PASHiOn

Personalised Against Standard High tibial Osteotomy, a prospective multi-centre randomised controlled trial

Statistical Analysis Plan

Version v1.0 – 23Apr2025

Based on Protocol version V6.0 – 04 Sep 2024
IRAS Project ID: 280261

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 Versus Arthritis funded prospective multi-centre randomised controlled trial comparing personalised against standard high tibial osteotomy (PASHiOn). 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 (1).

The analysis strategy will be available on request when the principal papers are submitted for publication in a journal. 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.

1.1 Key personnel

List of key people involved in the preparing, reviewing and approving the SAP and subsequent reports.

Author:
Alexander Thomas

Reviewers:
David Smith
David Beard
Alisdair MacLeod

Approvers:
Jonathan Cook
Richie Gill

1.2 Changes from previous version of SAP

A summary of key changes from earlier versions of SAP, with relevance to protocol changes that have an impact on the design, definition, sample size, data quality/collection and analysis of the outcomes will be provided. Include protocol version number and date.

Version number Issue date	Author of this issue	Protocol Version & Issue date	Significant changes from previous version together with reasons
V1.0_DDMonYYYY		V6.0 04 Sep 2024	Not applicable as this is the 1st issue

2. BACKGROUND AND OBJECTIVES

Up to 28% of the UK population over 40 have knee pain, with half of these people having radiographic osteoarthritis. This leads to a high demand for knee replacement surgeries. It is important to recognise that one third of knee replacements are now performed in patients younger than 65 indicating the disease burden and need for treatment in a younger group of osteoarthritis sufferers. The current treatment choice is a total knee replacement; however, it is not recommended for the earlier stages of knee osteoarthritis and it does not last as long in younger patients. Patients aged 64 years or less at time of primary surgery have double the revision rate of those aged between 65 and 74.

An alternative to knee replacement is high tibial osteotomy, commonly referred to as HTO. In HTO, the native joint is preserved by realigning the tibia to off-load the worn areas of the knee. A well performed HTO can delay the need for knee replacement by 10 years. There are barriers in place preventing wide adoption of HTO – including the challenging operative technique. The gold standard for assessing lower limb alignment in the coronal plane is to measure the hip-knee-ankle angle on full wight bearing x-rays. This measure forms our primary endpoint as we will analyse the absolute difference between planned and achieved coronal plane correction.

A list of objectives (as documented in the protocol) is given in Table *1*Table *1* along with the outcome measures that will be used to answer the objective.

Table 1 List of objectives, outcome measures and the time the measure will be taken

	Objectives	Outcome Measures	Time point
Primary	To establish if digitally planned personalised HTO surgery (TOKA) increases the accuracy of bone correction in comparison to conventional HTO surgery.	The difference between planned and achieved coronal plane correction in terms of the hip-knee-ankle angle, (as measured from full length weight bearing x-rays)	9 months post randomisation*
Secondary	To assess other markers of recovery and function to further inform the primary objective. Treatment arms will be compared in terms of:	Number of patients (%) achieving a pre-specified difference between planned and achieved coronal plane correction	9 months post randomisation*
		Change in coronal plane location of peak loading during gait, in comparison to baseline, measured by Video Vector	9 months post randomisation*

		Operative time in minutes	At procedure
		Patient reported quality of life: • KOOS • EQ-5D-5L • Visual analogue scale (VAS) pain score • UCLA Activity Score • Satisfaction and transition • Health resource use	9 months post randomisation*
	Examine relationship between correction, change in loading and clinical outcomes	Knee range of motion (ROM)	9 months post randomisation*
		To determine rates of AEs and ADEs	
		Calculated change in knee loading resulting from osteotomy correction from computer models of each patient	

*Whilst all analyses are formally defined from post randomisation a key trial conduct milestone is the time of surgery. Note patients will have variable waiting times for surgery after randomisation (up to 3 months within protocol) and assessments are required to align with routine post surgery time points within the established clinical pathway. A minimum of 4 months post surgery is set as the first acceptable (clinically valid) post surgery time point. This is accounted for in the design and analysis (see study design for further detail and assessment alignment). It allows for a ”window” of appropriate post randomisation data collection of about 3 months (between months 4, 5 and 6 post surgery) and can account for potential and unpredictable surgical delays whilst remaining faithful to follow up from randomisation.

