BONE JOINT JOURNAL

TITLE PAGE

TITLE: Outcomes in Ankle Replacement Study (OARS) – radiographic and patient reported outcomes in the first 12 months after ankle joint replacement

AUTHORS: Smith TO, Dainty J, Loveday D, Toms A, Goldberg A, Watts L, Pennington M, Dawson J, van der Meulen J, MacGregor A and the OARS Collaborators

AUTHORS & AFFILIATIONS

Author	Affiliations	Email	ORCID
Toby O Smith	Warwick Medical School, University of Warwick, Coventry	toby.o.smith@warwick.ac.uk	0000-0003-1673-2954
Jack Dainty	Norwich Medical School, UEA	Jack.dainty@uea.ac.uk	0000-0002-0056-1233
David T Loveday	Norfolk and Norwich University Hospitals, Norwich	david.loveday@nnuh.nhs.uk	0000-0002-6383-2037
Andoni Toms	Addenbrookes, Cambridge	andoni.toms@nhs.net	0000-0003-2847-020X
Andy Goldberg	UCL Division of Surgery, Royal Free Hospital, London. MSK Hub, Imperial College London, White City Campus, London.	andy.goldberg@nhs.net	0000-0002-8650-4503
Laura M Watts	Norwich Medical School, UEA	l.watts1@uea.ac.uk	0000-0002-1816-966X
Mark Pennington	Institute of Psychiatry, Psychology and Neurosciences, King's College London, London.	mark.w.pennington@kcl.ac.uk	0000-0002-1392-8700
Jill Dawson	Nuffield Department of Population Health, University of Oxford, Oxford	jill.dawson@ndph.ox.ac.uk	0000-0002-2358-1333
Jan van der Meulen	Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London	jan.vandermeulen@lshtm.ac.uk	0000-0002-9451-2335
Alex MacGregor	Norwich Medical School, UEA	a.macgregor@uea.ac.uk	0000-0003-2163-2325
And the OARS Collaborators			

Corresponding Author: Professor Alex MacGregor. Norwich Medical School, University of East Anglia, Norwich, NR4 7JT. Email: a.macgregor@uea.ac.uk

ABSTRACT

AIMS: To capture 12-month outcomes from a representative multicentre cohort of patients undergoing TAR, to describe the pattern of patient reported outcome measures (PROMs) at 12 months, and to identify predictors of these outcome measures.

METHODS: Patients listed for a primary TAR at 19 NHS hospitals between February 2016 and October 2017 were eligible. PROMs data were collected pre-operatively and at six and 12 months including: Manchester-Oxford Foot and Ankle Questionnaire (MOXFQ (foot and ankle) and the EQ-5D-5L. Radiological pre- and post-operative data included Kellgren-Lawrence score and implant position measurement. This was supplemented by data from the National Joint Registry through record linkage to determine: American Society of Anesthesiology (ASA) grade at index procedure; indication for surgery, index ankle previous fracture; tibial hind foot alignment, additional surgery at the time of TAR and implant type. Multivariate regression models assessed outcomes and the relationship between MOXFQ and EQ-5D-5L outcome with patient characteristics.

RESULTS: Data from 238 patients were analysed. There were significant improvements in MOXFQ and EQ-5D-5L among people who underwent TAR at six and 12-month assessments compared with pre-operatively scores ($p<0.001$). Most improvement occurred between pre-operative to six months, with little further improvement at 12 months. A greater improvement in MOXFQ outcome post-operatively was associated with older age and more advanced radiological signs of ankle osteoarthritis at baseline.

CONCLUSIONS: TAR significantly benefits patients with end-stage ankle disease. The lack of substantial further overall change between six and 12-months suggests that capturing PROMs at six months is sufficient to assess the success of the procedure. Older patients and those with advanced radiological disease who had the greater gains. These outcome predictors can be used to counsel younger patients and those with earlier ankle disease on the expectations of TAR.

INTRODUCTION

Ankle osteoarthritis (OA) is a common cause of ankle disability. An estimated 29,000 cases of symptomatic ankle OA are referred to UK specialists annually.[1] Ankle OA has similar impact on quality of life as hip OA,[2] end-stage kidney disease and heart failure.[3] Most ankle OA is post-traumatic.[3,4] It has been estimated that 56% of people who undergo ankle surgery for ankle OA are in employment.[5]

Surgical interventions in the form of arthrodesis and total ankle joint replacement (TAR) are part of the standard management of patients with end-stage ankle OA.[6] Whilst the use of TAR has been increasing in the UK in the last decade,[7] routine capture of patient reported outcomes does not take place including the National Joint Registry for England, Wales and the Isle of Man (NJR). Risk factors for revision after TAR have been examined. These have implicated a role for age, body mass index (BMI), activity, pre-operative pain, adjacent joint OA, post-operative radiological alignment and surgical technique.[8-10] However, the available data from studies investigating patient reported outcome measures (PROMs) following TAR are typically underpowered, based on limited follow-up, or derived from clinical trials conducted in selected groups.[11] There are currently no comprehensive data on patient outcomes following TAR from unselected populations to help guide patient and surgeon choice.

The objective of this study (the outcomes of ankle replacement study – OARS) was to capture 12-month outcomes from a representative multicentre cohort of patients undergoing TAR, to describe the pattern of PROMs at 12 months, and to identify predictors of these outcome measures.

METHODS

Participants and Recruitment

OARS was conducted at 19 NHS hospitals in England. All patients listed for a primary TAR between February 2016 and October 2017 were eligible for inclusion. Patients whose indication for surgery was acute trauma, and those who were undergoing a revision TAR were excluded. Recruits were identified prospectively at pre-operative assessment clinics held within two weeks prior to surgery. Following informed consent, each patient was asked to complete a pre-operative questionnaire. Consent was obtained to share radiographic data with the study team and for linkage to data held in the NJR. The study was performed contemporaneously with the TARVA trial which investigated clinical outcomes of TAR to ankle fusion.[11] This paper reports the data on 238 participants recruited to OARS solely, excluding 39 participants who were recruited into both OARS and TARVA.

Data Collection

The study was administered by the study team at the University of East Anglia. Baseline PROMs data were collected during the pre-operative appointment and posted to the study team. Post-operative questionnaires were mailed directly to patients at six and 12 months after the procedure. Reminder letters were sent to non-responders within six weeks of the planned follow-up interval.

Data collected in the baseline questionnaire included: age, gender, BMI, education status, housing status, whether participants lived alone or not, employment status and the duration of symptoms. PROMs data were collected in the baseline questionnaire including: a foot/ankle-specific PROMs (the MOXFQ (foot and ankle)).[12] The MOXFQ is a validated PROM for ankle surgery.[13,14] It contains 16-items, each with five response options, comprising three separate underlying dimensions: walking/standing problems (seven-items), foot pain (five-items), and issues related to social Interaction (four-items), including feelings of self-consciousness about foot/footwear appearance ('cosmesis'). Item responses were each scored from zero to four, with four representing the most severe state. Summed raw scale scores are then converted to a metric (zero to 100; 100=most severe). A generic quality of life/utility measure was assessed using the EQ-5D-5L.[15] Six and 12-month post-operative questionnaires collected data on the MOXFQ and EQ-5D-5L. The minimally clinically important difference (MCID) for the three subsections of the MOXFQ is 12.8, 4.6, and 20.3 respectively.[16]

The baseline and postal questionnaire data were supplemented by information obtained from the NJR through record linkage. This included information on American Society of Anesthesiology (ASA) grade[17] at index procedure; indication for surgery, index ankle previous fracture; tibial hind foot alignment, additional surgery at the time of TAR (either surgery to bone or soft tissue) and type of bearing used (fixed versus mobile TAR).

