BJO

■ **FOOT & ANKLE**

Radiological malunion after ankle fractures in older adults

DEFINITIONS AND NEW THRESHOLDS DERIVED FROM CLINICAL OUTCOME DATA FROM THE AIM TRIAL

R. Knight,
D. J. Keene,
S. J. Dutton,
R. Handley,
K. Willett

From University of Oxford, Oxford, UK

Aims

The rationale for exacting restoration of skeletal anatomy after unstable ankle fracture is to improve outcomes by reducing complications from malunion; however, current definitions of malunion lack confirmatory clinical evidence.

Methods

Radiological (absolute radiological measurements aided by computer software) and clinical (clinical interpretation of radiographs) definitions of malunion were compared within the Ankle Injury Management (AIM) trial cohort, including people aged ≥ 60 years with an unstable ankle fracture. Linear regressions were used to explore the relationship between radiological malunion (RM) at six months and changes in function at three years. Function was assessed with the Olerud-Molander Ankle Score (OMAS), with a minimal clinically important difference set as six points, as per the AIM trial. Piecewise linear models were used to investigate new radiological thresholds which better explain symptom impact on ankle function.

Results

Previously described measures of RM and surgeon opinion of clinically significant malunion (CSM) were shown to be related but with important differences. CSM was more strongly related to outcome (-13.9 points on the OMAS; 95% confidence interval (CI) -21.9 to -5.4) than RM (-5.5 points; 95% CI -9.8 to -1.2). Existing malunion thresholds for talar tilt and tibiofibular clear space were shown to be slightly conservative; new thresholds which better explain function were identified (talar tilt > 2.4°; tibiofibular clear space > 6 mm). Based on this new definition the presence of RM had an impact on function, which was statistically significant, but the clinical significance was uncertain (-9.1 points; 95% CI -13.8 to -4.4). In subsequent analysis, RM of a posterior malleolar fracture was shown to have a statistically significant impact on OMAS change scores, but the clinical significance was uncertain (-11.6 points; 95% CI -21.9 to -0.6).

Conclusion

These results provide clinical evidence which supports the previously accepted definitions. Further research to investigate more conservative clinical thresholds for malunion is indicated.

Introduction

The rationale for exacting restoration of skeletal anatomy after articular fractures is to improve outcomes by reducing complications from joint incongruence, particularly post-traumatic arthritis.[1,2] However, there is debate regarding the pathophysiology of post-traumatic

Table I. Details of malunion criteria met – comparison between clinically significant malunion and radiological malunion definitions.

CSM	RM			Participants, n (%)
	Talar tilt > 2°	Talar sublux > 2 mm	Tibfib clearspace ≥ 5 mm	
CSM and RM				**47 (8.53)**
Yes	Yes	Yes	Yes	18 (3.27)
Yes	Yes	Yes	No	20 (3.63)
Yes	Yes	No	Yes	2 (0.36)
Yes	Yes	No	No	1 (0.18)
Yes	No	Yes	Yes	1 (0.18)
Yes	No	Yes	No	4 (0.73)
Yes	No	No	Yes	1 (0.18)
No CSM and no RM				**278 (50.45)**
No	No	No	No	278 (50.45)
Total with CSM and RM in agreement				**325 (58.98)**
CSM and no RM				**6 (1.09)**
Yes	No	No	No	6 (1.09)
No CSM and RM				**220 (39.93)**
No	Yes	Yes	Yes	3 (0.54)
No	Yes	Yes	No	11 (2.00)
No	Yes	No	Yes	33 (5.99)
No	Yes	No	No	56 (10.16)
No	No	Yes	Yes	5 (0.91)
No	No	Yes	No	3 (0.54)
No	No	No	Yes	109 (19.78)
Total with CSM and RM in disagreement				**226 (41.02)**

CSM, clinically significant malunion; RM, radiological malunion; sublux, subluxation; Tibfib, tibiofibular.

osteoarthritis (OA). Some postulate that malunion of weightbearing joints leads to increased contact stresses on the articular surfaces,[3,4] in turn leading to OA, which can cause persistent symptoms and disability.[5,6] Others argue that post-traumatic OA may also be the result of genetic factors and direct damage caused by the initial trauma.[7] The contribution of adjustment remodelling of articular cartilage is recognized but poorly understood.[8] It has been suggested that function and pain are not as closely related to malunion as many clinicians believe.[9]