Table 2: Estimand-to-analysis table

Primary Objective: To establish if digitally planned personalised HTO surgery (TOKA) increases the accuracy of bone correction in comparison to conventional HTO surgery.	
Estimand: *Possible issues surrounding the waiting times for surgery and patients not receiving HTO prior to the date of primary outcome collection or not receiving the plate they were randomised to.*	
Treatment: *Control – Generic HTO device* *Treatment – Personalised HTO device*	
Estimand	***Analysis***
Target population *Patients in need of HTO that meet the eligibility criteria given in the protocol.*	***Analysis set***

	Intention-to-treat analysis set with available primary outcome data 9 months post randomisation. Though participants who did not undergo HTO will be excluded.
Variable	***Outcome measure***
Achieved coronal plane correction in terms of the hip-knee-ankle as measured on a full leg weight bearing X-Ray at 9 months post randomisation	*The absolute difference between planned and achieved coronal plane correction hip-knee-ankle angle and the true hip-knee-ankle angle at 9 months post randomisation. Measured in degrees.* *The amount of hip-knee-ankle correction required is decided by the surgeon using landmarks on the weight-bearing X-ray.*
Handling of intercurrent events	***Handling of missing data***
ICE 1: Surgery did not take place within 9 months post randomisation *Strategy 1: Treatment Policy* *ICE 2: Data is not available 9 months post randomisation* *Strategy 2: Principal stratum – excluded from analysis population* *ICE 3: Patient receives intervention that they were not randomised too.* *Strategy 3: Treatment Policy*	*There will be no imputation of missing data for the primary analysis of any outcome. Only those who have surgery will be included in the ITT population.* *Sensitivity analysis will be undertaken for the primary and key secondary outcome by imputing data under different missing not-at-random assumptions. Given the sample size only simple imputation approach will be used. A poor and a good outcome will be used where an absolute difference between planned and achieved is less than 3 degrees is considered good.*
Population-level summary measure	***Analysis approach***
The analysis set once adjusted for the ICEs is a subset of the intention-to-treat analysis set. It is those who have gone beyond 9 months post randomisation within the study.	*The summary method of interest will be a mean difference between the treatment groups estimated using a linear regression including all patients who receive surgery. The model will be adjusted for the randomisation factors (centre, sex, age, BMI) and baseline misalignment.* *A sensitivity analysis will be performed on the per protocol analysis set to assess the robustness of results but otherwise as above.*
[1]*Strategies defined in E9 (R1) include treatment policy, while on treatment, principal stratum and hypothetical* (2)	

3. STUDY METHODS

3.1 Trial Design/framework

PASHiOn is a multi-centre, blinded, superiority two arm, parallel group design randomised controlled trial of personalised versus conventional HTO surgery. Patients will be randomised in a 1:1 ratio. Patients and primary outcome assessors will be blinded to the allocation.

3.2 Embedded within the trial is a non-randomised pre-trial technology check and safety assessment, this is termed the Initial Safety Phase (Phase 1). Randomisation and Blinding

The patients within Phase 1, are not randomised as they all received HTO for initial safety checks to be undertaken. For this reason, there was also no blinding within Phase 1.

Randomisation to Phase 2 will be 1:1 by minimisation, using the following factors: trial centre, sex, age and BMI. We have chosen to separate patients younger than 50 years and 50 years and older, and patients with a BMI lower than 30 and those with a BMI of 30 or greater, as stratification cut-off points.

The minimisation algorithm will include a random element to reduce predictability. A small number of participants will be randomised using a simple randomisation schedule generated in advance to patient randomisation.

All patients and outcome assessors will be blinded to the treatment allocation.

3.3 Sample Size

The sample size was initially planned to include 93 patients (inclusive of 5 participants in Phase 1 and 88 in Phase 2), this was recalculated and the target sample size was reduced to 50. The 50 will be achieved over 2 stages: Phase 1 – initial safety phase (5 participants) and Phase 2 – RCT (45 participants).

- Phase 1 – Initial safety phase (5 participants)
- Phase 2 – RCT (45 participants)

This sample size calculation was based upon the primary outcome of value of absolute difference between achieved and planned correction. A total sample size of 45 participants will be sufficient to detect a standardised difference of 2.6, using 80% power, a 5% significance level and allowing for 10% loss to follow up.

