STATISTICAL AND HEALTH ECONOMICS ANALYSIS PLAN

ART: Ankle Recovery Trial

Does early mobilisation after Ankle fracture surgery enhance Recovery? A randomised controlled Trial comparing the use of plaster versus removable support boot.

ISRCTN Number: ISRCTN15497399

Funding Body: NIHR Research for Patient Benefit (PB-PG-0213-30021)

IRAS and HRA Number: 14/SC/1409

Sponsor: Poole Hospital NHS Foundation Trust (Ref: P140218)

Chief Investigator: Mr Bob Sangar, Poole Hospital NHS Foundation Trust

Current stage of SAP: First Draft

Statistical and Health Economic Analysis Plan Final Sign-Off:

	Name	Date	Signature
Author	Peter Thomas	28/11/2018	
Author (economics)	Elsa Marques	03/12/2018	
BUCRU sign-off	Sharon Docherty	28/11/2018	
Chief Investigator	Bob Sangar	14/12/2018	

Amendments:

Amendment Number	Date	Sign-off

1. Study summary

Study title	Does early mobilisation after **A**nkle fracture surgery enhance **R**ecovery? A pragmatic multi-centre randomised controlled **T**rial with qualitative component and health economic analysis comparing the use of plaster versus removable support boot **(ART)**.
Short study title:	The Ankle Fracture Recovery Trial **(ART)**
Study design	This study is designed as a pragmatic multi-centre randomised controlled trial with qualitative component and health economic evaluation comparing plaster cast and support boots as methods of post-operative ankle fracture management.
Study participants	***Inclusion Criteria*** • Received surgery for fixation of unstable ankle fracture • Provision of informed consent to participate ***Exclusion criteria*** • Under 16 year olds (skeletally immature). • Poor skin condition at operation site. • Serious concomitant disease (e.g. stroke, osteoporosis, arthritis). • Diabetic neuropathy/other sensory neuropathy (lack of sensation). • Non-ambulatory prior to injury. • Active leg ulceration. • Patients who are unable to understand the study information or unable to complete the outcome questionnaires. • Surgeon concerned about quality of fixation/integrity of wound. • Fracture requiring further stabilisation in/around the ankle (e.g. syndesmosis). • Open ankle fracture (bone broken through skin). • Participant is a participant in other concurrent interventional research which may over-burden the participant or confound data collection. • Concomitant injuries which will have a confounding effect on rehabilitation in the opinion of the investigator.
Number of participants	204 with complete data. Initially intended to recruit 246 (123 in each treatment group) to allow for incomplete data, later increased to 276.

Follow-up duration	Patients will be followed up in clinic at 4 weeks post-baseline (6 weeks post-operatively) and via questionnaires at 5 weeks post-baseline (7 weeks post-operatively) and 10 weeks post-baseline (12 weeks post-operatively). [Qualitative telephone interviews with up to 20 participants will take place after the 10-week follow up – these are not considered in this analysis plan].
Planned study period	36 months (24 month recruitment period). Recruitment period extended to 36 months in November 2017.
Study aim	To evaluate the relative effectiveness and cost-effectiveness of two methods of post-operative ankle fracture management (plaster versus removable support boot allowing range of movement) and to provide evidence-based recommendations for best care in clinical practice.
Study objectives	1. To determine whether there is a difference in ankle function between the two types of treatment. 2. To determine whether there is a difference in quality of life between the two treatments. 3. To estimate which is the cost-effective treatment option to inform decision-making. [4. To explore patient experiences and the psychological and social impact of their treatment – analysis of this qualitative component not considered in this analysis plan].
Outcome measure data	Primary outcome: The primary outcome measure for this study is the Olerud and Molander ankle score at five weeks after randomisation. Secondary outcomes: • Ankle functional data (range of movement, weight-bearing) • Standardised measure of general quality of life (EQ-5D-5L) • Healing status • Complications • Return to Usual Activities Other data collected; • Baseline characteristics • Healthcare resource use • Adverse events • Mobilisation and adherence to exercise

Interventions	Patients will be allocated in a 1:1 ratio with both groups weight-bearing as tolerated: 1. Plaster below knee i.e. immobilised for four weeks. 2. Removable support boot with range of movement for four weeks.

2. Aims and objectives for the statistical and health economic analyses:

Aim

To evaluate the relative effectiveness and cost-effectiveness of two methods of post-operative ankle fracture management (plaster versus removable support boot with range of movement) and to inform evidence-based care in clinical practice.

