CRAFFT – Children's Radius Acute Fracture Fixation Trial

A multi-centre prospective randomised non-inferiority trial of surgical reduction versus non-surgical casting for displaced distal radius fractures in children

Statistical Analysis Plan

Version 2.0

Linked to SAP - Data definitions and Tables
Version V2.0 - 05Aug2025

Based on Protocol Version 4.0 – 02Sep2024

Trial registration: ISRCTN 10931294

Oxford Clinical Trials Research Unit (OCTRU)
Centre for Statistics in Medicine (CSM)

CONTENTS

1. INTRODUCTION

This document details the proposed data presentation and analysis for the main paper(s) and final study reports from the ***NIHR HTA funded multicentre randomised controlled non-inferiority trial of surgical reduction versus non-surgical casting for displaced distal radius fractures in children (CRAFFT).*** The results reported in these papers should follow the strategy set out here. Subsequent analyses of a more exploratory nature or of extended study follow-up will not be bound by this strategy unless explicitly stated to be covered, though they are expected to follow the broad principles laid down here. The principles are not intended to curtail exploratory analysis (for example, to decide cut-points for categorisation of continuous variables), nor to prohibit accepted practices (for example, data transformation prior to analysis), but they are intended to establish the rules that will be followed, as closely as possible, when analysing and reporting the trial. This document follows published guidelines regarding the content of statistical analysis plans for clinical trial (Gamble et al).

Suggestions for subsequent analyses by journal editors or referees, will be considered carefully, and carried out as far as possible in line with the principles of this analysis strategy. If reported, the analyses will be marked as post-hoc; the source of the suggestion will be acknowledged, and the reader will be advised to rely primarily on the pre-specified analysis for the interpretation of the results.

Any deviations from the statistical analysis plan will be described and justified in the final report of the trial. The analysis should be carried out by an identified, appropriately qualified and experienced statistician, who should ensure the integrity of the data during their processing. Examples of such procedures include quality control and evaluation procedures.

Integral to this Statistical Analysis Plan (SAP) is the SAP – Data Definitions and Tables document which will include full detailed descriptions of all key outcomes, including their definition, generation and how they will be reported at the end of the study. These two documents should be read in tandem.

1.1 Key personnel

Author(s) (Trial Statistician)

Alex Zimmermann

Approver (Senior Statistician, Chief Investigator)

Daphne Kounali	OCTRU Lead Statistician
Daniel Perry	Chief Investigator

1.2 Changes from previous version of SAP

A summary of key changes from earlier versions of SAP, with particular relevance to protocol changes that have an impact on the design, definition, sample size, data quality/collection and analysis of the outcomes is provided below.

Version number Issue date	Author of this issue	Protocol Version & Issue date	Significant changes from previous version together with reasons
V1.0_22Dec2021	Ruth Knight	Protocol_V2.0_13Aug2021	Not applicable as this is the 1st issue
V2.0_05Aug2025	Alex Zimmermann	Protocol_V4.0_02Sep2024	*Amendment to how PROMIS Upper Extremity will be modelled.*

			Estimands framework introduced into this SAP. *The analysis population used for the primary objective has been changed from PP to ITT, and the per-protocol analysis set has been introduced.*

2. BACKGROUND AND OBJECTIVES

Fractures of the wrist (distal radius and ulna) are the most common fractures in children. For displaced (more serious) distal radius fractures the standard treatment in the UK is operative; however, this may not always be necessary since studies have consistently demonstrated that children's bones remodel without the need for manipulation. This study aims to investigate whether non-surgical casting is non-inferior to surgical fixation for these injuries. Doing so could avoid the risks for children undergoing surgery and reduce the burden on trauma operating theatre capacity.

Objectives	Outcome Measures	Time-point(s)
Primary Objective To determine whether non-surgical casting is non-inferior to surgical reduction, measured using observed differences in the PROMIS Upper Extremity Score at three months post-randomisation.	PROMIS Upper Extremity	3 months
Secondary objectives 1. To quantify and draw inferences from differences in function using the PROMIS Upper Extremity Score between non-surgical casting and surgical reduction during the first year post-randomisation.	PROMIS Upper Extremity	Baseline, 6 weeks, 6 and 12 months
2. To quantify and draw inferences from observed differences in pain scores between non-surgical casting and surgical reduction during the first year post-randomisation.	Wong-Baker Faces Pain Score	Baseline, 6 weeks, 3, 6 and 12 months
3. To quantify and draw inferences from observed differences in quality of life using EQ-5D-Y between the trial treatment groups during the first year post-randomisation.	EQ-5DY	Baseline, 6 weeks, 3, 6 and 12 months
4. To determine the complication rate up to 1-year post-randomisation, including re-fracture, the need for further operative fixation and the absence of radiographic remodelling.	Complications	Removal of the cast (clinical), 6 weeks, 3, 6 and 12 months
5. To estimate the cost-effectiveness of the treatments to the NHS and the broader economy, up to 1-year post-randomisation.	Healthcare resource use	6 weeks, 3, 6 and 12 months
6. To quantify and draw inferences from parental satisfaction with the cosmetic appearance of the arm between non-surgical casting and surgical reduction during the first year post-randomisation.	VAS Cosmesis	6 weeks, 3, 6 and 12 months
7. To quantify and draw inferences from patient satisfaction between non-surgical casting and surgical reduction during the first year post-randomisation.	Satisfaction score	12 months
8. To quantify and draw inferences from school attendance between nonsurgical casting and surgical reduction during the first year post-randomisation.	Bespoke 'School Attendance' questionnaire	6 weeks, 3, 6 and 12 months