3.4 Statistical Interim Analysis, Data Review and Stopping guidelines

No interim analysis is planned, similarly no formal subgroup analysis is planned either. One single formal analysis will be conducted after the final follow-up assessment has been completed and sufficient time allowed for data collection and data cleaning.

The independent DSMC reviewed the safety data for the 5 patients recruited into Phase 1. Based upon this review, the DSMC confirmed the trial could proceed to Phase 2.

During Phase 2, the DSMC and TSC have evaluated the risk of the trial and appropriate actions have been taken where necessary. For more details, see the reports and the charter.

3.5 Timing of Final Analysis

Final analysis is expected to begin after the final follow-up assessment is completed, allowing for a reasonable time to complete data collection and data cleaning.

3.6 Blinded analysis

There is no plan to perform any formal blinded analysis.

3.7 Statistical Analysis Outline as presented in the protocol

A separate statistical analysis plan (SAP) with full details of all statistical analysis planned for the data in this study will finalised prior to any primary outcome analysis. The SAP will be reviewed and receive input from the Trial Steering Committee (TSC) and Data and Safety Monitoring Committee (DSMC). Any changes or deviations from the original SAP will be described and justified in the protocol, final report and/or publications, as appropriate. It is anticipated that all statistical analysis will be undertaken using Stata (StataCorp LP, www.stata.com) or another well-validated statistical package.

A single formal analysis is planned for this trial, which will take place after the final follow-up assessment has been completed and sufficient time allowed for data collection and data cleaning. Analyses will be based on the intention-to-treat principle (participants analysed in allocated groups regardless of actual treatment received), though participants who did not undergo any HTO may be excluded, but will be repeated for the per-protocol population as a sensitivity analysis to test the robustness of the results.

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; with all tests carried out at a 5% two-sided significance level.

The primary endpoint, i.e. the absolute difference between achieved and planned correction at nine months post-randomisation (six months post-surgery), will be analysed using a linear regression model adjusted for the randomisation variables (centre, sex, age and BMI) , baseline misalignment and other baseline prognostic factors, if appropriate. Results from this model will be presented as adjusted mean difference in the absolute difference between achieved and planned correction (termed as “misalignment” from target value) between the groups and corresponding confidence intervals.

Given the overall size of the study no formal subgroup analyses are planned. Missing data pattern will be presented, and sensitivity analyses for missing data will be performed, including missing not random assumptions.

The key secondary outcome, i.e. number of patients (%) achieving a pre-specified difference between planned and achieved coronal plane correction at nine months post-randomisation (six months post-surgery), will be will be analysed by logistic regression model adjusted in line with the primary analysis. The results of this analysis will be presented as adjusted odds ratio and corresponding 95% confidence intervals. The unadjusted risk differences will also be provided to ensure that both relative and absolute effect sizes are presented. All

other secondary outcomes will be analysed using generalised linear models adjusted for the stratification factors and other important baseline prognostic factors.

4. STATISTICAL PRINCIPLES

There is a single primary outcome, so there is no concern regarding multiple testing in regards to multiple primary outcomes. A 2-sided significance level of 0.05 will be used, with 95% confidence intervals reported. All secondary analyses will be considered as supporting the primary analysis and will also be analysed using a 2-sided significance level of 0.05 with 95% confidence intervals.

Interim analyses of primary and secondary outcomes have not be carried out as was planned.

4.1 Definition of Analysis Populations

Populations for analysis are defined as follows:

- Intent-to-treat (ITT): all participants analysed in their randomised groups, regardless of actual treatment received as long as they received knee surgery of some kind.

