The FIRST study Statistical Analysis Plan Version 2 Based on version 4 of the FIRST study protocol (dated 017Jul2024)	
Study Title	Prospective randomised controlled trial comparing three splints for finger flexor tendon repairs – FIRST study
Funding body	NIHR HTA 133582

Authored by

Rombach

Ines Rombach (Senior Statistician) & — Date: 06/Mar/2025

CTRU, University of Sheffield

Approved by

S. Walters

Stephen Walters — Date: 7/Mar/2025

Senior Statistician, CTRU, University of Sheffield

E. Bamford

Emma Lucy Bamford — Date: 8/Mar/2025

Chief Investigator, Derby Hospitals NHS Foundation Trust

David J. Cooper

David Cooper — Date: 10/03/2025

FIRST Study TSC Statistician (University of Aberdeen)

Summary of changes

Version	Date approved	Modifications (with section)	Made by	Prior to/after blind/unblind review Prior to/after database lock
2.0	TBC	Addition of new section on handling missing baseline data (now section 8.6.2). It was always planned that all participants with at least one available follow-up measure should be included in analyses. Clarified aspects of the adverse event summaries that arose from blinded data cleaning. Specification of graphs to be presented. Clarification of how follow-up data will be handled that falls outside the protocol stipulated follow-up time windows. Minor clarification and correction of typographical errors.	Ines Rombach	Prior to database lock At the time of the change to the SAP, both Ines Rombach and Stephen Walters had attended Data Monitoring committee meetings where unblinded outcome data by treatment arm had been presented. Freya marks, who will be performing the analysis, had not been unblinded yet.

List of Abbreviations and Definitions of Terms

ROM	Active range of movement
AE	Adverse Event
BSSH	British Society of Surgery of the Hand
CRF	Case Report Form
CTRU	Clinical Trials Research Unit
EQ-5D-5L	EuroQol - 5 Dimensions - 5 Levels
GCP	Good Clinical Practice
GLM	Generalised Linear Model
ICC	Intra-class Correlation Coefficient
IQR	Inter Quartile Range
IRR	Incidence Rate Ratio
ISRCTN	International Standard Randomised Controlled Trial Number
ITT	Intention To Treat
NICE	National Institute for Health and Clinical Excellence
NHS	National Health Service
PEM	Patient Evaluation Measure
PROM	Patient reported outcome measure
PRWHE	Patient Rated Wrist and Hand Evaluation
PP	Per Protocol
QoL	Quality of Life
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
SAE	Serious Adverse Event
SD	Standard Deviation
SOP	Standard Operating Procedure
TMG	Trial Management Group
TSC	Trial Steering Committee
WPAI	Work productivity and activity impairment

Table of Contents

SUMMARY TABLE

This Statistical Analysis Plan (SAP) is based on version 2.1 of the FIRST protocol.

Study title	Prospective randomised controlled trial comparing three splints for finger flexor tendon repairs (FIRST study)
Trial design	A parallel group, superiority, analyst-blind, multi-centre, individual participant-randomised controlled trial.
Hypothesis, aims and objectives	**Hypothesis:** The trial hypothesis is that any one of the splints may be superior, in terms of mean post-randomisation scores (based on data collected at 6, 12, 26, and 52 weeks) for self-reported wrist/hand pain and functioning outcomes, to any of the others. **Aim:** To investigate the clinical and cost effectiveness of three splints in the repair of Zone I/II finger flexor tendons. **Study Objectives:** ● To determine if any splint is superior in terms of a patient-rated measure of pain and function ● To investigate how patient values and splint acceptability moderates objectively measured splint adherence, and how adherence mediates effectiveness ● To evaluate splint cost-effectiveness, from an NHS and societal perspective
Internal pilot/feasibility criteria	An 8-month internal pilot will assess the feasibility of the RCT.
Setting	Approximately 20 UK NHS Hospitals
Participants	Patients undergoing rehabilitation following the surgical repair of zone I/II flexor tendons will be recruited to the study.
Intervention & control groups	Three intervention groups will be compared: **1: Long splint:** custom-made, thermoplastic splint, with controlled early active movement. Covers whole hand and forearm; prevents motion of the wrist, allows controlled motion of the fingers. Worn continuously for 5 weeks and, intermittently, for 1 more week, whilst not using the hand for any activities. **2: Short splint:** custom-made, thermoplastic splint. Covers fingers, but allows motion at wrist. Worn at all times for 5 weeks; intermittently, for 1 more week, can use unaffected fingers for light activities only. **3: Mini splint:** custom-made finger-based splint preventing full extension of the injured fingers but allowing the hand and fingers to be used for daily activities with a wrist support. Finger element worn

	at all times for 5 weeks and intermittently for 1 more week. Wrist element worn at all times for the first 3 weeks of splint wear, and intermittently for 3 more weeks.
Primary outcome(s)	The primary outcome is the mean post-randomisation total score on the Patient Rated Wrist and Hand Evaluation (PRWHE), measured at baseline 6, 12, 26 and 52 weeks post-randomisation.
Secondary outcome(s)	The secondary outcomes consist of other patient-reported outcomes, and clinical outcomes.
Duration of follow-up	Participants will be followed up until 52 weeks post-randomisation.
Target sample size	429 participants
Definition of end of trial	The end of the trial is when the last recruited participant completes their 52 week follow up. Sites will be closed once data cleaning is completed and the ethics committee will be informed.

1 INTRODUCTION

This Statistical Analysis Plan (SAP) is written in conjunction with the International Conference of Harmonisation (ICH) topic E9 on statistical principles for clinical trials [1] guideline on clinical trials SAPs [2] applicable Standard Operating Procedures (SOPs) from the Sheffield Clinical Trials Research Unit (CTRU), particularly ST001 [3], and the trial protocol [4] and related amendments.

This SAP provides detailed guidance for the Trial Statistician undertaking the statistical analysis and reporting for the FIRST study. This section gives a brief background of the trial, the primary research question under investigation and the study design used to address the research questions.

1.1 Background

Hand injuries in the UK have increased 57% in 15 years, accounting for 20% of emergency presentations [5]. There were 7346 flexor tendon injuries in 2018-19; 75% were in working age men [6]. Such injuries most frequently occur from a direct laceration to the tendon in the finger or palm of the hand. Without surgical repair and rehabilitation, divided tendons do not heal, patients cannot bend fingers, grip objects or effectively care for themselves and others. Long-term prognosis can be poor: 50% of patients report pain and functional limitations 10 years post-injury [7]. Hand injuries may impact a patient's mental wellbeing and role in society, with increased dependence on others, and an inability to care for oneself or one's family [8] [9] . The economic impact of flexor injury is higher than for carpal tunnel release and distal radius fracture [10] with 80% indirect costs due to missed workdays. Patients with flexor tendon injuries had an average of 70 sick days [11] jeopardising financial stability [9]. Our PPI group had diverse experiences ranging from 'I'm self-employed, I drive diggers and if I don't go to work, I don't earn money', to, 'I was able to take as much paid time as I wanted off' [12].

British Society of Surgery of the Hand (BSSH) guidelines recommend primary end-to-end tendon repair with multi-strand locking sutures [13]. However, the outcome of flexor tendon surgery also relies on effective rehabilitation [14]. Patients routinely attend weekly appointments for up to three months with full recovery taking up to one year [15]. There are two components to rehabilitation: exercises to prevent hand stiffness and promote tendon glide/excursion, and custom-made splints to protect the repair. Early active mobilisation (EAM) exercise protocols are universally accepted and generally adopted in the NHS.

By comparison, there are three main custom-made splints used in the rehabilitation of zone I/II flexor repairs that aim to facilitate EAM, but also protect the newly repaired tendon:

1. Long - forearm and hand-based splint.
2. Short - hand based splint.
3. Mini - finger based splint.

Each splint is assumed to have different associated risks and benefits, but it is unclear which, if any, offers superior outcomes. As a result of the weak evidence base, clinical practice varies between hand centres, surgeons and therapists which reflects the clinicians' own preference, experience and beliefs. The three splints used in the NHS all differ in the amount of movement and function allowed, which probably affects initial choice, adherence and outcomes. Patients feel more protected from rupture

with the Long splint, but dislike its aesthetics, discomfort and functional restriction [16], which leaves 5- 10% with scar tissue adhesions, requiring further surgery [17]. The short and mini splints may be more aesthetically pleasing, offer a greater range of movement and lower risk of adhesions, but may come with an increased risk of rupture [18]. Patients will hold diverse opinion(s) on splint choice based on personal beliefs, circumstances and cultural values. Splint choice and adherence is also influenced by occupation and sick pay provision: observational studies confirm that injury compensation status impacts time off work in manual workers [19][20] . Thus, what patients value and, how they make trade-offs, may moderate splint adherence which, in turn may mediate clinical effectiveness.

1.2 Trial Design

The study is a UK multicentre, superiority, three-arm, parallel group, randomised, analyst-blind trial with a one-year follow-up. Participants will be randomised (1:1:1) to one of three intervention arms. The trial will be conducted in approximately 20 hospitals. An 8-month internal pilot will run at all sites planned to participate in the main trial, to assess the feasibility of the RCT. The progression criteria will be applied to data collected 8 months after the first site is opened. The progression criteria (site set up, participant recruitment, participant allocation per protocol and follow-up at 6 weeks) will be assessed by the Trial Steering Committee (TSC) at the end of the following month (see 2.1.1).

The primary outcome is the mean post-randomisation total score on the Patient Reported Wrist and Hand Evaluation (PRWHE). This outcome is measured repeatedly at 6, 12, 26 and 52 weeks post-randomisation.

1.2.1 Trial treatment

This study will assess the three main custom-made splints used in the rehabilitation of zone I/II flexor repairs; Long forearm and hand-based splint, Short hand based splint, and Mini finger based splint. For each type of splint, the splint will be fitted at the first hand therapy appointment post-surgery.

1. Long - forearm and hand-based splint: The forearm-based early active motion splint, or 'Long splint', is a custom-made, thermoplastic splint which allows controlled early active movement. It covers the dorsal aspect of the whole hand and forearm, thereby preventing motion of the wrist and allowing for controlled motion of the fingers. The Long splint will be prescribed for 5 weeks continuous wear, and intermittently for 1 more week. Patients are advised not to use the hand for any activities.

2. Short - hand based splint: The Manchester short splint, or 'Short splint', is a custom-made, thermoplastic splint which covers the dorsal aspect of the fingers, but allows motion at the wrist. The Short splint will be prescribed for 5 weeks continuous wear, and intermittently for 1 more week. Patients are advised to only use their unaffected fingers for light activities.

