

The Efficacy of Peroneal Nerve Functional Electrical STimulation (FES) for the Reduction of Bradykinesia in Parkinson's DiSease: An Assessor Blinded Randomised Controlled Trial

(STEPS II)



Statistical Analysis Plan

Version:

SAP V1.0 20-Aug-2026

This statistical analysis plan is based on protocol version 2.2 (26-May-2026)

Sponsor:	Salisbury NHS Foundation Trust
IRAS number	330866
Coordinating Centre	Peninsula Clinical Trials Unit (PenCTU); Statistics: Exeter CTU
ISRCTN number	13120555
Funder reference number	NIHR131791

KEY PERSONNEL INVOLVED IN THE PREPARATION OF THE STATISTICAL ANALYSIS PLAN

NAME	TRIAL ROLE	SIGNATURE	DATE
Adam Palmer-Welsh	Trial Statistician (author)	DocuSigned by: Adam Palmer-Welsh D953BABE77CD46C...	21 August 2026 09:29 BST
Ayesha S. Ali	Senior Trial Statistician (reviewer and approver)	Signed by: Ayesha Ali 21264704BD774F5...	20 August 2026 17:41 BST
Siobhan Creanor	Senior Statistician (reviewer and approver)	DocuSigned by: Siobhan Creanor 50999822B30D479...	27 August 2026 10:39 BST
Paul Taylor	Chief Investigator (reviewer and approver)	Signed by: Paul Taylor 3D041746168D49E...	25 August 2026 16:50 BST
Catriona Keerie	Independent Statistician on Trial Steering Committee (reviewer and approver)	Signed by: Catriona Keerie BA066E3B506C428...	21 August 2026 08:32 BST

DOCUMENT CONTROL SHEET

VERSION NUMBER	DATE	REASON FOR UPDATE	DESCRIPTION OF CHANGES

CONTENTS

Key personnel involved in the preparation of the Statistical Analysis Plan:2

 Name.....2

 Trial role.....2

 Signature2

 Date2

Document control sheet2

 Version number2

 Date2

 Reason for update.....2

 Description of changes.....2

1. Introduction6

 1.1 Purpose of the statistical analysis plan6

 1.2 Trial design and methods6

 1.2.1 Trial design6

 1.2.2 Primary objective and estimand6

 1.2.3 Secondary objectives.....7

 1.2.4 Exploratory objectives.....7

 1.3 Outcome Measures.....8

 1.3.1 Primary outcome8

 1.3.2 Secondary clinical outcomes.....8

 1.3.3 Secondary mechanistic outcomes8

 1.4 Non-comparative measures9

 1.5 Eligibility9

 1.6 Randomisation 10

 1.7 Sample size..... 10

 1.8 Blinding..... 11

 1.9 Trial schema..... 12

2. Trial reporting 13

 2.1 Final analyses..... 13

 2.2 Trial oversight committees 13

3. Data quality 14

 3.1 Data validation 14

 3.2 Data completeness 14

Statistical Analysis Plan

4. Preparation of datasets and documenting analyses 14

5. Defining analysis populations 15

6. Trial Population..... 15

6.1 Screening..... 15

6.2 Recruitment/Randomisation 16

6.3 Participants found to be ineligible post-randomisation 16

6.4 Change in participation status post-randomisation 16

6.5 Non-compliance (deviations, violations and serious breaches) 16

7. Baseline characteristics 17

8. Treatment/Intervention received 17

9. Safety analysis/harms 17

10. Analysis 18

10.1 Primary outcome..... 18

10.1.1 Null hypothesis 18

10.1.2 Methods 18

10.2 Subgroup analysis 20

10.3 Secondary clinical outcomes 20

10.4 Bayesian analysis of outcome measures 20

10.5 Additional analyses 21

10.5.1 Mechanistic Analysis 21

10.5.2 Non-comparative measures 22

10.5.3 Imputation of datasets for analysis 22

11. Economic evaluation..... 23

12. Statistical software 23

13. Storage and archiving 23

14. References 24

15. Appendices 27

15.1 Appendix 1: Abbreviations..... 27

15.2 Appendix 2: Outcome calculations 29

15.2.1 10mWS..... 29

15.2.2 Instrumented toe tap test 29

15.2.3 Stride length 29

15.2.4 MDS-UPDRS..... 29

15.2.5 N-FOG 30

Statistical Analysis Plan

15.2.6 Falls diary 30

15.2.7 FES-I 31

15.2.8 mini-BESTest..... 31

15.2.9 Daily step count..... 31

15.2.10 PDQ-39..... 31

15.2.11 EQ-5D-5L 32

15.2.12 LEDD 33

15.3 Appendix 3: Example Tables 34

1. INTRODUCTION

1.1 PURPOSE OF THE STATISTICAL ANALYSIS PLAN

This Statistical Analysis Plan (SAP) (1) provides guidelines for the analysis and presentation of results for the NIHR EME-funded STEPS II trial. This plan, along with all other documents relating to the analysis of this trial, will be stored in the ‘Statistics’ section of the (electronic) Trial Master File. The statistical analyses will be carried out by the Trial Statistician, with support and oversight from the Senior Trial Statisticians.

1.2 TRIAL DESIGN AND METHODS

1.2.1 TRIAL DESIGN

This is a multi-centre, two-arm, parallel, assessor-blinded, superiority, individually randomised controlled trial comparing Functional Electrical Stimulation (FES) in addition to usual care to usual care alone in adults with Parkinson’s disease (PD) who have trouble walking due to whole body bradykinesia. Participants will be randomised 1:1 via block randomisation (for more details on the randomisation procedure see section 1.6). The trial involves an internal pilot phase and a blinded sample size review re-estimation, after the first 12 months of participant recruitment.

1.2.2 PRIMARY OBJECTIVE AND ESTIMAND

The primary objective of the STEPS II trial is to compare change (from baseline) in whole body bradykinesia in people with PD (pwPD) at 18 weeks post-randomisation using FES with usual care versus usual care alone. The primary estimand (2,3) is presented in Table 1.

Table 1. Primary estimand framework

Population	People with idiopathic Parkinson’s disease who experience difficulty walking due to bradykinesia (defined as a measured 10m walking speed of less than 1.25ms ⁻¹ at screening)
Treatment conditions	- Intervention: Functional Electrical Stimulation (FES) in addition to usual care - Control: Usual care alone
Outcome variable	Change in walking speed (10mWS), assessed using the 10m walking test (10mWT) at 18 weeks post-randomisation and at baseline and then the difference in WS calculated between 18-weeks and baseline
Handling of intercurrent events ¹ (ICE) (not incorporated by the population or treatment condition)	- Death: principal stratum: participants who die will be excluded from the analysis population - Treatment discontinuation (e.g. participant decides to stop their FES intervention): treatment policy: analysed as though ICE did not occur
Population level summary	Between-group mean difference in the change of 10mWS between baseline and 18-weeks post-randomisation

¹ ICEs (and how they will be incorporated) are fully specified in section 10.1.2

1.2.3 SECONDARY OBJECTIVES

1. Determine if any effects in change of WS remain for up to four weeks after FES is withdrawn.
2. Investigate the effect of FES on other aspects of Parkinsonian gait and living life with PD, to inform the need and design of future research. To achieve this, we will produce the first estimates of effect size of FES on secondary outcome measures for:
 - I. Hypokinesia
 - II. Akinesia
 - III. Falls and balance
 - IV. Activities of daily living
 - V. Activity
 - VI. Quality of Life
 - VII. Cost/utility
 - VIII. The effect on gait while using FES.
3. Determine how pwPD and their carers perceive the usefulness and practical experience of FES use and its therapeutic effect.
4. Determine the safety of FES in pwPD.
5. Investigate potential mechanisms of action of FES in PD, by determining:
 - I. short-term (up to 6 weeks) changes in inter-limb coordination, APA and limb bradykinesia while stepping and walking
 - II. the link between putative mechanisms of action and their relationship to walking speed at 6 and 18 weeks, through causal mediation analysis
 - III. changes over time in the strength of muscles (a) directly targeted and (b) not targeted by the FES.