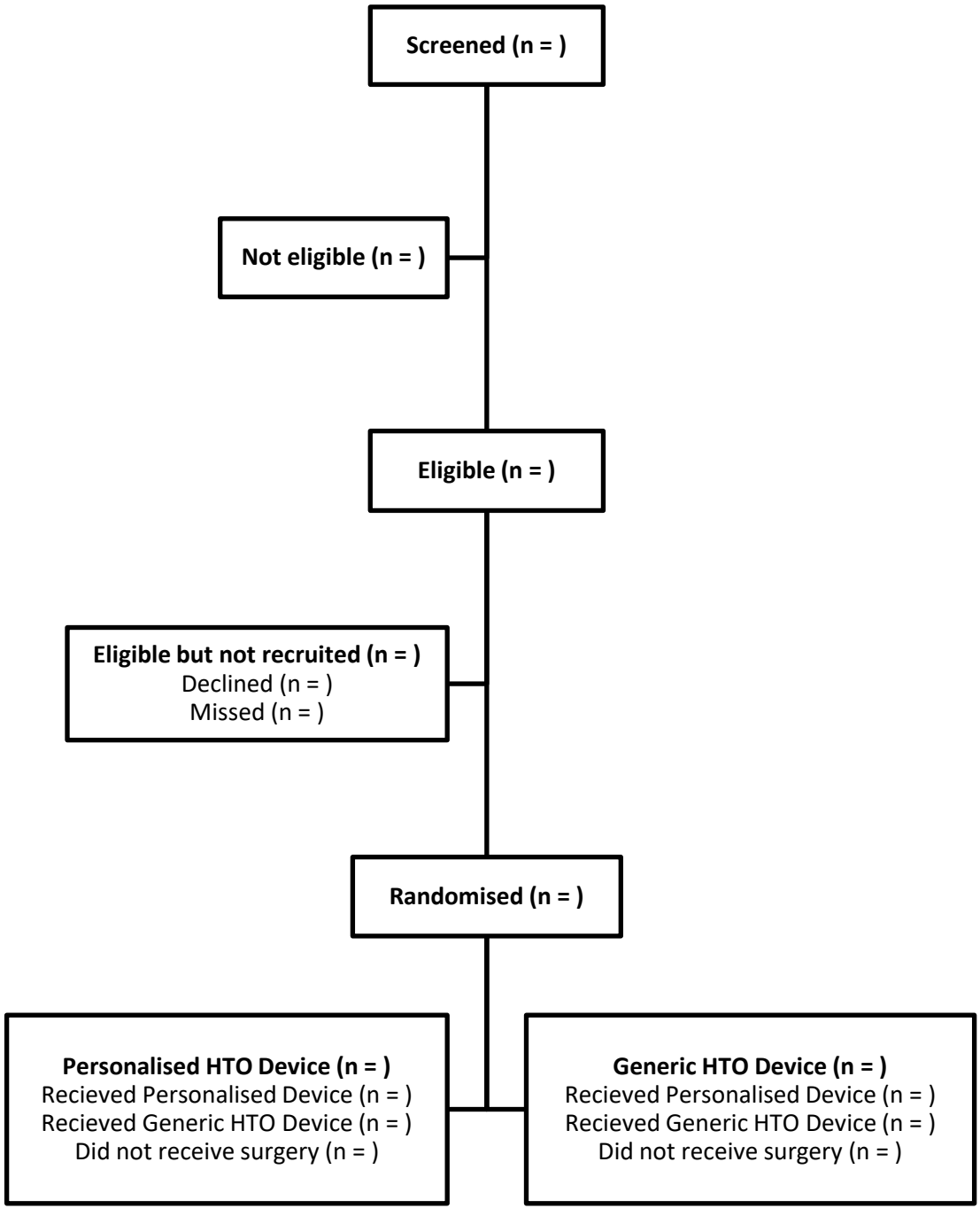
- Per-Protocol (PP): participants who received the intervention as intended will be analysed according to the treatment they actually received. Participants will be excluded from the per-protocol population if:
    - They did not receive the treatment to which they were randomised
    - They did not provide sufficient follow-up data analysis
    - They did not satisfy the eligibility criteria for the study
    - They did not adhere exactly to the protocol
        - Their preliminary scans were at least 6 months prior to surgery
        - Their surgery was at least 6 months post-randomisation

Exact exclusion criteria for PP analysis will be confirmed prior to conducting the final analysis (in that the estimates will not be generated and no results will be shared before this).

5. TRIAL POPULATION AND DESCRIPTIVE ANALYSES

Summary of flow of trial participants through the trial and baseline stratification, demographic and clinical characteristics of each group.

Figure 1 CONSORT Flow diagram



5.1 Representativeness of Study Sample and Patient Throughput

The flow of participants through the trial will be summarised as outlined in Figure 1. This shows the number of individuals screened, eligible, consented, randomised to each arm, receiving allocated treatment and included in primary analysis as recommended in the CONSORT guidelines (3). Reasons for ineligibility, withdrawals, and exclusions from primary analysis will be summarised.