3. Mini - finger based splint: The relative motion flexion splint, or 'Mini splint', is a custom-made finger-based splint which prevents full extension of the injured fingers but allows the hand and fingers to be used for daily activities with wrist support. Finger and wrist elements will be worn for the first 3 weeks, and the finger component only for between 5 and 8 weeks.

1.3 Aims and Objectives

1.3.1 Hypothesis

The trial hypothesis is that any one of the splints may be superior, in terms of the mean post-randomisation scores (based on data collected at 6, 12, 26, and 52 weeks) for self-reported wrist/hand pain and functioning outcomes, to any of the others.

1.3.2 Aims

To investigate the clinical and cost effectiveness of three splints, and mediators of effectiveness, in the repair of zone I/II finger flexor tendons.

1.3.3 Objectives

- To determine if any splint is superior in terms of a patient-rated measure of pain and function
- To investigate how patient values and splint acceptability moderates objectively measured splint adherence, and how adherence mediates effectiveness
- To evaluate splint cost-effectiveness, from an NHS and societal perspective

2 THE SCOPE OF THE SAP

As highlighted in Section 1.2, this trial has been designed with an internal pilot phase to assess the feasibility of a full-scale trial. This SAP focuses on addressing the clinical effectiveness related research questions of the trial. The internal pilot feasibility objectives highlighted in Section 2.1.1, and the health economic evaluation aspects, which are out of the scope of this SAP, will be addressed elsewhere. A separate analysis plan will also be written for the mediation analysis outlined in the protocol.

2.1 Outcome measures

2.1.1 Internal pilot Outcomes

An 8-month internal pilot will assess the feasibility of the RCT. This will include assessment of the following:

- Site set up
- Participant recruitment
- Participant allocation per protocol
- Follow-up at 6 weeks

Sheffield CTRU will aggregate study data to assess the feasibility of the research and intervention protocols based on the following feasibility outcomes:

Table 1: Internal pilot progression criteria

Criterion	Red (% complete)	Amber (% complete)	Green (% complete)
Number of Sites opened	<15 (75%)	>=15 (75%) and <20	20 (100%)
Rate/site/month	<0.9 (<60%)	>=0.9 (60%) and <1.4	1.5 (100%)
Number of participants recruited	<144 (<60%)	>=144 (60%) and <240	240 (100%)
Allocation per protocol	<90%	>=90% and <100%	100%
% FU (% of recruited)	<50%	>=50% and <75%	75%

The feasibility outcomes of the study will be reported to the funder, and are outside the scope of this document.

2.1.2 Primary outcome

The primary outcome is the mean post-randomisation total score of the Patient Rated Wrist and Hand Evaluation (PRWHE), measured at 6, 12, 26 and 52 weeks post-randomisation. The PRWHE is a 15-item patient reported outcome for assessing wrist and hand pain/disability on a scale of 0 to 100 (0 = no pain/disability) [21]. It is a modified version of the Patient Rated Wrist Evaluation (PRWE) to allow assessment of hand conditions, having the same 15-items and scoring system replacing the term “wrist” with “wrist/hand”.

Primary Estimand for the FIRST trial

Estimand attribute	Description
Population	Patients undergoing rehabilitation following the surgical repair of zone I/II flexor tendons.
Treatment(s)	Three intervention groups will be compared: **1: Long splint:** custom-made, thermoplastic splint, with controlled early active movement. Covers whole hand and forearm; prevents motion of the wrist, allows controlled motion of the fingers. Worn continuously for 5 weeks and, intermittently, for 1 more week, whilst not using the hand for any activities. **2: Short splint:** custom-made, thermoplastic splint. Covers fingers, but allows motion at wrist. Worn at all times for 5 weeks; intermittently, for 1 more week, can use unaffected fingers for light activities only. **3: Mini splint:** custom-made finger-based splint preventing full extension of the injured fingers but allowing the hand and fingers to be used for daily activities with a wrist support. Finger element worn at all times for 5 weeks and intermittently for 1 more week. Wrist element worn at all times for the first 3 weeks of splint wear, and intermittently for 3 more weeks.
Outcome (endpoint)	Mean post-randomisation total score on the PRWHE, measured at baseline 6, 12, 26 and 52 weeks post-randomisation.
Handling Intercurrent events	1. Stopping randomised treatment for any reason – treatment policy (as part of treatment) 2. Switching treatments – treatment policy (as part of treatment) 3. Changing randomised treatment or not receiving randomised treatment - treatment policy (as part of treatment) 4. Use of other medications/treatments/therapy - treatment policy (as part of treatment) 5. Death – if a participant dies during their trial follow-up, the 'while alive' strategy will be used See below for definitions of the different strategies for handling intercurrent events.
Summary measures	**Mean treatment differences in the average follow-up PRWHE** 1) Long splint vs. short splint 2) Long splint vs. mini splint 3) Short splint vs. mini splint

Strategies for handling intercurrent events:

Treatment policy - regardless of any post randomisation events, the treatment effect is described from the final outcome measure in all patients. Note that this approach cannot be used for truncated events, for example, where a variable cannot be measured due to death"[22].

While alive - data up to death will be included. No imputation for missing data due to death will be used in the main analyses.

Research questions answered by the estimands framework:

In patients undergoing rehabilitation following the surgical repair of zone I/II flexor tendons of the hand, what is the difference in mean post-randomisation quality of life (or pain/function) scores (as measured by the total score on the Patient Rated Wrist and Hand Evaluation (PRWHE) outcome) between participants randomised to the Long splint group in addition to usual care followed by any subsequent therapy/treatment (as needed); compared with participants randomised to Short splint in addition to usual care followed by any subsequent therapy/treatment (as needed) compared to participants randomised to the Mini splint group in addition to usual care followed by any subsequent therapy/treatment (as needed) up to 12-months from randomisation or death (whichever occurs first), regardless of study treatment discontinuation?

2.1.3 Secondary outcomes

Time points for secondary outcome data collection will be consistent with primary outcome data collection. Please see section 3.2.

Patient-reported Outcomes:

1. Patient Evaluation Measure (PEM) - patient reported measure of care received, function, pain and wellbeing;
2. Work productivity and activity impairment (WPAI);
3. EuroQoL EQ-5D-5L - health status questionnaire used to derive quality adjusted life years (QALYs) and used in the cost effectiveness analysis;
4. Details of any litigation/compensation for injury;
5. Global rating of change question;
6. Preferences for splint attributes (stated and revealed) and splint acceptability (see process evaluation).

Clinical outcomes:

1. Active range of movement (AROM): The AROM of the affected digit/digits will be measured with a finger goniometer according to a standardised protocol;
2. Grip Strength;
3. Splint adherence, assessed using a temperature sensor in the participants' splint;
4. Complications and adverse events.

2.1.4 Safety outcomes

Adverse events (AEs) and Serious Adverse Events (SAEs) will be recorded from the point a participant provides written informed consent for trial entry and up until participant's completion of the trial. Ongoing AE/SAEs will be followed up until the event has resolved or stabilised, or until the participant's involvement in the trial has ended. All SAEs occurring from the point of consent up to the end of involvement in the trial will be reported to the CTRU within 24 hours of recognition, unless exempt (see section 8.9). Delegated site trial staff will record all adverse events and make them known to the Principal Investigator and/or Co-Principal Investigator (see section 8.9).

3 ASSESSMENTS AND PROCEDURES

3.1 Data collection

All clinical data will be entered by research site staff onto the CTRU's in-house data management system (Prospect). Patient reported outcome measures (PROMs) data (to include the primary outcome assessment) will be completed online by the patient using a tablet in clinic at baseline, 6, 12 and 26 weeks and remotely via email, or over the phone where required, at 52 weeks, with paper copies available if this is not possible.

Complications and AEs/SAEs will be assessed at each clinic visit, and via phone call at 52 weeks by a delegated member of the research team at site. Site staff will remind participants to complete their questionnaires if they have not done so already. Non-responders to email questionnaires at the 52 week remote visit may also be followed up by the CTRU research team using contact details provided by the participant, to check that outcome measures have been received and to prompt them to return the outcome measures. Contact will be attempted on contact details provided by the patient, which may include telephone contact, email or text message. Up to three contact attempts will be made. At all contact points details of how the participant can contact the research team will be included and an offer to complete the questionnaires over the telephone will be made. Participants will be considered lost to follow-up if they have not returned the week 52 questionnaires at the point of study closure.

Data Management (CRF design, data cleaning and validation) will be provided by the CTRU. Project-specific procedures for data management will be detailed in a data management plan.

3.2 Study assessment schedule

The study assessment schedule below details the assessments required during the course of the study. All participants will undergo these assessments, regardless of which treatment they are randomised to.

For the purposes of the primary analysis all follow-up data will be included, based on the protocol specified follow-up time points under which it has been entered. Depending on the extent of data outside of the visit widows, a sensitivity analysis may be performed excluding these data, or reattributing them to a protocol stipulated follow-up visit that aligns with the date the assessment was performed, where appropriate.

Table 2: Data collection schedule

	Baseline (Clinic)	6 weeks (Clinic)	12 weeks (Clinic)	26 weeks (Clinic)	52 weeks (Remote)
Baseline and other covariates					
Pre-screening form/log (before baseline visit)	x	-	-	-	-
Eligibility form	x	-	-	-	-
Surgery details form	x	-	-	-	-
Informed consent form	x	-	-	-	-
Contact details	x	-	-	-	-
Demographics	x	-	-	-	-
Employment (including sick pay provision)	x	x	x	x	x
Vehicle use	x	-	-	-	-
Randomisation (at baseline)	x	-	-	-	-
Primary outcome					
Patient Rated Wrist and Hand Evaluation (PRWHE)	x	x	x	x	x
Patient reported measures					
Patient Evaluation Measure (PEM)	x	x	x	x	x
Work productivity and activity impairment (WPAI)	x	x	x	x	x
EuroQoL EQ-5D-5L	x	x	x	x	x
Litigation/compensation	-	-	-	-	x
Global rating of change (GRoC)	-	x	x	x	x
Preferences for splint attributes	x	x	-	-	-
Clinical outcomes					
Range of movement (Strickland score)	x	x	x	x	-
Grip strength	-	-	x	x	-
Splint adherence from heat sensor	-	x	-	-	-
Complications and AE/SAEs	x	x	x	x	x

4 SAMPLE SIZE ESTIMATION

The sample size was calculated using the methodology and formula for repeated outcome measures [23]. We assumed: i) 90% power; ii) 1.67% two-sided significance level (to allow for three head-to-head comparisons between the three randomised groups); iii) 1 baseline and 4 repeated assessments at 6, 12, 26 and 52 weeks post-randomisation; iv) a target difference of 6- points [24] in the post-randomisation mean PRWHE scores between any two of three groups; v) a standard deviation of 20 points for the PRWHE outcome at each post-randomisation time point [25][26]; vi) an exchangeable correlation or compound symmetry of 0.50 between the repeated PRWHE assessments at 6, 12, 26 and 52 weeks post-randomisation [25][27]; vii) 20% attrition. With these input parameters 114 subjects per

group are required (3 x 114 =342 in total). After allowing for 20 % attrition, we proposed to randomise 429 participants in a 1:1:1 ratio (143 Long splint: 143 Short splint: 143 Mini splint).