9. To determine the impact of injury, treatment and recovery on parent and child experience of daily life and the outcomes that are important to them	Child and parent experiences	3 and 12 months
10. To determine the barriers and facilitators to trial recruitment from parent/child and staff perspectives	Child, parent, and staff experiences	Pilot phase
Long-term outcomes. To be reported separately		
11. To quantify and draw inferences from longer-term pain, function & complications annually up until 3 years post-randomisation.	PROMIS Wong-Baker Faces Pain Score EQ-5DY VAS Cosmesis Complications	Annually (2 and 3 years)

The estimand for the primary objective (including the analysis of the primary outcome) is described in Table 1.

Table 1: Estimand-to-analysis table template

Primary Objective: *To determine whether non-surgical casting is non-inferior to surgical reduction, measured using observed differences in the PROMIS Upper Extremity Score at three months post-randomisation*

Estimand: *The mean difference in PROMIS Upper Extremity Score in children aged between 4 and 10 years old who undergo non-surgical casting or surgical reduction for a recently (within 7 days) broken wrist, 3 months after initiation of treatment, irrespective of any additional surgical intervention received after two weeks of initial treatment and which are part of standard care.*

Treatment:

Intervention: Non-surgical casting

Comparator: Surgical reduction

Estimand	***Analysis***
Target population	***Analysis set***
Children aged between 4 and 10 years old who undergo non-surgical casting or surgical reduction for a recently (within 7 days) broken wrist	*All randomised participants. Participants randomised to the non-surgical casting arm will be in the intervention group. Participants randomised to Surgical reduction will be the control group.* ***Sensitivity analysis:*** *The same analysis approach will be applied to a Per-Protocol analysis set.*
Variable	***Outcome measure***
PROMIS Upper Extremity Score at 3 months post-randomisation	*PROMIS Upper Extremity Score (defined as per PROMIS guidelines) using a computer adaptive testing (CAT) system*

	and proxy-reported by parent/guardian at 3 months post-randomisation.
Handling of intercurrent events	***Handling of missing data***
Treatment switching: The PROMIS Upper Extremity Score of a patient who receives the intervention/comparator/any other treatment instead of, or in addition to, their randomised treatment, within 14 days from randomisation (inclusive), is included regardless of attribution to treatment (treatment policy strategy). *A treatment policy estimand is the primary analysis.* *A secondary estimand will target the complier average treatment effect (CACE) and comparing those receiving non-surgical casting without any surgery with GA* or unconscious/procedural sedation *within 14 days post-randomisation and those controls who would have received non-surgical casting had they been randomised to it.*	*Missing outcome data (missing data for only some occasions across the 12 months post-randomisation for an individual) will be implicitly imputed in a mixed effects model using all data available at all time points This approach assumes that the missing data is missing at random (MAR) conditional on all the other data collected at other time points, and minimisation factors. An additional fully adjusted analysis will also include pre-randomisation covariates associated with both the outcome and missingness.* *Reference-based imputation will be used to address informative drop-out.*
Population-level summary measure	***Analysis approach***
Mean difference between treatments in PROMIS Upper Extremity Score 3 months after receiving treatment.	*Treatment groups will be compared using a multivariable mixed effects linear regression model adjusting for the minimisation factors. Time, recruiting centre, type of fracture (completely vs. incompletely off-ended), fracture location (physeal vs. metaphyseal) and age group (4-6 year olds vs. 7-10 year olds) will be adjusted for in the model.* *The treatment effects will be presented as an adjusted mean difference with a 95% confidence interval and corresponding 5% (2-sided) p-value.* ***Sensitivity analysis:*** *The secondary estimand (CACE) is targeted through an instrumental variable approach using randomisation as an instrument.* *The same analysis approach as the primary estimand will be applied to the Per-Protocol population to compare results with the primary estimand results and CACE results.*

[1]Strategies defined in E9 (R1) include treatment policy, while on treatment, principal stratum and hypothetical

3. STUDY METHODS

3.1 Trial Design/framework

CRAFFT is a multicentre randomised controlled trial using a non-inferiority two-arm parallel group design comparing surgical reduction versus non-surgical casting for severely displaced distal radius fractures in children aged 4-10 years. Participants were randomised in a 1:1 ratio to either surgical reduction or non-surgical casting. Participants were asked to complete further questionnaires on function, pain, quality of life, cosmesis and satisfaction at 6 weeks, 3 months, 6 months, and 12 months after randomisation. Data was collected electronically with email and/or text message prompts. This will be supplemented by telephone interviews as required. After the completion of the main phase of the study, patients will be followed up for an additional two years; since three years is known to be the period over which the bone can continue to change shape (remodel).