5.2 Withdrawal from treatment and/or follow-up

The numbers and percentages of participants who are lost to follow-up or withdrew are reported (Table *3*) by treatment allocation for each time point until the final time-point of 15 months post-randomisation. Reasons for withdrawal will also be summarised by treatment allocation in Table *4*.

Table 3 Details of withdrawal

Patient ID	Extent of withdrawal	Withdrawal Date (DDMonYYYY)	Withdrawal Reason	Last Completed Questionnaire	Data Available for Primary Outcome (Y/N)
ID 1 ID 2 ID 3 …	Intervention only vs. follow-up				

Table 4 Summary of reasons for withdrawal by treatment arm

Reason for Withdrawal	Extent of withdrawal	*Generic*		Personalised	
		n	%	n	%
Reason 1 Reason 2 No reason Other reason	Intervention only vs. follow-up				

5.3 Baseline Comparability of Randomised Groups

Baseline comparability of the randomised groups on the stratification factors (centre, sex, age, and BMI) will be presented in Table 5*. Baseline comparability of the randomised groups on other factors will be presented (*

Table *6* and Table *7*). Variables will be presented by numbers (with percentages) for binary and categorical variables. Means and standard deviations (SDs), or medians and interquartile ranges (IQRs) for continuous

variables will be presented; there will be no tests of statistical significance nor confidence intervals for differences between randomised groups.

Table 5 Stratification factors by treatment arm

	Generic		***Personalised***		***Total***	
	n	***%***	***n***	***%***	***n***	***%***
Centre						
Site 1						
Site 2						
Site 3						
Site 4						
Site 5						
Site 6						
Site 7						
Sex						
Male						
Female						
Age*						
<= 50						
> 50						
BMI*						
<= 30						
> 30						

Table 6 Categorical factors by treatment arm

	Generic		***Personalised***		***Total***	
	n	***%***	***n***	***%***	***n***	***%***
Study Knee						
Left						
Right						
UCLA Activity Score						
Wholly Inactive, dependent on others, and cannot leave residence						
Mostly Inactive or restricted to minimum activities of daily living						
Sometimes participates in mild activities, such as walking, limited housework and limited shopping						

	Generic		*Personalised*		*Total*	
	n	***%***	***n***	***%***	***n***	***%***
Regularly participates in mild activities *Sometimes participates in moderate activities such as swimming or could do unlimited housework or shopping* *Regularly participates in moderate activities* *Regularly participates in active events such as bicycling* *Regularly participates in active events, such as golf or bowling* *Sometimes participates in impact sports such as jogging, tennis, skiing, acrobatics, ballet, heavy labour or backpacking* *Regularly participates in impact sports*						

Table 7 Continuous factors by treatment arm

	Generic			*Personalised*			*Total*		
	n	***Mean (SD)***	***Median (IQR)***	***n***	***Mean (SD)***	***Median (IQR)***	***n***	***Mean (SD)***	***Median (IQR)***
Demographics *BMI*** *Height* *Weight* *Age* *Time from injury to randomisation** **Video Vector Analysis** *Pre-op Hip Knee Ankle angle (HKA)* *Planned Post Surgery HKA angle* *Difference between pre-op and planned HKA angle* **ROM** *Operative knee flexion*									

	Generic			*Personalised*			*Total*		
	n	***Mean (SD)***	***Median (IQR)***	***n***	***Mean (SD)***	***Median (IQR)***	***n***	***Mean (SD)***	***Median (IQR)***
Operative knee extension *Contralateral knee flexion* *Contralateral knee extension* **Calibration** *Operative knee width* *Contralateral knee width* **KOOS** *Symptoms* *Stiffness* *Pain* *Function – Daily living* *Function – Sports and recreational activities* *Quality of Life* *KOOS-5 summary score* **EQ-5D-5L** *Utility* *EQ-VAS* **Pain VAS** *Score*									

** Time is measured in days*
*** BMI (weight divided by height squared (kg/m^2))*

5.4 Unblinding

If patient unblinding is deemed necessary, it will be discussed on a case-by-case basis with the central study team and local PI. All unblinding will be at the discretion of the local investigators, when clinically indicated for the safety of the patient. There will be no formal analysis of the success of the blinding. All cases of treatment unblinding will be listed, together with who was unblinded, and reasons for unblinding summarised with numbers and percentages reported.