5 TRIAL FEATURES TO MINIMISE BIAS

5.1 Design, randomisation and concealment

Once eligibility has been confirmed, consent acquired, and baseline data taken, the participant will be randomly allocated to either the Long splint arm, the Short splint arm or the Mini splint arm on a 1:1:1 basis, using a web-based randomisation system provided by Sheffield CTRU. Patient details (ID, date of birth) and site will be entered into the randomisation system and the treatment allocation will be returned. Randomisation allocations will be based on computer-generated pseudo-random lists, stratified by site, with random permuted block sizes. Randomisation will be done by site staff during the clinic visit and participants will be informed of the outcome verbally. Their GP will also be informed of their participation in the trial, and their treatment allocation.

5.2 Blinding and trial integrity issues

In view of the nature of the intervention, patients and their treating clinicians cannot be blinded to the treatment allocation. Clinical assessors at sites measuring range of movement and grip strength will be blinded to trial allocation. Participants will be asked to remove their splint before each follow-up visit to ensure clinical assessors remain blind to trial allocation. The trial statistician(s) will remain blinded throughout the study, but will be unblinded at database freeze, for analysis.

6 TRIAL MONITORING AND INTERIM ANALYSIS

In compliance with Sheffield CTRU's SOPs, GV001[28][28], GV002 [29][29], and GV003 [30], trial protocol [4], and the DMEC Charter, the following committees will be established to govern the overall conduct and supervision of the trial:

- Trial Steering Committee (TSC)
- Data Monitoring and Ethics committee (DMEC)
- Trial Management Group (TMG)

The trial will be supervised on a day-to-day basis at Sheffield CTRU by the Trial Manager with supervision from the Chief Investigator.

The trial is a fixed sample size design with only one formal statistical analysis at the planned scheduled end when all participants are recruited and have completed the outcome assessments. However, the trial will be independently monitored by the DMEC within the premise of the DMEC Charter, which was agreed and signed by all the members.

6.1 Data sources and data management

All clinical data will be entered by research site staff onto the CTRU's in-house data management system (Prospect). Patient reported outcome measures (PROMs) data (to include the primary outcome assessment) will be completed online by the patient using a tablet in clinic at baseline, 6, 12 and 26

weeks and remotely via email, or over the phone where required, at 52 weeks, with paper copies available if this is not possible.

Complications and AEs/SAEs will be assessed at each clinic visit, and via phone call at 52 weeks by a delegated member of the research team at site. Site staff will remind participants to complete their questionnaires if they have not done so already. Non-responders to email questionnaires at the 52 week remote visit may also be followed up by the CTRU research team using contact details provided by the participant, to check that outcome measures have been received and to prompt them to return the outcome measures. Contact will be attempted on contact details provided by the patient, which may include telephone contact, email, or text message. Up to three contact attempts will be made. At all contact points details of how the participant can contact the research team will be included and an offer to complete the questionnaires over the telephone will be made. Participants will be considered lost to follow-up if they have not returned the week 52 questionnaires at the point of study closure.

Data Management (CRF design, data cleaning and validation) will be provided by the CTRU. Project-specific procedures for data management will be detailed in a data management plan.

7 ANALYSIS POPULATIONS AND SUBGROUPS

7.1 Analysis populations

Unless specified otherwise, analyses are based on an Intention-to-treat (ITT) population as defined in Table 3. Additional sensitivity analyses as described in Section 8.6.4 will be undertaken based on the per protocol (PP) population defined in Table 3 where appropriate. Clinical input has been sought through the TMG to help define the PP population.

Table 3: Definition of the analysis populations

Analysis Population	Outcomes	Participant Inclusion Criteria
Intention to treat (ITT)	All outcomes	Randomised participants, analysed according to their randomisation allocation, regardless of their adherence to the protocol. For analyses using repeated measures, all participants with at least one follow-up data point will be included.
Per protocol (PP)	Sensitivity analysis for the primary outcome	Participants are classed as adherent, and therefore part of the PP population if they: • Were fitted with their allocated splint after randomisation • Their splints were not removed early without immediate replacement (i.e. replacement on the same day) • Have a splint removal date recorded at least 5 weeks post randomisation (splints are to be worn for 6

		weeks, but intermittently in the last week, so the 5 weeks gives some leeway for slightly early removals) Note: inclusion in the PP population was initially planned to be guided the data obtained from the temperature sensors. Due to complexities around defining adherence based on the data generated by the sensors, and thresholds for minimum daily wear times, we opted for this pragmatic and intuitive definition of the PP population. Detailed information on splint wear based on the temperature sensor data will be presented separately. If more than 50% of participants do not meet the inclusion these criteria for the PP analysis, the threshold for adherence may be modified, such that the PP population has a sufficient number of participants in order to undertake formal statistical analysis and draw reliable conclusions.

7.2 Pre-specified subgroups

As stipulated in the protocol, subgroup analyses will be performed for the primary outcome. The objective is to explore whether there is heterogeneity in treatment effect across the following pre-specified subgroups:

1. Sex (male/female)
2. Employment type (not in paid employment, manual, non-manual, unknown)
3. Eligible for occupational sick pay (yes/no)
4. Age (<25, 25-65, over 65)
5. Zone of injury (I/II)
6. Tendons repaired (FDP, FDS main tendon, FDS radial slip, FDS ulna slip)

Details of statistical methods to undertake subgroup analyses and reporting are described in Section 8.6.6.

8 OUTLINE OF STATISTICAL ANALYSES

This section outlines the statistical analyses framework to be adopted, beginning with how trial data and results will be reported. The description of the statistical methods used to analyse outcomes to address trial research questions is provided in order of importance, starting with the primary outcome then followed by secondary and safety outcomes.

8.1 Reporting framework of trial data

Since this study is a parallel group, pragmatic, randomised controlled superiority trial, the analysis of trial data and reporting will be guided by the revised CONSORT statement for multi-arm parallel-group individually randomised trials [31] [32]. A detailed CONSORT flow diagram from screening to the end of

the trial will be constructed using the information summarised in Section 8.2 at the discretion of the Trial Statistician (e.g., in line with the preference of the target journal).

Baseline summary statistics will be reported by treatment arm and overall. Comparability between treatment arms will be descriptively reported without any statistical significance testing [33][34][35]. Binary and categorical variables will be presented with as the number of observations and proportion in each category, by treatment group and overall. Continuous variables will be summarised and presented by treatment group and overall, using the mean and standard deviation (SD) and/ or median, inter-quartile range (IQR), minimum and maximum.

8.1.1 Confidence Intervals and P Values

All statistical tests will be completed at the 1.67% significance level and estimates of the treatment effect will be reported with their associated 98.3% confidence intervals, to acknowledge the three head-to-head comparisons performed in this trial, and the adjustment for multiplicity in the sample size calculation. All tests completed will be two-sided. The results of the trial are focussed on the primary endpoint and principal follow-up metric (mean post-randomisation total score on the Patient Rated Wrist and Hand Evaluation (PRWHE), measured at baseline 6, 12, 26 and 52 weeks post-randomisation) for each of the three pairwise comparisons, and adjustment for multiple testing has been applied. No further adjustment for multiple testing is made for summaries at additional follow-up time points and secondary endpoints.

8.2 The CONSORT flowchart

Using guidelines from the CONSORT statement, the summaries outlined Table 4 will be calculated to construct a CONSORT flowchart. An example flowchart is shown in Section 10.1.

The CONSORT flowchart will include summaries of participant-requested withdrawals from the trial intervention and follow-up, including frequency and percentage of participants who have withdrawn from follow-up, and reasons for withdrawal. Clinician-led withdrawals from the randomised intervention will also be presented.

Table 4: CONSORT Summary

Screening Data	• Number initially identified • Number and proportion of patients provided with study information (with reasons if study population not provided) • Number of proposition of patients whose eligibility was assessed (with reasons if eligibility not assessed)
Eligibility	• Number and proportion eligible to take part • Number excluded due to failure to meet inclusion criteria with reasons.
Consent	• Number and proportion who consented. • Number not consented, with reasons.

Randomisation	• Number and proportion of participants randomised. • Number and proportion randomised to each of three intervention arms. • Number and percentage of those who received their randomised intervention.
Follow-up (6, 12, 26 and 52 week post-randomisation)	• Number and proportion followed-up[a], in each treatment arm. • Number and proportion discontinued with reason(s). • Number and percentage of participants who dropped out or withdrawn before the one year follow-up
Analysis Population Data	• Number of those with available primary endpoint data at each post-randomisation follow-up visit in the ITT population • Number of those with available primary endpoint data at each post-randomisation follow-up visit in the PP population

[a] *These numbers will be presented for each of the four follow-up time points and include all participants for whom any follow-up data at the relevant time point are available.*

8.3 Data manipulation and definitions

This section details definitions and calculations needed for certain outcomes, but it does not include all outcome measures. Most of the outcomes are directly recorded on CRFs so no additional data manipulations are required. The following sub-sections describe the details of any calculations and definitions which are not directly recorded on the CRFs.

8.3.1 Primary outcome definition

The primary outcome is the mean of the Patient Reported Wrist and Hand Evaluation (PRWHE) collected over the follow-up. Protocol-specified follow-up time points are 6, 12, 26 and 52 weeks after randomisation.