The study consists of two phases: Phase 1 (internal pilot) confirmed the expected rate of recruitment and test data collection procedures in a large-scale multi-centre randomised controlled trial. Phase 2 (main RCT) took place in a minimum of 32 UK centres. The trial progressed to the main phase, and internal pilot phase patients will be included in the final analysis, as planned.

3.2 Randomisation and Blinding

Consented participants were randomised to one of two intervention groups (1:1) using a computer randomisation service provided by OCTRU. Randomisation allocation was implemented using a minimisation algorithm with minimisation factors: centre, type of fracture translation (completely off-ended versus incompletely off-ended), fracture location (metaphyseal or physeal) and age group (4-6 years, 7-10 years). The minimisation algorithm was seeded with a number of allocations and a non-deterministic probabilistic element was introduced in order to prevent predictability of the treatment allocation. Participants and their parents cannot be blinded to their treatment. The treating clinician also cannot be blinded to the treatment they are providing.

3.3 Sample Size

674 participants providing data on the PROMIS Upper Extremity Score for children at 3 months post-randomisation (337 in each group) will provide 90% power and 2.5% (1-sided) significance to detect whether non-surgical casting for the treatment of displaced wrist fractures is non-inferior to surgical reduction assuming a non-inferiority margin of -2.5 points, a standard deviation of 10 and no difference between the two groups (PASS 16, 2018). The choice of the non-inferiority margin and the baseline standard deviation have been based on discussions with patients, their parents and the literature validating the PROMIS Upper Extremity Score in a range of different diseases.

Allowing for 10% loss to follow-up, would result in an overall target of 750 patients (375 per group). Given that the primary outcome is at 3-months post-randomisation, it may be necessary to adapt the loss to follow-up inflation based on observed follow-up rates to ensure that the study recruits effectively and efficiently.

As degree of translation has the potential to influence outcome, this has been incorporated as a stratification factor to ensure that it is balanced across the treatment groups, and we will assess for differential outcomes in the important subgroups using treatment-by-subgroup interactions. Based on site audits we expect approximately one third to one quarter of the fractures will be completely off-ended, i.e., 200-250 patients. Two-hundred patients in this subgroup would allow non-inferiority to be concluded with 90% power at the 2.5% (1-sided) significance level using a non-inferiority margin of between -4.5 and -5 points and assuming a standard deviation of 10. Therefore, we plan to continue recruiting until a minimum of 200 patients in the completely off-ended subgroup have been randomised.

3.4 Statistical Interim Analysis, Data Review and Stopping guidelines

The Data and Safety Monitoring Committee (DSMC) is a group of independent experts external to the trial who assess the progress, conduct, participant safety, and, if required, critical endpoints of a clinical trial.

The DSMC will follow the charter as described in the document stored in the TMF. They will not be asked to perform any formal interim analyses of effectiveness. They will, however, review accruing data, summaries of the data presented by treatment group, and will assess the screening algorithm against the eligibility criteria. In the internal pilot phase, they will closely monitor recruitment based on the stop/go criteria. They will also consider emerging evidence from other related trials or research and review related SAEs that have been reported. They may advise the chair of the Trial Steering Committee at any time, if, in their view, the trial should be stopped for ethical reasons, including concerns about participant safety. DSMC meetings will be held at least annually during the recruitment phase of the study.

The internal pilot was designed to have a minimum of 15 centres recruiting over a 9 month period. Stop/go criteria for the internal pilot are as follows:

- Red: recruitment is less than 100 participants
- Amber: recruitment between 100 and 130 participants
- Green: recruitment exceeds 130 participants

3.5 Timing of Analysis

The analysis of outcomes up to 12 months after randomisation will be conducted once all participants have reached this time-point (short-term follow-up). This SAP concerns the methods which will be used to analyse these outcomes.

This trial also includes long-term follow-up (annually at 2 and 3 years post-randomisation). The analysis of these outcomes will be performed and reported separately. An additional analysis plan will be prepared detailing the methods which will be used to analyse the long-term outcomes as an update to this SAP, following the same analyses principles.

3.6 Blinded analysis

A blinded analysis of the data (not separated by treatment arm) will be undertaken prior to the final data lock to investigate the distribution of variables, missing data distributions, and to finalise the per protocol population.

3.7 Statistical Analysis Outline

Standard descriptive statistics will be used to describe the demographics between the treatment groups reporting means and standard deviations or medians and interquartile ranges as appropriate for continuous variables and numbers and percentages for binary and categorical variables. All comparative outcomes will be presented as summary statistics and reported together with 95% confidence intervals, and P-values.