The AIM trial, a pragmatic equivalence randomized controlled trial with blinded outcome assessors, was designed to compare two treatments, surgical fixation, and close contact casting for ankle fractures.[10] Participants were adults aged 60 years or older with acute, overtly unstable ankle fractures. The primary outcome was the Olerud-Molander Ankle Score (OMAS; 0 to 100, higher scores = better ankle function)[11] at six months, and equivalence was prespecified as ± six points. In total, 620 patients were recruited to the trial, and equivalence of the two treatments at six months was demonstrated (OMAS score 66.0 (95% confidence interval (CI) 63.6 to 68.5) for surgery vs 64.5 (95% CI 61.8 to 67.2) for casting). Results from the primary analysis of the AIM trial indicated that radiological malunion was more common in the casting group (38/249; 15%) than in

the surgery group (8/274; 3%).[10] The same cohort was then followed for a minimum of three years (median 3 years, range 2.9 to 9.5).

The original AIM trial definition of malunion was based on prespecified radiological anatomical measurement thresholds, which were then confirmed by surgeon experts (KW and RH) as to whether there was 'clinically significant' malunion (see Methods). The variation in radiological projections, individual patient ankle joint morphology, and bone size were additional factors taken into account by the surgeons. The prespecified thresholds for each linear and angular measurement were considered the consensus definitions from the accepted research literature.[3,12] The lack of confirmatory clinical evidence for these definitions has previously been discussed.[13]

Data from the long-term follow-up of AIM trial participants[14] provided an opportunity to investigate the long-term effects of radiological malunion at six months, on participant function as measured by the OMAS score at three years follow-up, and to explore the relationship between each anatomical measurement (as a continuous variable) and patient-reported function.

The objectives of the present analysis were, firstly, to compare the definition of clinically significant malunion adopted in the trial with absolute radiological malunion

Table II. Radiological characteristics of individuals with a clinically significant malunion but not a radiological malunion.

Talar tilt, °	Talar subluxation, mm	Tibiofibular clear space, mm	Posterior malunion, Y/N
2.00	0	4.00	No
1.52	0	4.13	No
0.05	0	4.70	No
0.32	2	3.33	No
1.18	0	3.85	Yes

Note: cut-points for radiological malunion are talar tilt > 2°, talar subluxation > 2 mm, and/or tibiofibular clear space > 5 mm.

Table III. Relationship between presence of malunion and Olerud-Molander Ankle Score change scores.

Type of malunion	N	Impact on OMAS change score		
		Effect estimate	95% CI	p-value*
RM	422	-5.5	-9.8 to -1.2	0.012
CSM	422	-13.9	-21.9 to -5.8	0.001
PMM	174	-11.6	-22.6 to -0.6	0.039

*Linear regression.
CI, confidence interval; CSM, clinically significant malunion; OMAS, Olerud-Molander Ankle Score; RM, radiological malunion.

measurements; secondly, to investigate the impact of malunion on change in function to three years; and thirdly, to explore the possibility of a better definition of malunion based on radiograph data. The contribution of posterior malleolar fractures (size and displacement) was also considered.

Methods

Definitions of malunion. Three definitions of malunion were of interest: radiological malunion (RM) based on computer software measurement of radiograph data; clinically significant (radiological not necessarily symptomatic) malunion (CSM) based on surgeon expert opinion; and malunion of a posterior malleolar fracture (PMM) based on radiograph data. All definitions of malunion were determined at six months after randomization.

RM was defined based on computer calculation of radiograph continuous linear and angular measurements derived directly from image data. Radiographs were assessed within a purpose-built software using MatLab (The MathWorks, USA). The literature-sourced consensus thresholds for malunion were used.[3,12] Any measurement satisfying any one of the following three criteria (3 C) was considered a malunion: 1) talar tilt > 2°; 2) talar subluxation > 2 mm; or 3) tibiofibular clear space ≥ 5 mm. This definition of malunion was determined using the anteroposterior radiographs, or the mortise radiographs if available.

The definition of CSM was based on two surgeons' (KW and RH) expert opinions of participants' whole radiological series. Malunion was defined as a radiograph which surpassed the radiological consensus (using the thresholds outlined in the definition above), and was also considered to capture true displacement that was likely to be clinically significant (not apparent displacement due to, for example, angle of projection). This was the definition of malunion which was used in the original analyses of the AIM trial.