Primary objective

1. To determine whether there is a difference in function between the two types of treatment.

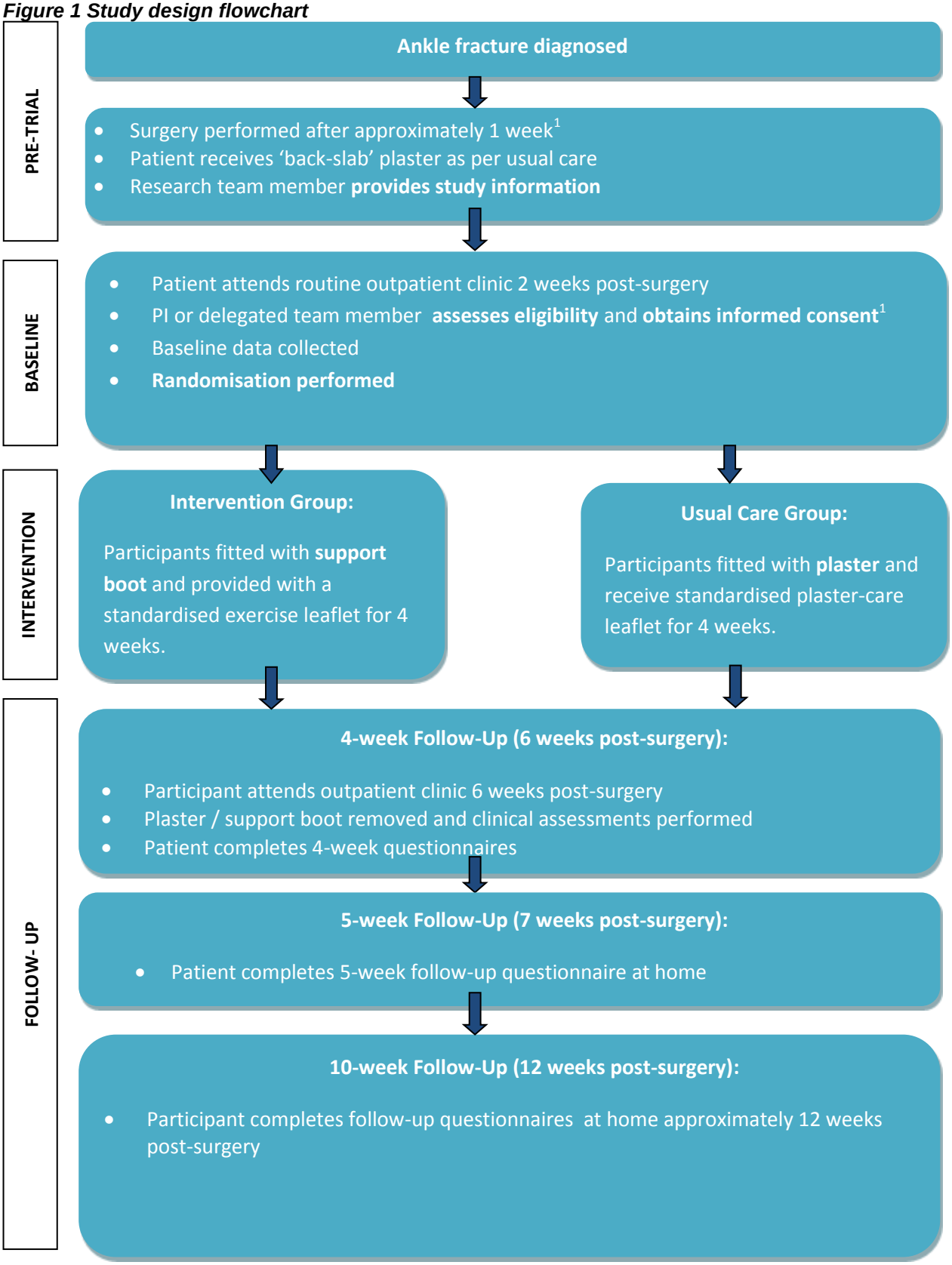
Secondary objectives

1. To determine whether there is a difference in quality of life between the two treatments.
2. To determine whether there is a difference in healing, complications and adverse events between the two treatments.
3. To determine whether there is a difference between return to work, driving and usual activities between the two treatments
4. To determine which is the most cost-effective treatment option if provided in the NHS.
5. [To explore patient experiences and the psychological and social impact of their treatment. These qualitative data are not considered further in the statistical analysis plan].

3. Overall design and analysis:

This study is designed as a pragmatic multi-centre randomised controlled trial with qualitative component and health economic evaluation comparing plaster cast and removable support boots as methods of post-operative ankle fracture management.

Figure 1 Study design flowchart



[1] Patients not requiring surgery, or who are ineligible or unwilling to participate in the trial will be excluded from the trial and managed as per usual care.

4. Participant recruitment:

4.1 Summary of sample size considerations:

The study initially aimed to recruit 246 patients in total (123 in each group) over a 2 year recruitment period, but this was later changed to 276 (to allow for a slightly higher than expected rate of missing/ incomplete data) over 3 years in November 2017 (due to slow recruitment).

The study is powered to detect a 10-point difference in ankle score at 5 weeks post-operatively since patient and public involvement indicated that walking without aids and getting back to work were the most important issues. A change from “support with a stick/crutch” to “no support” and a change to “getting back to normal work/activities of daily life” from “simpler or part-time work” are both associated with a 10-point change on the Olerud and Molander subjective scale (primary outcome).

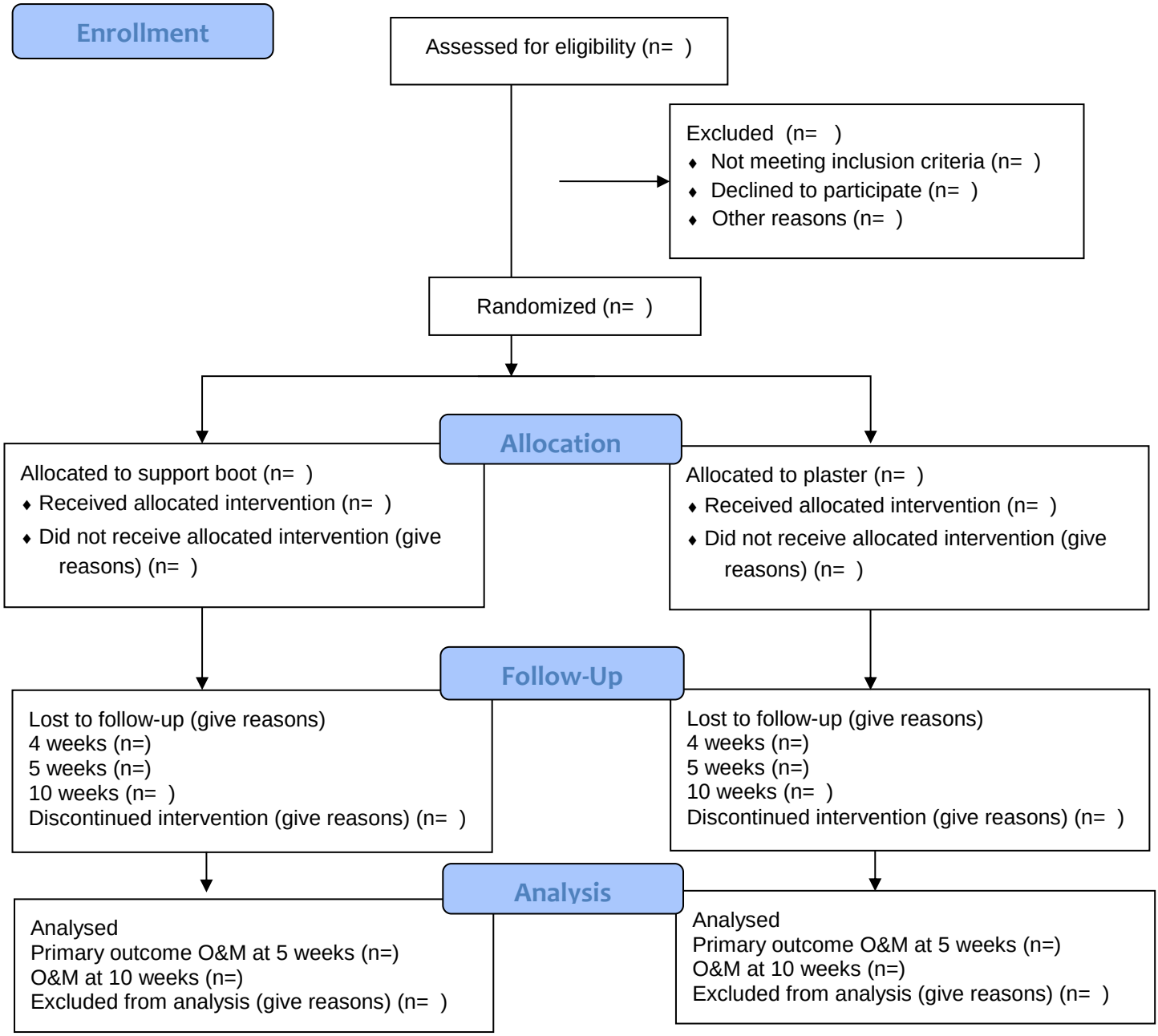
An estimate of standard deviation of the ankle score six weeks post-surgery was derived from an audit of 18 patients having plaster casts at Poole Hospital. This figure of 21.9 is very similar in magnitude to those reported in the Cochrane review. Based on an unpaired t-test with two groups of equal size, assuming 90% power, a two-sided 0.05 significance level, a standard deviation of 21.9 and a mean difference of 10 points between the plaster and boot groups, a total sample size of 204 (102 in each group) is required.