1.2.4 EXPLORATORY OBJECTIVES

To understand the experiences and views of individuals with PD using FES, and any potential perceived impact or carryover effect of FES, an embedded qualitative study will be carried out. FES participants, and family members/carers of said individuals, will be asked to take part in semi-structured interviews. The planned analyses for the qualitative component are not covered in this SAP.

1.3 OUTCOME MEASURES

1.3.1 PRIMARY OUTCOME

The primary outcome is the change in walking speed (10mWS), measured in ms^{-1} between baseline and 18 weeks post-randomisation. Walking speed (WS) will be recorded using the 10m walking test (10mWT) at baseline and 18 weeks post-randomisation and the change in WS will be calculated as the absolute difference between baseline and 18 weeks post-randomisation.

1.3.2 SECONDARY CLINICAL OUTCOMES

1. Time course and carry over of whole-body bradykinesia, measured using 10mWS, at baseline and 2 weeks, 6 weeks, 18 weeks (primary outcome) and 22 weeks post-randomisation.
2. Bradykinesia, measured using an instrumented toe tap test (part of MDS-UPDRS (4) section 3), at baseline and 2 weeks and 6 weeks post-randomisation.
3. Hypokinesia, measured using stride length (in metres), at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.
4. Hypokinesia, measured using MDS-UPDRS (4) section 3, at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.
5. Akinesia, measured via N-FOG (5), at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.
6. Falls, measured using falls diary (6), at baseline and 6 weeks, 18 weeks and 22 weeks post-randomisation.
7. Fear of falling, measured via FES-I (7), at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.

8. Balance, measured using the mini-BESTest (8), at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.
9. Activities of daily living, measured using MDS-UPDRS (4) sections 1A, 1B and 2, at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.
10. Activity level, measured as daily step count averaged over 7 days recorded using a StepWatch prior to baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.
11. Quality of life, measured using PDQ-39 (9) and EQ-5D-5L (10), at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.
12. Normal care, recorded using Levodopa equivalent daily dose (LEDD) (11) and duration of participation in therapeutic activities, at baseline and 2 weeks, 6 weeks, 18 weeks and 22 weeks post-randomisation.

1.3.3 SECONDARY MECHANISTIC OUTCOMES

1. Changes in inter-limb coordination, assessed using the PCI, at baseline and 2 weeks and 6 weeks post-randomisation.
2. Change in anticipatory postural adjustments (APAs), assessed using IMUs on the feet and lumbar sections of the back, at baseline and 2 weeks and 6 weeks post-randomisation.
3. Change in limb bradykinesia while stepping and walking, assessed using IMUs attached to the foot (concurrently with MDS-UPDRS toe tapping test) and through the maximum velocity of foot motion while walking (concurrently with 10mWS), at baseline and 2 weeks and 6 weeks post-randomisation.
4. Changes over time in the strength of muscles directly targeted and not targeted by the intervention, assessed using hand-held dynamometry and maximum voluntary torque on both the FES and non-FES sides, at baseline and 2 weeks and 6 weeks post-randomisation.

1.4 NON-COMPARATIVE MEASURES

The following measures are collected in this trial in the FES arm only:

1. Cost, measured using records of clinical contacts and consumables related to FES usage, at baseline and 1 week, 2 weeks, 6 weeks and 18 weeks post-randomisation.
2. Bradykinesia, measured using 10mWS with and without FES (including the short term carry over effect (12)), collected at baseline and 1 week (without FES only), 2 weeks, 6 weeks and 18 weeks (with and without FES) post-randomisation.
3. Hypokinesia, measured using stride length calculated during the 10mWT with and without FES, collected at baseline and 1 week (without FES only), 2 weeks, 6 weeks and 18 weeks (with and without FES) post-randomisation.
4. Effort of walking, measured using the Borg Rating of Perceived Exertion (RPE) (13) during the 10mWT with and without FES, collected at baseline and 1 week (without FES only), 2 weeks, 6 weeks and 18 weeks (with and without FES) post-randomisation.
5. Akinesia, measured using the timed up and go (TUG) test with and without FES, collected at baseline and 1 week (without FES only), 2 weeks, 6 weeks and 18 weeks (with and without FES) post-randomisation.
6. Experience of FES, measured using the FES experience questionnaire, collected at 18 weeks post-randomisation.
7. FES usability, measured using the Systems Usability Scale (SUS), collected at 18 weeks post-randomisation.
8. Intervention compliance, assessed from the number of steps taken and the total stimulation time with FES (measured using an ODFS pace device), collected at 2 weeks, 6 weeks and 18 weeks post-randomisation.

1.5 ELIGIBILITY

Individuals must satisfy the following criteria to be enrolled in the study:

Inclusion criteria

1. Aged ≥ 18 years at the time of informed consent
2. Diagnosed with idiopathic Parkinson's disease
3. Hoehn and Yahr stages I to IV
4. Difficulty in walking due to Parkinson's Disease bradykinesia, defined as a measured 10mWS of less than 1.25ms^{-1} (time to complete 10mWT $>8\text{s}$) at screening. Participants asked to walk "briskly but safely"
5. Able to walk 50m with appropriate walking aids, but without assistance from another person. Appropriate aids include walking sticks, tri or quad sticks, Ankle Foot Orthoses (AFOs) and similar devices
6. Able to obtain standing from sitting without the assistance of another person
7. Able to understand and comply with the treatment and assessment procedures

Exclusion criteria

1. Receiving, or scheduled to start, deep brain stimulation, within the next 6 months
2. Scheduled to start apomorphine or duodopa within the next 6 months (those who are currently taking duodopa and apomorphine are eligible)
3. Pyramidal and/or extrapyramidal systems injuries
4. Untreated or refractory epilepsy with seizures in the last 3 months
5. Pregnancy or planned pregnancy
6. Cardiac pacemaker, or other active medical implanted devices
7. Denervation of the common peroneal nerve, or other neurological condition known to cause dropped foot
8. Severe osteoarticular pathology that involves the calf bones, knee and tibio-tarsal joints, or other conditions that significantly affect walking
9. Malignancy or dermatological conditions in the leg that would be stimulated
10. Major cognitive impairment; dementia
11. Under treatment for an unresolved deep vein thrombosis in the leg that would be stimulated
12. Participating in another interventional clinical trial (observational studies are permitted)

1.6 RANDOMISATION

Participants will be randomly assigned, on a 1:1 basis, to the FES with usual care arm or usual care alone arm. To accomplish this, and to ensure reasonable balance in the two arms across sites, randomisation will be based on a block randomisation algorithm, which will stratify participants by site.

The randomisation sequence, using variable block sizes, will be generated by a statistician independent to the trial team and implemented through a secure web-based system, ensuring allocation concealment.

Full details of the randomisation method, including confidential information on block size, can be found in the Randomisation Requirements Document for the trial.

1.7 SAMPLE SIZE

The STEPS II trial aims to recruit a total of 234 participants (~117 per arm). Assuming a mean difference from baseline of 0.13ms^{-1} between the two arms, and a standard deviation (SD) of 0.28ms^{-1} , a total of 198 participants will be required to give the study 90% power to detect this difference at 18 weeks, with a two-sided significance level of 5%. To account for participants who are randomised but do not reach the primary endpoint, a 15% loss to follow-up rate has been added, which increases the required sample size to 234 participants. The required sample size is based on data from the feasibility RCT, where the between-group mean difference in change in walking speed was estimated as 0.14ms^{-1} with a 95% confidence interval (CI) of

(0.03 to 0.26). Full details of the power scenarios based on the planned recruitment target can be seen in Appendix 1 of the STEPS II Protocol version 2.0.

There is a planned blinded review of the assumption of the SD used in the sample size calculation. This will happen at the same time at which the recruitment and retention progression criteria are assessed, after the first 12 months of participant recruitment. At this time, the observed SD of the change in walking speed between baseline and week 18 will be calculated by the trial statisticians, pooled across all participants with available data. The observed SD (and corresponding confidence interval) will be compared with the assumed 0.28ms^{-1} used in the sample size calculation and reviewed by the trial oversight committees, Sponsor and funder. If the observed SD is larger than expected implications on the required sample size will be discussed, considering other factors such as the emerging retention rate at the primary end point and rates of completeness of the primary outcome, and remedial action will be considered if deemed necessary.