PMM, that is malunion of a posterior malleolar fracture, was defined as the presence of a posterior malleolar fracture of more than 5% of the articular surface as measured on a lateral radiograph and an articular

step > 2 mm.[15] Measurements of radiograph data using purpose-built software within MatLab were again used to determine which radiographs satisfied these criteria. This definition of malunion was determined using the lateral radiograph.

Long-term function. Ankle function was measured in the AIM trial using the participant-reported Olerud-Molander Ankle Score (OMAS).[11] The OMAS ranges from 0 to 100 with higher scores indicating better function. In this analysis, change from baseline to long-term follow-up was of interest and was calculated as the difference between these values. The baseline OMAS score was a recall of ankle function pre-injury recorded at trial recruitment.

Statistical analysis. The relationship between RM and CSM was first investigated by cross-tabulating these variables. Cohen's Kappa[16] was then used to assess the agreement between RM and CSM. Characteristics of those presenting with malunion under only one of the definitions were investigated to identify patterns within these individuals.

Linear regression models were used to explore the relationship between OMAS change scores and both RM and CSM. These models were adjusted for allocated treatment, participants' baseline OMAS scores, and age at baseline. Additional linear regression models were also used to explore the relationship between OMAS change scores and each of the variables used to determine RM (talar tilt, talar subluxation, and tibiofibular clear space) both individually and jointly.

For those participants with a posterior malleolar fracture, linear regression was used to explore the relationship between PMM and OMAS change scores. Additional linear regressions explored the relationship between articular step and surface and OMAS change scores.

The impact of missing data was explored by comparing baseline characteristics of the full AIM trial dataset and of those who provided sufficient radiological and long-term OMAS data.

The possibility of different thresholds for talar tilt, talar subluxation, and tibiofibular clear space which better explain changes in ankle function was investigated.

Table IV. Comparison of baseline characteristics and six-month malunion rates of the whole trial population compared to those with long-term follow-up data.

Characteristic	All (n = 620)	RM analysis (n = 422)	CSM analysis (n = 424)	All posterior fractures (n = 226)	PMM analysis (n = 174)
Median baseline age, yrs (IQR)	70 (65 to 76)	69 (64 to 74)	69 (64 to 74)	70 (64 to 75)	68 (63 to 74)
Median baseline OMAS (IQR)	100 (80 to 100)	100 (90 to 100)	100 (90 to 100)	100 (90 to 100)	100 (90 to 100)
Allocated to CCC, n (%)	311/620 (50.2)	209/422 (49.5)	210/424 (49.5)	112/226 (49.6)	82/174 (47.1)
RM, n (%)	268/554 (48.4)	201/422 (47.6)	n/a	N/A	N/A
CSM, n (%)	53/556 (9.5)	n/a	34/424 (8.0)	N/A	N/A
PMM, n (%)	N/A	N/A	N/A	40/226 (17.7)	22/174 (12.6)

CCC, close contact casting; CSM, clinically significant malunion; IQR, interquartile range; N/A, not applicable; n/a, not available; OMAS, Olerud-Molander Ankle Score; PMM, malunion of a posterior malleolar fracture; RM, radiological malunion.

Each variable was plotted against OMAS change scores, and non-parametric smoothing[17] of this relationship was conducted. Following this, piecewise linear models with one change point – that is linear models which allow two different trajectories, one below a threshold and one above – were used to determine new thresholds for each variable in turn. These models were fitted using a profile likelihood approach[18] to identify the change point, with the change point representing the new threshold. A profile likelihood approach fits piecewise linear models with a series of different change points, compares the log-likelihoods of each of these models, and selects the value for the change point which maximizes this. Since new thresholds were identified, the previous analyses (comparing malunion definitions and exploring the relationship between malunion and change in function) were repeated for these new thresholds. Similar methods were also used to search for different thresholds for PMM which better explained changes in function.

Each of these analyses was conducted for all participants from the AIM trial cohort with available data only, and no imputation of missing data was used. Analyses were conducted using Stata 15 (StataCorp, USA) and R (R Foundation for Statistical Computing, Austria).