Radiological assessment

Details of the radiographic assessments included in OARS have been published previously.[18] All measurements were performed in participating hospitals, using electronic goniometers and calipers on a diagnostic PACS workstation (Synapse® PACS, Fujifilm Healthcare, Singapore) with 3MP diagnostic monitors (Barco, Kortrijk, Belgium). Measurements on all patients were performed by a musculoskeletal radiologist (AT) with 25 years' experience. Inter and intra-rater reliability was assessed on a subset of 62 OARS patients and demonstrated to be good.[18]

The analysis included data from pre-operative and post-operative prior to hospital discharge, anteroposterior (AP) and lateral weight-bearing radiographs. Based on an earlier examination of the data, the pre-operative measurements included: distal tibial articular angle, talocalcaneal and calcaneal inclination angles, minimal joint space width and the Kellgren-Lawrence score (**Figure 1**).[18-21] Post-operative measurements included angles α and β to define the angulation of the tibial component relative to the tibia and γ which defined the angular position of the talar component (**Figure 1**).[18]

Sample Size

A review of registrations to the NJR prior to the study indicated that OARS could recruit appropriately 250 participants over a 20-month period from participating sites. This sample size was expected to deliver a power of 95% to detect a standardised difference of 0.23 (SD: 1.00) between the mean value of the total MOXFQ score at baseline and at six months (MCID for MOXFQ is reported as one standard deviation (SD)).[16]

Data Analysis

Analyses were conducted using R open-source software (www.r-project.org). Within the MOXFQ, scores were derived for each of the three subdomains: (1) walking/standing problems (combination of seven items – higher scores indicate more problems); (2) foot pain (combination of five items – higher scores indicate more pain); and (3) social interaction (combination of four items – higher scores indicate less social interaction). EQ-5D-5L scores were converted to a single utility index using country specific value sets using the 'eq5d' package in R, setting the country to 'England'.[22] The EQ-5D-5L

Health visual analogue scale (VAS) score (0 (worst health) to 100 (perfect health)) were included as a separate outcome.

Baseline characteristics were analysed using descriptive statistics. The univariate association with baseline characteristics were examined in patient strata by comparing point estimates and their confidence intervals (CI).

Effect sizes (ES) were calculated by dividing the mean change in scores (post-surgical minus pre-surgical scores), by the SD at baseline (pre-surgical score). An ES of 1.0 is thus equivalent to a change of one SD in the sample. Values of 0.2, 0.5 and 0.8 are typically regarded as indicating small, medium and large degrees of change.[23]

To accommodate potential confounding, a multivariate regression model was used to model the change in scores from baseline.[24] As there was little difference in the outcome data at 12 months in the data, the model considered only the change in values between baseline and the six-months. The model was constructed using a backwards elimination process, starting with the full variable set in **Table 1**. Independent variables were checked for correlation with each other. Within the radiological measurements, as expected, minimum joint spacing correlated strongly with Kellgren-Lawrence scores. Accordingly, only Kellgren-Lawrence score was retained in the analysis. The following variables were considered as explanatory: age (65 years or older defined as 'older',[25] gender, problem length, BMI, associated procedure on bone at time of surgery, associated procedure on soft tissue at time of surgery, indications for implantation, index ankle previous fracture, tibia-hindfoot alignment, bearing type, Kellgren-Lawrence score, distal tibial articular angle, talo-calcaneal angle and calcaneal inclination. The independent variables were entered as single variables. Interactions were not considered given the limited statistical power. The analyses were conducted separately for MOXFQ and EQ-5D-5L. The models were fitted and then iteratively pruned by backwards elimination of least significant variables (one at a time and model refitted) until only significant ($p<0.05$) explanatory variables remained. In this analysis, missing values in BMI and Kellgren-Lawrence score were imputed in R using the 'mice' package.[26] The number of imputations was n=100; Predictive Mean Matching (PMM) was performed.

RESULTS

Cohort characteristics

As summarised in **Figure 2**, 290 participants were recruited. From these, data was available from 238 participants to include in the analysis. There were no important differences between the analysed and non-analysed cohorts (**Supplementary File 1**). The cohort characteristics are summarised in **Table 1**. The 238 patients analysed had a mean (SD) age of 65.2 (11.1) years, and BMI 29.6 (4.8) kg/m^2. Most were male (58%); 55% had retired. Most were coded as having ankle OA (78%) with 78% having the highest grade of OA (Kellgren-Lawrence=4). In total, 21 different lead surgeons performed the 238 TARs. Ninety-seven percent of cases were performed by a surgeon of consultant-level. A summary of the TAR implants used is presented in **Supplementary Table 2**.

Clinical outcome of TAR

MOXFQ improved significantly across all three domains (foot pain, walking and social interactions) among people who underwent TAR at both six months and 12-month assessments compared with pre-operatively scores ($p<0.001$; **Figure 3**). The same pattern of improvement was seen for the EQ-5D-5L (**Figure 4**). The most improvement in MOXFQ and EQ-5D-5L score occurred between pre-operative and six months timepoints, with little further improvement at 12 months (**Table 2**; **Table 3**). This change of 37 was above the MCID for the MOXFQ.[16] Effect sizes are shown in **Figure 3** and **Figure 4**. All ES are deemed to be 'large' (>0.8) apart from the ES for EQ-5D-5L Health which was 'small'. Generally, the ESs for the MOXFQ variables were larger than for the EQ-5D-5L variables.

Radiological outcomes of TAR

Table 4 and **Table 5** indicate how MOXFQ and EQ-5D-5L varied with three important radiological outcomes (angles: α, β and γ). There was no relationship between radiological variables and outcome in these data ($p>0.05$).

Relationship between MOXFQ and EQ-5D-5L and patient characteristics

Table 2 and **Table 4** show the relationship between patient (baseline) characteristics with MOXFQ and EQ-5D-5L at six and 12 months. In general, patients younger than 65 years, female or had BMI>30 kg/m^2 reported poorer baseline MOXFQ total scores ($p<0.05$). A trend towards a greater improvement in outcome was observed for those who had ankle problems for more than five years and for those whose Kellgren-Lawrence score was four (the highest). The choice of bearing type (mobile or fixed) did not appear to have been influenced by baseline MOXFQ (**Table 2**) or EQ-5D-5L (**Table 4**) or influence change over six and 12 months.

Table 6 shows the pre-operative variables that were significantly associated with MOXFQ and EQ-5D-5L outcomes in multivariate analysis for the difference between baseline and six months. A greater improvement in perceived MOXFQ outcome post-operatively was associated with an older age and more advanced radiological signs of ankle OA (indicated through Kellgren-Lawrence score) at baseline. No variables were significantly associated with EQ-5D-5L outcome ($p>0.05$).

DISCUSSION

The findings demonstrate that clinically important improvements in foot pain, function, social interaction and HRQoL are seen over the first 12 post-operative months following TAR. Outcomes, when measured by PROMs, do not substantially improve between six to 12-months. Increasing age and more advanced pre-operative Kellgren-Lawrence scores are predictors of improvement in MOXFQ outcome.