1.8 BLINDING

This trial uses a mixture of blinded and unblinded assessors. All clinical outcome measures for the between-group comparative analyses are collected by blinded assessors at the blinded assessment visits. Intervention participants will be asked not to use FES on the day of the blinded assessment and reminded that they should not bring the device with them to blinded assessments. Participants will be asked not to discuss their group allocation with the blinded assessor or any other pwPD who are participating in the study.

Measures taken while FES is being used will be made by an unblinded treating FES clinician on a separate day, as part of the routine care FES clinical visits.

A Blinding of Trial Statisticians (BOTS) risk assessment has been completed giving full details of the blinding status of every statistician involved with the trial. This is stored in the statistical section of the Trial Master File (on SharePoint). The senior statisticians will be unblinded throughout the trial. All non-statistical members of the central coordinating trial team will be unblinded throughout the trial.

1.9 TRIAL SCHEMA



Figure 1: Participant pathway flowchart

*Routine FES treatment pathway

Trial Duration

Participants were originally intended to be recruited over a two-year period across at least seven secondary care trusts/hospitals in the UK. Following the internal pilot phase and review of recruitment, a nine-month extension to recruitment was agreed. However, not all eight care trusts will continue to recruit participants through to the end of the extended recruitment period. All participants are planned to be followed up until 22 weeks post-randomisation.

2. TRIAL REPORTING

2.1 FINAL ANALYSES

The final statistical analyses will be conducted after the final participant's 22-week follow-up visit. Following final data entry and cleaning, the database will be locked, and a blinded version of the dataset will be sent to the blinded statistician for analysis. The primary analyses (see section 10.1.2) for the primary outcome and all secondary clinical outcomes (see section 10.3) will be performed by a statistician who is blinded to allocated group. Once the results of the blinded analyses have been discussed and the interpretations determined, the remaining analyses will be performed after unblinding.

Undertaking of the planned sample size review will be by an unblinded statistician and any decisions to change the sample size will be based on discussions and recommendations from the unblinded Data Monitoring Committee and the blinded Trial Steering Committee (see section 1.7).

2.2 TRIAL OVERSIGHT COMMITTEES

Trial oversight will be provided by an independent Data Monitoring Committee (DMC) and a Trial Steering Committee (TSC).

The TSC is an executive oversight body operating on behalf of the Sponsor and will make decisions as to the future continuation (or otherwise) of the trial. They will monitor ongoing data for recruitment, baseline characteristics, outcome measure completion, safety and changes in participant status.

The DMC will review similar data to the TSC but will also have access to unblinded data to better monitor and identify any ethical or safety concerns. Additionally, the trial DMC will provide feedback on the planned sample size review (see section 1.7).

The TSC will meet at least annually in accordance with an agreed set of terms of reference detailed in the TSC Charter, to review the progress of the trial and will report to the Sponsor and Funder. DMC meetings will be scheduled to precede TSC meetings so that DMC recommendations can be considered by the TSC as appropriate.

3. DATA QUALITY

3.1 DATA VALIDATION

The trial data management team will carry out any necessary monitoring and cleaning as recruitment/follow-up occurs. Queries will be raised if missing data or irregularities are identified. Details of the checks to be performed and related procedures can be found in the trial's data management plan.

Basic data validation checks on the primary outcome measure will be carried out before the final analysis by the analysing statistician. Any identified queries will be reported to the data manager (DM) to be addressed prior to commencing the final analyses. In addition, any outcome measures that require calculating/deriving will be coded by the statistician from the participant data instead of using any calculated variables originating

from the database (see appendix 2 for details on calculation methods). Note codifying of free text variables (such as those who select “Other” for categorical questions) is not the responsibility of the analysing statistician and will be performed by the relevant member of the trial team. The primary analysis of the primary outcome will be double coded by a second statistician to ensure reproducibility.

3.2 DATA COMPLETENESS

Completeness rates (received/expected) of each outcome will be reported at each visit, overall and by allocated group. This data will also be reported by site. This will be used to inform the planned multiple imputation (see section 10.5.3).

4. PREPARATION OF DATASETS AND DOCUMENTING ANALYSES

After final follow-up data collection has occurred, data entry and cleaning complete, the locked dataset will be exported and sent to the unblinded statistician. This export will include identifiable marked variables relating to the following demographic data:

- date of birth
- sex at birth
- gender
- ethnicity

Data from these variables are required to calculate the summary statistics of baseline characteristics of the trial sample. All other identifiable data will be excluded from the exported dataset. The unblinded statistician will then prepare a blinded dataset and send this to the analysing statistician for analysis. Once the blinded analyses are complete and discussed, the unblinded dataset will also be sent to the analysing statistician to undertake the remaining analyses.

The data will be provided in long format, therefore, a long-to-wide reshape will be performed on the data necessary for any analysis requiring wide format data. This data will be reshaped using the participant ID and timepoint (encompassing event, instrument and instance as applicable) data. Given the data will be sent in a single dataset, no merging is expected but, if necessary, the participant ID and timepoint data will be used to ensure a 1:1 merge.

Before analysis starts, the analysing statistician will perform IQ checks to ensure the statistical software used is fit for purpose. The analysis will use a “main” file (in the form of a computable script) showcasing the order of analysis and use reproducible environment methods (such as the renv package in R). All scripts will be filed in line with the standard ExeCTU analysis structure.

5. DEFINING ANALYSIS POPULATIONS

The following analysis sets will form the basis of the analysis:

1. Randomised set (RS) - A treatment policy strategy will be used as the basis of the primary analysis of the primary outcome and analyses of most of the secondary outcomes. This population will include all randomised participants according to their allocated treatment regardless of the treatment received (following intention to treat (ITT) methods).
2. Primary analysis set (PAS) – A subset of the RS used to perform the primary analysis of the primary outcome. This set includes all alive randomised participants who provided at least one post-baseline 10mWS measurement.
3. Per protocol set (PPS) - A principal stratum strategy will be used in a sensitivity analysis of the primary outcome. This will include the subset of participants from the RS who did not experience any important

protocol deviations. As an example, participants who fully withdrew from the trial or participants allocated to the intervention group who stopped their FES intervention will be excluded from this set.

4. Treated set (TS) - Safety/harm outcomes will be reported by “as treated” groups for the safety population: categorised as those participants who received any form of FES vs those participants who did not receive any form of FES (rather than as randomly allocated).

Note that further exclusions can occur in some analyses due to strategies employed to account for ICEs. These strategies are detailed in Section 10.1.2.

The data for the primary analyses:

1. The primary analysis of the primary outcome will use the observed data only

The data for the sensitivity analyses not using the dataset above:

2. A while-on-treatment policy will be implemented by setting any outcomes after treatment discontinuation to missing
3. “Best/Worst” and “Worst/Best” imputation (15)
4. Multiple Imputation (MI) (14,15)

For details on imputation methods see section 10.5.3.

6. TRIAL POPULATION

A CONSORT (16) flow diagram will be used to summarise participant flow through the trial.

6.1 SCREENING

The data collected for patients who are screened will be retained. The number of patients registered, screened and the proportion of patients continuing with the study will be reported. Additionally, reasons for non-eligibility and non-progression (for eligible patients) will be reported. Patient characteristics for those screened but not randomised will be reported. These characteristics will include, but will not be limited to:

- screening site
- age
- sex at birth
- ethnicity
- employment status
- marital status
- independence level

Median, minimum, maximum and interquartile ranges, or means, standard deviations and ranges will be presented for continuous variables, and numbers and percentages for categorical variables. No statistical hypothesis testing will be carried out on any screening characteristics.

6.2 RECRUITMENT/RANDOMISATION

The date the first site opened to recruitment, number of sites open to recruitment and date current data lock was taken will be included in all reports, along with when sites close to recruitment.

