gait analysis. *J Bone Joint Surg Am.* 1979;61-A(7):964–975.
25. Bagley C, Say ML, Low Choy N, et al. The McGuire Score for Clinical Assessment of Clubfoot: Inter-Rater Reliability of Clinicians. Abstract from APA Physiotherapy Conference 2011; 2011, Brisbane, Queensland, Australia.
26. Domsic RT, Saltzman CL. Ankle osteoarthritis scale. *Foot Ankle Int.* 1998;19(7):466–471.
27. Budiman-Mak E, Conrad KJ, Roach KE. The Foot Function Index: a measure of foot pain and disability. *J Clin Epidemiol.* 1991;44(6):561–570.
28. Marson BA, Craxford S, Deshmukh SR, Grindlay D, Manning J, Ollivere BJ. Outcomes reported in trials of childhood fractures: a systematic review. *Bone Jt Open.* 2020;1(5):167–174.
29. No authors listed. Guidance for Industry - Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. U. S. Food & Drug Administration. <https://www.fda.gov/downloads/drugs/guidances/UCM193282.pdf> (date last accessed 24 August 2022).
30. Coles TM, Hernandez AF, Reeve BB, et al. Enabling patient-reported outcome measures in clinical trials, exemplified by cardiovascular trials. *Health Qual Life Outcomes.* 2021;19(1):164.
31. McPhail SM, Williams CM, Schuetz M, Baxter B, Tonks P, Haines TP. Development and validation of the ankle fracture outcome of rehabilitation measure (A-FORM). *J Orthop Sports Phys Ther.* 2014;44:1–2.
32. McKeown R, Ellard DR, Rabiu A-R, Karasouli E, Kearney RS. A systematic review of the measurement properties of patient reported outcome measures used for adults with an ankle fracture. *J Patient Rep Outcomes.* 2019;3(1):70.
33. Ng R, Broughton N, Williams C. Measuring recovery after ankle fractures: a systematic review of the psychometric properties of scoring systems. *J Foot Ankle Surg.* 2018;57(1):149–154.
34. Calvert M, Blazeby J, Altman DG, et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA.* 2013;309(8):814–822.
35. van Gerven P, Weil NL, Termaat MF, et al. Routine follow-up radiographs for ankle fractures seldom add value to clinical decision-making: a retrospective, observational study. *J Foot Ankle Surg.* 2018;57(5):957–960.
36. Pearson NA, Tutton E, Joeris A, et al. Co-producing a multi-stakeholder Core Outcome Set for distal Tibia and Ankle fractures (COSTA): a study protocol. *Trials.* 2021;22(1):443.
37. Williamson PR, Altman DG, Blazeby JM, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials.* 2012;13(1):132.
38. Ramiro S, Page MJ, Whittle SL, et al. The OMERACT core domain set for clinical trials of shoulder disorders. *J Rheumatol.* 2019;46(8):969–975.
39. Goldhahn J, Beaton D, Ladd A, et al. Recommendation for measuring clinical outcome in distal radius fractures: a core set of domains for standardized reporting in clinical practice and research. *Arch Orthop Trauma Surg.* 2014;134(2):197–205.
40. Haywood KL, Griffin XL, Achten J, Costa ML. Developing a core outcome set for hip fracture trials. *Bone Joint J.* 2014;96-B(8):1016–1023.
41. Aquilina A, Costa M, Griffin X, Tutton L, Fitzpatrick R, Khan U. Core Outcomes for research on Open Lower Limb Fractures (CO-OLLF): Development of a core outcome set and core outcome measurement set. <http://www.comet-initiative.org/studies/details/688> (date last accessed 24 August 2022).

Author information:

- N. A. Pearson, PhD, MSc, BSc (Hons), Research Fellow
- K. L. Haywood, DPhil, BSc (Hons), PGCertLaTHE, MCSP, SRP, Professor of Health Outcomes

Warwick Research in Nursing, Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK.

- E. Tutton, PhD, MSc, BSc (Hons), PGCEA, RN, Senior Research Fellow, Warwick Research in Nursing, Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK; Kadoorie Centre, Oxford Trauma and Emergency Care, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Services, University of Oxford, Oxford, UK; Oxford University Hospitals NHS Foundation Trust, Oxford, UK.
- A. Joeris, MD, MSc HEPM, Head of Clinical Research, AO Foundation, Davos, Switzerland.
- S. E. Gwilym, MBBS, BSc, DPhil (Oxon), FRCS (Tr&Orth), Associate Professor in Trauma and Orthopaedics
- D. J. Keene, BSc (Hons), MSc, PGCert Clinical Trials, DPhil, MCSP, MMACP, University Research Lecturer in Trauma Rehabilitation
- Kadoorie Centre, Oxford Trauma and Emergency Care, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Services, University of Oxford, Oxford, UK.
- R. Grant, BEd (Hons), Patient Research Partner, National Institute for Health Research, Applied Research Collaboration, West Midlands, Warwick Medical School, University of Warwick, Coventry, UK; Musculoskeletal Patient and Public Involvement Group, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Services, University of Oxford, Oxford, UK; Health Protection Research Unit in Gastrointestinal Infections, University of Liverpool, Liverpool, UK.

Author contributions:

- N. A. Pearson: Project administration, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing.
- E. Tutton: Methodology, Supervision, Formal analysis, Writing – review & editing.
- A. Joeris: Conceptualization, Funding acquisition, Formal analysis, Writing – review & editing.
- S. E. Gwilym: Conceptualization, Funding acquisition, Formal analysis, Writing – original draft, Writing – review & editing.
- R. Grant: Formal analysis, Writing – review & editing.
- D. J. Keene: Conceptualization, Formal analysis, Writing – review & editing.
- K. L. Haywood: Conceptualization, Methodology, Funding acquisition, Supervision, Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing.

Funding statement:

- The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: this research was supported by a research grant awarded to Warwick Medical School, University of Warwick by the AO Foundation (no grant number assigned). Professor Haywood was chief investigator for the study.

ICMJE COI statement:

- D. J. Keene was a co-investigator for the AIM trial and is lead investigator for the AFTER trial, both are NIHR-funded ankle fracture trials in the UK. A. Joeris is employed by the AO Foundation.

Acknowledgements:

- We would like to thank AO Trauma for funding this project and Sam Johnson for her expertise in supporting development of the literature search strategies. We also thank our COSTA collaborators Andrew Metcalf, Joel Humphrey, Marilyn Heng, Marinos Piripiris, Martin Underwood, Prakash Jayakumar, Rebecca Kearney, Rodrigo Pesantez, and Simon Lambert for their contribution to the study design. We would also like to thank the study patient and public involvement group for their contributions to the study.

Ethical review statement:

- No ethical approval was required for the purposes of completing this systematic review of published work.

Open access funding

- The open access fee for this study was funded by a research grant awarded to Warwick Medical School, University of Warwick by the AO Foundation.

© 2022 Author(s) et al. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives (CC BY-NC-ND 4.0) licence, which permits the copying and redistribution of the work only, and provided the original author and source are credited. See <https://creativecommons.org/licenses/by-nc-nd/4.0/>