Our findings are consistent with the results of smaller published studies. Agarwalla et al[27] reported significant improvements up to six months after TAR using the American Orthopaedic Foot and Ankle Society Ankle Hindfoot Score and VAS and by the Short Musculoskeletal Function Assessment Dysfunction and Bother subsections at 12 months. Similarly, King et al[28] reported improvements in both HRQoL assessed with the EQ-5D-5L and clinical outcomes with the MOXFQ for a cohort of 20 patients two years after TAR. The OARS study was conducted contemporaneously with a randomised clinical trial of ankle joint replacement compared with fusion (the TARVA trial[11]) at similar sites. OARS was designed as an observational study, with broader selection criteria from a wider base and more reflective of real-world clinical practice. No patient from TARVA is included in the present analysis. The outcomes recorded in OARS and those in the interventions arm in TARVA are nevertheless complementary and give the message that TAR can offer significant benefit in symptoms within the short-term.

Our analysis showed that increasing severity of pre-operative joint degeneration, as determined by the Kellgren-Lawrence scale, was associated with greater improvement. This is consistent with previously reported findings indicating that radiological severity may be related to post-operative prognosis.[29] Chambers et al[29] hypothesised that patients with more 'undamaged' parts of the joint have greater proprioceptive feedback pre-operatively. This feedback is lost after TAR. They suggested patients may find a painful joint, with largely healthy areas, is preferable to the 'artificial' joint with lost proprioceptive capabilities. This provides a possible explanation for these findings.

A number of potentially important variables showed no relationship with change in MOXFQ or EQ-5D-5L in the present study. These included BMI, gender, previous ankle fracture or associated procedures at the time of TAR. This is important because variables such as BMI have been previously used to help determine eligibility for surgery and 'ration' treatment.[8] Furthermore, the finding that associated procedures and previous ankle fracture was not associated with outcome indicates that complexity of TAR surgery may be of less prognostic importance than previously suggested.[8] We acknowledge that these observations need to be tempered by statistical power considerations. In the case of BMI, the relatively narrow range of BMI also limits conclusions over outcomes in the most overweight or obese patients. Further the focus here has been on the analysis on PROMs rather than survivorship, which is clearly an equally important consideration on TAR decision-making.

Older people reported better outcomes when measured using the walking/standing subdomain of the MOXFQ. This lends support to the previously held view that TAR may offer greater benefit for those with lower physical demand[30] on the assumption that age is a surrogate for physical activity. Whilst it was not the purpose of this analysis to explore whether choice of implant (particularly fixed versus mobile) or level of physical demand were important factors for younger versus older patients, further exploration may be helpful to counsel younger patients on whether TAR is the right surgical choice for them.

Our analysis found no relationship between outcome and pre-operative x-ray malalignment. Previous studies have acknowledged some limitations to the reliability of ankle radiological measures for this population.[18] Whilst poor measurement precision may be one explanation for why no relationship was identified, this was mitigated against by excluding those measurements with documented poor reliability. We also note that the majority of patients in the present study had near normal alignment (within 10 degrees of neutral).

A strength of this study is that it includes data from 19 hospitals in England, from a wide and representative sample from 19 hospitals in England. Whilst previous study[12] has presented MOXFQ scores in relation to a wide range of foot and ankle surgical procedures, where scores for all procedures were amalgamated, the current study presents pre- and post-operative MOXFQ scores specifically for ankle replacement. Accordingly, these data offer novel reference data useful to both surgeons and researchers. The data contributes to a growing literature offering MOXFQ reference scores for particular conditions or procedures. These data are extremely useful for designing future comparative studies. Furthermore, as MOXFQ scores are patient-centred, these data are potentially also useful for informing expectations of outcomes following surgical procedures, in prior decision-making discussions with patients. Finally, the NJR does not capture PROMs for TAR. As PROMs collection is routine for other arthroplasties such as TKR and THR, the data are the first to provide nationally-representation of data.

The study presents with three key limitations should be acknowledged. Firstly, there were missing data for radiological data due to challenges in data collection, limited data on potentially important factors on outcome such as ethnicity and limited information on occupational profiles pre- and post-operatively. Similarly, whilst we explored global outcomes with the PROMs, we were unable to report comprehensively on wider health status and any medical co-morbidities. Such information may provide greater context to inform the generalisability of this cohort. Secondly, whilst a 12-month follow-up provides important indicative findings, a longer follow-up period, including data on implant survival is clearly needed to fully to provide a comprehensive assessment of the outcomes of TAR. Finally, this analysis was designed to detect changes in MOFXQ and EQ-5D-5L data over the follow-up period following TAR. We excluded data on 12 participants (4%) who underwent additional surgery with the aim of exploring the outcomes of TAR alone. This may have created bias towards better outcomes for the TAR cohort as a whole.

CONCLUSIONS

This is the first study to assess 12-month PROM TAR outcomes in a representative, national cohort. The study demonstrates how changes in PROMs after TAR are dependent on patient characteristics. TAR significantly improved patient reported outcomes over the first 12 post-operative months compared to pre-operative outcomes. Older age and worse pre-operative Kellgren-Lawrence score were predictors of a greater improvement in outcome at six months. There was no substantial difference in outcome between six and 12-months suggesting ceasing PROMs surveillance at six-months may appropriately predict outcome.

FIGURE AND TABLE LEGENDS

Figure 1: Diagram illustrating the radiological measures used for anteroposterior (A,C) and lateral (B,D) radiographs of the ankle.

Figure 2: OARS participant flow-chart illustrating recruitment, attrition and numbers included in the final analysis.

Figure 3: Preoperative and postoperative (six and 12 month) MOXFQ (bars represent mean value with SE)

Figure 4: Preoperative and postoperative (six and 12 month) EQ5D (bars and data points represent mean value with SE)

Table 1: Total ankle replacement patient characteristics at baseline

Table 2: Pre-operative variables: change from baseline at six months and 12 months for the MOXFQ Total Score

Table 3: Pre-operative variables: change from baseline at six months and 12 months for the EQ-5D-5L utility score

Table 4: Post-operative variables: change from baseline at six months and 12 months for the MOXFQ Total Score

Table 5: Post-operative variables: change from baseline at six months and 12 months for the EQ-5D-5L utility score

Table 6: Statistical association of TAR study outcomes with baseline explanatory variables using linear regression models

Supplementary Table 1: Total ankle replacement (TAR) patient characteristics at baseline for those that did not complete the MOXFQ at six months.

Supplementary Table 2: Summary of frequency of TAR implants across 260 participants.

DECLARATIONS

OARS Collaborators: Stephen Bendall (Brighton & Sussex University Hospitals NHS Trust), James McKenzie (The Royal Orthopaedic Hospital NHS FT), Ian Sharpe (Royal Devon & Exeter NHS Foundation Trust), Henry Willmot (East Sussex Healthcare NHS Trust), Chris Blundell (Sheffield Teaching Hospitals), Andrew Bing (Robert Jones & Agnes Hunt Orthopaedic Hospital NHS FT), Mike Karski (Wrightington, Wigan & Leigh NHS Foundation Trust), Atef Hakmi (East and North Hertfordshire NHS Trust), Andy Goldberg (Royal National Orthopaedic Hospital NHS Trust), Dave Townshend (Northumbria Healthcare NHS Foundation Trust), Martin Raglan and Sunil Dhar (Nottingham University Hospitals NHS Trust), David Loveday (Norfolk and Norwich University Hospitals NHS Foundation Trust), Aamir Zubairy (East Lancashire Hospitals NHS Trust), Mark Farndon and David Lavalette (Harrogate & District NHS Foundation Trust), Harish Krupp (United Lincolnshire Hospitals NHS Trust), Viren Mishra and Javid Salim (Hull and East Yorkshire Hospitals NHS Trust), Malik Siddique and Paulo Torres (Newcastle upon Tyne Hospitals NHS FT).