Numbers of patients consented and randomised by site and overall will be presented (including information on date of first/last participant randomised and date site opened/closed to recruitment). Randomisation over time (monthly) by site will also be provided graphically. Numbers who drop out between consent and

randomisation will be presented overall, and by sex at birth (17), and their reasons for dropping out (where provided) will be tabulated.

6.3 PARTICIPANTS FOUND TO BE INELIGIBLE POST-RANDOMISATION

The number of participants found to be ineligible post-randomisation will be reported along with reasons for ineligibility. Any participant found to be ineligible at the point of randomisation during the data validation process (see section 3.2) who was subsequently randomised and continued in the trial will be included in the analyses. If significant numbers of participants become ineligible at or post-randomisation, a sensitivity analysis excluding these participants will be performed. As eligibility for WS is determined at screening, any value of the 10mWS collected at baseline is allowed and does not make the participant ineligible. See section 10.1.2 for more detail on how these intercurrent events will be handled during analysis.

6.4 CHANGE IN PARTICIPATION STATUS POST-RANDOMISATION

Randomised participants can change their participation status during the trial in the following ways:

1. Full withdrawal from all aspects of the study
2. Partial withdrawal from one or more of the following aspects:
 - a. Treatment (intercurrent event)
 - b. Follow up visits
 - c. Falls and exercise diary

The number of participants in each type of change in participation status, and reasons, will be reported in all analyses by randomised group, sex at birth (17) and overall. See section 5 for details on how these participants will be included in the various analysis populations.

The number (percentage) of participants lost to follow up will be reported overall and split by randomised allocation and sex at birth (17). The median length of follow up will be presented by treatment group in the final analysis. The number (percentage) of participants who provide 10mWS data at each follow up will be reported by randomised group. Graphical representations of the participation status of those randomised participants who experienced a change in participation status (as defined above) against the time since randomisation will be presented.

6.5 NON-COMPLIANCE (DEVIATIONS, VIOLATIONS AND SERIOUS BREACHES)

Non-compliances will be defined as either a deviation, violation or serious breach and reported on the Non-compliance form. The frequency of each category of deviation/violation/serious breach will be presented at both site and participant level. Treatment related non-compliance will be summarised, with further detail given in section 8.

7. BASELINE CHARACTERISTICS

A descriptive comparison of participant demographics and clinical characteristics at baseline will be presented, overall and split by randomised treatment arm and sex at birth (following the SAGER (17) guidelines). Additionally, these characteristics will also be reported for subset of participants who are retained and provide primary outcome data. These characteristics will include, but will not be limited to:

- recruitment site (stratification factor)
- age
- sex at birth
- ethnicity
- employment status

- marital status
- independence level

Median, minimum, maximum and interquartile ranges, or means, standard deviations and ranges will be presented for continuous variables, and numbers and percentages for categorical variables. No statistical hypothesis testing will be carried out on any baseline characteristics.

8. TREATMENT/INTERVENTION RECEIVED

To assess treatment compliance in the FES arm, the following information will be recorded and reported descriptively by sex at birth (following the SAGER (17) guidelines) and overall:

- The number of participants in the FES arm who did not start their allocated treatment, including the reason(s) for not starting treatment
- The number of participants in the FES arm who completed their allocated treatment
- The number of treatment delays in the FES arm and the reason(s) for these delays, if participants experience multiple delays, summary statistics for the number of delays experienced per participant will also be reported
- The number of participants in the FES arm who discontinue treatment and the reason(s) for the discontinuations
- The weekly step counts, calculated using steps recorded by the ODFS pace device
- The dose, measured using the time spent with stimulation as recorded by the ODFS pace device
- The number and percentage of participants meeting the adherence threshold will be reported overall and by sex at birth.

Adherence to treatment will be assessed using device-recorded step counts. Participants recording a mean of at least 1,000 device-recorded steps per week during the treatment period will be considered adherent. This threshold was informed by data from the feasibility study, in which the lower quartile of weekly device-recorded step counts was 1,152 steps per week, and was selected to reflect the anticipated variability in physical fitness and mobility within the study population. No statistical significance testing will be performed.

9. SAFETY ANALYSIS/HARMS

See section 13.1 of the trial protocol for details of how *device-related adverse reactions (DARs)* and *serious adverse events (SAEs)* are defined. All safety analysis will utilise the TS with observed data only. Safety outcomes will be descriptively summarised. The number of *DARs*, alongside the number of participants experiencing these *reactions*, will be reported on an as treated basis, for all oversight committee reports and at the end of the trial. All *SAEs*, device related or otherwise, will be reported similarly. In addition, odds ratios and risk differences (with 95% confidence interval) will be reported for deaths across the follow up period in accordance with CONSORT Harms 2022 (18). All safety outcomes will also be presented disaggregated by sex at birth as per the SAGER (17) guidelines.

10. ANALYSIS

10.1 PRIMARY OUTCOME

10.1.1 NULL HYPOTHESIS

There is no difference in the mean change of 10mWS (between baseline and 18 weeks post-randomisation) between FES and usual care versus usual care alone in pwPD.

10.1.2 METHODS

TMP_005_ST V2 10 Jul 2025

STEPS II, IRAS ID: 330866

SAP V1.0, 20 Aug 2026

RESTRICTED

Page 16 of 35

PRIMARY ANALYSIS METHODS FOR PRIMARY OUTCOME

Details of how 10mWS will be calculated can be found in appendix 2. The primary analysis of the primary outcome measure will compare the change in 10mWS between allocated groups using a restricted maximum likelihood (REML) approach using a mixed model with repeated measures (MMRM). The model will account for degrees of freedom using the Kenward-Roger method. It will adjust for the fixed effects of allocated treatment arm, baseline 10mWS values, ankle proprioceptive sensory discrimination (a baseline assessment of ankle sensory function measured using a two-point discrimination) and recruiting site (stratification factor) (19). Timepoints will be treated as the repeated measure using an unstructured covariance structure to model within-patient measurements. The adjusted mean-difference (and two-sided 95% CI) for the change in 10mWS at 18 weeks post-randomisation will be presented. Additionally, the adjusted mean change (and two-sided 95% CI) from baseline will be calculated from the model for each treatment group. This is a valid and unbiased approach when these data are missing at random (MAR) or missing completely at random (MCAR). The primary analysis will be independently coded by a second statistician to verify the results. This analysis will use the PAS and only observed data will be supplied to the model.

ANALYSIS METHODS FOR HANDLING INTERCURRENT EVENTS: PRIMARY OUTCOME

The primary analysis will account for the pre-specified intercurrent events (ICEs) (2,3) using the following methods:

- Changes to participant health status
 - Death – principal stratum: data from participants who die during the trial will be excluded from the primary analysis. A secondary analysis (see following section) will be conducted to test the assumptions from this method of handling the ICE.
 - Pregnancy – see *Events related to trial procedures* point below.
- Treatment discontinuation
 - Full withdrawal – while on treatment: participant data collected prior to withdrawal from the trial will be included in analyses. Data following withdrawal will be considered as ‘non-existent’ and will not be explicitly imputed during the planned imputation processes (see section 10.5.3).
 - Partial withdrawal - treatment policy: participant data will be included in the analysis as if the ICE did not occur. A sensitivity analysis following the primary analysis method will be undertaken using the principal stratum strategy in the sub-population who do not experience treatment discontinuation.
- Events related to trial procedures
 - Found to have been ineligible at point of randomisation – treatment policy: participants who are found to be ineligible at the point of randomisation during the data validation process will be included in the primary analysis as if the ICE did not occur (20). No sensitivity analysis is planned for this ICE.
 - Found to be ineligible post-randomisation – treatment policy: in the primary analysis as if the ICE did not occur (20). A sensitivity analysis following the primary analysis method will be undertaken following the principal stratum method in the sub-population who did not experience this ICE.
 - Intervention adherence – treatment policy: participant data will be included in the primary analysis regardless of the level of adherence. An exploratory analysis of the primary outcome targeting the sub-population who meet the prespecified minimum adherence criterion ($\geq 1,000$ device-recorded steps per week on average during the treatment period; see section 8 for details) will be assessed using a complier-average causal effect (CACE) analysis, which will utilise a two-stage least squares instrumental variable approach to provide an unbiased estimate of the efficacy of FES.