5.5 Treatment Compliance with Details of Interventions

Compliance with the interventions according to random allocation in this study will be presented as numbers and percentages of participants receiving their allocated treatment. The investigational device is a single-use implant, so the proportion of participants that received their allocated treatment defines treatment compliance. As such, participant compliance with the device is not relevant. If, for any reason, the randomised treatment (device) is not used this will be recorded and presented.

Operation and theatre times will be summarised descriptively. Overall data per arm and per surgeon/site will be summarised as appropriate, including exploring any indication of learning.

5.6 Reliability

Sites will take the responsibility to double check and confirm the data collected. This includes registration data, the primary outcome (hip-knee-ankle angle measurement), and any withdrawal/death notifications. This data will be collected from trial sites directly and will be sent to the trial team on a regular basis.

It is worth noting that the target HKA alignment is predetermined for participants receiving the personalised HTO device in advance of surgery allowing for time to design the device. Conversely, participants receiving the generic HTO device have their target alignment determined on the date of surgery and completed in the CRF post-surgery. No methods are planned to counteract this potential bias in favour of the generic device.

6. ANALYSIS

6.1 Outcome Definitions

Primary Outcome Definition:
A full leg weight bearing X-Ray will be taken at baseline and at 9 months post randomisation. The images will be used to calculate the difference between planned and achieved coronal plane correction. Note for all measures 9 months post randomisation are intended to align with the 6 months post-surgery within the standard protocol (surgery within 3 months of randomisation and no delays). These time points can therefore be used interchangeably for the purposes of trial management.

The amount of Hip Knee Ankle (HKA) correction required is decided by the surgeon using landmarks on the weight-bearing X-ray. The Hip-Knee-Ankle (HKA) angle is defined as the lateral angle between two lines: one line from the centre of the femur head using Mose circles to the middle of the distance between the tibial spines, and a second line from the centre of the ankle to the centre of the tibial spines. An angle of more than 180 degrees denotes a varus alignment. The baseline HKA angle and the planned HKA correction angle is recorded pre-operatively. The HKA correction angle achieved is recorded post-operatively.

The primary outcome measure compares the planned correction against the actual HKA correction achieved, determined by the difference in degrees, between actual pre-op alignment and actual post-op alignment.

The follow-up X-ray imaging taken 9 months post randomisation will be anonymised and sent to two blinded reviewers to independently measure the difference between planned and achieved correction. If a consensus cannot be reached between the two reviewers, a third blinded reviewer will be consulted.

Secondary Outcome Definitions:

- **Number of patients achieving a pre-specified difference, 3 degrees, between planned and achieved coronal plane correction of the HKA**. The HKA is calculated as per the primary outcome and converted into a binary variable difference (with 3 degrees) or not.

- **Change in coronal plane location of peak loading during gait:** Video Vector Analysis (VVA) will be used at baseline and at 6 months post-surgery. VVA simultaneously records ground reaction force data from a force plate and video data from a video camera focussed on the lower limb of a

participant. This enables the magnitude, orientation, and location of the ground reaction force to be determined relative to the limb under investigation.

- **Range of Motion (ROM):** ROM flexion and extension will be measured (in degrees) at baseline and nine months post randomisation.

- **Operative time in minutes**: The difference in operative time and theatre time between the treatment arms. Operative time is defined as time from 'knife to skin' to time dressings have been applied; theatre time is defined as the time the participants enters the anaesthetic room to when they leave the operating theatre.

- **KOOS**: Knee injury and osteoarthritis Outcome score (KOOS) is a patient reported outcome measure derived from 5 subscales; symptoms (including stiffness), pain, function (daily living), function (sports and recreation activities) and quality of life with scores ranging from 0 – 100, a higher score indicating better health.
    - **Scoring:** Each subscale score is calculated separately, calculate mean score (x100) of individuals items of each subscale and divide by 4 (the highest possible score for a single question). Then take 100 – the score calculated.
      While there is no overall KOOS score the summary score using an average of the 5 domains has been suggested and used in previous clinical trials.
    - **Missing data rules:** If a mark is placed outside a box, the closest box is chosen. If two boxes are marked, that which indicates the more severe problem is chosen. As long as at least 50% of the subscale items are answered for each subscale, a mean score can be calculated. If more than 50% of the subscale items are omitted, the response is considered invalid and no subscale score should be calculated. For the subscale Pain, this means that 5 items must be answered; for Symptoms, 4 items; for ADL, 9 items; for Sport/Rec, 3 items; and for QOL, 2 items must be answered in order to calculate a subscale score. Subscale scores are independent and can be reported for any number of the individual subscales, i.e. if a particular subscale is not considered valid (for example, the subscale Sport/Rec 2 weeks after total knee replacement), the results from the other subscale can be reported at this time-point.