Acknowledgements: We would like to thank Dr Helen Doll who was a co-applicant on the original OARS funding application. We would also like to thank the following NHS organisations who provided x-rays used in the radiological analysis: Brighton & Sussex University Hospitals NHS Trust; Derby Hospitals NHS Foundation Trust; East and North Hertfordshire NHS Trust; East Lancashire Hospitals NHS Trust; East Sussex Healthcare NHS Trust; Frimley Health NHS Foundation Trust; Harrogate and District NHS Foundation Trust; Hull and East Yorkshire Hospitals NHS Trust; Newcastle Upon Tyne Hospitals NHS Foundation Trust; Norfolk and Norwich University Hospital NHS Foundation Trust; Northumbria Healthcare NHS Foundation Trust; Nottingham University Hospitals NHS Trust; Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust; Royal Devon and Exeter NHS Foundation Trust; Royal National Orthopaedic Hospital NHS Trust; Royal Orthopaedic Hospital NHS Foundation Trust; Sheffield Teaching Hospitals NHS Foundation Trust; United Lincolnshire Hospitals NHS Trust; Wrightington, Wigan & Leigh NHS Foundation Trust.

We thank the patients and staff of all the hospitals in England, Wales and Northern Ireland who have contributed data to the National Joint Registry. We are grateful to the Healthcare Quality Improvement Partnership (HQIP), the NJR Research Committee and staff at the NJR Centre for facilitating this work. The authors have conformed to the NJR's standard protocol for data access and publication. The views expressed represent those of the authors and do not necessarily reflect those of the National Joint Registry Steering Committee or the Healthcare Quality Improvement Partnership (HQIP) who do not vouch for how the information is presented.

Permission (licence) to use the MOxFQ was granted by the Clinical Outcomes Team at Oxford University Innovation, owners of the MOxFQ. Further information on the MOxFQ, including access to a sample copy, links to key publications, list of available language versions and other valuable information can be found on the Oxford University Innovation website at https://innovation.ox.ac.uk/outcome-measures/manchester-oxford-foot-questionnaire-moxfq/

Ethical Approval: East Midlands - Nottingham 1 Research Ethics Committee (Ref: 15/EM/0153) Approval: 27 March 2015.

Funding: The study was funded by a Versus Arthritis project grant (reference: 20284).

REFERENCES

1. Goldberg AJ, MacGregor A, Dawson J, et al. The demand incidence of symptomatic ankle osteoarthritis presenting to foot & ankle surgeons in the United Kingdom. *Foot* 2012;22(3):163-166.

2. Glazebrook M, Daniels T, Younger A, et al. Comparison of health-related quality of life between patients with end-stage ankle and hip arthrosis. *J Bone Joint Surg (Am)* 2008;90(3):499-505.

3. Paget LDA, Tol JL, Kerkhoffs GMMJ, Reurink G. Health-related quality of life in ankle osteoarthritis: a case-control study. *Cartilage* 2021;13(1 Suppl):S1438-1444.

4. Jaleel A, Golightly YM, Alvarez C, Renner JB, Nelson AE. Incidence and progression of ankle osteoarthritis: The Johnston county osteoarthritis project. *Semin Arthritis Rheum* 2021;51(1):230-235.

5. Gagné OJ, Veljkovic A, Glazebrook M, et al. Prospective cohort study on the employment status of working age patients after recovery from ankle arthritis surgery. *Foot Ankle Int* 2018;39(6):657-663.

6. Clough TM, Ring J. Total ankle arthroplasty. *Bone Joint J* 2021;103(4):696-703.

7. NJR. National Joint Registry 18th Annual Report 2021. National Joint Registry. Available at: https://reports.njrcentre.org.uk/Portals/0/PDFdownloads/NJR 18th Annual Report 2021.pdf Accessed on: 04 May 2022

8. Hermus JPS, van Kuijk SMJ, Spekenbrink-Spooren A, et al. Risk factors for total ankle arthroplasty failure: a Dutch Arthroplasty Register study. *Foot Ankle Surg* 2021:S1268-7731(21)00243-5.

9. Cody EA, Bejarano-Pineda L, Lachman JR, et al. Risk factors for failure of total ankle arthroplasty with a minimum five years of follow-up. *Foot Ankle Int* 2019;40(3):249-258.

10. Halai MM, Pinsker E, Mann MA, Daniels TR. Should 15° of valgus coronal-plane deformity be the upper limit for a total ankle arthroplasty? *Bone Joint J* 2020;102(12):1689-1696.

11. Goldberg AJ, Chowdhury K, Bordea E, et al. Total ankle replacement versus arthrodesis for end-stage ankle osteoarthritis. A randomized controlled trial. *Ann Intern Med* 2022; In Press: doi:10.7326/M22-2058

12. Dawson J, Boller I, Doll H,et al. The MOXFQ patient-reported questionnaire: assessment of data quality, reliability and validity in relation to foot and ankle surgery. *Foot* 2011;21(2):92-102.

13. Morley D, Jenkinson C, Doll H, et al. The Manchester-Oxford Foot Questionnaire (MOXFQ): Development and validation of a summary index score. *Bone Joint Res* 2013;2(4):66-69.

14. Jia Y, Huang H, Gagnier JJ. A systematic review of measurement properties of patient-reported outcome measures for use in patients with foot or ankle diseases. *Qual Life Res* 2017;26(8):1969-2010.

15. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16(3):199-208.

16. Dawson J, Doll H, Coffey J, Jenkinson C; Oxford and Birmingham Foot and Ankle Clinical Research Group. Responsiveness and minimally important change for the Manchester-Oxford foot questionnaire (MOXFQ) compared with AOFAS and SF-36 assessments following surgery for hallux valgus. *Osteoarthritis Cartilage* 2007;15(8):918-931.

17. Saklad M. Grading of patients for surgical procedures. *Anesthesiology* 1941;2(3):281-284.

18. Low SBL, Kim M, Smith T, Loveday D, MacGregor A, Toms AP. The reliability of radiographic measures of total ankle replacement position: an analysis from the OARS cohort. *Skeletal Radiol* 2021;50(7):1411-1417.

19. Kellgren JH, Lawrence J. Radiological assessment of osteo-arthrosis. *Annal Rheum Dis* 1957;16(4):494.

20. Lee OS, Lee SH, Lee YS. Does coronal knee and ankle alignment affect recurrence of the varus deformity after high tibial osteotomy? *Knee Surg Relat Res* 2018;30(4):311.

21. Waldt S, Woertler K. Measurements and classifications in musculoskeletal radiology. Thieme, Stuttgart 2014. ISBN 978-3-13-169271-9

22. Morton F, Nijjar JS. eq5d: Methods for analysing 'EQ-5D' data and calculating 'EQ-5D' index scores. R package version 0.10.1: 2021. Accessed: 08 November 2022. Available at: https://CRAN.R-project.org/package=eq5d

23. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;27(3 Suppl): S178-S189.

24. van Breukelen GJP. ANCOVA Versus CHANGE from baseline in nonrandomized studies: the difference. *Multivariate Behavioral Res* 2013;48(6):895-922.

25. Merkt H, Haesen S, Meyer L, Kressig RW, Elger BS, Wangmo T. Defining an age cut-off for older offenders: a systematic review of literature. *Int J Prison Health* 2020;16(2):95-116.

26. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Software* 2011;45(3):1-67.