The study initially aimed to over-recruit by 20% to allow for non-responders, or missing data, not to contribute data to the main analysis, giving a recruitment target of 246. Prior to study commencement it was anticipated that the true percentage may be reduced by offering flexible methods of completion; either online, post or by phone. As the study progressed, the percentage with missing primary outcome data was monitored and discussed at regular trial management group meetings and reductions/ increases in over-recruitment were considered so that the target of 204 with completed data would be met. Late in 2017 it became apparent that data on primary outcome was available on 74% and so the recruitment target was increased to 276.

4.2 CONSORT flow chart:

A CONSORT flow chart will be produced showing the flow of recruitment into the RCT (numbers available, approached, eligible, randomised, along with reasons if not approached or not eligible) and through the study (numbers with outcome data, reasons for withdrawing etc.).

Figure 2: CONSORT Flow Diagram



5. Trial data collected

Variable	Purpose	Source	Level of measurement	Recoding	Analysis assumptions
Stratification variables					
Study site	Stratification variable and covariate in main analyses	Screening	Nominal	Poole Basingstoke Portsmouth Torbay Taunton Peterborough Salisbury Yeovil	
Participant descriptives at baseline					
Date of baseline visit	Baseline characteristic	Case Report Form	Date		
Date of Birth	Baseline characteristic	Case Report Form	Date		
Age	Baseline characteristic and covariate in main analyses		Scale	Calculated from DoB and date of randomisation	Normal distribution
Gender	Baseline characteristic	Case Report Form	Nominal	Male Female	
Height (cm)	Baseline characteristic	Case Report Form	Scale	Used only to calculate BMI	
Weight (kg)	Baseline characteristic	Case Report Form	Scale	Used only to calculate BMI	
BMI	Baseline characteristic	Case Report Form	Ratio	Either entered as this or calculated from height/weight^2	Normal distribution

Date of fracture	Baseline characteristic	Case Report Form	Date		
Fracture side	Baseline characteristic	Case Report Form	Nominal	Left Right	
Fracture complexity	Baseline characteristic and covariate in main analyses	Case Report Form	Nominal	Simple Comminuted	
Fracture classification (Weber A, B, C)	Baseline characteristic	X-ray assessment by Chief Investigator	Nominal	A B C	
Medial malleolar involvement (fracture pattern)	Baseline characteristic	Case Report Form	Nominal	No Yes	
Current relationship status (categories)	Baseline characteristic	Questionnaire booklet (Baseline)	Nominal	Single Married Civil partnership Divorced/civil partnership dissolved Widowed/surviving civil partner	
Living alone prior to ankle injury	Baseline characteristic	Questionnaire booklet (Baseline)	Nominal	No Yes	
Qualifications (categories)	Baseline characteristic	Questionnaire booklet (Baseline)	Nominal	None GCSE A/AS level First degree Higher degree Other	

Employment status prior to injury (categories)	Baseline characteristic (and also used to value time off work due to injury for the economic evaluation).	Questionnaire booklet (Baseline)	Nominal	Full time paid Part time paid Retired Volunteer Unemployed Looking after home Full time education Other	
Type of paid employment (only available for those in paid employment) (categories)	Baseline characteristic (and also used to determine inter-sectoral transfers and private losses for economic evaluation).	Questionnaire booklet (Baseline)	Nominal	Self-employed Employed by other Other	
Olerud and Molander adapted for Baseline					
Experience of pain prior to injury (categories)	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal (multiple responses possible – worst response used)	Never Walking on uneven surface Walking on even surface outdoors Walking indoors Constant and severe	Only used to derive total O&M score
Stiffness prior to injury	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal	None Stiffness	Only used to derive total O&M score
Swelling prior to injury (categories)	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal	None Only evenings Constant	Only used to derive total O&M score

Stair climbing prior to injury (categories)	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal	No problems Impaired Impossible	Only used to derive total O&M score
Running prior to injury	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal	Possible Impossible	Only used to derive total O&M score
Jumping prior to injury	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal	Possible Impossible	Only used to derive total O&M score
Squatting prior to injury	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal	Possible Impossible	Only used to derive total O&M score
Support used prior to injury (categories)	Adapted for Baseline descriptive	Questionnaire booklet (Baseline)	Ordered nominal	None Taping/wrapping Stick/crutch	Only used to derive total O&M score
Total O&M Score	Adapted for baseline descriptive		Scale with potential range 0 – 100, low scores indicate worse symptoms	Coding as specified in "Olerud C, Molander H. A scoring scale for symptom evaluation after ankle fracture. Arch Orthop Trauma Surg. 1984;103:190–194." The only deviation is that item 9 is not relevant at baseline, and so was not included on the questionnaire. A score of 20 is assumed for this item for everyone	Normally distributed and interval scaled
EQ-5D-5L derived index	Baseline descriptive	Questionnaire booklet (baseline)		See under "Secondary outcomes" for details of individual items and derivation of single index.	
EQ-5D Health today	Baseline descriptive	Questionnaire booklet (baseline)	Scale	Visual analogue scale ranging from 0 = worst health can imagine to 100 = best health can imagine	Assumed normally distributed and interval scaled
Primary outcome measure					
Olerud and Molander Ankle Scale					

Experience of pain (categories)	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal (multiple responses possible – worst category only coded)	Never Walking on uneven surface Walking on even surface outdoors Walking indoors Constant and severe	Only used to derive total O&M score
Stiffness (categories)	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	None Stiffness	Only used to derive total O&M score
Swelling (categories)	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	None Only evenings Constant	Only used to derive total O&M score
Stair climbing (categories)	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	No problems Impaired Impossible	Only used to derive total O&M score

Running	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	Possible Impossible	Only used to derive total O&M score
Jumping	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	Possible Impossible	Only used to derive total O&M score
Squatting	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	Possible Impossible	Only used to derive total O&M score
Support used (categories)	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	None Taping/wrapping Stick/crutch	Only used to derive total O&M score