Results

Comparing malunion definitions. A total of 551 individuals provided data on both their RM status and their CSM status. The agreement between these two definitions was moderate (59.0%); however, this was significantly larger than would be expected if the two definitions were unrelated (p < 0.001, Cohen's kappa). Table I summarizes the number and proportion meeting each of the RM criteria with and without CSM. Six individuals who had a CSM but not a RM were identified. Many of these individuals were close to the malunion cut-off for at least one of the variables, and one had a PMM (Table II). The group of individuals satisfying RM criteria only and not the CSM criteria was larger (220 participants). Many of these individuals failed to meet the

talar subluxation criteria, but met the exacting computer software determined thresholds for talar tilt or tibiofibular space through ankle joint posture or normal variations in bone morphology.

Impact of malunion on function. A total of 422 individuals provided radiological data and long-term follow-up on the OMAS. For these individuals, adjusted linear regression of OMAS change scores on RM was found to result in a statistically significant reduction in OMAS change scores (p = 0.012, see Table III); the estimated size of this change was -5.5 points (95% CI -9.8 to -1.2) and since the 95% CI included values below the minimum clinically important difference (MCID) for the OMAS (6 points),[19] this indicated little clinical importance. No substantial differences were found in the baseline characteristics and six-month malunion rates of those randomized in the AIM trial and those included in the model (see Table IV). A similar number of individuals had data on CSM status and long-term follow-up on the OMAS, and adjusted linear regression of OMAS change scores on CSM also resulted in a statistically significant reduction in OMAS change scores (p = 0.001; see Table III). The estimated size of this change was larger (-13.9 points) and the 95% CI only just crossed the MCID threshold (-21.9 to -5.8). Again, no substantial differences between the overall AIM trial population and those included in the model were identified (Table IV).

A total of 174 individuals had a posterior malleolar fracture and long-term follow-up on the OMAS. For these individuals, PMM had a statistically significant impact on OMAS change scores (p = 0.039) (Table III). The estimated size of this change was large (-11.6 points), but the CI contained the MCID (-22.6 to -0.6). More of those with a posterior malunion were missing long-term follow-up data than of those with a posterior fracture and no malunion (Table IV).

Linear regressions of OMAS change scores on talar tilt, talar subluxation, and tibiofibular clear space were also fitted (Table V). Both talar tilt (p < 0.001) and talar subluxation (p < 0.001) had a statistically significant

Table V. Relationship between Olerud-Molander Ankle Score change from baseline and malunion variables.

Fitted model	Coefficient	95% CI	p-value*
Model 1			
Talar tilt, °	-1.83	-2.71 to -0.96	< 0.001
Model 2			
Talar subluxation, mm	-3.42	-5.19 to -1.65	< 0.001
Model 3			
Tibiofibular clear space, mm	-1.39	-3.00 to 0.22	0.090
Model 4			
Talar tilt, °	-1.27	-2.38 to -0.17	0.024
Talar subluxation, mm	-1.84	-4.07 to 0.40	0.107
Model 5			
Post malleolar articular step, mm	-0.12	-2.90 to 2.66	0.933
Model 6			
Post malleolar articular surface, %	-0.21	-0.53 to 0.12	0.216
Model 7			
Post malleolar articular step, mm	-0.01	-2.79 to 2.77	0.993
Post malleolar articular surface, %	-0.21	-0.53 to 0.12	0.218

*Linear regression.
CI, confidence interval.
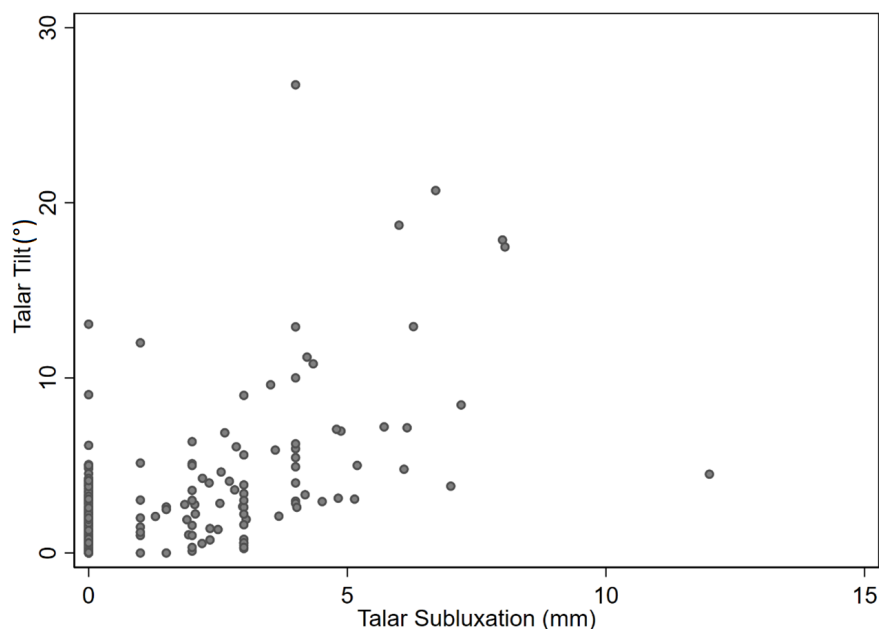
Fig. 1
Scatter plot of talar tilt against talar subluxation.