The PRWHE scores are calculated in line with the Patient-Rated Wrist Evaluation (PRWE), from which it was adapted. The following rules are used to generate the scores:

- Firstly, individual pain and function scores are generated.
    - The pain score is generated from the five pain questions (pain at rest, pain when doing tasks with repeated wrist movement, pain when lifting a heavy object, pain when it is at its worst, frequency of pain).
    - The five items are rated on a Likert scale ranging from 0 to 10, indicating no pain/ never and worst ever/ always, respectively.
    - The five items are summed up, generating a pain score ranging from 0 to 50
    - The function score is generated from the ten questions on specific activities (fasten buttons on shirt, cut meat (or vegetables), turn doorknob, use hand to push up from chair, carry 10lb object, use bathroom tissue) and usual activities (personal care, household work, work, recreational activities).
    - The ten items are rated on a Likert scale ranging from 0 (no difficulty) to 10 (unable to do).

- The ten items are summed up and divided by 2, generating a function score ranging from 0 to 50.
- The total PRHWE score is generated by summing the pain and function scores, and ranges from 0 to 100, with lower values indicating better pain and function outcomes.

In the scoring manual [26], MacDermid indicates that missing items are commonly replaced by the rounded mean of the relevant subscale. For this trial, we will replace missing items by the arithmetic mean for the relevant subscales if no more than half of the items are missing. This means that up to two missing items for the questions on pain, and up to five missing items for the questions on function can be replaced by the mean of the available items for the relevant subscales. This approach has been confirmed with the author of the PRWHE.

The PRWHE questionnaire contains an additional section titled 'other concerns'. These questions do not contribute to the overall score or subscales. Responses to these questions will be presented descriptively by treatment arm, without comparative analyses. Responses to the question 'Do you have any other concerns' will be listed in the appendix of the final report.

8.3.2 Patient Evaluation Measure (PEM)

The PEM consists of three subscales (treatment (only asked at follow-up), hand health profile, overall satisfaction), comprising 5, 11 and 3 questions respectively.
Possible responses to each item range from 1 to 7, with lower numbers indicating more favourable outcomes. For the purposes of calculation summary scores, 1 is subtracted from all responses, so that each item has an integer score between 0 and 6.
The overall score (referred to as PEM – Hand Health and Overall Assessment Questionnaire [14 items]) will be calculated by summing the values for the available items in parts two and three and expressing it as a percentage of the maximum possible score [36]. Higher scores indicate higher levels of disability.
To align the reporting of outcomes with other trials in similar patient populations, the PEM - Hand Health Questionnaire (11 items), based on part two only, will also be generated, in line with the calculations described above. Again, higher scores indicate higher levels of disability.
We will handle missing data in the PEM in line with other trials, including the DISC study. As such, we will replace missing items in part 2 (11 questions) with the mean of the non-missing item in that part of the questionnaire if up to two items are missing. Scores will not be calculated if more than two items in part 2, or any items in part three are missing (for the calculation of the 14-item score).
The first part, which assessed the patient's view of the consultation, will be excluded from the calculation [36]. Summaries of the data from part 1 will be descriptively presented.

8.3.3 Work productivity and activity impairment (WPAI)

This questionnaire consists of six questions on employment status, hours of work missed, and productivity over the past seven days. Questions are presented separately (a proportion of hours of work lost due to health problems may be calculated), and no overall score is generated.

8.3.4 EuroQoL EQ-5D-5L

Health status questionnaire used to derive quality adjusted life years (QALYs) and used in the cost effectiveness analysis. The EQ-5D-5L consists of a utility score, generated from the responses to the five questions on mobility, self-care, usual activities, pain/ discomfort, and anxiety/ depression, as well as a separate visual analogue scale (VAS), ranging from 0-100.

The utilities will be generated based on the current NICE (National Institute for Health and Clinical Excellence) guidance [37]. At the time of writing this SAP, NICE recommended that the mapping function developed by Hernández Alava et al [38] is used for economic evaluations.

8.3.5 Active range of movement (AROM)

The AROM of the affected digit/digits will be measured with a finger goniometer according to a standardised protocol. The Total Active Motion (TAM) will be calculated as the total active flexion of the proximal interphalangeal joint (PIPj) and distal interphalangeal joint (DIPj) motion in a composite fist position minus the extension deficit.

The Strickland score for active range of motion will then be calculated from this measurement, defined as:

$$Strickland\ Score = \frac{(Active\ flexion\ PIPj\ +\ DIPj) - (Extension\ defecit\ PIPJ\ + DIPj)}{175} \times 100$$

No values for AROM can be generated if any of the required components are missing. The AROM will be calculated and summarised for all affected digits, however for analysis purposes, the digit closest to the 'little finger' will be selected for the analysis model detailed in section 8.7.

AROM data for the unaffected hand are collected at baseline only.

8.3.6 Grip strength

Grip strength will be calculated as the average of the three grip strength recordings performed with the GripAble handheld dynamometer using a standardised protocol. This average is calculated and presented directly on the CRF.
Grip strength is collected for the affected and unaffected hand.

8.3.7 Splint adherence

Splint adherence is collected using temperature sensors embedded within the trial splints. This will record the temperature, at 10-minute intervals, for the full 6 week period that the participants will be wearing their splint.

For the secondary outcome analysis of splint adherence, we will base adherence on the first 5 weeks of the splint wear period, see section 8.7. However, further exploratory analysis will be conducted using the temperature sensor data, which will include all 6 weeks' data, the details of this are outlined in section 8.10.

We will use either a fixed temperature threshold (estimated at 27.5 degrees for the mini splint and 28.5 for the short/long splints) or an individualised temperature threshold, to define whether the splint is being worn or not. The threshold will be chosen based on experiments conducted by our collaborators and current literature and the results of ongoing experiments. The thresholds will be specified prior to the final analysis being performed.

Note: A separate definition of splint adherence will be used in the definition of the per-protocol population (see Table 3).

8.4 Demographics and baseline characteristics

Baseline characteristics will be summarised by treatment arm and overall. The variables shown in Table 4 will be presented. Categorical variables will be presented using counts and percentages, continuous variables will be presented with means and standard deviations or median and inter-quartile ranges as appropriate. No statistical significance testing will be used to test baseline imbalances between groups, but any noteworthy differences will be descriptively reported.

Table 4: Baseline characteristics

Variable	Scoring	Long splint (n=xx)	Short splint (n=xx)	Mini splint (n=xx)	Total (n=xx)
Site	Sheffield	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	….	….	….	….	….
Sex	Male	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Female	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Ethnicity[a]	White [b]	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Mixed/multiple ethnic groups [c]	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Asian/Asian British [d]	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Black/African/Caribbean/Black British [e]	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Other ethnic group [f]	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Prefer not to say	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Employment type (based on employment CRF)	Manual[g]	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Non-manual	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Not in paid employment				
	Unknown	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Dominant hand	Left	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Right	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)

	None	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Dominant hand is treated in trial	Yes	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	No	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Unknown	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Practice a religious faith that requires handwashing	Yes	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	No	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Prefer not to say	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Highest level of educational qualification	Primary or less	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Secondary	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Higher	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
	Other	xx(xx%)	xx(xx%)	xx(xx%)	xx(xx%)
Age (years)	Mean (sd)	xx(xx)	xx(xx)	xx(xx)	xx(xx)
	Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx	xx to xx
Time from injury to surgery (days)	Mean (sd)	xx(xx)	xx(xx)	xx(xx)	xx(xx)
	Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx	xx to xx
Time from surgery to day of randomisation/ day splint being fitted (days)[h]	Mean (sd)	xx(xx)	xx(xx)	xx(xx)	xx(xx)
	Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max	xx to xx	xx to xx	xx to xx	xx to xx

[a] *The main ethnic groups could be collapsed depending on the observed distribution.*

[b] *White: English/Welsh/Scottish/Northern Irish/British, Irish, Gypsy or Irish Traveller, Roma and Any other White background;*

[c] *Mixed/multiple ethnic groups: White and Black Caribbean, White and Black African, White and Asian, and Any other mixed/multiple ethnic groups background;*

[d] *Asian/Asian British: Indian, Pakistani, Bangladeshi, Chinese, and Any other Asian background;*

[e] *Black/African/Caribbean/Black British: African, Caribbean, and Any other Black/African/Caribbean/Black British background;*

[f] *Other ethnic group: Arab, and Any other ethnic group; Prefer not to say.*

[g] *For those in manual employment, details on the extent of their manual work will also be provided (0-25%, 26-50%, 51-75%, 76-100%)*

[h] *The date of randomisation and splint fitting is expected to be the same for most participants. Where this is not the case, the earlier date of randomisation or the day the splint was fitted will be used. We will provide details on the number of participants where these dates do not coincide, and by how many days they differ.*

8.4.1 Surgery details

Surgery details as outlined in
Table 5 will be summarised by treatment arm.

Table 5: Surgery details

Surgery details	• Injured hand (Left/Right) • Number of digits with flexor tendon injury (one/two) Details of each flexor tendon injury (the denominator will be the total number of injuries) • Injured digit (Index, middle, ring, little) • Zone of injury (I or II) • Tendon repaired[a] (FDP, FDS main tendon, FDS radial slip, FDS ulna slip) • Number of core strands • Gauge of core suture strand (3/0, 4/0, Other) • Technique (Adelaide, Cruciate, Modified Kessler, other) • Epitendinous suture (yes/no) • Associated injuries[a] (yes- Nerve, vessel, pulley, skin loss, fracture, non-repaired injured tendon(s)/no) • Structures repaired[a] (yes- RDN, UDN, RDA,UDA, Pulley/ no) Any other digit injuries in the same hand (yes/no) Number of other injured digits (1,2,3 or 4) Details of other injured digit (the denominator will be the total number of injuries) • Digit (thumb, index, middle, ring, little) • Type of injury (Laceration, extensor tendon, nerve, vessel, pulley, fracture, flexor tendon, skin loss) • Additional intervention (yes[a]- tendon repair, nerve repair, vessel repair, pulley repair, fracture stabilisation, wound closure only/no)

[a] *More than one option can be ticked, so the denominator here will be different to the rest of the tables.*

8.4.2 Employment at baseline

The details of each participant's employment status at baseline will be descriptively summarised in the baseline summaries section. At subsequent follow-ups the participants are asked about employment details and absence from work, this will be summarised separately from the baseline summaries.

Table 6: Employment details at baseline

Employment details at baseline	Currently in paid employment (yes/no)
	• Participant returned to work following their hand surgery (yes/no)
	  ○ Participant has returned to work on reduced hours (yes/no)
	• Participant considers their job to be manual work, involving their hands[a]? (yes/no)
	  ○ Percentage of the participant's time at work which includes manual work involving their hands? (0-25%, 26-50%, 51-75%, 76-100%)
	Participant is eligible for occupational sick pay? (Yes/No)

[a]*A person whose job involves physical work using their hands e.g. landscaping / construction / restaurant work / hotel work /childcare / manufacturing / healthcare professional.*

8.5 Intervention details

In this section details of the splint fitted will be summarised, by treatment arm. Together with summaries of any additional splints, splint removal and any additional exercises prescribed, as shown in Table 7. Follow-up data will also be summarised in this section, as shown in Table 8. No formal statistical testing will be done, only descriptive statistics will be reported here.