27. Agarwalla A, Gowd AK, Cody EA, et al. Timeline for maximal subjective outcome improvement following total ankle arthroplasty. *Foot Ankle* Surg 2021;27(3):305-310.

28. King A, Bali N, Kassam AA, Hughes A, Talbot N, Sharpe I. Early outcomes and radiographic alignment of the Infinity total ankle replacement with a minimum of two year follow-up data. *Foot Ankle Surg* 2019;25(6):826-833.

29. Chambers S, Ramaskandhan J, Siddique M. Radiographic severity of arthritis affects functional outcome in total ankle replacement (TAR). *Foot Ankle Int* 2016;37(4):351-*35*4

30. Samaila EM, Bissoli A, Argentini E, Negri S, Colò G, Magnan B. Total ankle replacement in young patients. *Acta Biomed* 2020;91(4-S):31-35.

Figure 1: Illustration of the radiological measures used for anteroposterior (A,C) and lateral (B,D) radiographs of the ankle.

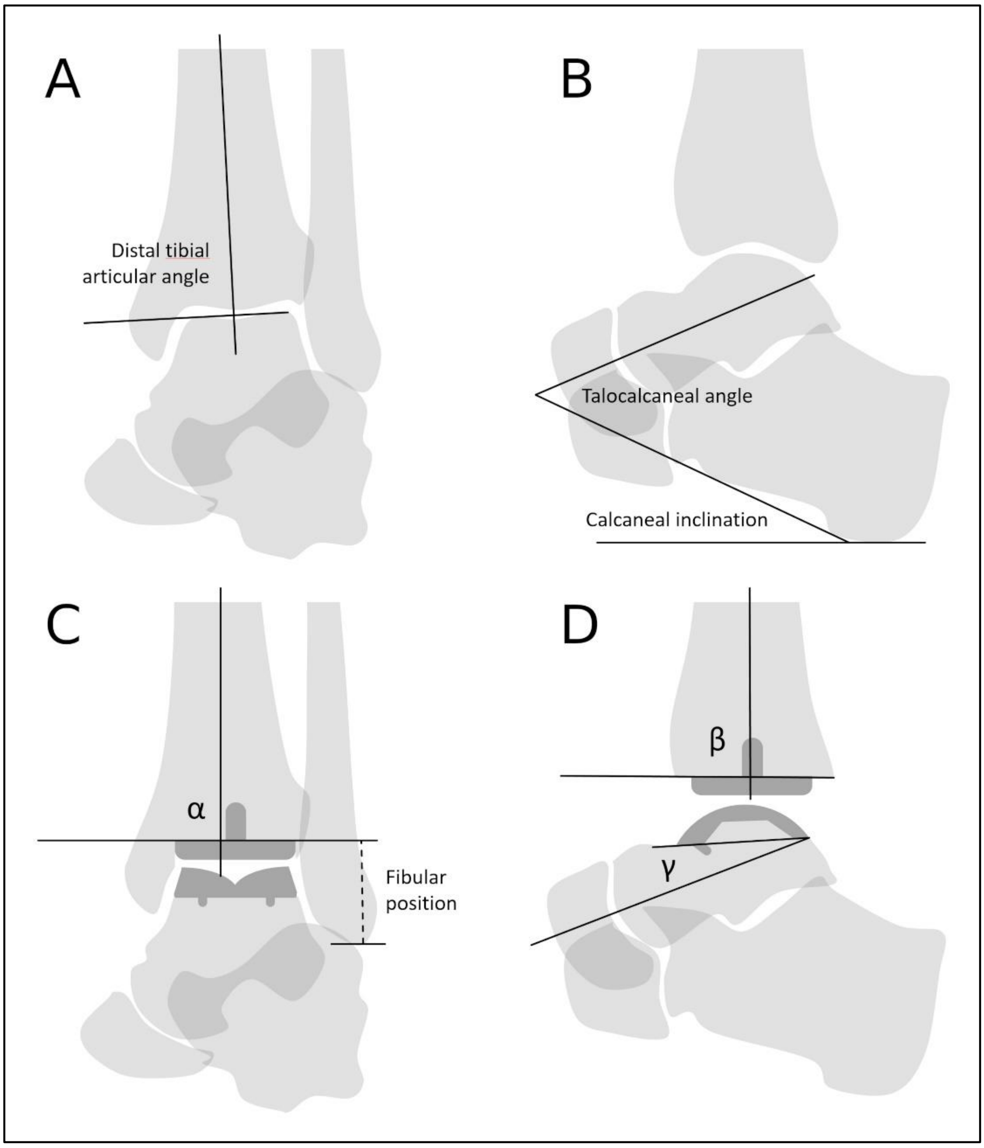


Figure 2: OARS participant flow-chart illustrating recruitment, attrition and numbers included in the final analysis.