impact on OMAS change scores; however, tibiofibular clear space did not ($p = 0.090$). An additional model was fitted including both talar tilt and talar subluxation; only talar tilt remained statistically significant in this case ($p = 0.024$). This finding is likely due to the correlation between the two variables (Figure 1). The relationships between articular step and surface and OMAS change scores were also investigated for the 174 individuals with a PMM; neither was found to have a significant impact ($p = 0.933$ for articular step, $p = 0.216$ for articular surface) (Table V).
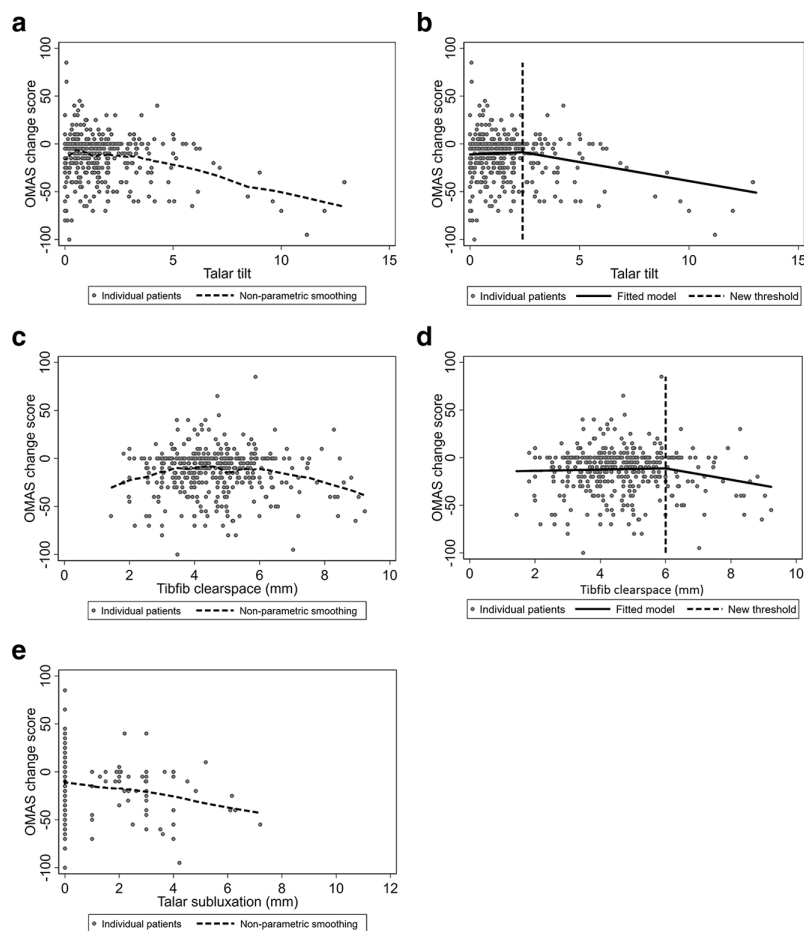
Fig. 2

a) Non-parametric smoothing of the relationship between talar tilt and Olerud-Molander Ankle Score (OMAS) change scores. b) Piecewise linear modelling of talar tilt and OMAS change to identify new threshold. c) Non-parametric smoothing of the relationship between tibiofibular clear space and OMAS change scores. d) Piecewise linear modelling of tibiofibular clear space and OMAS change to identify new threshold. e) Non-parametric smoothing of the relationship between talar subluxation and OMAS change scores. Tibfib, tibiofibular.