Table 6: Intervention details

Intervention details	Study splint fitting[a] • Grade of person fitting the splint (Band 5,6,7,8a, 8b or above) • Material used for splint (Ezeform, Manosplint (Carolina), Manosplint (Ohio), Orficast, Orfit, Polform, San splint, Wood cast, X-lite, other) • Time to fabricate and fit splint (minutes)- mean (SD) • Fitted with temperature sensor (yes/no) Replacement splint fitting • Main reason for replacement splint (Broke, caused pain, hygiene issues, irritated skin, lost, too tight, too loose, other)

	Additional splint outside of study protocol • Type of splint (long splint, short splint, mini splint, hand based POSI, forearm based POSI, dorsal PIPJ extension splint, volar digit gutter, other) • Main reason for splinting (facilitate tendon glide, maintain/improve extension, optimise alignment of affected digit, pain management, protect tendon, skin/scar contracture, to help swelling, to stretch a joint, other) • Grade of person fitting the splint (Band 3, 4, 5,6,7,8a, 8b or above) • Material used for splint (Ezeform, Manosplint (Carolina), Manosplint (Ohio), Orficast, Orfit, Polform, San splint, Wood cast, X-lite, other) • Time to fabricate and fit splint (minutes)- mean (SD) • Duration of splint wear (days)- mean (SD)

	Study splint removal
	• Splint removed before the prescribed length of time (yes/no)
	  o Reason for early removal (participant choice, surgical reason, investigator decision, other)
	  o Details will be provided where available

Follow-up summaries (for each follow-up time point)	• Method of follow-up (clinic, telephone)
	• Experienced problems with splint (yes/no) [six weeks only]
	  o Problems[b] (broke, caused pain, hygiene, irritated skin, too tight, too loose, other)
	• Participant had any additional surgery relating to their injury since the last study visit (Yes/No). Details of the additional surgeries will be provided.
	• Has the participant had any imaging scans relating to their injury since the last study visit (yes/no)
	  o Ultrasound, X-ray, MRI, other
	• Has the participant had any additional hand related appointments, since the last study visit? (yes/no). The number of appointments with doctors, nurses, hand therapists, hand therapy specialists and others will be summarised)

[a] *This information will be reported for each replacement splint a participant requires, with additional information for main reason for replacement splint.*

[b] *Multiple options can be ticked here, so please note the denominator will not be the same as the other categories.*

An additional table summarising the information surrounding splint wear, additional splints and splint removal will be summarised as shown in Table 7.

Table 7: Details of splint wear

	Long splint n=xx	Short splint n=xx	Mini splint n=xx	Total n=xx
Participants fitted for a splint[a]	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
-with additional splint(s)				
Study splint worn for 6 weeks	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
-with overlapping additional splint(s)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
-with additional splint(s) after	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Study splint removed early	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
-with overlapping additional splint(s)	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
-with additional splint(s) after	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Number of additional splints				
0	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
1	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
2	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)
Total number	xx (xx%)	xx (xx%)	xx (xx%)	xx (xx%)

[a] *This is the number of participants fitted for a splint from those who were randomised to a splint.*

8.6 Analysis of the primary outcome

The primary outcome is the mean post-randomisation PRWHE score, which is calculated as described in section 8.3.1. The principal analysis will be based on the ITT population.

Summary statistics (means and standard deviations, with median, interquartile range and range) for the PRWHE score at each time point will be provided by treatment arm. These summary statistics will be based on the observed data, without imputation for missing data.

We will graphically present the PRWHE data, using line graphs with the mean PRWHE over time (baseline, 6, 12, 26 and 52 weeks) by randomised group for all available data (also reporting the number of observations at each time point) and for the subset of participants with complete data (completed all 5 PRWHE assessments) to show how the outcome data changes over time for the three groups.

For completeness, we will also present descriptive statistics for the mean of the average post-randomisation PRWHE scores for each participant. Specifically, we will calculate a mean post-randomisation score (i.e. the average of all observed follow-up PRWHE scores) for each participant with at least one post-randomisation assessment. We acknowledge that these summaries will not exactly align with the way the differences in means are calculated in the primary analysis.

Treatment effects will be obtained from a linear mixed effects model incorporating all post randomisation PRWHE scores (at 6, 12, 26 and 52 weeks) as outcomes, with random effects for centre and subject (to account for the repeated observations per patient), and fixed effects for randomised group, time post-randomisation and baseline PRWHE score [27]. The regression coefficients will be estimated by restricted maximum likelihood (REML) with a small sample Kenward-Roger correction to the degrees of freedom for the hypothesis tests for the fixed effects.. Specifically, the model fitted will be a 3-level mixed effects model, where PRWHE scores are our outcome variable (level 1), which are repeated over the follow-up period for each individual (level 2), and the participants are nested within each site (level 3).

An interaction term between randomised group and time will be included. Time will be included in the model as a categorical (i.e. factor) variable, based on the protocol-stipulated time point for which data were collected. We shall assume an exchangeable correlation between the repeated measurements. We shall assume an exchangeable covariance structure between the random effects of subject and site. As specified above, in the analysis model, time will be based on the protocol stipulated follow-up time points (i.e. 6, 12, 26 and 52 weeks), rather than the actual time point at which the questionnaires were completed.*

Three treatment effect contrasts (marginal means, i.e., the differences in mean follow-up PREWHE scores between the groups) will be estimated and reported from the linear mixed model: 1) Long vs Short splint; 2) Long vs Mini splint; 3) Mini vs Short splint. We will estimate 98.3% confidence intervals for the three treatment effects in line with the sample size calculation. This model will include all patients who provide valid PRWHE data for at least one post-randomisation follow-up time point The handling of missing baseline data is described below.

The difference in PRWHE scores at each time point, 6, 12, 26 and 52 weeks (i.e. marginal effects at each time point) will also be derived from the same analysis model and presented with corresponding 98.3% confidence intervals.

In the event that the pre-specified analysis model does not converge, a simpler model will be used. Firstly, a simpler covariance structure between the random effects of subject and sites will be used (i.e. an independent covariance structure). If convergence issues persist, a model without the clustering within sites will be used.

*Note: Where data have been completed outside a protocol stipulated follow-up time point, data may be reassigned to another protocol stipulated follow-up time point if no other data has been completed for this time point. Specifically, if a PRWHE had entered into the database for the 12-week follow-up, but the completion date indicates that the observation does actually fall more appropriately into the 26-week follow-up window, then this data point will be analysed as 26-week data, provided no other data are available for the 26-week time window. If other data are available, then the data entered at 12 weeks will be reported and analysed for that time point.

Summary statistics (descriptive, as well as graphical) will be provided for the timing of questionnaire completion from randomisation (in days) to show timing of the data used in the analysis for each protocol-stipulated time point.

8.6.1 Missing primary outcome data

The number and proportion of available PRWHE scores at baseline and each follow-up time point will be presented overall and by treatment arm.

The number and proportion of participants with different missing data patterns will be presented overall and by treatment arm. The different patterns include:

- PRWHE score only available at baseline
- PRWHE score available at baseline and 6 weeks
- PRWHE score available at baseline, 6 weeks, and 12 weeks
- (all other possible combinations will also be presented)

Key baseline characteristics (including age, sex, baseline PRWHE score, number of digits with flexor tendon injury (one/two)) will be compared between participants with complete PRHWE follow-up vs. participants with some missing follow-up data. This may be further categorised based on the missing data patterns described above.

For the primary endpoint, all available data will be used to generate the mean PRWHE score over the follow-up, and all participants with at least one follow-up measure will be included in the analysis model. The linear mixed effects model will assume that data are missing at random, i.e., that the likelihood of data being missing is unrelated to the underlying value of the missing observations, conditional on the data included in the analysis model. If a participant does not have outcome data available at all protocol stipulated follow-up time points, the mean PRWHE score will be generated from the available follow-up time points for this participant, and implicit imputation of their missing data, based on the model parameters.

Additional analyses of the primary endpoint are listed below

8.6.2 Missing baseline data

For the primary analysis model, and all other statistical models, missing baseline data will be handled by mean imputation for continuous data, and mode imputation for categorical data, done by recruiting centre to ensure all participants with available outcome data are included in the analyses.
Summary statistics of baseline data will be restricted to the observed data only. Analysis models will be adjusted for baseline data within which missing data have been mean-imputed, as described above. This approach will be clarified in all relevant tables and other study output.

8.6.3 Varying the analysis population

We will repeat the primary analysis model for the PP population. Between group differences for both the primary endpoint (mean post-randomisation PRWHE score) and the results at each time point will be shown for the PP population.

8.6.4 Sensitivity analysis for missing data

It is also important to assess the potential impact of missing data on the trial conclusions. The primary analyses assumed a missing at random mechanism, and therefore missing not at random (MNAR) scenarios will be investigated in the sensitivity analysis. No imputation will be performed for participants that have died before the relevant post randomisation follow-up assessment. Specifically, we will examine if trial conclusions would change if we assumed that missing outcomes are up to 6 points on the PRWHE score worse (i.e., higher) than expected under a missing at random scenario. Six-points were chosen as this was the clinically important difference used in the sample size calculation.

Specifically, we will use a multiple imputation (multiple imputation by chained equations – MICE – with predictive mean matching) model for missing PRWHE outcomes that uses the same covariates as the primary analysis model and produces imputations separately by treatment arm. 30 imputations will be generated. Imputed data will be replaced by values 6 points higher than their original value. Treatment estimates for the mean PRWHE scores will be obtained from these data.

The focus of this analysis is to examine if the trial conclusions change as a result of the different assumptions about the missing data mechanism. We will not investigate scenarios where the MNAR scenarios are only applied to one trial arm, as these seem clinically less plausible. The proposed sensitivity analysis for missing data is expected to differ most from the primary analysis if there are differential rates of missing data in the different trial arms.

The primary analysis model will be used to derive treatment effects and corresponding 98.3% confidence intervals under this MNAR scenario for the primary endpoint. These will be compared to the results from the primary analysis. If the trial conclusions, i.e., the assessment of the comparative effectiveness of the trial splints, are consistent between the primary and MNAR sensitivity analysis, we will conclude that the conclusions of the trial are robust to the assumptions made about missing data.