- **EQ-5D-5L**: To measure quality of life the EQ-5D-5L will be collected at baseline, 6 and 12 months. The EQ-5D-5L is a validated, generalised, health related quality of life questionnaire consisting of 5 domains related to daily activities with a 5-level answer possibility, which will be converted into multi-attributed utility scores using established algorithms. The EQ-5D-5L will be scored in line with current NICE guidance, which currently stipulate that the response will be mapped to EQ-5D-3L utilities using the Hernandez Alava method (e.g. via eq5dmap in Stata and eq5d in the R software) (4).

- **Pain and irritation VAS score:** A VAS score used to measure patient reported knee pain and irritation. Patients will be requested to mark on a scale of 0 (no pain) to 10 (worst possible pain).

- **UCLA Activity Score:** The UCLA Activity score is a scale ranging from 1 to 10. The patient indicates their most appropriate activity level, with 1 defined as "wholly inactive, dependent on others, and cannot leave residence" and 10 defined as "regularly participates in impact sports". Treat as categorical.

- **Satisfaction and Transition questionnaire:** To measure patient satisfaction and transition, participants will be asked how satisfied they were with the treatment they received and based on their experience, how willing they would be to have the operation again.
    - This questionnaire consists of 5 questions, all in a likert scale with between 3 and 6 answers. Responses to each question will be presented separately:
        - How satisfied are you with your knee after your HTO surgery?
        - How are the problems related to your knee now, compared with before your knee surgery?
        - If you could go back in time, would you still choose to have the knee operation?
        - In general, would you say that your health is:
        - Compared to one year ago, how would you rate your health in general now?

- **Health Resource Use:** No formal health economic evaluation will be performed in the remit of this trial. This study will collect information on participants' health resource use, including time in operating theatre, visits to primary care, and hospital care services, during the follow-up. Summaries will be presented by trial arm and mean differences with 95% confidence intervals.

- **Safety:** Safety data will be collected throughout the duration of the trial, to determine the rates of Adverse Events and Adverse Device Effects.

6.2 Analysis Methods

6.2.1 Assessments of normality

Given the sample size and the methods being used no formal checks of assumptions will be carried out. As the sample size is small, a non-parametric methods (Mann-Whitney), will be used and medians and IQRs will be also reported for each treatment allocation (in addition to mean and SD) as a sensitivity analysis.

6.2.2 Primary Outcome - The absolute difference between achieved and planned correction at nine months post-randomisation (six months post-surgery)

Unadjusted summary statistics will be displayed by treatment allocation using means, standard deviations, and IQR. Linear regression model will be used to compare the treatment groups including all patients. The model will be adjusted for baseline misalignment only. Results from this model will be presented as adjusted mean difference in the absolute difference between achieved and planned correction (termed as "misalignment" from target value) between the groups; with corresponding 95% confidence intervals and p-values.

As a supporting analysis, a Mann-Whitney test comparing the primary outcome at 9 months will also be presented. These analyses will be performed on the ITT population, and no imputation of missing data is planned.

The primary analysis of absolute misalignment will be measured separately by two blinded reviewers, an Altman-Bland plot will be used to assess the agreement and discrepancies.

6.2.3 Secondary outcomes

The key secondary outcome, i.e. number of patients (%) achieving a pre-specified difference between planned and achieved coronal plane correction at nine months post-randomisation, will be analysed by logistic regression model. The model will be adjusted for the randomisation factors (sex, age, and BMI). The results of this analysis will be presented as adjusted odds ratio, with corresponding 95% confidence intervals and p-values. The unadjusted risk differences will also be provided to ensure that both relative and absolute effect sizes are presented.

Continuous clinical secondary outcomes (change in coronal plane location of peak loading, range of motion, calibration) will be analysed using a linear regression model to compare treatment groups, adjusting for the randomisation factors (sex, age, and BMI) and baseline misalignment.