All analyses will be carried out under a treatment policy strategy (that is, all patients will be analysed in the group they were randomised to regardless of actual treatment received).

The PROMIS Upper Extremity Score for children is the primary outcome of the study and will be compared between treatment groups as the dependent variable in a multivariable mixed-effects linear regression model, adjusting for the stratification factors. The treatment difference prescribed in the primary estimand will be the estimated treatment difference at 3 months since randomisation from the multivariable mixed-effects linear regression model.

Subgroup analyses will be undertaken using the same methodology by incorporating a treatment by subgroup variable interaction. Multi-level, mixed effects repeated measures linear regression models will be used to analyse continuous secondary outcomes, if appropriate; otherwise, appropriate non-parametric alternatives will be used. Complications will be reported by type for each treatment group as per randomisation and their received treatment. Counts of patients with complications from initial treatment to 12 months will be descriptively in the ITT and as-treated populations. If a sizeable counts of complications in each arm are observed which are clinically meaningful to aggregate together, then consideration will be given towards using a adjusted mixed effects Poisson regression model.

It is anticipated that all statistical analyses will be undertaken using Stata (Release 17, College Station, TX: StataCorp LLC, www.stata.com) and/or R (Version 4.3.1).

4. STATISTICAL PRINCIPLES

4.1 Statistical Significance and Multiple Testing

There is no correction for multiple testing. The significance level used will be 0.05 and 95% confidence intervals will be reported. All secondary analyses will be considered as supporting the primary analysis and will also be analysed using a significance level of 0.05 with 95% confidence intervals (CI).

No interim analyses of primary and secondary outcomes were carried out and none was requested by the DSMC.

Exploratory outcome will be reported using a significance level of 0.01 with 99% confidence intervals.

4.2 Definition of Analysis Populations

Populations for analysis are defined as follows:

Intent-to-treat (ITT): all participants will be analysed according to the treatment group they were allocated to at randomisation.

Per-protocol (PP): includes participants who received the intervention as intended and they will be analysed according to their randomised treatment. Participants will be excluded from the per-protocol population if:

- They did not receive their randomised treatment as intended:
  - if a patient randomised to the surgery group did not receive surgery with GA or unconscious/procedural sedation within 14 days inclusive since randomisation, or
  - if a patient randomised to the non-surgery group received surgery with GA or unconscious/procedural sedation within 14 days inclusive since randomisation
- They were randomised in error (e.g. patients were not eligible to be randomised or were randomised using inaccurate data for stratification factors)
- Major protocol deviations as assessed during a clinical review by the trial management group prior to final analysis

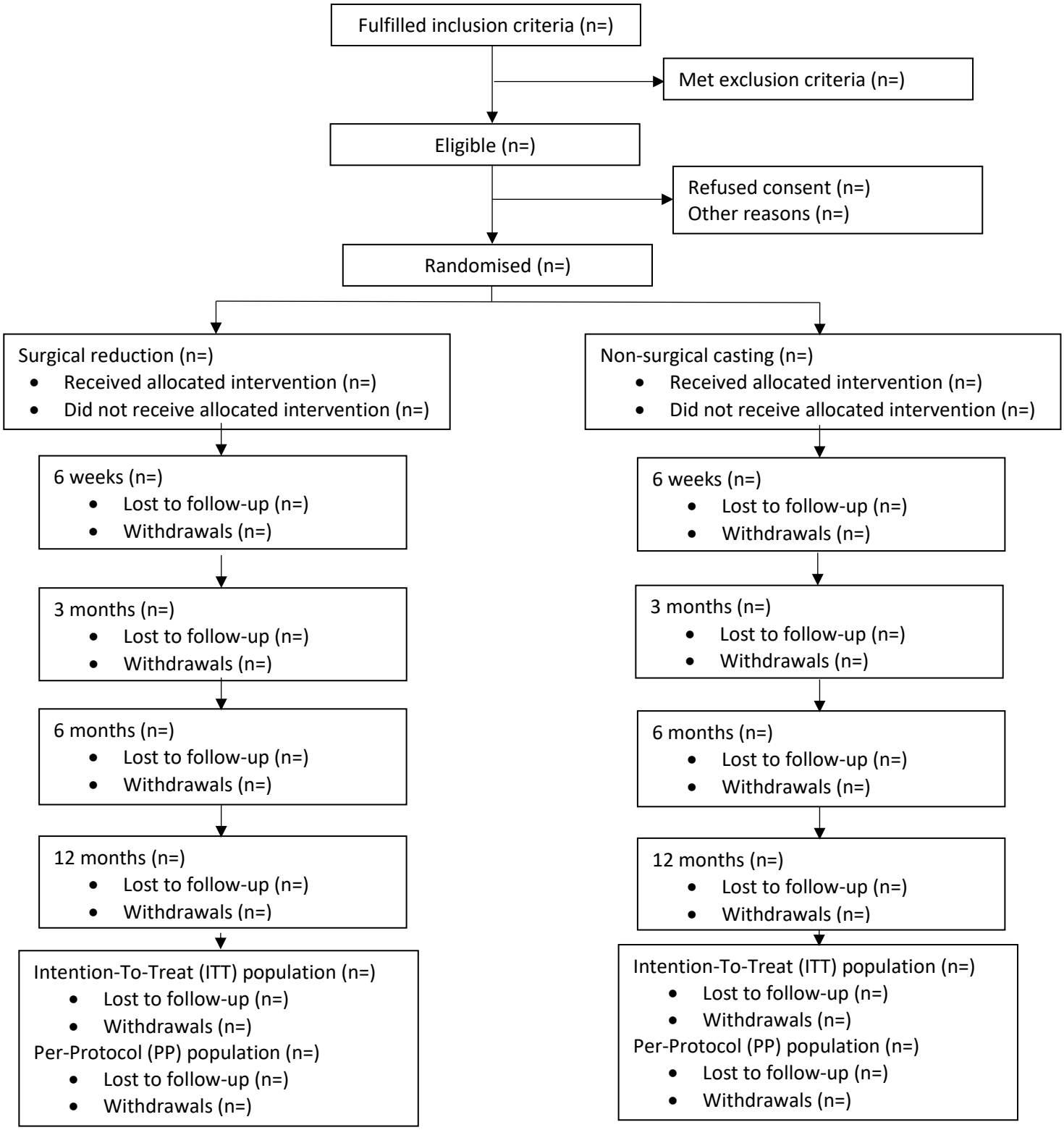
The Chief Investigator will perform a blinded analysis of the protocol deviations listed in REDCap prior to the final data lock and establish whether participants with protocol deviations received their randomised treatment as intended. Participants who did not receive their randomised treatment as intended will be excluded from the per-protocol intervention population.