If the trial conclusions differ between the primary and MNAR sensitivity analysis, we will conclude that the conclusions may not be robust to the assumptions made about missing data and interpret the trial accordingly.

8.6.5 Sensitivity analysis considering the timing of assessments

The primary analysis uses the data according to the protocol stipulated visits they have been entered for. If more than 25% of assessments fall outside reasonable visit windows (+/- 1 week for the 6 week visit, +/- 2 weeks for the 12 week visit, +/- 3 week for the 12 week visit, +/- 4 weeks for the 52 week visit), then additional a sensitivity analysis will be performed excluding these data.

8.6.6 Sub-group analysis

The main objective of this section is to explore heterogeneity in the intervention effects in pre-specified subpopulations described in Section 7.2. The sub-group analysis will be completed on the ITT population only. The analysis will be based on the primary analysis model with the addition of an interaction term between the treatment and subgroup to assess the stability of the result in different populations. For simplicity, the interaction between treatment and time will be excluded from this model. Treatment effect estimates for the mean PRWHE post-randomisation scores with 98.3% confidence intervals will be calculated for each sub-group. The results will be displayed graphically using Forest plots. No p-values will be presented, as it is acknowledged that the trial is not powered for these subgroup analyses, and they are considered exploratory.

8.7 Analysis of secondary outcomes

The scores on the repeated continuous secondary outcomes (PEM, EQ-5D-5L, Range of movement, and Grip strength) will be compared between the randomised groups using a similar longitudinal mixed effects linear regression model as described for the analysis of the primary outcome. Treatment effects and corresponding 98.3% confidence intervals for each protocol stipulated follow-up points will also be presented.
Note that no overall score exists for the WPAI, and relevant questions will be analysed using descriptive summaries only.

The categorical responses from the global rating of change question at each post randomisation time-point will be summarised by randomised group and compared between the groups using a chi-squared test or Fisher's exact test, depending on the data distribution.

Adherence to the randomised splint treatment based on the temperature sensors during the first 5-weeks post randomisation, when the splints are to be worn continuously, will be estimated from the heat sensor in the splint. Adherence will be summarised for each randomised group using the mean number of hours per day wearing the splint, and mean adherence compared between the groups using a linear regression model. As with the primary outcome, three treatment effect contrasts, and their associated confidence intervals will be estimated and reported from the model: 1) Long vs Short splint; 2) Long vs Mini splint; 3) Mini vs Short splint. Adherence data will be presented over the first five weeks,

and separately for each week. Section 8.3.7 outlines the details of the methods used to define adherence from the temperature sensor data. In addition the following analyses may also be presented:

- Splint adherence (mean daily wear time) over the full 6-week follow-up period, for each trial arm
- Graphical summaries of wear time over the 6-week follow-up period, for each trial arm
- Summaries of adherence by arm, for each week during the 6-week follow-up period

The proportion of participants meeting different adherence pattern (to be defined and approved by the Trial Management Group) will also be presented.

Details of any litigation/compensation for injury will be a binary (yes/no) response at the 52-week follow-up visit. This will be summarised as the number and percentage in each arm and overall.

Preferences for splint attributes (stated and revealed) and splint acceptability. These questions consist of a series of questions, with responses rated from 0 to 3. These data will be summarised in a tabular form as categorical variables, with the number and percentage for each rating for each question.

The details of changes to participants' employment status will be summarised descriptively, in the same way as described in section 8.4.2. An additional summary of any absences from work and reduced hours as detailed on the employment log CRF will also be descriptively summarised.

8.8 Protocol non-compliances

The number (and percentage) of patients with major and minor protocol deviations will be summarised by treatment group with details of type of deviation provided. No formal statistical testing will be undertaken between the groups.

8.9 Recording and Analysis of Safety outcomes

All AEs will be assessed by site staff for relatedness and seriousness (see seriousness criteria in section 10.1 of the study protocol [4]). Non-serious AEs will only be recorded where they involve the injured hand/ upper limb or are considered possibly related to the injury or its treatment. AEs will be recorded on the adverse event form within the participant CRF, and in the medical notes.

All AEs which meet the criteria for seriousness (see section 10.1 of the study protocol [4]) will be recorded in the adverse event form and in the medical notes, regardless of relatedness. SAEs will require more detailed information to be recorded. For the purposes of this study, flexor tendon rupture is considered a medically significant event, and any incidents will be recorded as SAEs.

Study specific exemptions

The following events are expected and, should they meet the criteria for seriousness, do not require reporting to CTRU within 24 hours, but should be reported within the time frames specified below.

Within 72 hours, for ongoing safety monitoring purposes:

1. Flexor tendon rupture

Before the participants next scheduled follow-up visit:

2. Local pressure areas as a result of the splint, plaster of paris or dressings.
3. Infection leading to:
    a. Treatment with oral antibiotics.
    b. Treatment with intravenous antibiotics either as an in-patient or outpatient.
    c. Requiring surgical washout.
4. Stiffness of the affected hand requiring surgery e.g. tenolysis/arthrolysis.
5. Scar issues e.g. hypersensitivity/ hypertrophic scars.
6. Delayed wound healing requiring an extended period of dressing.
7. Complex regional pain syndrome.
8. Fixed flexion deformity of the proximal interphalangeal joint (PIPj) or distal interphalangeal joint (DIPj) requiring additional splintage.

In the statistical report, adverse events from the day of randomisation will be recorded.

Summary measures will be presented by treatment group as the number and percentage of participants reporting an AE/SAE, as well as the total number of AE/SAEs reported. This will be conducted on the ITT population, as shown in Table 8 to Table 10. No formal statistical testing will be undertaken. The summaries in in Table 8 to Table 10 will be repeated on only the adverse events categorised as 'reasonable possibility of being related' to the original injury/treatment. Additional graphical summaries of the AEs/SAEs may be presented, this is left to the discretion of the trial statisticians.
The proportion of participants with at least one AE/SAE, as well as the number of participants with AEs of special interest (Flexor tendon rupture of primary repair, Fixed Flexion Deformity, Stiffness, Infections) may be presented, if sufficient numbers of such events are recorded. Unadjusted odds ratios and risk differences will be obtained, with corresponding confidence intervals.

Table 8: Summary of adverse events

Variable	Scoring		Mini splint (n=xx)	Short splint (n=xx)	Long splint (n=xx)
Number of all AEs	Including repeated events		xx	xx	xx
Participants with ≥1 AE	Any expected AE		xx(xx%)	xx(xx%)	xx(xx%)
AEs per participant	0		xx(xx%)	xx(xx%)	xx(xx%)
	1		xx(xx%)	xx(xx%)	xx(xx%)
	2		xx(xx%)	xx(xx%)	xx(xx%)
	3		xx(xx%)	xx(xx%)	xx(xx%)
	4		xx(xx%)	xx(xx%)	xx(xx%)
	5+				
AEs per participant	Mean (sd)		xx(xx)	xx(xx)	xx(xx)
	Median (IQR)		xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
	Min to Max		xx to xx	xx to xx	xx to xx
AE category⁺	Complex regional pain syndrome		xx(xx%)	xx(xx%)	xx(xx%)
	Fixed flexion deformity		xx(xx%)	xx(xx%)	xx(xx%)
	Scar issues^		xx(xx%)	xx(xx%)	xx(xx%)
	Delayed wound healing		xx(xx%)	xx(xx%)	xx(xx%)
	Infection		xx(xx%)	xx(xx%)	xx(xx%)
	Stiffness		xx(xx%)	xx(xx%)	xx(xx%)
	Flexor tendon rupture of primary repair*		xx(xx%)	xx(xx%)	xx(xx%)
	Local pressure areas		xx(xx%)	xx(xx%)	xx(xx%)
	Other^		xx(xx%)	xx(xx%)	xx(xx%)
Outcome	Recovered		xx(xx%)	xx(xx%)	xx(xx%)
	Recovering		xx(xx%)	xx(xx%)	xx(xx%)
	Not recovered		xx(xx%)	xx(xx%)	xx(xx%)
	Recovered with Sequelae		xx(xx%)	xx(xx%)	xx(xx%)
	Fatal		xx(xx%)	xx(xx%)	xx(xx%)
	Unknown		xx(xx%)	xx(xx%)	xx(xx%)
Relationship to original injury/treatment	Reasonable possibility of being related		xx(xx%)	xx(xx%)	xx(xx%)
	No reasonable possibility of being related		xx(xx%)	xx(xx%)	xx(xx%)
	Not assessable		xx(xx%)	xx(xx%)	xx(xx%)
	Where reasonable possibility of AE being related to original injury/ treatment⁺:	Splint			
		Surgery	xx(xx%)	xx(xx%)	xx(xx%)
		Injury	xx(xx%)	xx(xx%)	xx(xx%)
	Where no reasonable possibility of being related	Involves affected	xx(xx%)	xx(xx%)	xx(xx%)

		hand/upper limb			
		Doesn't involve the affected hand/upper limb	xx(xx%)	xx(xx%)	xx(xx%)
AE grading[$]	Rating provided on score of 0-7[$]	Mean (SD)	xx(xx)	xx(xx)	xx(xx)
		Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
		Min to Max	xx to xx	xx to xx	xx to xx

[+]*Categories are not mutually exclusive.*

**Details of flexor tendon rupture of primary repairs will be summarised in addition to this table, see Table 9 below.*

[$]*AE grades are obtained via a central review. Higher scores indicate more severe outcomes. The exact definition of the scores will be provided in the statistical report. The scores may also be presented as categorical variables where appropriate.*

^Scar issues will be broken down further into scar sensitivity issues and tendon tethering during a central review of the adverse events. Further subcategories may be added as a result of this review. Additional categories will be used if deemed appropriate in the central review of events reported in this category.

Table 9: Summary of flexor tendon rupture of primary repairs

Variable			Mini splint (n=xx)	Short splint (n=xx)	Long splint (n=xx)
Time from randomisation to rupture (days)		Mean (SD)	xx(xx)	xx(xx)	xx(xx)
		Median	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
		Min to Max	xx to xx	xx to xx	xx to xx
Flexor tendon rupture of primary repairs	Digit	Index	xx(xx%)	xx(xx%)	xx(xx%)
		Middle	xx(xx%)	xx(xx%)	xx(xx%)
		Ring	xx(xx%)	xx(xx%)	xx(xx%)
		Little	xx(xx%)	xx(xx%)	xx(xx%)
	Tendon[a]	FDP	xx(xx%)	xx(xx%)	xx(xx%)
		FDS	xx(xx%)	xx(xx%)	xx(xx%)

[a] *More than one tendon can be ticked, so these categories are not mutually exclusive.*

Depending on the data, this table may need amending to allow for ruptures involving more than one digit.