Continuous patient reported outcomes (KOOS sub-scale and summary scores, EQ-5D-5L utility score, EQ-5D-5L VAS, pain and irritation VAS) will be analysed using a linear regression model. The model will include fixed effects to adjust for sex, age, BMI, and baseline misalignment. The adjusted mean differences will be presented with 95% confidence intervals and p-values.

The UCLA Activity Score and the satisfaction and transition questionnaire will be analysed using a linear regression model, adjusting for sex, age, BMI, and baseline misalignment. Adjusted odds ratios, 95% confidence intervals and p-values will be presented.

6.2.4 Safety outcomes

Intra-operative complications, device related complication, general AEs and SAEs will be presented as the numbers and percentages of participants who experienced at least one. Safety data will be analysed using a logistic regression model adjusting for age, sex, BMI and baseline misalignment, and will be presented as adjusted odds ratios with 95% confidence intervals and p-values. The unadjusted risk differences will also be provided.

6.3 Missing Data

The number and percentage of individuals with missing data for each outcome at each time point will be summarised by intervention. Missing data will be assumed to be missing at random, but the nature of this assumption will be explored, and if judged appropriate, missing data for the primary outcome and any key secondary outcomes will be imputed under a variety of different missing data assumptions as part of sensitivity analysis.

6.4 Sensitivity Analysis

The primary outcome analysis will be redone with the addition of the randomisation factors. The model will adjust for baseline misalignment, sex, age, and BMI.

The primary outcome (absolute difference) analysis, and key secondary outcome (% participants achieving pre-specified difference) analysis will be repeated for the per-protocol population.

Sensitivity analysis will be undertaken for the primary and key secondary outcome by imputing data under different simple imputation missing not-at-random assumptions.

6.5 Pre-specified Subgroup Analysis

There are no sub-group analyses planned for this study.

6.6 Supplementary/Additional Analyses and Outcomes

There are no planned supplementary analyses for this study.

6.7 Health Economics and Cost Effectiveness (where applicable)

Health resource use data will be presented with numbers and percentages. This will cover operation times, inpatient visits and outpatient visits during the follow-up period. The number of GP clinic visits, home visits and telephone consultations will also be recorded alongside the number of practice nurse visits, practice nurse consultations, district nurse visits, physiotherapy visits, occupational therapy visits and counsellor/psychiatrist/psychologist visits.

Device deficiencies will also be descriptively summarised. This includes deficiencies caused by identity, quality, durability, reliability, safety, performance, malfunction, user error and inadequate labelling.

7. VALIDATION OF THE PRIMARY ANALYSIS

To validate the primary outcome (absolute difference between planned and achieved correction between pre-op and post-op alignment) and key secondary outcomes (number (%) achieving a pre-specified difference between planned and achieved correction) a statistician not involved in the trial will independently repeat the analyses detailed in this SAP, by using different statistical software (if possible). The results will be compared and any discrepancies will be reported in the Statistical report (See OCTRU SOP STATS-005 Statistical Report).

8. SPECIFICATION OF STATISTICAL PACKAGES

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

9. REFERENCES

1. *Guidelines for the Content of Statistical Analysis Plans in Clinical Trials.* **Gamble, C, et al.** 2017, JAMA; 318(23):2337-2343. doi: 10.1001/jama.2017.18556.
2. *Addendum on estimands and sensitivity analysis in clinical trials.* **International Council for Harmonisation of Technical Requirement for Pharmaceuticals for Human use.** 2019, Vol. E9 (R1).
3. *CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.* **Schulz, K, Altman, D and Moher, David.** 2010, BMJ.
4. *A Comparison of Direct and Indirect Methods for the Estimation of Health Utilities from Clinical Outcomes.* **Hernández Alava, Mónica, et al.** 2014, Vol. 34.

10. APPENDIX: GLOSSARY OF ABBREVIATIONS

SAP	Statistical Analysis Plan
DSMC	Data and Safety Monitoring Committee
TSC	Trial Steering Committee
CI	Chief Investigator
HTO	High Tibial Osteotomy
TOKA	Tailored Osteotomy Knee Alignment
ICE	Intercurrent Event
KOOS	Knee injury and Osteoarthritis Outcome Score
PROM	Patient Reported Outcome Measure
ROM	Range of Motion
UCLA Activity Score	University of Carolina, Los Angeles Activity Score
VVA	Video Vector Analysis
HKA	Hip Knee Ankle
ITT	Intention-to-Treat
PP	Per Protocol