- Data collected out of window – treatment policy: participant data will be included in the primary analysis model regardless of when they provided outcome data.

Note for all ICE related sensitivity analyses, only the ICE of interest will be changed; all other ICEs will be treated as in the primary analysis estimand.

ADDITIONAL ANALYSIS METHODS FOR THE PRIMARY OUTCOME

Descriptive statistics for baseline demographics will be assessed for any imbalances between the two allocated groups. In the case a significant imbalance is found in any baseline participant characteristics thought to be predictive of the primary outcome, sensitivity analyses with additional adjustment for these characteristics will be performed using the same methods described for the primary analysis above.

The assumptions of the analysis model for the primary outcome (linearity, normality of residuals, random effects independent of fixed effects, data missing at random) will be assessed graphically and if there is evidence of significant violation of assumptions, alternative methods will be explored, such as transforming the data or bootstrapping methods.

Pre-specified sensitivity analyses will also be undertaken to assess the robustness of the primary analysis of the primary outcome measure results, specifically the missing at random (MAR) assumption. To explore the impact of this assumption and impact of the missing data, a “best-worst” and “worst-best” case (15) sensitivity analysis will be used (see section 10.5.3 for more detail on these imputations).

An adaptive trimmed-means approach (21), which is valid when data are missing not at random (MNAR), may be undertaken as a secondary analysis if the data does not meet the MAR assumption. This approach targets a different estimand to that of the primary analysis, by using a composite strategy to handle the intercurrent event of participants leaving the study early, therefore discontinuing their allocated treatment.

The primary analysis will be repeated using imputed data using multiple imputation (see section 10.5.3 for details on the MI process) if the data meets the MAR assumption.

A supplementary analysis will be conducted with deaths handled using the while-alive approach. All other intercurrent events will be handled in the same way as the primary estimand. To use the while-alive approach, repeated measures mixed models cannot be used due to the implicit imputation that occurs when using these models. Instead, a linear regression model will be used having imputed missing 10mWS data (as advised from person communication with Brennan Kahan). The first step will be to impute all missing 10mWS data in the PAS using multiple imputation following the methods specified in Section 10.5.3 (post-death missing 10mWS data and standard missing 10mWS data are treated the same). Following the imputation, all post-death missing 10mWS data will be set to missing again and a new while-alive outcome variable will be created. This will be a single outcome per participant that will be equal to the week 18 10mWS if the participant was alive at 18 weeks, equal to the week 6 assessment if the participant died between week 6 and week 18, equal to the week 2 assessment if the participant died between week 2 and week 6 and the baseline assessment if the participant passed away between randomisation and week 2 assessment. The change in 10mWS will be calculated from this “while-alive” outcome. The model will adjust for the fixed effects of allocated treatment arm, baseline 10mWS values and ankle proprioceptive sensory discrimination and recruiting site (stratification factor). The estimated between-group difference in change in absolute 10mWS will be extracted from the model, together with the 95% confidence interval. It is expected the estimate should be broadly similar to the primary estimand.

All additional analyses of the primary outcome will use the PAS with the various imputations implemented as necessary.

10.2 SUBGROUP ANALYSIS

The primary analysis for the primary outcome will be repeated for the following subgroups by including an interaction with allocation for the relevant subgroup in the primary analysis model: however, this analysis will be exploratory only, and results will be interpreted cautiously, as the study is not powered to detect any subgroup effects:

- Hoen and Yahr groups (1 and 2 vs 3 and 4).
- Tremor dominance (Yes vs No) - based on clinical opinion
- Slow and fast walkers (Slow vs Fast) – participants will be categorised as slow (<0.8m/s) or fast (0.8-1.25m/s) according to their baseline walking speed.
- Sex at birth (Male vs Female) (17) – participants who choose not to disclose their sex at birth will be set as missing for all analyses where sex at birth is used

Results will be presented using confidence intervals, as well as a global p value for the interaction between the allocation and each of the identified sub-groups of interest. All subgroup analyses will utilise the PAS and only observed data will be supplied to the model.

10.3 SECONDARY CLINICAL OUTCOMES

PRIMARY ANALYSIS METHODS FOR SECONDARY CLINICAL OUTCOMES

See section 1.3.2 for the list of secondary clinical outcomes. Details of how these outcomes will be calculated can be found in appendix 2.

Continuous secondary clinical outcomes will be descriptively summarised and analysed using the following approaches (i) using simple two-sample t-test approach to calculate the unadjusted between-group difference and two-sided 95% confidence interval and (ii) using a MMRM (19) approach to calculate the adjusted mean-difference (and two-sided 95% CI) following the method outlined for the primary analysis (see section 10.1.2). Estimates and two-sided 95% CIs will be presented for all timepoints. This a valid and unbiased approach when these data are missing at random (MAR) or missing completely at random (MCAR).

The potential carryover effect of FES will also be investigated by comparing the between-group difference in the change at week 18 and week 22, with the estimated effect (and corresponding 95% confidence interval) calculated from the mixed repeated measures (19) model.

Binary secondary clinical outcomes, including the proportions of participants with an increase in walking speed of at least 0.13ms^{-1} , will be analysed using logistic regression models, with adjustment for treatment arm, recruiting site, baseline values (where applicable) and ankle proprioceptive sensory discrimination.

Falls rate will be analysed by calculating the relative risk using a negative binomial model with appropriate adjustment for treatment arm, recruiting site, history of falls (categorized as falls less than once a year, once a year and more than once a year) and ankle proprioceptive sensory discrimination.

Intercurrent events will be accounted for in the same manner as the primary analysis estimand (see section 10.1.2) for all secondary outcomes. All secondary clinical outcomes will be analysed using the PAS and only observed data will be supplied to the models.

10.4 BAYESIAN ANALYSIS OF OUTCOME MEASURES

There is no planned Bayesian analysis of the primary or secondary outcomes. Bayesian methods are used in the analysis of the mechanistic outcomes (see section 10.5).

10.5 ADDITIONAL ANALYSES

10.5.1 MECHANISTIC ANALYSIS

DEFINITION AND CALCULATION OF MECHANISTIC OUTCOME MEASURES

The following mechanistic outcome measures will be used to explore hypothesised pathways between treatment allocation and walking speed. Full computational details are presented in Appendix 2.

- 1. Inter-limb coordination, measured using the Phase Coordination Index (PCI), at baseline, 2 weeks and 6 weeks post-randomisation.
- 2. Whole-body coordination, measured using the mediolateral anticipatory postural adjustments (APA) at baseline, 2 weeks and 6 weeks post-randomisation.
- 3. Muscle strength, measured using maximal voluntary isometric contraction at baseline, 2 weeks and 6 weeks post-randomisation.
- 4. Foot velocity during swing phase (walking), measured using step-sensor and motion-capture data at baseline, 2 weeks and 6 weeks post-randomisation.
- 5. Foot-tapping velocity (seated; 10 taps), measured using step-sensor and motion-capture data at baseline, 2 weeks and 6 weeks post randomisation.

ANALYSIS OVERVIEW

Changes in muscle strength, targeted (ankle dorsiflexors and everters) and not targeted (hip flexors), will be assessed at each time point of interest following the methods for the primary analysis of the secondary outcomes. The relationship between the change in muscle strength and the change in walking speed and duration (as assessed by activity monitors), over time, will be determined.

To evaluate the prespecified mechanistic hypotheses, a two-stage, single-group joint Structural Equation Modelling (SEM) framework will be implemented in a Bayesian estimation paradigm. This framework enables simultaneous estimation of (i) the effects of treatment on mechanistic outcomes and (ii) the contribution of mechanistic changes to subsequent walking speed improvements.

STRUCTURAL MODELLING FRAMEWORK

Stage 1: Mechanistic Processes

At week 2, each mechanistic outcome (PCI, APA magnitude, swing phase foot velocity, and foot-tapping velocity) will be regressed on:

- randomised treatment allocation
- the corresponding baseline value of the mechanistic outcome
- baseline ankle proprioceptive sensory discrimination score
- recruiting site

The residual variation across the five week 2 mechanistic outcomes will be allowed to correlate freely, capturing shared unmeasured physiological influences.