Work and activities of daily life	Used to derive primary outcome at week 5, secondary outcome at week 10	Questionnaire booklet (weeks 5 & 10)	Ordered nominal	Same as before injury Loss of tempo Change to simpler or part time job Severely impaired work capacity	Only used to derive total O&M score
Total O&M score	Primary outcome at week 5, secondary outcome at week 10		Scale with potential range 0 – 100, low scores indicate worse symptoms	Coding as specified in "Olerud C, Molander H. A scoring scale for symptom evaluation after ankle fracture. Arch Orthop Trauma Surg. 1984;103:190–194."	Assumed normal distribution and interval scaled
Secondary Outcomes					
Ankle Function					
Dorsiflexion angle non-injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	See row below	Only used to derive difference score
Dorsiflexion angle injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	Difference between uninjured and injured foot derived (Angle in uninjured foot – angle in injured foot)	Difference values are normally distributed
Plantarflexion angle non-injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	See row below	Only used to derive difference score
Plantarflexion angle injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	Difference between uninjured and injured foot derived (Angle in uninjured foot – angle in injured foot)	Difference values are normally distributed

Ankle inversion non-injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	See row below	Only used to derive difference score
Ankle inversion injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	Difference between uninjured and injured foot derived (Angle in uninjured foot – angle in injured foot)	Difference values are normally distributed
Ankle eversion non-injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	See row below	Only used to derive difference score
Ankle eversion injured (degrees)	Secondary outcome week 4	Case Report Form	Scale	Difference between uninjured and injured foot derived (Angle in uninjured foot – angle in injured foot)	Difference values are normally distributed
Circumference ankle, non-injured (cm)	Secondary outcome week 4	Case Report Form	Scale	See row below	Only used to derive difference score
Circumference ankle, injured (cm)	Secondary outcome week 4	Case Report Form	Scale	Difference between uninjured and injured foot derived (circumference in uninjured foot – circumference in injured foot)	Difference values are normally distributed
Circumference calf, non-injured (cm)	Secondary outcome week 4	Case Report Form	Scale	See row below	Only used to derive difference score
Circumference calf, injured (cm)	Secondary outcome week 4	Case Report Form	Scale	Difference between uninjured and injured foot derived (circumference in uninjured foot – circumference in injured foot)	Difference values are normally distributed
Early weight bearing	Secondary outcome week 4	Questionnaire booklet at 4 weeks	Nominal	None Touch weight bearing Partial weight bearing Full weight bearing	

Date of full weight bearing without plaster or boot	Secondary outcome at 10 weeks	Questionnaire booklet at 10 weeks	dd/mm/yy	Derive (a) indicator variable: Full weight bearing Not full weight bearing (b) time to event variable containing (i) number of days between randomisation and full weight bearing for those who are full weight bearing and (ii) number of days between randomisation and 10 week questionnaire completion for those still not weight bearing (censored observations)	
Use of walking aids	Secondary outcomes at 4 and 10 weeks	Questionnaire booklet at 4 and 10 weeks	Nominal	No walking aids Use of one or more walking aids (any of sticks, crutches, frame, rollator)	
EQ-5D-5L					
Mobility today	Used to derive quality adjusted life years (QALYs) at 12 weeks	Questionnaire booklet (Baseline, weeks 4, 5 & 10)	Ordered nominal	No problems Slight problems Moderate problems Severe problems Unable to	Only used to create index value
Self-care today	Used to derive QALYs at 12 weeks	Questionnaire booklet (Baseline, weeks 4, 5 & 10)	Ordered nominal	No problems Slight problems Moderate problems Severe problems Unable to	Only used to create index value
Usual activities today	Used to derive QALYs at 12 weeks	Questionnaire booklet (Baseline, weeks 4, 5 & 10)	Ordered nominal	No problems Slight problems Moderate problems Severe problems Unable to	Only used to create index value
Pain or discomfort today	Used to derive QALYs at 12 weeks	Questionnaire booklet (Baseline, weeks 4, 5 & 10)	Ordered nominal	None Slight Moderate Severe Extreme	Only used to create index value

Anxiety or depression today	Used to derive QALYs at 12 weeks	Questionnaire booklet (Baseline, weeks 4, 5 & 10)	Ordered nominal	Not Slightly Moderately Severely Extremely	Only used to create index value
EQ-5D-5L derived index	Used to derive QALYs at 12 weeks			Singe index of health mapped from 3L value set using Van Hout algorithm as per NICE statement Van Hout B, Janssen M, Feng Y et al. (2012) Interim scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health, 15: 708-15. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf Potential values up to 1 with high scores indicating better health	Assumed to be normally distributed and interval scaled
Health today	Secondary outcome weeks 4, 5 & 10	Questionnaire booklet (Baseline, weeks 4, 5 & 10)	Scale	Visual analogue scale ranging from 0 = worst health can imagine to 100 = best health can imagine	
Healing status					
Evidence of healing	Secondary outcome at 4 weeks	X-ray assessed by Chief Investigator at 4 weeks	Nominal	No Yes	
Fracture displacement	Secondary outcome at 4 weeks	X-ray assessed by Chief Investigator at 4 weeks	Nominal	Undisplaced Displaced	
Complications					

Pulmonary embolism	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No embolism in 10 weeks Yes, embolism in 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Wound breakdown	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No breakdown in 10 weeks Yes, breakdown in 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Compartment syndrome	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No syndrome in 10 weeks Yes, syndrome in 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Blisters	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No blister in 10 weeks Yes, minor blister only in 10 weeks Yes, major blister in 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important

DVT	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No DVT in 10 weeks Yes, DVT (below or above knee or not investigated) in 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Nerve injury	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No nerve injury in 10 weeks Yes, nerve injury (sensory, motor or both) in 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Pressure sore	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No pressure sore in 10 weeks Yes pressure sore grade 1 Yes pressure sore grade 2-4	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Wound infection	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data: No infection in 10 weeks Yes, minor infection in 10 weeks Yes, major infection in 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important

Other complications	Secondary outcome at 4 weeks and 10 weeks	Case Report Form	Nominal	4 and 10 week data combined. Include cases with valid 4 week data for all complications listed below: No within 10 weeks Yes, within 10 weeks	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Overall complications	Secondary outcome at 4 weeks and 10 weeks		Nominal	4 and 10 week data combined: No, none of the complications listed below Yes, one or more of (a) pulmonary embolism, (b) wound breakdown, (c) compartment syndrome, (d) blisters (major), (e) DVT (any), (f) nerve injury (any), (g) pressure sore (grade 2 or more), (h) wound infection (major)	Assumes most complications occur within 4 weeks and missing 10 week data is minimally important
Return to usual activities					
Returned to work (even if only partially)	Used to derive secondary outcome measure at 4 and 10 weeks	Questionnaire booklet at 4 and 10 weeks	Nominal	Not in paid employment prior to injury Not returned Returned fully or partially	Used to derive secondary outcome and economic outcome (see below)