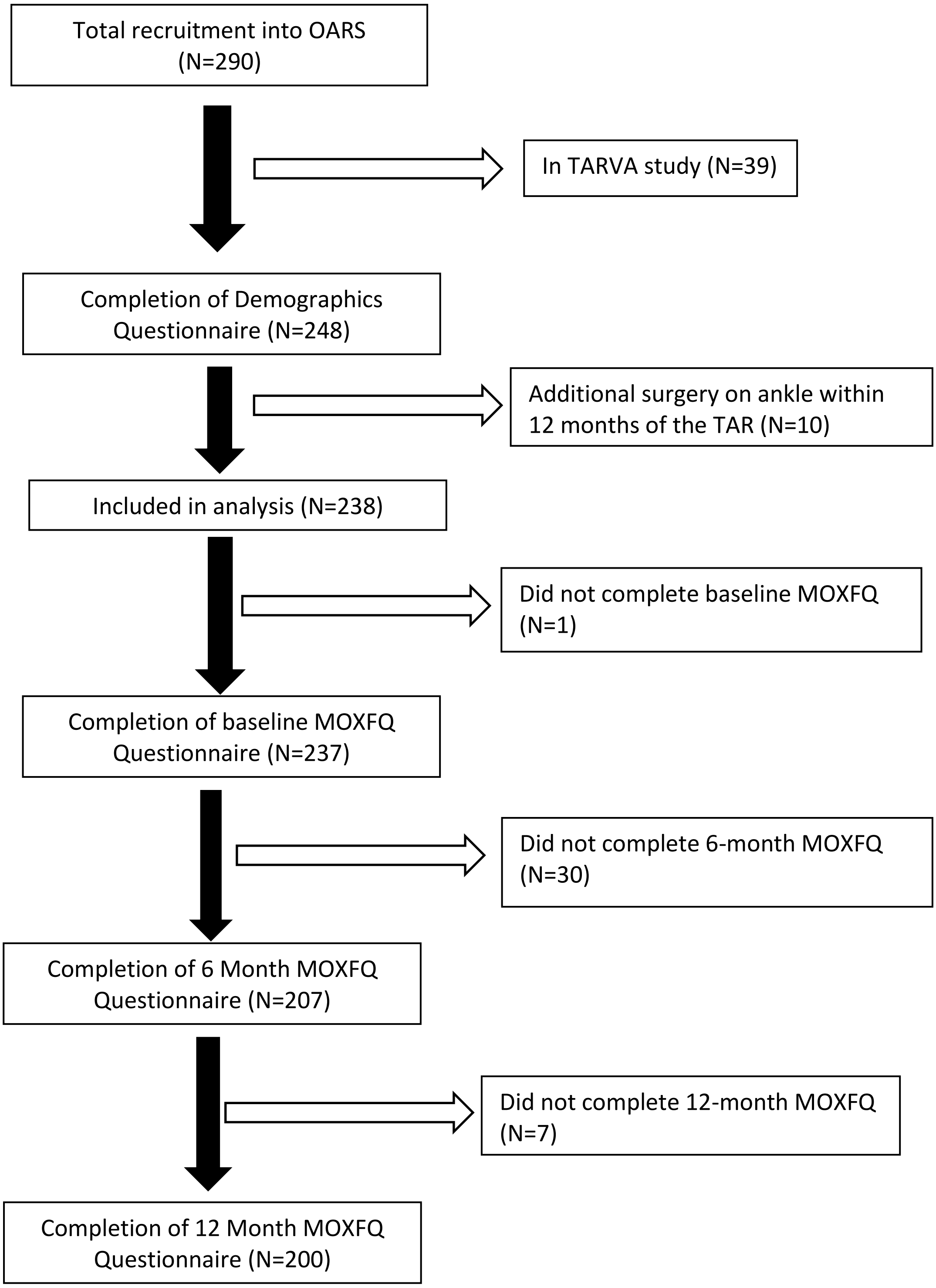


Figure 3: Preoperative and postoperative (six and 12 month) MOXFQ (bars represent mean value with SE)

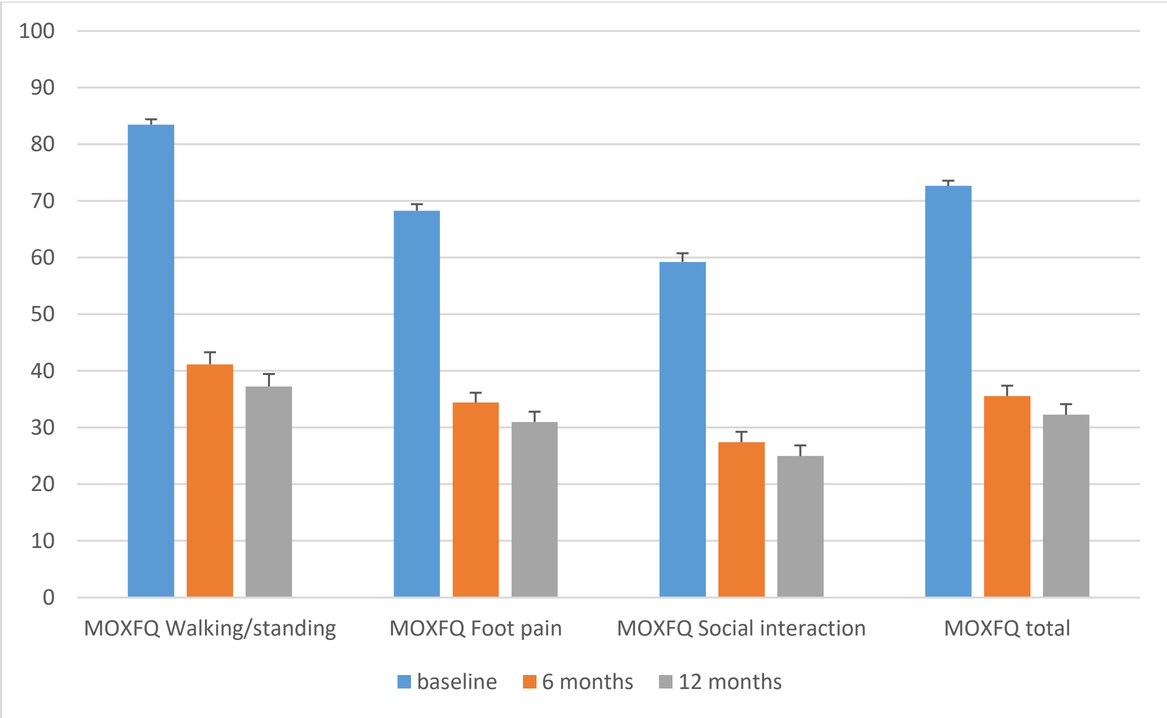
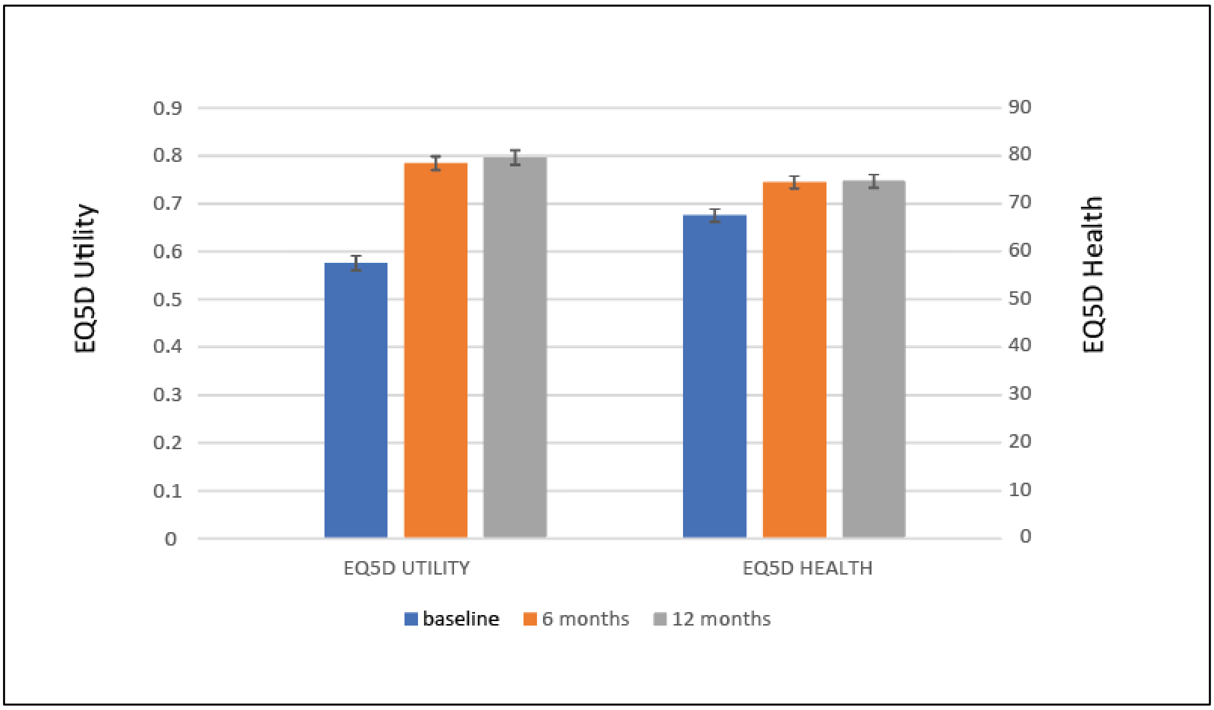


Figure 4: Preoperative and postoperative (six and 12 month) EQ-5D (bars and data points represent mean value with SE).