5. TRIAL POPULATION AND DESCRIPTIVE ANALYSES

5.1 Representativeness of Study Sample and Patient Throughput

A CONSORT flow chart using the CONSORT PRO [5] and detailing the number of participants screened, screening failures, the number of participants randomised to each arm, the number of participants providing primary outcome data and the number of participants who withdrew or were lost to follow-up will be created. A participant lost to follow-up is a participant from whom no primary outcome was received beyond the respective timepoint (inclusive). Participants who withdrew were not considered lost to follow-up.

Figure 1: Example CONSORT diagram



5.2 Withdrawal from treatment and/or follow-up

Withdrawals and losses to follow-up at each time-point will be reported as numbers and percentages for each treatment group (**Table 1**). In addition, total withdrawals, reasons for withdrawal and time to withdrawal will be summarised by treatment group (**Table *2***). If differential losses are identified, the reasons for these will be explored further. No deaths are anticipated in this study; however, if any do occur these will be reported separately along with the reason.

Table 1: Withdrawals and losses to follow-up post-randomisation by allocated treatment

		Surgical reduction (n, %)	Non-surgical casting (n, %)
6 weeks	**Missing data**		
	Withdrawn		
	Lost to follow-up		
3 months	**Missing data**		
	Withdrawn		
	Lost to follow-up		
6 months	**Missing data**		
	Withdrawn		
	Lost to follow-up		
12 months	**Missing data**		
	Withdrawn		
	Lost to follow-up		

Table 2: Summary of reasons for withdrawal by allocated treatment

	Surgical reduction	Non-surgical casting
Total withdrawals (baseline to 12 months)[1]		
Time to withdrawal (days)[2]		
Reasons for withdrawal[1]		
Parent/guardian doesn't like idea of being part of research		
Child doesn't like idea of being part of research		
Parent/guardian doesn't want to complete questionnaires		
Child doesn't want to complete questionnaires		
No reason		
Other		

[1] *Summaries are n (%)*

[2] *Summaries are median (IQR)*

5.3 Baseline Characteristics

Numbers (with percentages) for binary and categorical variables and mean (and standard deviation), or median (with lower and upper quartiles) for continuous variables will be presented; there will be no tests of statistical significance nor confidence intervals for differences between randomised groups on any baseline variable. The tables presented cover minimisation factors, baseline characteristics and injury details, and baseline patient/proxy reported outcome measures by allocated treatment.

5.4 Unblinding

This is not a blinded trial and as such it is not possible for unblinding to occur.

5.5 Treatment Compliance with Details of Intervention

The interventions compared in this trial are:

Non-surgical casting: application of a plaster cast to hold the bone fragments in the optimal possible position without giving medication to deliberately alter the conscious level of the child. Usual practice is for the plaster cast to be used for 4-6 weeks.

Surgical reduction: with or without fixation. The bones will be realigned under general anaesthesia or sedation altering the conscious state of the child. Following surgery, usual practice is for the arm to be immobilised in cast for 4-6 weeks.