Current work situation compared to pre-injury (only completed by those in paid employment pre-injury)	Used to derive secondary outcome measure at 4 and 10 weeks	Questionnaire booklet at 4 and 10 weeks	Nominal	Same duties, same hours. Same duties, reduced hours. Modified duties, same hours. Modified duties, reduced hours.	Only used to derive secondary outcome and economic outcome (see below)
Summary of paid employment situation	Secondary outcome measure at 4 and 10 weeks	Derived from questionnaire booklet at 4 and 10 weeks	Nominal	Uses information from 4 and 10 weeks combined (needs both in order to be included): Not in paid employment prior. Back at work doing same duties and same hours at 4 weeks Back at work doing same duties and same hours at 10 weeks (but not at 4 weeks) Not back at work at 10 weeks, or doing reduced hours, or doing modified duties.	
Date returned to work (available only for those in paid employment)	Secondary outcome measure at 4 and 10 weeks	Questionnaire booklet at 4 and 10 weeks	dd/mm/yy	Derive (a) indicator variable: Returned to work No returned to work. (b) time to event variable containing (i) number of days between randomisation and return to work for those who have returned to work and (ii) number of days between randomisation and 10 week questionnaire completion for those still off work (censored observations)	Used to derive lost productivity

Absenteeism (available only for those in paid employment pre-injury)	Secondary outcome measure at 4 and 10 weeks	Questionnaire booklet at 4 and 10 weeks	Ratio	Question could be answered in hours or days. If answered in days, convert to hours by multiplying by 7. Combine number of hours from 4 and 10 week data.	1 working day = 7 hours. Assume data normally distributed. Used to derive lost productivity
Presenteeism (available only on those in paid employment and who have returned to work)	Secondary outcome measure at 4 and 10 weeks	Questionnaire booklet at 4 and 10 weeks	Scale	11 point scale. 0 = injury had no effect on work, 10 = injury prevented working.	Assumed normal distribution and interval scaled. Used to derive lost productivity
Impact on everyday activities	Secondary outcome measure at 4 and 10 weeks	Questionnaire booklet at 4 and 10 weeks	Scale	11 point scale. 0 = injury had no effect on daily activities, 10 = injury completely prevented daily activities	Assumed normal distribution and interval scaled. Used to derive lost leisure
Date returned to driving (include only those who were driving prior to the ankle injury)	Secondary outcome at 10 weeks	Questionnaire booklet at 10 weeks	dd/mm/yy	Derive (a) indicator variable: Returned to driving Not returned to driving. (b) time to event variable containing (i) number of days between randomisation and return to driving for those who have returned to driving and (ii) number of days between randomisation and 10 week questionnaire completion for those still not driving (censored observations)	
Resource Use (see section on economic evaluation)					

Serious adverse events	**As reported to PenCTU**		Count	Defined as any untoward and unexpected medical occurrence that: Results in death Is life-threatening Requires hospitalisation or prolongation of existing inpatients´ hospitalisation, Results in persistent or significant disability or incapacity Is a congenital anomaly or birth defect Coded as: None One or more	
Mobilisation and adherence to exercise					
Adherence to physiotherapy exercises	Adherence to intervention	Questionnaire booklet at 4 weeks	Nominal	3+ times per day 1-2 times per day <1 times per day	
Reasons for not adhering to exercise (only available for those not completing exercises 3+ times per day)	Adherence to intervention	Questionnaire booklet at 4 weeks	Nominal	Multiple responses possible: Pain (yes) Swelling (yes) No time (yes) Didn't understand (yes) Forgot (yes) Other (yes)	

6. Missing data:

Outcome and resource use data will be sought for all randomised participants even if they weren't given or didn't use the boot/ plaster. All analyses will be performed blinded to trial allocation. Variables will be tabulated or plotted to identify implausible values and missing items. For missing items and implausible values, the original data will be checked at the source. A decision will be made, informed by clinical opinion, on whether to drop implausible values (i.e. consider values missing).

No imputation methods will be used for the main analysis of the clinical outcomes (though see section on sensitivity analysis). We will assume that the missing data mechanism is "Missing Completely at Random" (MCAR). Further consideration of this assumption will be made in a set of additional analysis (see section 9.6).

For the economic outcomes, we anticipate that there will be a larger quantity of missing data given the nature of patient reported cost data collection. There is often a relationship between cost categories and clinical outcomes, and these data may not be missing completely at random, but dependant on baseline characteristics and observed clinical outcomes (e.g. patients with poorer health states may have used more or fewer health care resources for the duration of the trial). We will assume that the economic data is "Missing at Random" (MAR), where we have observed all the other outcomes and characteristics associated with the missing values, and the primary economic analysis will include complete sets with imputed data.

We will be compiling all sources of resource use data collected to derive cost categories. This will include patient reported questionnaires at 4 and 10 weeks and the review of hospital medical records at the end of the study. We will use answers to signalling questions (for example, if a patient reported he has used a resource but failed to quantify it) to define whether the cost category is missing or not. When a self-reported resource use question is partially completed and has few missing values, we will consider the use of simple imputation methods when appropriate (e.g. weighted average for the trial arm for patients of similar characteristics with same completed data pattern).

Missing data will be imputed using multiple imputation methods with chained equations, in a two-stage ordinary least squares model, adjusting for socio-demographic characteristics collected at baseline. We will jointly impute all cost categories and quality of life utility scores and derive a minimum of 20 imputation sets. If the imputation model allows, we will also jointly impute the primary outcome data and provide imputation estimates for the additional analyses in section 9.6.

7. Interim analysis:

No interim analyses are planned.

8. Blinding:

The statistical analysis of clinical outcomes will be conducted by the trial statistician/ data analyst blind to treatment arm. The results of the statistical analysis will be presented to the rest of the trial team blinded to treatment arm. Once the interpretation of the results has been agreed within the trial team then the treatment arms will be un-blinded to the whole trial team by PenCTU.

9. Main analysis of clinical and cost-effectiveness outcomes:

Participants will be analysed in the group they were randomised to, and (with the consent of participants) we will attempt to collect complete data on everyone and use those data in the analyses.

Baseline descriptive data on demographics (age, gender, current relationship status, whether living alone, highest educational level, employment status prior to injury, type of employment), fracture characteristics (fracture side, classification, pattern, complexity), BMI, Olerud and Molander ankle score pre-injury, and EQ-5D (derived index and health thermometer) will be presented overall and for both groups separately. This will help with (a) assessment of external validity of the trial, and (b) to see whether the 2 groups were comparable at baseline (no significance tests will be conducted). The Olerud and Molander scale includes a question on activity compared to before injury. This would make this question confusing to complete at baseline, and so for baseline only this question has been dropped and all participants will be given the maximum score of 20 for this question (i.e. the score for work and activities same as before injury).

9.1 Primary clinical outcome

The primary clinical outcome is the Olerud and Molander ankle score (measured out of 100) five weeks post-baseline. Questionnaires that were not completed in the 3-10 day time window will be treated as missing data.