Note: Effect sizes (ES) were as follows: EQ5D Utility (ES=1.0 at 6 months; 1.1 at 12 months); EQ5D Health (ES=0.3 at 6 month

Table 1: Total ankle replacement patient characteristics at baseline

Variable	Value
TAR patients, n	238
Age, years [mean (SD)]	65.2 (11.1)
BMI, kg/m^2 [mean (SD)]	29.6 (4.8)
Problem length, months [median (IQR)]	58 (24, 120)
Gender	No missing
Male	137 (58%)
Female	101 (42%)
Education	6 missing (3%)
No qualifications	60 (25%)
Secondary school	90 (38%)
Higher education	82 (34%)
Housing	1 missing (<1%)
Home owner	191 (80%)
Privately renting	17 (7%)
Council renting	19 (8%)
Other	10 (4%)
Live alone?	19 missing (8%)
Yes	58 (24%)
No	161 (68%)
Employment	No missing
Full time	57 (24%)
Part time	29 (12%)
Homemaker/caregiver	3 (1%)
Retired	132 (55%)
Student	0 (0%)
Unemployed	17 (7%)
Ankle to be operated on	No missing
Left	116 (49%)
Right	122 (51%)
Problem with other ankle?	No missing
Yes	70 (29%)
No	168 (71%)
Primary ASA	25 missing (11%)
Fit and healthy	36 (15%)
Mild disease not incapacitating	142 (60%)
Incapacitating systemic disease	35 (15%)
Indications for Implantation	25 missing (11%)
Osteoarthritis	185 (78%)
RA	17 (7%)
Other inflammatory arthritis	5 (2%)
Other	6 (3%)
Index ankle previous fracture	31 missing (13%)
Yes	74 (31%)
No	133 (56%)
Tibia Hindfoot Alignment	39 missing (16%)
5-15º Varus	63 (26%)
5-15º Valgus	26 (11%)

Physiological Neutral	87 (37%)
16-30º Valgus	7 (3%)
16 - 30º Varus	16 (7%)
Associated procedure on bone at time of surgery	25 missing (11%)
Yes	17 (7%)
No	196 (82%)
Associated procedure on soft tissue at time of surgery	25 missing (11%)
Yes	41 (17%)
No	172 (72%)
Joint replacement bearing type	31 missing (13%)
Fixed	139 (58%)
Mobile	68 (29%)
Radiological Variable [mean; SD]	28-40% missing
Minimal Joint Space, mm [median (IQR)]	0.0 (0.0, 0.8)
Distal tibial articular angle	89.7 (8.8)
Talo calcaneal angle	44.3 (7.4)
Calcaneal inclination	21.7 (8.4)
KL score [n (%)]	No missing
1	2 (1%)
2	14 (6%)
3	36 (15%)
4	186 (78%)

ASA-American Anaesthesiological Society; BMI – Body Mass Index; IQR – Inter-Quartile Range; KL - Kellgren Lawrence; SD - Standard Deviation; TAR – Total Ankle Replacement

Table 2: Pre-operative variables: change from baseline at six months and 12 months for the MOXFQ Total Score

	Baseline MOXFQ Total		**6 Months – Baseline MOXFQ Total**		**12 Months – Baseline MOXFQ Total**	
	n	*Mean (95% CI)*	*n*	*Mean (95% CI)*	*n*	*Mean (95% CI)*
All patients	237	72.7 (70.9, 74.5)	207	-37.2 (-40.8, -33.6)	200	-40.4 (-44.1, -36.8)
Age						
<=65 years	114	74.8 (72.4, 77.2)	97	-33.3 (-38.3,-28.2)	93	-36.6 (-42.0, -31.2)
>65 years	123	70.7 (68.1, 73.3)	110	-40.7 (-45.6,-35.7)	107	-43.8 (-48.7,-38.9)
Gender						
Male	137	70.7 (68.2, 73.1)	117	-37 (-41.7,-32.3)	111	-40.3 (-45.0, -35.7)
Female	100	75.4 (72.8, 78)	90	-37.4 (-43,-31.9)	89	-40.5 (-46.4,-34.7)
BMI						
<=30kg/m^2	178	71.5 (69.5, 73.5)	156	-36.3 (-40.4,-32.2)	151	-38.8 (-43.0, -34.6)
>30kg/m^2	59	76.1 (72.6, 79.6)	51	-39.9 (-47.1,-32.7)	49	-45.3 (-52.7,-37.9)
Problem length						
<=5 years	116	70.3 (67.6, 73.1)	105	-39.9 (-44.6,-35.2)	101	-43.5 (-48.4, -38.7)
>5 years	119	75 (72.7, 77.4)	101	-34.7 (-40.1,-29.3)	98	-37.5 (-42.9,-32.1)
Diagnosis						
Osteoarthritis	184	72.3 (70.2, 74.4)	170	-38 (-41.9,-34.1)	165	-41.7 (-45.8, -37.7)
Other	28	76.4 (71.9, 80.9)	24	-37.8 (-47.8,-27.8)	23	-36.1 (-46.5,-25.7)
Bearing						
Fixed	138	72.7 (70.3, 75.1)	130	-38.1 (-42.3,-33.8)	125	-42.8 (-47.2, -38.5)
Mobile	68	72.6 (69.2, 76)	59	-35.4 (-42.8,-28.1)	58	-36.1 (-43.4,-28.8)
KL score						
1 to 3	75	74 (70.9, 77)	65	-30.5 (-36.6,-24.5)	63	-38.8 (-45.3, -32.4)
4	162	72 (69.8, 74.2)	142	-40.2 (-44.6,-35.9)	137	-41.1 (-45.6,-36.7)
Tibia Hindfoot alignment						
Neutral	87	73.7 (70.8, 76.6)	83	-36.6 (-42.3,-30.9)	81	-41.5 (-47.4,-35.6)
Varus	79	71.7 (68.5, 74.9)	70	-36.7 (-42.9,-30.5)	67	-38.8 (-45.5,-32.1)
Valgus	33	75.1 (70.9, 79.3)	32	-46.2 (-54.6,-37.7)	31	-45 (-53.6,-36.4)

BMI – body mass index; CI – confidence intervals; Kg/m^2 – kilograms per meter squared; KL – Kellgren Lawrence; MOXFQ – Manchester-Oxford Foot and Ankle Questionnaire; n – number of participants

Table 3: Pre-operative variables: change from baseline at six months and 12 months for the EQ-5D-5L utility score

	Baseline EQ-5D-5L Utility		6 Months – Baseline EQ-5D-5L Utility		12 Months – Baseline EQ-5D-5L Utility	
	n	*Mean (95% CI)*	*n*	*Mean (95% CI)*	*n*	*Mean (95% CI)*
All patients	193	0.58 (0.55, 0.60)	139	0.21 (0.16, 0.25)	139	0.23 (0.19, 0.28)
Age						
<=65 years	92	0.55 (0.51, 0.6)	72	0.22 (0.16, 0.27)	72	0.22 (0.16, 0.28)
>65 years	101	0.59 (0.56, 0.63)	67	0.2 (0.14, 0.25)	67	0.24 (0.19, 0.3)
Gender						
Male	108	0.57 (0.53, 0.61)	84	0.21 (0.16, 0.26)	84	0.23 (0.18, 0.29)
Female	85	0.58 (0.53, 0.62)	55	0.21 (0.14, 0.27)	55	0.23 (0.16, 0.3)
BMI						
<=30kg/m^2	148	0.58 (0.55, 0.62)	105	0.18 (0.13, 0.22)	105	0.22 (0.17, 0.26)
>30kg/m^2	45	0.55 (0.49, 0.6)	34	0.31 (0.22, 0.39)	34	0.27 (0.18, 0.36)
Problem length						
<=5 years	98	0.59 (0.55, 0.63)	67	0.22 (0.16, 0.28)	67	0.25 (0.19, 0.31)
>5 years	94	0.56 (0.52, 0.6)	72	0.2 (0.14, 0.25)	72	0.21 (0.16, 0.27)
Diagnosis						
Osteoarthritis	159	0.58 (0.55, 0.61)	118	0.21 (0.17, 0.26)	118	0.23 (0.19, 0.28)
Other	23	0.5 (0.4, 0.6)	14	0.23 (0.1, 0.36)	14	0.26 (0.12, 0.4)
Bearing						
Fixed	120	0.57 (0.53, 0.61)	82	0.21 (0.16, 0.26)	82	0.23 (0.18, 0.28)
Mobile	58	0.58 (0.53, 0.63)	46	0.2 (0.12, 0.28)	46	0.23 (0.15, 0.31)
KL score						
1 to 3	62	0.57 (0.52, 0.62)	48	0.24 (0.18, 0.29)	48	0.25 (0.18, 0.33)
4	131	0.58 (0.54, 0.61)	91	0.19 (0.14, 0.24)	91	0.22 (0.17, 0.27)
Tibia Hindfoot alignment						
Neutral	79	0.56 (0.51, 0.6)	59	0.2 (0.14, 0.25)	59	0.2 (0.14, 0.26)
Varus	64	0.59 (0.54, 0.64)	44	0.22 (0.14, 0.3)	44	0.25 (0.17, 0.32)
Valgus	30	0.52 (0.45, 0.59)	23	0.28 (0.19, 0.38)	23	0.35 (0.25, 0.45)