Stage 2: Walking Speed

Change in 10-metre walking speed (10mWS) at week 6 will be jointly modelled as a function of:

- all five week 2 mechanistic outcomes
- treatment allocation

- baseline 10mWS
- baseline ankle proprioceptive sensory discrimination score
- recruiting site

The modelling described in Stage 1 will be repeated for mechanistic outcomes measured at week 6 using the same adjustment set. Change in 10mWS at week 18 will be modelled analogously to Stage 2, using the four week 6 mechanistic outcomes alongside treatment allocation and baseline covariates. Residual correlations among week 6 mechanistic outcomes will again be freely estimated.

ESTIMATION METHOD

All SEMs will be estimated under a Bayesian framework using Markov Chain Monte Carlo (MCMC) sampling. This approach yields posterior distributions for all parameters, accommodates non-normal sampling distributions for indirect effects, and provides robust inference for mediated pathways.

Likelihood Specification

All mechanistic outcomes and walking speed outcomes will be modelled using Gaussian likelihoods with unknown means and variances estimated within the model.

Prior Distributions

Weakly informative priors will be used to allow the data to dominate inference while ensuring numerical stability:

- Regression coefficients (all structural paths): Normal(0, 1,000)
- Intercepts: Normal(0, 1,000)
- Residual standard deviations: Half-Cauchy(0, 5)
- Residual correlation matrix for mechanistic outcomes: LKJ(1) prior

These priors are consistent with recommended practice for Bayesian SEMs.

Posterior Sampling and Convergence

Posterior estimation will use multiple MCMC chains, each with:

- an initial burn-in period
- a large number of retained samples
- thinning as required

Convergence will be assessed using:

- trace-plot inspection
- the Gelman–Rubin statistic (R-hat <1.01)
- effective sample sizes
- absence of divergent transitions

Posterior summaries will include posterior means, posterior medians (as needed), and 95% credible intervals (CrIs).

Missing Data

Missing values will be addressed using Bayesian full information estimation, equivalent to FIML in the frequentist setting, under a Missing at Random (MAR) assumption.

EFFECTS TO BE REPORTED

The following effects will be extracted directly from the joint Bayesian SEM and summarised using posterior means and 95% credible intervals:

TMP_005_ST V2 10 Jul 2025

STEPS II, IRAS ID: 330866

SAP V1.0, 20 Aug 2026

RESTRICTED

Page 21 of 35

- Treatment effects on each mechanistic outcome at week 2 and 6
- Effects of each mechanistic outcome on walking speed at weeks 6 and 18
- Direct treatment effects on walking speed at weeks 6 and 18
- Indirect (mediated) effects through each mechanistic pathway
- Total mediated effects across all mechanistic pathways
- The proportion of the total treatment effect mediated by the combined mechanisms

Indirect and proportion mediated effects will be obtained directly from their posterior distributions, avoiding reliance on normal-theory approximations. Intercurrent events will be accounted for in the same manner as the primary analysis estimand (see section 10.1.2).

These mechanistic analyses are exploratory only and will be interpreted with appropriate caution.

10.5.2 NON-COMPARATIVE MEASURES

Non-comparative measures (as defined in section 1.4) will be reported descriptively at each timepoint. Median, minimum, maximum and interquartile ranges, or means, standard deviations and ranges will be presented for continuous variables, and numbers and percentages for categorical variables. In addition, mean change scores will be reported alongside 95% CIs where applicable. Summary statistics and change scores (with 95% CIs) will be plotted for each variable by timepoint. Where applicable, measures taken with FES and without FES will be presented graphically.

10mWS (as part of the bradykinesia secondary clinical outcome) is measured four times at each timepoint. The first is a warmup walk and will not be used in these analyses. The second is representative of walking without FES, while the third is representative of walking with FES. The fourth walk is without FES and represents a short-term carryover effect. (12). Change scores for the short-term carryover effect will be presented in addition to the FES/non-FES scores and presented following the methods for other non-comparative measures. The observed data only will be used for all non-comparative measures.

Analyses of the non-comparative measures will utilise the PAS and use the observed data only.

10.5.3 IMPUTATION OF DATASETS FOR ANALYSIS

MULTIPLE IMPUTATION

Multiple imputation (14) will be used for various analyses in this trial. The multiple imputation model will include all variables specified for inclusion in the primary analysis model (see section 10.1.2), along with any auxiliary variables deemed to be predictive of the missing variable data. We will create binary indicator variables for the outcomes of interest that indicate missingness (e.g. 0 indicates missing and 1 indicates non-missing). Logistic regressions will be used to assess the mechanisms of missingness that may be present in the baseline data.

This model will follow Multiple Imputation by Chained Equations (MICE) methods (25). Continuous variables will be imputed using predictive mean matching (PMM) (26,27). This prevents the potential imputation of impossible values that can occur with other methods and is still valid for non-normally distributed variables. Binary and categorical variables will be imputed using logistic regression. Data will be imputed separately for each randomised group. A passive approach will be utilised for any derived outcomes (i.e. imputation will happen prior to calculating derived variables). This imputation will assume MAR. At least m imputed data sets will be imputed where m is the percentage of participants with incomplete data (i.e. if 10% of participants have incomplete data then 10 data sets will be imputed). This number will be increased if stable estimates of the standard error are not produced. In the event of non-convergence, the maximum number of iterations will be increased.

Summary statistics (mean and SD) will be plotted for each imputed variable against iteration number to check convergence of the imputation model. Additionally, density plots will be used to compare the compatibility of the observed and imputed distributions.

The analysis model will be run for each dataset separately and pooled using Rubin's (28) rules. The analysis model will be run for each dataset separately and pooled using Rubin's (28) rules.

BEST/WORST AND WORST/BEST CASE IMPUTATION

Best/Worst and worst/best case imputation (15) will be used in a sensitivity analysis of the primary outcome only. Missing values will be imputed using the group mean $\pm$ the group standard deviation (depending on if calculating worst or best case) for continuous variables. A passive approach will be taken where imputation will occur before the calculation of the derived outcome.

11. ECONOMIC EVALUATION

There is no formal economic evaluation in the STEPS II efficacy trial. Intervention health resource cost will be collected and summarised. This information will be used to inform future studies.

12. STATISTICAL SOFTWARE

Statistical software, including version number, will be documented on all reports and publications. All analysis will be carried out in appropriate statistical software such as STATA (29) or R (30). If other specialist programs are used these will also be stated within the report.

13. STORAGE AND ARCHIVING

Final data extracts used to analyse the trial will be stored in a restricted section of the trial SharePoint. All data extract filenames will contain the date and time stamps of extraction. Final analysis code will be stored in the Statistics section of the trial SharePoint. All scripts will be logged with date and time stamps, and these will be stored with the analysis code.