Table 10: Serious adverse event summary

Variable	Scoring		Mini splint (n=xx)	Short splint (n=xx)	Long splint (n=xx)

Number of all SAEs	Including repeated events		xx(xx%)	xx(xx%)	xx(xx%)
Participants with ≥1 SAE	Any expected AE		xx(xx%)	xx(xx%)	xx(xx%)
AEs per participant	0		xx(xx%)	xx(xx%)	xx(xx%)
	1		xx(xx%)	xx(xx%)	xx(xx%)
	2		xx(xx%)	xx(xx%)	xx(xx%)
	3		xx(xx%)	xx(xx%)	xx(xx%)
	4		xx(xx%)	xx(xx%)	xx(xx%)
	5+				
AEs per participant	Mean (sd)		xx(xx)	xx(xx)	xx(xx)
	Median (IQR)		xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
SAE category	Complex regional pain syndrome		xx(xx%)	xx(xx%)	xx(xx%)
	Fixed flexion deformity		xx(xx%)	xx(xx%)	xx(xx%)
	Scar issues^		xx(xx%)	xx(xx%)	xx(xx%)
	Delayed wound healing		xx(xx%)	xx(xx%)	xx(xx%)
	Infection		xx(xx%)	xx(xx%)	xx(xx%)
	Stiffness		xx(xx%)	xx(xx%)	xx(xx%)
	Flexor tendon rupture of primary repair*		xx(xx%)	xx(xx%)	xx(xx%)
	Local pressure areas		xx(xx%)	xx(xx%)	xx(xx%)
	Other^		xx(xx%)	xx(xx%)	xx(xx%)
Relationship to original injury/treatment	Reasonable possibility of being related		xx(xx%)	xx(xx%)	xx(xx%)
	No reasonable possibility of being related		xx(xx%)	xx(xx%)	xx(xx%)
	Not assessable		xx(xx%)	xx(xx%)	xx(xx%)
	Where reasonable possibility of AE being related to original injury/ treatment[+]:	Splint			
		Surgery	xx(xx%)	xx(xx%)	xx(xx%)
		Injury	xx(xx%)	xx(xx%)	xx(xx%)
	Where no reasonable possibility of being related	Involves affected hand/upper limb	xx(xx%)	xx(xx%)	xx(xx%)
		Doesn't involve the affected hand/upper limb	xx(xx%)	xx(xx%)	xx(xx%)
Assessment	Intensity	Mild	xx(xx%)	xx(xx%)	xx(xx%)
		Moderate			
		Severe			
	Action taken	None	xx(xx%)	xx(xx%)	xx(xx%)
		Stop splinting	xx(xx%)	xx(xx%)	xx(xx%)
		Other	xx(xx%)	xx(xx%)	xx(xx%)
		Specific treatment	xx(xx%)	xx(xx%)	xx(xx%)
	Outcome	Recovered	xx(xx%)	xx(xx%)	xx(xx%)
		Recovering	xx(xx%)	xx(xx%)	xx(xx%)
		Not recovered	xx(xx%)	xx(xx%)	xx(xx%)
		Recovered with Sequelae	xx(xx%)	xx(xx%)	xx(xx%)
		Fatal	xx(xx%)	xx(xx%)	xx(xx%)
		Unknown	xx(xx%)	xx(xx%)	xx(xx%)

	Seriousness	Death	xx(xx%)	xx(xx%)	xx(xx%)
		Life threatening	xx(xx%)	xx(xx%)	xx(xx%)
		Inpatient hospitalisation	xx(xx%)	xx(xx%)	xx(xx%)
		Prolonged hospitalisation	xx(xx%)	xx(xx%)	xx(xx%)
		Persistent or significant disability or incapacity	xx(xx%)	xx(xx%)	xx(xx%)
		Considered medically significant by the investigator	xx(xx%)	xx(xx%)	xx(xx%)
	Frequency	Isolated	xx(xx%)	xx(xx%)	xx(xx%)
		Intermittent	xx(xx%)	xx(xx%)	xx(xx%)
		Continuous	xx(xx%)	xx(xx%)	xx(xx%)
		Unknown	xx(xx%)	xx(xx%)	xx(xx%)
Hospitalisation	Participant hospitalised	Yes	xx(xx%)	xx(xx%)	xx(xx%)
		No	xx(xx%)	xx(xx%)	xx(xx%)
Central review	Expected SAE	Yes	xx(xx%)	xx(xx%)	xx(xx%)
		No	xx(xx%)	xx(xx%)	xx(xx%)
AE grading[$]	Rating provided on score of 0-7[$]	Mean (SD)	xx(xx)	xx(xx)	xx(xx)
		Median (IQR)	xx(xx to xx)	xx(xx to xx)	xx(xx to xx)
		Min to Max	xx to xx	xx to xx	xx to xx

**Scar issues will be broken down further into scar sensitivity issues and tendon tethering during a central review of the adverse events. Further subcategories may be added as a result of this review. Additional categories will be used if deemed appropriate in the central review of events reported in this category.*
[$]AE grades are obtained via a central review. Higher scores indicate more severe outcomes. The exact definition of the scores will be provided in the statistical report. The scores may also be presented as categorical variables where appropriate.

For AE and SAE categories of special interest (Flexor tendon rupture of primary repair, Fixed Flexion Deformity, Stiffness, Infections), information on events per participant, outcome, relationship to original injury/treatment, rating will be presented separately.
Separate summaries will be presented for all observed events, and the most severe event per participant.

8.10 Exploratory analysis

In addition to the analysis outlined in the previous sections, further exploratory analysis will be completed. This will be conducted on the ITT population.

8.10.1 The primary outcome and adherence data

This section details analyses which will be used to explore the effect of adherence, with the trial treatment, on the primary outcome. This analysis will be exploratory and will include, but is not limited to the following summaries:

- Graphical summaries presenting mean daily splint wear time (over the 5-week follow-up period) as measured by the temperature sensors against mean PRWHE score (over the post-randomisation follow-up period) for each treatment arm.

8.10.2 Splint acceptability/attributes and adherence data

Summaries of splint adherence against splint acceptability and splint attributes will be produced. The details of these summaries will depend on the distribution of the data, however no formal statistics will be done on these data as part of this analysis plan. These data will be used in the mediation analysis, which is described separately.

8.10.3 Staff demographic analysis

As detailed in the protocol, anonymised sets of summary demographics for site staff within hand units will be requested from recruiting sites, to explore whether the diversity of recruiting staff demographics has an impact on the diversity of participants recruited to the study. This will include staff ethnicity and sex. The summaries will be presented in a similar manner to the baseline demographics as shown in Table 4.

8.10.4 Social deprivation of participants

Consented participants postcodes will be collected to demonstrate the demographic spread of participants and to explore the impact of social deprivation on the study outcomes. The postcodes will be used to calculate the Index of Multiple Deprivation Decile, which is scored from 1 (most deprived) to 10 (least deprived), this will then be summarised appropriately depending on the distribution of the data.

8.11 Mediation analysis

In addition to the analysis outlined for the process evaluation in section 12.1 of the study protocol, a mediation analysis will be performed, which will be detailed in a separate analysis plan. A secondary mediation analysis will investigate putative mediation factors (stated and revealed preferences, adherence) using direct acyclic graphs and structural equation models to test for mediation of splint on pain and function (PRWHE) through the factors. Analyses will adjust for baseline measures of the factor and possible measured confounders/moderators (e.g. age, sex). Possible mediation factors will be tested for by testing interactions between baseline factors and treatment on treatment response and safety outcomes. The full details of this analysis will be outlined in a separate mediation analysis plan.

8.12 Model checking

Data will be visually inspected prior to analysis, and residuals inspected after model fit where appropriate. Model assumptions will be checked via graphical methods (e.g. histograms of residuals and scatterplots of residuals vs. covariates). Influential observations and outliers will also be investigated and sensitivity analyses at the discretion of the trial statistician will be undertaken and reported.

Any changes to the model specification and their justification will be described in the final report.

8.13 Statistical Software

The final analysis will be carried out using any suitable packages such as R or STATA.