Multiple regression including study site, age and fracture complexity as “fixed effect” factors will be used to compare mean ankle score at five weeks post-baseline between the plaster and boot groups.

Study site is a design (stratification) variable and so included in the statistical model. Recognising that small numbers of participants randomised within a site may adversely affect the model, we will try and ensure that each site has a minimum of 10 participants. Any site for which this doesn’t occur will be combined, for statistical purposes, with the site that is closest in distance.

Instead of implementing stratification of age and fracture complexity at randomisation (which in any case could be problematic if there is a large number of strata with a relatively small sample size), the potential issue of imbalance in age and fracture complexity between the two

arms of the trial will be addressed by adjusting (stratifying) for the covariates of age (as a continuous variable) and fracture complexity in the analysis.

Two groups are proposed regarding fracture complexity: Simple fibular fracture and comminuted fracture. If the fibular fracture has a single fracture line with two discrete fragments only, it comprises a 'simple' fracture. If the fibula has more than one fracture line and therefore more than two discrete fragments that require accurate reduction to achieve stability, it comprises a comminuted fracture. The pattern and complexity of the fracture can influence the choice of hardware used by the surgeon to achieve stability. Simple fractures are fixed with unlocked plates whereas complex fractures can require more sophisticated and expensive locked plates.

The pre-baseline measure of the O&M score is collected retrospectively and so has not been included in the primary analysis. By ten weeks post-baseline, it is anticipated that both groups will be similar on the primary outcome, and so this has not been included in an overall repeated measures type analysis. The five week post-baseline data collection period is the critical period, as this is when it is anticipated that the group who had boots will have a greater level of function.

9.2 Primary economic outcome

The primary economic evaluation will compare the costs and health benefits of the support boot compared to plaster in short-term ankle fracture management from an NHS and Personal Social Services perspective. This evaluation will be a cost-utility evaluation at 10 weeks post-randomisation, 12 weeks after surgery. This analysis will take a health and social care payer perspective in relation to a 10 week QALY difference between arms.

(i) Derivation of Quality Adjusted Life Years (QALYs).

For this trial we have used the 5 level version of the EuroQoL's generic quality of life tool, the EQ-5D-5L.[1] More recently, work undertaken to value 5 level EQ-5D health states could not be reconciled with the valuations for the previous 3 level EQ-5D health states and NICE issued a statement recommending the use of the 3L version. NICE also issued guidance on methodology to derive UK valuations from the 5L version. The profiles from the patients' answers to the EQ-5D-5L will be weighted using the EuroQol's published United Kingdom value set mapped from the 3L version using Van Hout's algorithm[2] to produce a single composite, utility based quality of life score, as per NICE position statement of August 2017 [3].

Quality Adjusted Life Years (QALYs) at 10 weeks will then be created from the four time point utility scores assuming a linear change between the time points and using the area under the curve approach. Multiple regression will be used to investigate differences in QALYs between the two groups. The regression will control for baseline utility, the trial stratification variable (study site) and pre-specified characteristics (age and fracture complexity) [4].

(ii) Identifying, measuring, valuing resources and creating cost categories

Section 7.4 of the protocol identifies the resource use categories we are collecting for this trial at 5 and 10 weeks, and details data collection methods. Resource use will be valued using unit costs for health and social care, Department of Health reference costs or other national routine sources whenever available. Trust finance departments will be contacted for procurement costs on the different brands of boots and plaster materials.

Resource use for the economic result at 10 weeks will be composed by adding secondary care resource use collected from the review of medical notes at 10 weeks and the two resource use questionnaires at 4 weeks and 10 weeks. Cost categories will be aggregated in several categories: Costs related to delivering the interventions (boot and plaster), follow-up secondary care costs; community based resources (including medications); home changes and equipment; and use of social services. Together these categories produce the health and social care payer costs.

(iii) Analysis of costs and outcomes:

Costs and QALYs (mean and confidence intervals) will be estimated using regression analysis, adjusting for site, age, and fracture complexity in addition to baseline utility for QALYs). We will ascertain whether the treatment arm would be dominant, i.e. less costly and more effective, or dominated, i.e. more costly and less effective, with respect to QALYs and the Olerud and Molander scale.

Cost is cumulative and additive: missing data in one cost component at one follow-up time point means that the full cost per patient cannot be computed. The strategy to deal with missing data in cost categories is detailed in section 6. The main economic analysis will report the primary economic result using the completed imputed dataset, with the complete case scenario as the secondary analysis.

QALY gains between arms will be compared to identify whether any treatment arm is dominant (i.e. less costly and more effective). If no arm is dominant, results will be presented as a bootstrapped incremental cost-effectiveness ratio (ICERs) with respect to QALYs gained. Bootstrapped estimates of costs and effects will be plotted in the cost-effectiveness plane. Incremental net monetary benefit statistics (INMB) will be derived using the £20,000 and £30,000 societal willingness to pay thresholds for a QALY. Cost-Effectiveness Acceptability Curves (CEACs) will be plotted to show the probability of the intervention being cost-effective at a range of societal willingness to pay thresholds, thus addressing uncertainty around the adoption decision. If one arm is dominant, ICERs are not meaningful and we will only report INMB statistics and plot the cost-effectiveness planes and CEAC.

9.3 Secondary clinical outcomes

Multiple regression will be used to investigate differences between the two groups in the other continuous outcomes measured at four weeks post-baseline (e.g. EQ-5D health thermometer, degrees of range of movement [difference between uninjured and injured foot in dorsiflexion, plantarflexion, inversion and eversion], circumferential swelling [ankle and calf], 5 weeks post baseline (EQ-5D health thermometer) and 10 weeks (e.g. EQ-5D thermometer, O&M score, absenteeism, presenteeism, impact on daily activities).

The overall summary of complications (consisting of pulmonary embolism, wound breakdown, compartment syndrome, blisters, DVT, nerve injury, pressure sore, wound infection and other complications), healing status (displacement and evidence of post-operative fracture healing), use of walking aids and summary of serious adverse events will be compared between the two trial arms using logistic regression (binomial or multinomial depending on the number of categories), again taking study site, age and fracture complexity into account. To reduce the risk of type 1 error, individual complications will be reported descriptively but will not be compared unless the overall summary is statistically significant.

Survival analysis will be used to investigate the time from randomisation to starting to drive again, time from randomisation to full weight bearing and time from randomisation to returning to work.

Adherence to physiotherapy exercises will be analysed using multinomial logistic regression adjusting for site, age and fracture complexity.

9.4 Secondary economic outcomes

Given this young and active population, it is possible that differences in ankle management will generate differences in return to work and usual activities. If that is the case, a potentially interesting secondary economic result would be to report the productivity losses accruing from treating patients with the boot compared with the cast. Productivity losses will be reported separately to the health and social care payer perspective economic result.