14. REFERENCES

1. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017 Dec 19;318(23):2337.
2. European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials [Internet]. 2020 [cited 2025 Nov 13]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf
3. Kahan BC, Hindley J, Edwards M, Cro S, Morris TP. The estimands framework: a primer on the ICH E9(R1) addendum. *BMJ*. 2024 Jan 23;e076316.
4. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders*. 2008 Nov 15;23(15):2129–70.
5. Nieuwboer A, Rochester L, Herman T, Vandenberghe W, Emil GE, Thomaes T, et al. Reliability of the new freezing of gait questionnaire: Agreement between patients with Parkinson's disease and their carers. *Gait Posture*. 2009 Nov;30(4):459–63.
6. Taylor PN, Sampson T, Beare B, Donavon-Hall M, Thomas PW, Marques E, et al. The effectiveness of peroneal nerve functional electrical stimulation for the reduction of bradykinesia in Parkinson's disease: A feasibility study for a randomised control trial. *Clin Rehabil*. 2021 Apr 1;35(4):546–57.
7. Yardley L, Beyer N, Hauer K, Kempen G, Piot-Ziegler C, Todd C. Development and initial validation of the Falls Efficacy Scale-International (FES-I). *Age Ageing*. 2005 Nov 1;34(6):614–9.
8. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. *J Rehabil Med*. 2010 Apr;42(4):323–31.
9. JENKINSON C, FITZPATRICK R, PETO V, GREENHALL R, HYMAN N. The Parkinson's Disease Questionnaire (PDQ-39): development and validation of a Parkinson's disease summary index score. *Age Ageing*. 1997;26(5):353–7.
10. Devlin N, Pickard S, Busschbach J. The Development of the EQ-5D-5L and its Value Sets. In: *Value Sets for EQ-5D-5L*. Cham: Springer International Publishing; 2022. p. 1–12.
11. Jost ST, Kaldenbach M, Antonini A, Martinez-Martin P, Timmermann L, Odin P, et al. Levodopa Dose Equivalency in Parkinson's Disease: Updated Systematic Review and Proposals. *Movement Disorders*. 2023 Jul 5;38(7):1236–52.
12. Taylor PN, Seary C, Sampson T, Beare B, Stevenson VL, Donavon-Hall M, et al. The orthotic and therapeutic effects of peroneal nerve functional electrical stimulation in Parkinson's disease: Post hoc analysis from a feasibility study for a randomized control trial. *Artif Organs* [Internet]. 2022 Nov 29 [cited 2025 Dec 11];46(11):264–7. Available from: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/aor.14408>
13. Scherr J, Wolfarth B, Christle JW, Pressler A, Wagenpfeil S, Halle M. Associations between Borg's rating of perceived exertion and physiological measures of exercise intensity. *Eur J Appl Physiol*. 2013 Jan 22;113(1):147–55.

14. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009 Sep 1;338(jun29 1):b2393–b2393.
15. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol*. 2017 Dec 6;17(1):162.
16. Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *BMJ*. 2025 Apr 14;389:e081123.
17. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. *Res Integr Peer Rev*. 2016 May 3;1(1):2.
18. Junqueira DR, Zorzela L, Golder S, Loke Y, Gagnier JJ, Julious SA, et al. CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *BMJ*. 2023 Apr 24;381:e073725.
19. Schuler A. Mixed Models for Repeated Measures Should Include Time-by-Covariate Interactions to Assure Power Gains and Robustness Against Dropout Bias Relative to Complete-Case ANCOVA. *Ther Innov Regul Sci*. 2022 Jan 21;56(1):145–54.
20. Yelland LN, Sullivan TR, Voysey M, Lee KJ, Cook JA, Forbes AB. Applying the intention-to-treat principle in practice: Guidance on handling randomisation errors. *Clinical Trials*. 2015 Aug 1;12(4):418–23.
21. Wang M, Liu J, Molenberghs G, Mallinckrodt C. An evaluation of the trimmed mean approach in clinical trials with dropout. *Pharm Stat*. 2018 May 6;17(3):278–89.
22. Hair JF, Hult GTM, Ringle CM, Sarstedt M, Danks NP, Ray S. An Introduction to Structural Equation Modeling. In 2021. p. 1–29.
23. Streiner DL. Finding Our Way: An Introduction to Path Analysis. *The Canadian Journal of Psychiatry*. 2005 Feb 1;50(2):115–22.
24. Burton PR. Helping doctors to draw appropriate inferences from the analysis of medical studies. *Stat Med*. 1994 Sep 15;13(17):1699–713.
25. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011 Feb 20;30(4):377–99.
26. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. *BMC Med Res Methodol*. 2014 Dec 5;14(1):75.
27. Little RJA. Missing-Data Adjustments in Large Surveys. *Journal of Business & Economic Statistics*. 1988 Jul;6(3):287–96.
28. Rubin DB. Multiple Imputation for Nonresponse in Surveys. In Wiley; 1987. p. 29–62.
29. StataCorp. Stata Statistical Software: Release 18.5. College Station, TX: Stata Press; 2024.

Statistical Analysis Plan

30. R Core Team. R Foundation for Statistical Computing, Vienna, Austria. 2024 [cited 2025 May 15]. R: A Language and Environment for Statistical Computing. Available from: <https://www.R-project.org/>
31. Goetz CG, Luo S, Wang L, Tilley BC, LaPelle NR, Stebbins GT. Handling missing values in the MDS-UPDRS. *Mov Disord*. 2015 Oct;30(12):1632–8.
32. National Institute for Clinical Excellence (NICE). Position statement on use of the EQ-5D-5L value set for England (updated October 2019) [Internet]. 2019 [cited 2025 Nov 24]. Available from: <https://www.nice.org.uk/position-statements/position-statement-on-use-of-the-eq-5d-5l-value-set-for-england-updated-october-2019>
33. van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets. *Value in Health*. 2012 Jul;15(5):708–15.

15. APPENDICES

15.1 APPENDIX 1: ABBREVIATIONS

Table 2 Table of abbreviations

Abbreviation	Meaning
10mWS	10m walking speed
10mWT	10m walking test
AE	Adverse Event
AFO	Ankle Foot Orthoses
APA	Anticipatory Postural Adjustment
AR	Adverse Reaction
BOTS	Blinding of Trial Statisticians
CI	Confidence Interval
COM	Centre of mass
CONSORT	Consolidated Standards of Reporting Trials
CrI	Credible interval
DMEC	Data monitoring committee
ExeCTU	University of Exeter Clinical Trials Unit
EQ-5D-5L	EuroQol 5 dimension 5 level health related quality of life scale
FES	Functional electrical stimulation
FES-I	Falls efficacy scale - international
ICE	Intercurrent event
ISRCTN	International Standard Registered Clinical/social sTudy Number
IMU	Inertial Measurement Unit
LEDD	Levodopa Equivalent Daily Dose
MAR	Missing at Random
MCAR	Missing Completely at Random
MDS-UPDRS	Movement Disorder Society - Unified Parkinson's Disease Rating Scale
MI	Multiple Imputation
MICE	Multiple imputation by chained equations
mini-BESTtest	Mini Balance Evaluation System Test
MNAR	Missing Not at Random
N-FOG	New freezing of gait questionnaire
NICE	National Institute for Health and Care Excellence
ODFS	Odstock Dropped Foot Stimulator
PCI	Phase Coordination Index
PD	Parkinson's disease
PDQ-39	Parkinson's Disease Questionnaire 39 (quality of life)
PenCTU	Peninsula Clinical Trials Unit
PMM	Predictive mean matching
pwPD	People living with Parkinson's disease
QoL	Quality of Life
RCT	Randomised Controlled Trial
SAE	Serious Adverse Event
SAP	Statistical analysis plan
SD	Standard deviation

Statistical Analysis Plan

STEPS	The Effectiveness of Peroneal Nerve Functional Electrical ST imulation (FES) for the Reduction of Bradykinesia in Parkinson's Disease: A Pragmatic Feasibility Study for a Single Blinded Randomised Control Trial
TMF	Trial master file
TSC	Trial steering committee
TUG	Timed Up and Go test
WS	Walking speed

15.2 APPENDIX 2: OUTCOME CALCULATIONS

Calculations for the primary and secondary clinical outcome measures are presented below. Any strategies for handling missing items in a PROM will be used prior to any planned imputation.

15.2.1 10MWS

Used for the primary outcome, the whole-body bradykinesia secondary clinical outcome and the bradykinesia secondary non-comparative measure. The 10mWS will be calculated from the time recorded for the completion of the 10mWT following the formula below:

$$10mWS = \frac{10}{10mWT \text{ time}}$$

The 10mWS values will be calculated in metres per second (ms⁻¹). As distance is fixed and the time taken is a single value this outcome data will have no strategies for managing partially missing outcome data.

15.2.2 INSTRUMENTED TOE TAP TEST

Used for the bradykinesia secondary clinical outcome. This outcome is calculated during task 3.7 of the MDS-UPDRS (4). Speed of movement is calculated using an attached IMU during the task. Speed of movement (foot tapping velocity) will be measured in ms⁻¹ and averaged over 10 taps. The mean foot tapping velocity will be calculated separately for the right and left foot. As these outcomes are derived as single summary values for each foot, no specific strategies for managing partially missing outcome data are required.