New malunion thresholds. Increasing talar tilt led to decreasing OMAS change scores, with non-parametric smoothing indicating that this relationship is stronger above a certain threshold (Figure 2a). A piecewise linear model with one change point demonstrated that the talar tilt threshold above which this relationship is altered is 2.4°, with little relationship between talar tilt and OMAS change score below this threshold (Figure 2b). This threshold is close to the existing accepted threshold of 2°. Non-parametric smoothing also indicated that the relationship between tibiofibular clear space and OMAS change scores varied depending on the size of the tibiofibular clear space (Figure 2c), with a piecewise linear model determining 6 mm to be a suitable value for this change point. Below this value tibiofibular clear space has little impact on OMAS change scores, while above the threshold increasing tibiofibular clear space results in a reduction in OMAS change score (Figure 2d). This threshold is close to the existing accepted value (5 mm). Figure 2e demonstrates that

increasing talar subluxation decreases OMAS change scores; however, no change in the trajectory for this relationship could be determined.

New thresholds for RM were determined to be any individual satisfying one of the following criteria: 1) talar tilt > 2.4°; 2) talar subluxation > 2 mm; or 3) tibiofibular clear space > 6 mm. This new definition was compared with CSM and led to an increased agreement, 413 out of 551 classifications match (74.95%, p < 0.001, Cohen's kappa). Linear regression demonstrated a stronger relationship between the new definition of RM and OMAS change scores, as would be expected, with the presence of a RM using the new definition leading to a mean difference in OMAS change scores of -9.1 points (p < 0.001; 95% CI -13.80 to -4.41).

For those with a posterior malleolar fracture, the relationships between OMAS change scores and articular step (Figure 3a) and articular surface (Figure 3b) were explored using scatterplots and non-parametric smoothing. No clear indication of a relationship between
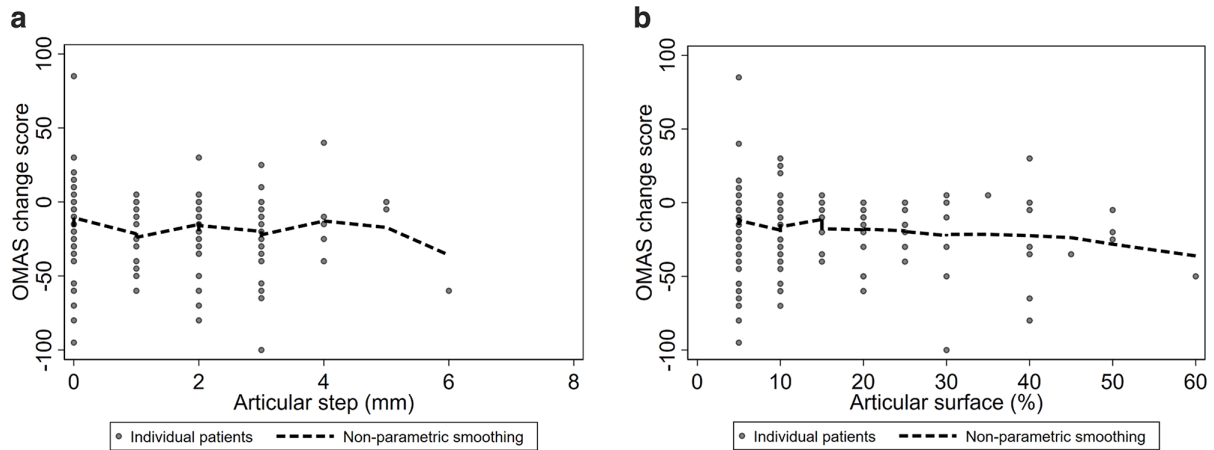
Fig. 3

a) Non-parametric smoothing of the relationship between: a) posterior malleolar fracture articular step and Olerud-Molander Ankle Score (OMAS) change scores; and b) posterior malleolar fracture articular surface and OMAS change scores.

these variables was identified as anticipated based on linear regressions. It is possible that the discrete nature of these variables played a role in these results.

Discussion

These analyses demonstrate that while the definitions of RM (based on absolute radiological measurements aided by computer software) and CSM (based on surgeon expert opinion of whole series radiographs) are related, there are individuals who satisfy only one of the two definitions. In addition, both of these measures were shown to have a statistically significant relationship to long-term function measured using the OMAS. The clinical impact of malunion was uncertain with the MCID for the OMAS (six points) contained in the 95% CI for the estimated impact of malunion on OMAS change score. CSM was more strongly related to outcome than RM with the 95% CI for the effect only just overlapping the MCID in this case (upper limit -5.8 points), confirming that the clinician-based definition does provide greater functional significance in terms of patient-reported outcome.