Productivity losses will measure absenteeism and presenteeism in the workplace, informal care from family and friends, and lost leisure time. Absenteeism and presenteeism will be measured using the questions described in page 23 of this analysis plan, which mimic the questions in the validated Work and Productivity and Activity Impairment questionnaire[3]. Absenteeism and presenteeism will be valued using weekly average earnings and the human capital approach due to the short-term follow-up of this trial. Estimations will be adjusted for pre-randomisation work status and return to work. We will consider valuing some productivity losses (e.g. volunteer work) at minimum wage rates. If useful, the productivity losses can be included in a societal perspective result, which would also include private patient expenses reported in the 10-week period.

A further secondary economic analysis would be a cost-effectiveness evaluation at 5 weeks post-randomisation, to coincide with the timing of the primary clinical outcome analysis. This

analysis would take a health and social care payer perspective in relation to the 5-week QALY and the change in the O&M scale at 5 weeks. Depending on the result for the primary economic evaluation, it is possible that the 5-week economic result would be superseded by the 10-week primary result, for the purpose of informing decision-making and commissioning of services. If that is the case, this analysis would be redundant and will not be performed.

9.5 Sub-group analyses

The effectiveness of the intervention may vary across different subgroups (strata) of patients (e.g. age and fracture complexity). Therefore for each outcome supplementary statistical analyses are proposed in which it is tested whether there is a statistical interaction between age and treatment arm and between fracture complexity and treatment arm. It is however acknowledged that statistical power will be reduced for these significance tests and they will be considered to be exploratory. Age will be grouped as: under 65 years of age at the time of fracture and 65 + years.

If relevant, subgroups identified in the primary clinical results will be replicated for the primary economic results.

9.6 Additional analyses:

Clinical outcomes:

Some additional analyses (including sensitivity analyses) on primary outcome will also be conducted:

(a) It is acknowledged that some participants (in either group) will leave the 4 week appointment with a plaster cast or boot, and so will complete the scale whilst still wearing the plaster/ boot. In a supplementary sensitivity analysis these participants will be excluded from the analysis to see if the results change.

(b) Pre-injury Olerud and Molander scale will be included as a covariate.

(c) Whether or not the patient had a medial malleolar fracture will be included as a covariate

(d) Any baseline variables that appear, by chance, to differ between the groups will be added in as covariates.

(e) The main analysis will analyse participants in the group to which they were randomised ("intention to treat" approach). As a supplementary analysis two sets of "per protocol" analyses will also be conducted. In the first of these, patients will be analysed according to the method (boot or plaster) that they actually wore. If patients crossed over from one to the other during the 4 week period they will be excluded from the per protocol analysis. If patients didn't wear the boot/ plaster for the full 4 weeks they will be excluded. In the second analysis participants

will additionally be excluded if they didn't perform the exercises recommended. Thus if they performed the exercises less than once per day they will be excluded

(f) Data on primary outcome completed outside of the 3-10 day window will be included in the analysis

(g) The robustness of the results to missing data will be assessed by repeating the analysis incorporating the imputed values that were derived as part of the economic analysis. The main clinical outcome analysis will assume data is missing completely at random. The economic analyses uses a less stringent missing at random (MAR) assumption as detailed in section 6. The multiple imputation model will include site and fracture complexity, and the baseline value of the O&M score. If possible, these will be included within the imputation model for the cost and utility variables to best utilise all available information to predict missing values. The purpose of this analysis is to see how sensitive the clinical results might be to the missing data assumption

(h) The study will also compare the ten week Olerud and Molander Ankle score between groups to check whether any differences still remain at 10 weeks.

(i) To see whether patients in both groups have returned to pre-injury levels of function we will calculate the mean difference (95% CI) between pre-injury and 10 week values for each group, and the proportion of patients whose score is worse by 10 points or more. These statistics will also be compared between groups.

Supplementary analysis (b), (c) (d) will also be conducted for secondary outcomes.

Sensitivity and additional analysis on the cost-effectiveness outcome:

One way and scenario sensitivity analysis will be used to address uncertainty of costing and methodological assumptions. These may include, for example:

a) Varying the range of procurement costs of the boot.
b) Using different costing assumptions to the added 1 week of self-report use in questionnaires or other resource use valued.
c) Varying multiple imputation model assumptions

As analysis progresses, it is possible that other sources of uncertainty, not yet anticipated, might arise, which would warrant sensitivity analysis.

In sensitivity analysis to the secondary analysis, we can value productivity losses using a different weekly earning rates and potentially the friction cost approach.

Imputation and regression models for costs and outcomes will be rerun for the different sensitivity and additional analyses scenarios. Results will be presented disaggregated by perspective as per primary analysis for clear decision-making information.

Templates for tables of results

Table 1: Descriptive statistics

	Plaster (n=??)	Support boot (n=??)
Site n (%) Poole Basingstoke Portsmouth Torbay Taunton Peterborough Salisbury Yeovil		
Age mean (SD)		
Gender n (%) Male Female		
BMI mean (SD)		
Fracture side n(%) Left Right		
Fracture complexity n(%) Simple Comminuted		
Fracture classification n(%) A B C		
Fracture pattern (Medial malleolar involvement) n(%) No Yes		
Current relationship status n(%) Single Married Civil partnership Divorced/ partnership dissolved Widowed/surviving civil partner		
Living alone prior to injury n(%) No Yes		

Highest education qualification n(%) None GCSE A/AS level First degree Higher degree Other		
Employment status n (%) Full time paid Part time paid Retired Volunteer Unemployed Looking after home Full time education Other		
Type of paid employment n(%) Not in paid employment Self-employed Employed by other Other		

Table 2: Baseline outcome measures

	Plaster (n=??)	Support boot (n=??)
Total score of Olerud and Molander Scale mean (SD) Potential range 0-100, low scores indicate worse symptoms		
EQ-5D-5L derived index mean (SD). Potential range -0.281 to 1, lower scores indicate worse health		
EQ-5D Health Today mean (SD. Potential range 0-100, lower scores indicate worse health		

Table 3: Primary and secondary outcome measure at 5 weeks

		5 week follow-up
Primary		
O&M score Potential range 0-100. Lower values indicate more symptoms	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Secondary		
EQ-5D Health Today Potential range 0-100, lower scores indicate worse health	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	

Table 4: Secondary outcome measures at 4 weeks

		4 week follow-up
Secondary		
Dorseflexion angle in degrees (difference non-injured ankle - injured ankles).	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Plantarflexion angle in degrees (difference non-injured ankle - injured ankles).	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Ankle inversion angle in degrees (difference non-injured ankle - injured ankles).	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Ankle eversion angle in degrees (difference non-injured ankle - injured ankles).	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Ankle circumference in cms (difference non-injured ankle - injured ankles).	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	