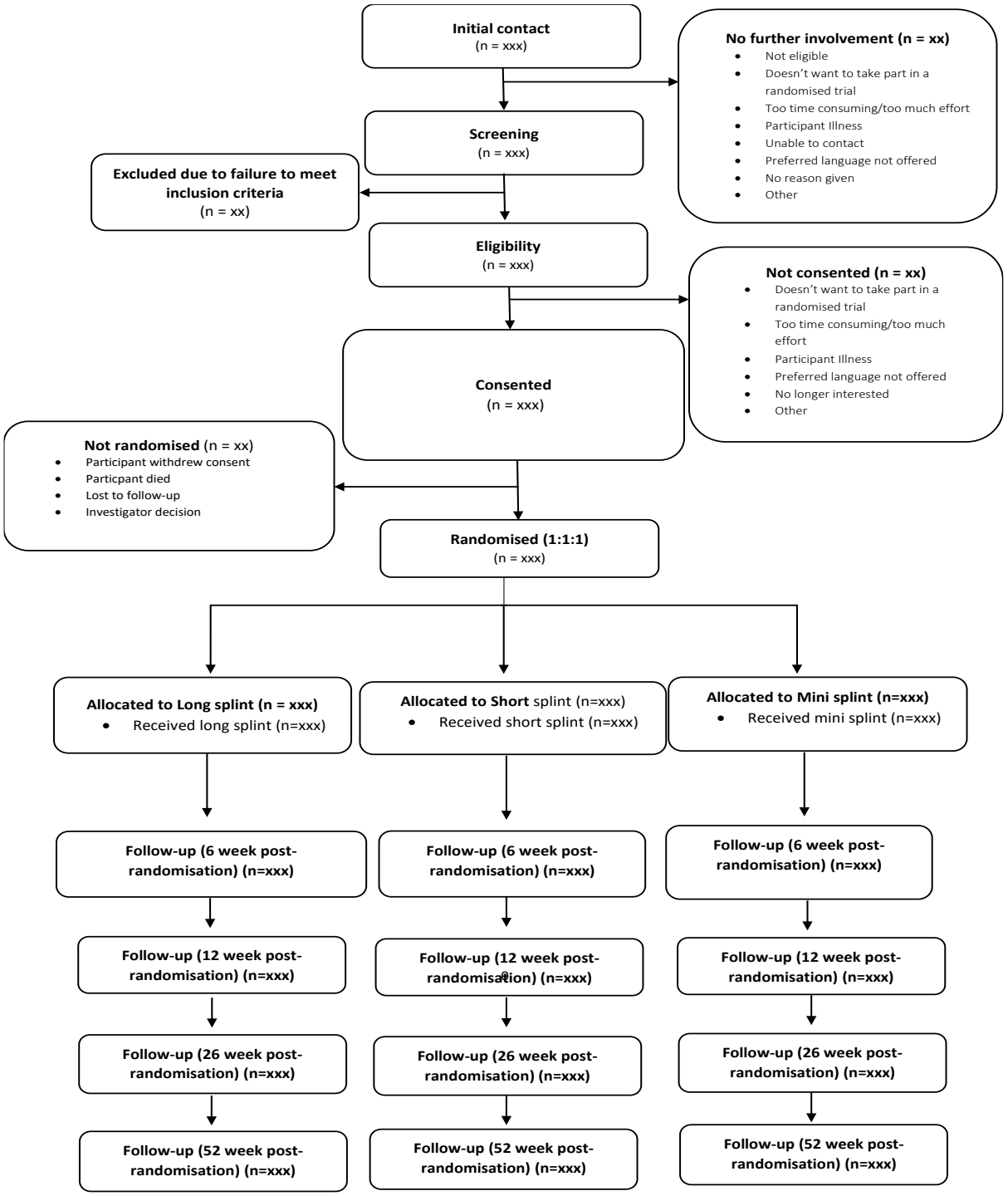
9 IMPLEMENTATION OF THE SAP

This SAP will be used as a work description of the Trial Statistician in consultation with the Senior Trial Statistician. No analyses by treatment arm will be undertaken by the trial statistician until after the sign-off of this SAP by relevant personnel. Data will be released by the data management after sign-off to the Trial Statistician, given a window period to query any spurious data and initiate data lock before actual analysis. At this point, no changes will be allowed on the database. Unblinded DMEC reports will be produced by the Sheffield CTRU data management team.

10 APPENDIX

10.1 Example consort diagram

Figure 1: Template consort diagram



NB: The number of participants at each follow-up are based on the number of valid PRHWE scores available. We will also indicate how many participants have at least one post-randomisation PRWHE score available. Reasons for unavailable scores will be presented where available.

The number of participants in the ITT and PP analysis populations will be presented.

11 REFERENCES

[1] ICH, "ICH E9: Statistical Principles for Clinical Trials," 1998.

[2] C. Gamble *et al.*, "Guidelines for the content of statistical analysis plans in clinical trials," *JAMA - J. Am. Med. Assoc.*, vol. 318, no. 23, 2017, doi: 10.1001/jama.2017.18556.

[3] "ST001 The Statistical Analysis Plan _v6.0 2021".

[4] T. Grieve and E. Bamford, "FIRST trial protocol version 2.0," pp. 1–39, 2022.

[5] L. S. Robinson, M. Sarkies, T. Brown, and L. O'Brien, "Direct, indirect and intangible costs of acute hand and wrist injuries: A systematic review," *Injury*, vol. 47, no. 12. 2016. doi: 10.1016/j.injury.2016.09.041.

[6] "Https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity, Hospital admitted patient care activity. In: NHS Digital [Internet]." available: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-admitted-patient-care-activity (accessed Jan. 21, 2022).

[7] M. Gustafsson, L. Hagberg, and M. Holmefur, "Ten years follow-up of health and disability in people with acute traumatic hand injury: Pain and cold sensitivity are long-standing problems," *J. Hand Surg. Eur. Vol.*, vol. 36, no. 7, 2011, doi: 10.1177/1753193411408186.

[8] E. Ladds, N. Redgrave, M. Hotton, and M. Lamyman, "Systematic review: Predicting adverse psychological outcomes after hand trauma," *J. Hand Ther.*, vol. 30, no. 4, 2017, doi: 10.1016/j.jht.2016.11.006.

[9] J. S. Schier and J. Chan, "Changes in Life Roles After Hand Injury," *J. Hand Ther.*, vol. 20, no. 1, 2007, doi: 10.1197/j.jht.2006.10.005.

[10] R. Mehrzad, V. Mookerjee, S. Schmidt, C. C. Jehle, E. Kiwanuka, and P. Y. Liu, "The Economic Impact of Flexor Tendon Lacerations of the Hand in the United States," *Ann. Plast. Surg.*, vol. 83, no. 4, 2019, doi: 10.1097/SAP.0000000000001950.

[11] H. E. Rosberg, K. S. Carlsson, S. Höjgård, B. Lindgren, G. Lundborg, and L. B. Dahlin, "What determines the costs of repair and rehabilitation of flexor tendon injuries in zone II? A multiple regression analysis of data from southern Sweden," *J. Hand Surg. Am.*, vol. 28 B, no. 2, 2003, doi: 10.1016/S0266-7681(02)00352-2.

[12] "Current Research. In: Pulvertaft Hand Centre within the Royal Derby Hospital - Derby Hospitals NHS [Internet]. 31 Aug 2017".

[13] "BSSH standards of care in hand trauma. [cited 21 Jan 2022].".

[14] E. Deniz, K. Ayse, K. Sehim, D. Mehmet, S. Aysun, and E. Süreyya, "Postoperative Management of Flexor Tendon Repair in Zone 2.," *J. Phys. Ther. Sci.*, vol. 12, no. 1, 2000, doi: 10.1589/jpts.12.63.

[15] E. J. May and K. L. Silfverskiöld, "Rate of recovery after flexor tendon repair in zone II a prospective longitudinal study of 145 digits," *Scand. J. Plast. Reconstr. Surg. Hand Surg.*, vol. 27, no. 2, 1993, doi: 10.3109/02844319309079789.

[16] F. Sandford, N. Barlow, and J. Lewis, "A Study to Examine Patient Adherence to Wearing 24-Hour Forearm Thermoplastic Splints after Tendon Repairs," *J. Hand Ther.*, vol. 21, no. 1, 2008, doi: 10.1197/j.jht.2007.07.004.

[17] J. B. Tang, "Recent evolutions in flexor tendon repairs and rehabilitation," *Journal of Hand Surgery: European Volume*, vol. 43, no. 5. 2018. doi: 10.1177/1753193418773008.

[18] M. J. Hirth, J. W. Howell, and L. O'Brien, "Relative motion orthoses in the management of various hand conditions: A scoping review," *J. Hand Ther.*, vol. 29, no. 4, 2016, doi: 10.1016/j.jht.2016.07.001.

[19] S. L. Filan, "The effect of workers' or third-party compensation on return to work after hand surgery," *Med. J. Aust.*, vol. 165, no. 2, 1996, doi: 10.5694/j.1326-5377.1996.tb124853.x.

[20] J. C. MacDermid, A. Donner, R. S. Richards, and J. H. Roth, "Patient versus injury factors as predictors of pain and disability six months after a distal radius fracture," *J. Clin. Epidemiol.*, vol.

55, no. 9, 2002, doi: 10.1016/S0895-4356(02)00445-6.
[21] J. C. MacDermid, "The PRWE/PRWHE update," *J. Hand Ther.*, vol. 32, no. 2, 2019, doi: 10.1016/j.jht.2019.01.001.
[22] T. P. Clark, B. C. Kahan, A. Phillips, I. White, and J. R. Carpenter, "Estimands: Bringing clarity and focus to research questions in clinical trials," *BMJ Open*, vol. 12, no. 1, 2022, doi: 10.1136/bmjopen-2021-052953.
[23] G. F. Moore *et al.*, "Process evaluation of complex interventions: Medical Research Council guidance," *BMJ*, vol. 350, 2015, doi: 10.1136/bmj.h1258.
[24] F. Davidoff, "Understanding contexts: How explanatory theories can help," *Implement. Sci.*, vol. 14, no. 1, 2019, doi: 10.1186/s13012-019-0872-8.
[25] Machin D, Campbell MJ, Tan SB, *Repeated Outcome Measures. Sample Sizes for Clinical, Laboratory and Epidemiology Studies.*
[26] "MacDermid JC. The patient-rated wrist evaluation (PRWE)© user manual. Hamilton: McMaster University."
[27] Walters. SJ., *Quality of Life Outcomes in Clinical Trials and Health-Care Evaluation: A Practical Guide to Analysis and Interpretation.* John Wiley & Sons; 2009.
[28] Sheffield CTRU, "GV001: Trial Management Group," 2016.
[29] Sheffield CTRU, "GV002: Trial Steering Committee," 2016.
[30] Sheffield CTRU, "GV003: Data Monitoring and Ethics Committee," 2016.
[31] E. Juszczak, D. G. Altman, S. Hopewell, and K. Schulz, "Reporting of Multi-Arm Parallel-Group Randomized Trials: Extension of the CONSORT 2010 Statement," *JAMA - J. Am. Med. Assoc.*, vol. 321, no. 16, 2019, doi: 10.1001/jama.2019.3087.
[32] M. Zwarenstein *et al.*, "Improving the reporting of pragmatic trials: An extension of the CONSORT statement," *BMJ*, vol. 337, no. 7680, 2008, doi: 10.1136/bmj.a2390.
[33] S. Senn, "Testing for baseline balance in clinical trials," *Stat. Med.*, vol. 13, no. 17, 1994, doi: 10.1002/sim.4780131703.
[34] M. J. Knol, R. H. H. Groenwold, and D. E. Grobbee, "P-values in baseline tables of randomised controlled trials are inappropriate but still common in high impact journals," *European Journal of Preventive Cardiology*, vol. 19, no. 2. 2012. doi: 10.1177/1741826711421688.
[35] M. R. de Boer, W. E. Waterlander, L. D. J. Kuijper, I. H. M. Steenhuis, and J. W. R. Twisk, "Testing for baseline differences in randomized controlled trials: An unhealthy research behavior that is hard to eradicate," *Int. J. Behav. Nutr. Phys. Act.*, vol. 12, no. 1, 2015, doi: 10.1186/s12966-015-0162-z.
[36] J. J. Dias, B. Bhowal, C. J. Wildin, and J. R. Thompson, "Assessing the outcome of disorders of the hand. Is the patient evaluation measure reliable, valid, responsive and without bias?," *J. Bone Jt. Surg. - Ser. B*, vol. 83, no. 2, 2001, doi: 10.1302/0301-620X.83B2.10838.
[37] "https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l."
[38] "Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English Population Study," 2020.