Separately, we have demonstrated that PMM also has a statistically significant effect on long-term function; however, there were fewer participants with posterior malleolus malunion than the 3 C criteria, so there was less information about this type of malunion and more uncertainty of clinical importance regarding longer term outcomes. This result may have been further impacted by substantial amounts of missing long-term follow-up data for participants with a radiological posterior malunion.

Previous thresholds for restoring precise congruence of the ankle articulations after fracture have been based on two widely cited cadaveric studies.[3,12] This evidence has a substantial influence on clinical practice, despite methodological limitations and a lack of confirmatory clinical evidence of short- or long-term functional outcomes.[13]

The analyses presented here indicate that for adults aged 60 years and over with an overtly unstable ankle fracture, these previously accepted clinical definitions of malunion are reasonable, if perhaps slightly conservative. The new thresholds generated in this study indicate that in this patient group somewhat higher thresholds for both talar tilt (2.4° instead of 2°) and tibiofibular clear space (6 mm instead of 5 mm) may be appropriate. If thresholds were to change in practice, this should ideally be done in the context of large-scale prospective observational studies in order to evaluate the implications of these changes in practice.

One important issue when interpreting our findings, highlighted in a previous annotation in *The Bone & Joint Journal*,[20] is that an association between radiological malunion and functional outcome was identified in the AIM cohort, but other factors may explain this relationship. During early clinic follow-up, surgeons and patients may have been more likely to accept a degree of malalignment if comorbidities or other clinical issues indicated that the impact of further intervention would not be warranted, given the trade-off between potential benefits and risks. All analyses described here have been adjusted for treatment allocation.

One strength of the study is the use of a patient-reported outcome (OMAS); this means that the associations observed are of importance to patients. In addition, the data source in these analyses was a large clinical trial cohort with limited loss of clinical follow-up data; however, the generalizability of these results is limited to the population eligible for inclusion in the clinical trial. Patients with diabetes mellitus and neurovascular compromise were excluded from the AIM trial and so the results cannot be generalized to these groups. An important limitation, related to some of the issues with OMAS baseline data, is that it was apparent that although

recalled pre-injury ankle function status was asked for in the baseline questionnaires, a small number of individuals scored themselves in the range expected if one were acutely injured. The result was that in these individuals their change from baseline to three-year follow-up was an improvement in ankle function post-injury. This is a limitation of using patient-reported outcomes and highlights the challenges of ensuring that framing of questionnaires is very clear in the acute setting. In addition, while use of a patient-reported outcome means the associations are of importance to patients, the lack of results based on clinical outcomes to support these findings can also be considered a limitation of this study.

Another potential limitation is that the radiographs were prescribed to be load-bearing; we recognize this may have affected the findings. However, taking non-weightbearing views is consistent with most common orthopaedic routine follow-up practice and in this context our findings are relevant and applicable. Surgeons will seek further diagnostic imaging where uncertainty exists. This is the standard approach to interpreting a radiograph, and it is important for the study results to be relevant and translatable into clinical practice that this was the method adopted. The risk of potential bias is accepted, as in the process of measurement noting the presence of metal implants cannot be avoided.

These results indicate that the use of patient-reported clinical outcomes to generate thresholds for malunion in younger adults and in individuals with other types of fracture would be of value to further support, or perhaps challenge, the clinically accepted definitions of malunion which are currently based on limited clinical evidence.

Take home message
- These results provide clinical evidence which supports the previously accepted definitions of malunion.
- The data indicate that existing thresholds for talar tilt and tibiofibular clearspace may be slightly conservative, and this should be investigated further in future studies.