Treatment switching is defined as when either of the following are true:

- a patient randomised to the surgery group did not receive GA surgery with sedation/anaesthesia within 14 days inclusive since randomisation, or
- a patient randomised to non-surgery group received surgery with sedation/anaesthesia within 14 days inclusive since randomisation

Details of how the data collected will be used to identify treatment switching is provided in the CRAFFT SAP DD&T. The same document also presents the tables that will be used to: (i) summarise treatment switching; (ii) provide further details of the initial treatment received; and (iii) present details of how long the cast was worn for and any cast changes that occurred.

5.6 Reliability

To ensure consistency, validation checks of the data will be conducted. This will include checking for duplicate records, checking the range of variable values and validating potential outliers where possible (referring back to sites if necessary). As the data is collected electronically, many of these checks will be implemented automatically as part of the data entry procedure and data collection instruments have been validated prior to data entry commencing.

For each variable, missing value codes will be checked for consistency and proportion of missing values per variable will be presented. Patterns of missing data will be explored.

6. ANALYSIS

The statistical methods to be used to compare groups for primary and secondary outcomes and methods for point and interval estimation are summarised below. This includes methods for additional analyses, such as adjusted analyses and subgroup analyses.

6.1 Outcome Definitions

A summary of the outcomes collected and the timepoints at which these are collected are provided below. Full details of the definitions of these outcomes and how they will be generated is provided in the CRAFFT SAP DD&T Section 2.

Outcome	Baseline	6 weeks	3 months	6 months	12 months	2 years	3 years
PROMIS UE Score	x	x	x	x	x	x	x
Wong-Baker FACES	x	x	x	x	x	x	x
EQ-5D utility	x	x	x	x	x	x	x
EQ-5DY VAS	x	x	x	x	x	x	x
Complications	x	x	x	x	x	x	x
VAS Cosmesis		x	x	x	x	x	x
Satisfaction					x		
School Attendance		x	x	x	x		

6.2 Analysis Methods

Primary outcome

PROMIS upper extremity scores for children will be summarised by treatment group by timepoint using raw means and SDs and will also be presented graphically. These summaries will be presented for both the PP and ITT populations. Adjusted treatment differences and associated 95% CIs will be reported and the lower limit of the 95% CIs will be compared against the pre-specified non-inferiority margin (-2.5 points). The primary estimand will be based on the ITT population, however supporting analyses will be undertaken on the PP population.

Treatment groups will be compared using a multivariable multilevel mixed effects linear regression model which adjusts for the minimisation factors. Type of fracture (completely vs. incompletely off-ended), fracture location (physeal vs. metaphyseal) and age will be included as fixed effects and recruiting centre will be included as a random effect. A treatment by time interaction (as a categorical variable) will be included, however constrained to no between group treatment difference at baseline. The random part of the model will include random intercept terms for centres (level 3) and the individual participant (level 2) to account for the dependence of observations of participants within the same centre and observations within the same participant over time. At the individual participant level, it will also include a random coefficient for the effect of time (measurement occasion) and an unstructured covariance structure for the association between the random intercept and occasion. If there are low numbers of participants in some centres, centres with less than 6 participants randomised will be grouped. If after grouping small centres the model still experiences estimating problems, centres will be included as fixed effects or will be removed from the model.

Additionally, treatment effect in the primary outcome will be assessed in a PP population analysis as well as a CACE analysis. We will also assess the potential for differential treatment effect in the subgroup defined by type of fracture (off-ended) using treatment by subgroup (off-ended or not) interactions in both the ITT and PP populations.

A fully adjusted analysis using the ITT population will introduce pre-randomisation covariates associated with the outcome. These covariates will be those from the following list which are found to be associated with the outcome as well as missingness:

- Sex
- Mechanism of injury (high/low energy fall)
- Dominant arm (left/right/unsure)

Secondary outcomes

Continuous secondary outcomes (Wong-Baker FACES, EQ-5D and VAS, and VAS cosmesis) will be summarised by treatment group at each time-point using raw means and SD, as well as being summarised graphically. These outcomes will be analysed using the same multivariable multi-level mixed effects linear regression models as the model for the primary estimand. The UK EuroQol 5 dimensions 3 levels (EQ-5D-3L) value sets will be used for all participants instead of the originally planned EQ-5DY for the EQ-5D outcome.
Examples of how these results will be presented are provided in the CRAFFT SAP DD&T.

For each of type of complication that occurs at least once, the number and proportion of participants experiencing that complication will be summarised by treatment group. Complications up to 8 weeks post-randomisation will be summarised as outlined in the CRAFFT SAP DD&T. Those between 8 weeks and 12 months post-randomisation will be summarised as described in the same document.

Secondary objective 5 are health economics outcome and is outside the scope of this document. Details regarding analysis of these objectives will be contained within the CRAFFT Health Economics and Analysis Plan.

Satisfaction will be summarised descriptively between treatment arms.