BMI – body mass index; CI – confidence intervals; Kg/m^2 – kilograms per meter squared; KL – Kellgren Lawrence; n – number of participants

Table 4: Post-operative variables: change from baseline at six months and 12 months for the MOXFQ Total Score

	Baseline MOXFQ Total		**6 Months – Baseline MOXFQ Total**		**12 Months – Baseline MOXFQ Total**	
	n	*Mean (95% CI)*	*n*	*Mean (95% CI)*	*n*	*Mean (95% CI)*
Alpha						
<90	120	73.6 (71.4, 75.9)	111	-35.8 (-40.9, -30.7)	107	-37.7 (-44.9, -30.6)
>=90	65	70.3 (66.5, 74.1)	59	-39.2 (-45.7, -32.6)	58	-42.5 (-49.1, -35.9)
Beta						
<90	108	72.4 (69.8, 75.1)	101	-36.6 (-42.3, -31.0)	98	-39.3 (-44.8, -33.8)
>=90	77	72.5 (69.5, 75.4)	69	-37.4 (-43.0, -31.9)	67	-39.5 (-45.8, -33.2)
Gamma						
<20	79	73.2 (70.3, 76.1)	74	-38.7 (-44.3, -33.1)	73	-42.7 (-48.2, -37.1)
>=20	106	71.9 (69.2, 74.6)	96	-35.6 (-41.3, -29.9)	92	-36.8 (-42.7, -30.9)

CI – confidence intervals; n – number of participants

Table 5: Post-operative variables: change from baseline at six months and 12 months for the EQ-5D-5L utility score

	Baseline EQ-5D-5L Utility		**6 Months – Baseline EQ-5D-5L Utility**		**12 Months – Baseline EQ-5D-5L Utility**	
	n	*Mean (95% CI)*	*n*	*Mean (95% CI)*	*n*	*Mean (95% CI)*
Alpha						
<90	102	0.55 (0.51, 0.58)	102	-0.33 (-0.41 ,-0.25)	102	-0.32 (-0.26 ,-0.38)
>=90	57	0.62 (0.57, 0.67)	57	-0.44 (-0.53 ,-0.34)	57	-0.43 (-0.53 ,-0.33)
Beta						
<90	93	0.56 (0.52, 0.61)	93	-0.37 (-0.45 ,-0.28)	93	-0.36 (-0.44 ,-0.28)
>=90	66	0.58 (0.54, 0.63)	66	-0.37 (-0.46 ,-0.29)	66	-0.36 (-0.45 ,-0.27)
Gamma						
<20	71	0.55 (0.51, 0.60)	71	-0.33 (-0.41 ,-0.24)	71	-0.32 (-0.41 ,-0.22)
>=20	88	0.59 (0.55, 0.63)	88	-0.40 (-0.49 ,-0.32)	88	-0.39 (-0.48 ,-0.31)

CI – confidence intervals; n – number of participants

Table 6: Statistical association (regression coefficient (95% CI) and p-value) of TAR study outcomes with baseline explanatory variables using linear regression models

	MOXFQ*				EQ-5D-5L*	
	Total	**Walking/standing**	**Foot pain**	**Social interaction**	**Utility**	**Health**
Age[a]	0.36 (0.04, 0.67); p=0.028	0.44 (0.05, 0.83); p=0.029	-	-	-	-
Kellgren-Lawrence score=4[bc]	-10.14 (-17.62, -2.66); p=0.009	-12.57 (-21.86, -3.29); p=0.009	-9.39 (-16.94, -1.84); p=0.016	-	-	-

*Relationship of MOXFQ and EQ-5D-5L outcomes with explanatory variables are estimated using linear regression models. Each outcome (e.g. MOXFQ Total) is calculated as the difference between the 6-month and the baseline value. Initially, each model contained the same explanatory baseline variables and was refined using backwards elimination until all the remaining explanatory variables had p<0.05 (see Methods for details). Table shows the regression coefficient (95% CI) and p-value. Blank cells indicate that variable was not significant.

[a] Being older leads to a greater improvement in outcome post-operation
[b]The worse the ankle is damaged (as seen via x-ray), the greater improvement in perceived outcome post-operation
[c]With reference to Kellgren-Lawrence score=1-3

Supplementary Table 1: Total ankle replacement (TAR) patient characteristics at baseline for complete data (all MOXFQ complete) versus missing (either due to incomplete MOXFQ (n=38) or exclusion due to additional surgery (n=10)).

Variable	Missing Data Cohort	Complete Data
TAR patients, n	48	200
Age, years [mean (SD)]	62.5 (11.0)	65.7 (10.9)
BMI, kg/m^2 [mean (SD)]	30.2 (4.4)	29.5 (3.6)
Problem length, months [median (IQR)]	60 (333, 120)	55 (24, 120)
Gender		
Male	33 (69%)	111 (56%)
Female	15 (31%)	89 (45%)
Education		
No qualifications	19 (40%)	46 (23%)
Secondary school	11 (23%)	80 (40%)
Higher education	16 (33%)	70 (35%)
Housing		
Homeowner	34 (71%)	165 (83%)
Privately renting	8 (17%)	10 (5%)
Council renting	4 (8%)	16 (8%)
Other	2 (4%)	8 (4%)
Live alone		
Yes	7 (15%)	51 (26%)
No	35 (73%)	136 (68%)
Employment		
Full time	13 (27%)	48 (24%)
Part time	5 (10%)	25 (13%)
Homemaker/caregiver	0 (0%)	3 (2%)
Retired	25 (52%)	111 (56%)
Student	0 (0%)	0 (0%)
Unemployed	5 (10%)	13 (7%)

BMI – Body Mass Index; CI – confidence intervals; IQR – Interquartile Range; N – number of participants; SD – Standard Deviation; TAR – Total Ankle Replacement

Supplementary Table 2: Summary of frequency of TAR implants across 207 participants.

Components (Brand/Manufacturer)	Bearing Type	Frequency
Infinity – Wright Medical UK Ltd	Fixed	113
Inbone – Wright Medical UK Ltd	Fixed	25
Salto – Wright Medical UK Ltd	Mobile	1
Box – MatOrtho Ltd	Mobile	28
Zenith – Corin	Mobile	22
Star – Stryker	Mobile	8
Hintegra – DT Medtech International	Mobile	9
Mobility - DePuy	Mobile	1