Twitter
Follow D. J. Keene @davidkeenePT

References
1. **Marsh JL**, **Buckwalter J**, **Gelberman R**, **et al**. Articular fractures: does an anatomic reduction really change the result? *J Bone Joint Surg Am*. 2002;84-A(7):1259–1271.
2. **Dirschl DR**, **Marsh JL**, **Buckwalter JA**, **et al**. Articular fractures. *J Am Acad Orthop Surg*. 2004;12(6):416–423.
3. **Ramsey PL**, **Hamilton W**. Changes in tibiotalar area of contact caused by lateral talar shift. *J Bone Joint Surg Am*. 1976;58-A(3):356–357.
4. **McKinley TO**, **Rudert MJ**, **Tochigi Y**, **et al**. Incongruity-dependent changes of contact stress rates in human cadaveric ankles. *J Orthop Trauma*. 2006;20(10):732–738.
5. **Brown TD**, **Johnston RC**, **Saltzman CL**, **Marsh JL**, **Buckwalter JA**. Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease. *J Orthop Trauma*. 2006;20(10):739–744.
6. **Horisberger M**, **Valderrabano V**, **Hintermann B**. Posttraumatic ankle osteoarthritis after ankle-related fractures. *J Orthop Trauma*. 2009;23(1):60–67.
7. **Furman BD**, **Olson SA**, **Guilak F**. The development of posttraumatic arthritis after articular fracture. *J Orthop Trauma*. 2006;20(10):719–725.
8. **Kraeutler MJ**, **Kaenkumchorn T**, **Pascual-Garrido C**, **Wimmer MA**, **Chubinskaya S**. Peculiarities in ankle cartilage. *Cartilage*. 2017;8(1):12–18.
9. **Giannoudis PV**, **Tzioupis C**, **Papathanassopoulos A**, **Obakponovwe O**, **Roberts C**. Articular step-off and risk of post-traumatic osteoarthritis. Evidence today. *Injury*. 2010;41(10):986–995.
10. **Willett K**, **Keene DJ**, **Mistry D**, **et al**. Close contact casting vs surgery for initial treatment of unstable ankle fractures in older adults: A randomized clinical trial. *JAMA*. 2016;316(14):1455–1463.
11. **Olerud C**, **Molander H**. A scoring scale for symptom evaluation after ankle fracture. *Arch Orthop Trauma Surg (1978)*. 1984;103(3):190–194.
12. **Yablon IG**, **Heller FG**, **Shouse L**. The key role of the lateral malleolus in displaced fractures of the ankle. *J Bone Joint Surg Am*. 1977;59-A(2):169–173.
13. **Smith G**. The isolated lateral malleolar fracture: where are we and how did we get here? *Surgeon*. 2013;11(1):6–9.
14. **Keene DJ**, **Lamb SE**, **Mistry D**, **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. *JAMA*. 2018;319(12):1274–1276.
15. **Drijfhout van Hooff CC**, **Verhage SM**, **Hoogendoorn JM**. Influence of fragment size and postoperative joint congruency on long-term outcome of posterior malleolar fractures. *Foot Ankle Int*. 2015;36(6):673–678.
16. **Cohen J**. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37–46.
17. **Cleveland WS**. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc*. 1979;74(368):829–836.
18. **Montoya JA**, **Díaz-Francés E**, **Sprott DA**. On a criticism of the profile likelihood function. *Stat Papers*. 2009;50(1):195–202.
19. **Nilsson GM**, **Eneroth M**, **Ekdahl CS**. The Swedish version of OMAS is a reliable and valid outcome measure for patients with ankle fractures. *BMC Musculoskelet Disord*. 2013;14(1):109.
20. **Keene DJ**, **Willett K**. Implications of the Ankle Injury Management (AIM) trial: close contact casting or surgery for older adults with an unstable ankle fracture? *Bone Joint J*. 2019;101-B(12):1472–1475.

Author information:
- R. Knight, MMath, PhD, Senior Medical Statistician
- S. J. Dutton, BSc, MSc, PGCE, OCTRU Lead Statistician, Associate Professor
  Oxford Clinical Trials Research Unit, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK.
- D. J. Keene, BSc (Hons), MSc, DPhil, University Research Lecturer in Trauma Rehabilitation
- K. Willett, MBBS, Professor of Trauma Surgery
  Kadoorie Centre for Critical Care Research and Education, John Radcliffe Hospital, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK.
- R. Handley, FRCS, Consultant Trauma & Orthopaedic Surgeon, Oxford Trauma Service, John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK.

Author contributions:
- R. Knight: Formal analysis, Methodology, Writing – original draft.
- D. J. Keene: Conceptualization, Methodology, Writing – original draft.
- S. J. Dutton: Supervision, Writing – review & editing.
- R. Handley: Conceptualization, Writing – review & editing.
- K. Willett: Conceptualization, Writing – review & editing.