Calf circumference in cms (difference non-injured ankle - injured ankles).	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Early weight bearing	Plaster n (%) None Touch weight bearing Partial weight bearing Full weight bearing Support boot n (%) None Touch weight bearing Partial weight bearing Full weight bearing Adjusted odds ratios (95% CI) None Touch weight bearing Partial weight bearing Full weight bearing p-value	Reference
Use of walking aids	Plaster (n(%) using aids) Support boot (n(%) using aids) Adjusted odds ratio (95% CI) p-value	
EQ-5D Health Today Potential range 0-100, lower scores indicate worse health	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	

Evidence of healing	Plaster (n(%) with evidence) Support boot (n(%) with evidence) Adjusted odds ratio (95% CI) p-value	
Fracture displacement	Plaster (n(%) displaced) Support boot (n(%) displaced) Adjusted odds ratio (95% CI) p-value	
Presenteeism Potential range 0-10 with higher scores meaning greater effect on work	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Impact on everyday activities Potential range 0-10 with higher scores meaning greater effect activities	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	

Table 5: Secondary outcome measures at 10 weeks

		10 week follow-up
Secondary		
Days to full weight bearing	Plaster (median) Support boot (median) Adjusted hazard ratio (95% CI) p-value	
Use of walking aids	Plaster (n(%) using aids) Support boot (n(%) using aids) Adjusted odds ratio (95% CI) p-value	
EQ-5D Health Today Potential range 0-100, lower scores indicate worse health	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Overall complications	Plaster (n(%)) Support boot (n(%)) Adjusted odds ratio (95% CI) p-value	
Pulmonary embolism[1]	Plaster (n(%)) Support boot (n(%))	
Wound breakdown[1]	Plaster (n(%)) Support boot (n(%))	

Compartment syndrome[1]	Plaster (n(%)) Support boot (n(%))	
Blisters[1]	Plaster (n(%)) Support boot (n(%))	
DVT[1]	Plaster (n(%)) Support boot (n(%))	
Nerve injury[1]	Plaster (n(%)) Support boot (n(%))	
Pressure sore[1]	Plaster (n(%)) Support boot (n(%))	
Wound infection[1]	Plaster (n(%)) Support boot (n(%))	
Other complications[1]	Plaster (n(%)) Support boot (n(%))	

Paid employment	Plaster n (%) Not in paid employment prior. Back at work doing same duties and same hours at 4 weeks Back at work doing same duties and same hours at 10 weeks (but not at 4 weeks) Not back at work at 10 weeks, or doing reduced hours, or doing modified duties. Support boot n (%) Not in paid employment prior. Back at work doing same duties and same hours at 4 weeks Back at work doing same duties and same hours at 10 weeks (but not at 4 weeks) Not back at work at 10 weeks, or doing reduced hours, or doing modified duties. Adjusted odds ratios (95% CI) Not in paid employment prior. Back at work doing same duties and same hours at 4 weeks Back at work doing same duties and same hours at 10 weeks (but not at 4 weeks) Not back at work at 10 weeks, or doing reduced hours, or doing modified duties. p-value	
Days to return to work	Plaster (median) Support boot (median) Adjusted hazard ratio (95% CI) p-value	

Absenteeism (hours)	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Presenteeism Potential range 0-10 with higher scores meaning greater effect on work	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Impact on everyday activities Potential range 0-10 with higher scores meaning greater effect activities	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	
Days to return to driving (includes only those driving prior to injury, n=??)	Plaster (median) Support boot (median) Adjusted hazard ratio (95% CI) p-value	
O&M score Potential range 0-100. Lower values indicate more symptoms	Plaster (mean(SD)) Support boot (mean(SD)) Adjusted mean difference (95% CI) p-value Standardised effect size	

[1] Logistic regression analysis only carried out if there is a significant difference in overall complications

Table 6: Adherence to exercise

		10 week follow-up
Adherence to physiotherapy exercise	Plaster (n(%)) 3+ times per day 1-2 times per day <1 times per day Support boot (n(%)) 3+ times per day 1-2 times per day <1 times per day Adjusted odds ratio (95% CI) 3+ times per day 1-2 times per day <1 times per day p-value	 - - -

Table 7: Cost consequences table (all available data at 12 weeks)

Note: do not report difference between arms here, nor p-values, nor CIs.

	Plaster				Boot			
Health Outcomes	N (number of patients with data available)			Mean (SD)	N (number of patients with data available)			Mean (SD)
QALY gained from baseline								
[Other potential clinical outcomes of interest]								
Resource category	N (number of patients with data available)	N>0 (number of patients using this resource)	Units of resource use (SD)	Mean cost (SD)	N (number of patients with data available)	N>0 (number of patients using this resource)	Units of resource use (SD)	Mean cost (SD)
Secondary care								
Ankle stabilization treatment cost								
Physiotherapy contacts								
Other OP contacts								
A&E visits								
Hospital admissions								
(…)								
Primary care								
GP contacts								
Nurse contacts								
(…)								
Total NHS cost								
Social care								
Walking aids								
(…)								
Total social care cost								
Productivity losses (reported in secondary analysis only)								
Days off work								
Days lost due to presenteeism								
Informal care days								
(…)								

Table 8: Economic results

All estimates reported are adjusted for trial stratification variables (centre, fracture complexity and age) and other variables deemed important as per clinical and economic analysis plan.

All results would report estimates with complete datasets (using imputed data) in primary analysis, and complete cases in secondary analysis.

Note: do not report p-values

	Difference (Boot-Plaster)		
	N	Mean	(95 % CI)
Health Outcomes			
QALY gain (adjusted, imputed data)			
QALY gain (adjusted, complete data)			
Costs (all estimates adjusted and reporting imputed and complete data results)			
Ankle stabilisation treatment			
Secondary care costs			
Primary care costs			
Social Care costs			
Total NHS+Social care costs			
(…) (Potential secondary analysis i.e. report productivity losses)			
Main Result			
Incremental net monetary benefit using societal WTP £20,000/QALY			
Incremental net monetary benefit using other societal WTP thresholds			
Incremental cost-effectiveness ratio in relation to QALY gain (if appropriate to report – non-dominance)			
Sensitivity Analysis			
Scenario A - Incremental net monetary benefit (£20,000/QALY)			
Scenario B - Incremental net monetary benefit (£20,000/QALY)			
(…)			

Appendix

	Uninjured side Mean (SD)	Injured side Mean (SD)	Difference (uninjured – injured) Mean (SD)
Dorsiflexion angle (degrees)			
Plantarflexion angle (degrees)			
Ankle inversion angle (degrees)			
Ankle eversion angle (degrees)			
Ankle circumference (cms)			
Calf circumference (cms)			

References

1. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011; 20(10): 1727-36.

2. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012; 15(5): 708-15.

3. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf

4. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Econ 2005; 14(5): 487-96.

References

1. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; **20**(10): 1727-36.
2. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012; **15**(5): 708-15.
3. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics* 1993; **4**(5): 353-65.