School Attendance will be summarised by treatment group over 12 months, and the difference in the mean number of days missed between treatments in the 12 months since randomisation will be estimated using a linear regression model adjusted for the recruiting centre, type of fracture (completely vs. incompletely off-ended), fracture location (physeal vs. metaphyseal) and age group (4-6 year olds vs. 7-10 year olds), similarly to our primary analysis model. If the number of school days missed up to 12 months is missing for more than 40% of participants, no modelling will be performed, and the results will be presented descriptively.

Additionally, the secondary outcomes (Wong-Baker FACES, EQ-5D and VAS, VAS cosmesis, Satisfaction, School Attendance and complications) will be assessed in the same manner as described, however including a type of fracture (completely off-ended or not) interaction term with treatment in order to estimate the treatment difference in participants with completely off-ended fracture, only in the ITT population.

Safety

It is not anticipated that there will be any serious adverse events (SAEs) in this trial; however, if any do occur the total number of SAEs, the number of participants with SAEs and the number of SAEs per participant will be summarised as outlined in the CRAFFT SAP DD&T. Line listings of all SAEs by treatment group will also be provided.

6.3 Missing Data

The number and percentage of individuals with available data for each outcome at each time-point will be summarised by treatment group. Following the guidelines for each of the PROMs used in this trial, no item-level imputation is planned for any of the outcomes, if one response is missing the whole item will be considered missing.

Table 3: Data availability by treatment group

	Surgical reduction					Non-surgical casting				
	Baseline	**6w**	**3m**	**6m**	**12m**	**Baseline**	**6w**	**3m**	**6m**	**12m**
Demographic details		NA	NA	NA	NA		NA	NA	NA	NA
Injury details		NA	NA	NA	NA		NA	NA	NA	NA
Treatment details										
PROMIS Upper Extremity										

Wong-Baker FACES									
EQ-5D utility									
EQ-5D VAS									
VAS cosmesis	NA								
Complications	NA								
Resource use	NA								
Satisfaction	NA	NA	NA	NA		NA	NA	NA	NA
School attendance	NA					NA			

All the main analyses in this trial will be performed on all available data.

6.4 Sensitivity Analysis

Missing outcome data: The primary analysis method proposed is reasonably robust to missing at random (MAR) data. Treatment effects will be compared to estimates resulting from the same model used in the primary (ITT) analyses after further adjusting for additional factors that are both prognostic of the outcome and associated with drop-out.

Informative drop out: Sensitivity analysis will also be undertaken on the primary analysis by imputing missing data under different missing not at random (MNAR) assumptions for the model targeting the ATE estimand and using the ITT population. This will be achieved using reference-based imputation following Cro et al. where we build upon the MAR-based MI under a pattern-mixture modelling framework, where different distributions are specified for fully and partially observed cases with reference to the main analysis (MAR) such that the overall outcomes distribution is a mixture of the two. This is done to assess the impact of unobserved participants having a worse or better response than those anticipated under MAR assumptions (e.g. implemented through the mimix Stata command) and is consistent with an ITT primary analysis. The scenarios that will be considered are described in Cros 2016 will. These sensitivity analyses maybe supplemented with additional analyses guided but the findings in the previous section describing dependencies of missing data on observed covariates.

6.5 Pre-specified Subgroup Analysis

Fracture type (completely vs. incompletely off-ended), fracture location (Metaphyseal fracture and Physeal fractures), and age have the potential to influence the outcome following this injury and will be considered for subgroup analyses.

The purpose of subgroup analyses is to investigate if the estimated treatment effects are relatively consistent across subgroup and for this extent will be viewed as exploratory. None of the included subgroups are based on post-randomisation patient characteristics or events If there are too few participants in a categorical subgroup (i.e. <=15) or one treatment arm of a subgroup (i.e. <=5), the analysis will not be conducted.

Therefore, two subgroup analyses, and a further analysis to produce treatment effect estimates at the 25th, 50th, and 75th percentiles (i.e. high-density areas) of the age range, of the PROMIS upper extremity scores at 3 months are planned. The subgroup and treatment effect by age analyses will use the same model as for the primary estimand, with the addition that the treatment effects will be estimated for the three analyses by including a treatment by timepoint and treatment by subgroup/age variable interaction terms in the models and reporting the associated adjusted differences with 95% Confidence Intervals). The model to estimate the treatment effect by age, will not include the age stratification variable.

PROMIS scores will be summarised by treatment group separately for the two subgroups using raw means and SDs.

6.6 Supplementary/ Additional Analyses and Outcomes

There are two additional analyses planned in this trial:

- Analysis of radiographic data
- Analysis of long-term follow-up data

Separate analysis plans will be drawn up for each of these analyses and may be considered appendices to this SAP.

6.7 Health Economics and Cost Effectiveness (where applicable)

The statistician is not undertaking this analysis. A separate Health Economics Analysis Plan (HEAP) will be drafted by the trial health economist.

7. VALIDATION OF THE PRIMARY ANALYSIS

To validate the primary outcome and EQ5D outcome in the ITT population, a statistician not involved in the trial analysis will independently repeat the analyses detailed in this SAP. This may be by using different statistical software. The results will be compared, and any unresolved discrepancies will be reported in the Statistical report.

8. SPECIFICATION OF STATISTICAL PACKAGES

All analysis will be carried out using appropriate validated statistical software such as STATA, SAS, SPLUS or R. The relevant package(s) and version number(s) will be recorded in the Statistical report.

9. PUBLICATION

A statement of Clinical Trials Unit involvement should be including in any publication of the SAP should be included. For example:

This study will be/has been conducted as part of the portfolio of trials in the registered UKCRC Oxford Clinical Trials Research Unit (OCTRU) at the University of Oxford. It will follow/has followed their Standard Operating Procedures ensuring compliance with the principles of Good Clinical Practice and the Declaration of Helsinki and any applicable regulatory requirements.

10. REFERENCES

1 Gamble, C., et al., Guidelines for the content of statistical analysis plans in clinical trials. Jama, 2017. 318(23): p. 2337-2343.

2 Kang M, Kendall MA, Ribaudo H, Tierney C, Zheng L, Smeaton L, Lindsey JC. Incorporating estimands into clinical trial statistical analysis plans. Clin Trials. 2022 Jun;19(3):285-291. Doi: 10.1177/17407745221080463. Epub 2022 Mar 8. PMID: 35257600; PMCID: PMC9232859.

3 ICH E9(R1) Addendum on estimands and sensitivity analysis in clinical trials

4 Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1. doi: 10.1186/2046-4053-4-1

5 Calvert, M., et al., Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. Jama, 2013. 309(8): p. 814-822.

6 Excellence, N.I.f.C., NICE guide to the methods of technology appraisal. NICE, London, 2004.

7 Wille, N., et al., Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. Quality of life research, 2010. 19(6): p. 875-886.

8 Eidt-Koch, D., T. Mittendorf, and W. Greiner, Cross-sectional validity of the EQ-5D-Y as a generic health outcome instrument in children and adolescents with cystic fibrosis in Germany. BMC pediatrics, 2009. 9(1): p. 55.

9 Kreimeier, S., et al., Valuation of EuroQol five-dimensional questionnaire, youth version (EQ-5D-Y) and EuroQol five-dimensional questionnaire, three-level version (EQ-5D-3L) health states: the impact of wording and perspective. Value in Health, 2018. 21(11): p. 1291-1298.

10 Curtis, L. and A. Burns, Unit Costs of Health and Social Care: PSSRU, University of Kent at Canterbury. 2018.

11 Department of Health. NHS Reference Costs 2015-16. 05 March 2020]; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016

12 White, I.R., P. Royston, and A.M. Wood, Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine, 2011. 30(4): p. 377-399.

13 White, I.R., et al., Strategy for intention to treat analysis in randomised trials with missing outcome data. Bmj, 2011. 342: p. d40.

14 Sterne, J.A., et al., Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj, 2009. 338: p. b2393.

15 Manca, A., N. Hawkins, and M.J. Sculpher, Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health economics, 2005. 14(5): p. 487-496.

16 Sullivan, T.R., et al., *Should multiple imputation be the method of choice for handling missing data in randomized trials?* Statistical methods in medical research, 2018. 27(9): p. 2610-2626.

17 Cro, S., et al., *Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: A practical guide.* Statistics in Medicine, 2020. 39(21): p. 2815-2842.

18 Cro S, Morris TP, Kenward MG, Carpenter JR. Reference-based sensitivity analysis via multiple imputation for longitudinal trials with protocol deviation. Stata J. 2016 Apr;16(2):443-463. PMID: 29398978; PMCID: PMC5796638.

APPENDIX: GLOSSARY OF ABBREVIATIONS

CI	Confidence Interval
CONSORT PRO	Consolidated Standards of Reporting Trials – Patient-Reported Outcomes
CRAFFT	Children's Radius Acute Fracture Fixation Trial
CSM	Centre for Statistics in Medicine
CTU	Clinical Trials Unit
DSMC	Data and Safety Monitoring Committee
EQ-5DY	EuroQoL 5 Dimensions – youth version
GA	General Anaesthetic
HEAP	Health Economics Analysis Plan
IQR	Interquartile range
ITT	Intent-to-treat
MNAR	Missing not at random
NIHR HTA	National Institute for Health Research Health Technology Assessment
OCTRU	Oxford Clinical Trials Research Unit
PP	Per protocol
PROMIS	Patient-Reported Outcome Measurement Information System
PSS	Personal Social Services
QALY	Quality Adjusted Life Years
RCT	Randomised Controlled Trial
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAP DD&T	Statistical Analysis Plan – Data Definitions and Tables
SD	Standard deviation
SOP	Standard Operating Procedures
TMF	Trial Master File
TSC	Trial Steering Committee
VAS	Visual Analogue Scale
SD	Standard Deviation