15.2.3 STRIDE LENGTH

Used for the hypokinesia secondary clinical outcome. Stride length will be calculated from the 10mWT from the number of strides taken as per the formula below:

$$\text{Stride length} = \frac{10}{\text{number of strides taken}}$$

The values of stride length will be calculated in metres (m). As distance is fixed and stride length is a single value, this outcome data will have no strategies for managing partially missing outcome data.

15.2.4 MDS-UPDRS

Each parkinsonian sign or symptom is rated on a 5-point Likert-type scale (ranging from 0 (no problems) to 4 (severe problems)), with higher scores indicating more severe impairment. The MDS-UPDRS (4) is split into four sections:

1. Non-Motor Aspects of Experiences of Daily Living (13 items)
2. Motor Aspects of Experiences of Daily Living (13 items)
3. Motor Examination (33 items)
4. Motor Complications (6 items)

Sections 1 and 2 are used for calculating the activities of daily living secondary clinical outcome. This results in a range of scores from 0 to 52 for each section. As recommended, each section will be analysed and reported separately (4).

Section 3 is used for the hypokinesia secondary clinical outcome. This results in a range of scores from 0 to 132.

No units are used for outcomes using MDS-UPDRS data. Any items marked as “not applicable” will be scored as 0. Missingness will be assessed for each section separately as each section allows for different number of missing items (31) as follows:

1. Non-Motor Aspects of Experiences of Daily Living (1 item allowed to be missing)
2. Motor Aspects of Experiences of Daily Living (1 item allowed to be missing)
3. Motor Examination (<4 items allowed to be missing)
4. Motor Complications (no items allowed to be missing)

Where responses contain an acceptable level of missing data the total score will be imputed (31) as follows:

1. Sum the non-missing items
2. Multiply by number of items in section (e.g. multiply by 13 in section 1)
3. Divide by number of non-missing items

15.2.5 N-FOG

Used for the akinesia secondary clinical outcome. The N-FOG (5) consists of 9 items, consisting of a binary question to determine if FOG occurred and then 8 items rated on a 4 or 5-item Likert-type scale (ranging from 0 (no problems) to 3/4 (severe problems)), with higher scores indicating more severe impairment. This gives a range of scores from 0 to 28 (the sum of the 8 Likert-style questions) for those whom have experienced FOG and 0 for those who answered “no” to the first item. No units are used for this outcome. For those who answer “yes” to the first question, the total score will be set as missing if any of the 8 likert style items are missing. If the first item is missing, the total will also be set as missing.

15.2.6 FALLS DIARY

Used for the falls secondary clinical outcome. The falls diary (6) records the number of falls and fall related injuries. Both of these will be used as outcomes in this trial. Both outcomes will be measured from baseline until 18 weeks post-randomisation and from 18 weeks post-randomisation until the 22 week post-randomisation endpoint. Note as we are not collecting falls in the period pre-baseline, no baseline scores will be used as a adjusting factor in the analysis of these outcomes. Missingness for the number of falls and number of injury related falls will be assessed separately. Any missing data will result in the final outcome being classed as missing.

Falls rate (measured as number of falls per year) will be used to measure the number of falls outcome measure.

Injury related falls will instead be categorised into a binary indicator of any injury related falls for each timepoint (6). No units will be used for this outcome measure.

15.2.7 FES-I

Used for the fear of falling secondary clinical outcome. The FES-I (7) is a 16-item questionnaire where each response is rated on a 4-point Likert style scale (1 = Not at all concerned, 2 = Somewhat concerned, 3 = Fairly concerned, 4 = Very concerned). The total score is the sum of the responses ranging from 16 to 64 with higher scores indicating greater fear of falling. No units are used for this outcome.

Responses with >4 missing items will have their total set to missing (7). Where there are ≤4 missing values, the total score will be imputed as follows:

1. Non-missing items summed
2. Divide sum by number of non-missing items
3. Multiply by 16 and round up to nearest integer if necessary

15.2.8 MINI-BEST TEST

Used for the balance secondary clinical outcome. The mini-BESTest (8) consists of 14 items each rated on a 3-point Likert-type scale (ranging from 0 (severe) to 2 (normal)). The total score is the sum of the responses ranging from 0 to 28, with higher scores indicating better balance. No units are used for this outcome. If any of the 14 items are missing, the total score will be set to missing.

15.2.9 DAILY STEP COUNT

Used for the activity level secondary clinical outcome. This outcome is measured using a StepWatch prior to the participant attending their blinded assessment. The mean value of the seven daily step counts will be calculated and used as the measure of comparison. In the case that more than seven values are provided, the final seven counts will be used (as these are closest to the assessment). For example, if nine days of data are provided, then data from days 3 to 9 inclusive will be used. If there are less than 7 days of data provided, the mean value will be set to missing at that timepoint.

15.2.10 PDQ-39

Used for the quality-of-life secondary clinical outcome. The PDQ-39 (9) consists of 39 items each rated on a 5-point Likert-type scale (ranging from 0 (never) to 4 (always)), with higher scores indicating lower QoL. The outcome is split into eight sections that each focus on a different aspect of QoL (number of items, item numbers):

1. Mobility (10, #1-10)
2. Activities of daily living (ADL) (6, #11-16)
3. Emotional well-being (6, #17-22)
4. Stigma (4, #23-26)
5. Social support (3, #27-29)
6. Cognition (4, #30-33)
7. Communication (3, #34-36)
8. Bodily discomfort (3, #37-39)

The score for each section is calculated as follows:

1. Calculate the sum of the scores
2. Divide by the maximum score of that section (e.g. for section 1 you would divide by 40)
3. Multiply by 100 to get a total score

This will give a score between 0 and 100 for each section. Summary statistics of the scores for these sections will be presented. The overall total score will be calculated by taking the mean value of the eight section total scores. Missingness will be calculated separately for each section, where any missing item will result in that section being classed as missing. If any section is missing, the overall total will also be classed as missing.

15.2.11 EQ-5D-5L

Used for the quality-of-life secondary clinical outcome. This questionnaire is divided into two parts:

- A 5-item descriptive system providing a possible 3125 health states consisting of five dimensions which are scored between 1-5 levels of perceived problems (1 = no problem through to 5 = unable to/extreme problems:
 - Mobility
 - Self-care
 - Usual activities
 - Pain/discomfort
 - Anxiety/depression
- A VAS with a range of 0-100

Statistical Analysis Plan

Following current NICE guidance (32), the 5-level descriptive system data will be mapped onto the 3-level value set for the UK (33). Index scores range from -0.59 to 1; 1 is the best possible health state. Negative values represent health states perceived as worse than dead, which is equal to 0.

No units are used for this outcome. Missingness for the VAS and descriptive system will be calculated separately. The VAS is a single value, and no strategies will be employed for partial missingness. The total for the descriptive system will be set as missing if any of the five items are missing.

15.2.12 LEDD

The LEDD (11) is used for the normal care secondary clinical outcome. Standard conversion formulae (11) are used to convert the dosage of various PD medications into a single LEDD. The following medications (and associated scale factors) are captured:

Table 3. Scale factors for LEDD conversion

Drug name	Scale factor	Notes
Co-beneldopa	1	The larger of co-beneldopa and co-careldopa is taken as the levodopa dose (L).
Co-careldopa	1	The larger of co-beneldopa and co-careldopa is taken as the levodopa dose (L).
Pramipexole	100	
Ropinirole	20	
Rotigotine	30	
Apomorphine	10	Dose is per ml (e.g. 3ml would be 10x3)
Rasagiline	100	
Selegiline (oral)	10	
Safinamide	1.25	
Entacapone	$(L \times 0.33) + L$	
Co-careldopa and entacapone	1.33	
Opicapone	$(L \times 0.5) + L$	
Amantadine (oral capsules)	1	If taken in solution dose is per ml
Procyclidine	20	
Trihexyphenidyl or benzhexol	100 per dose	The strength of the dose does not matter (e.g. both 3mg three times a day and 4mg three times per day will result in a LEDD of 300mg)
Other		

Up to ten medications can be recorded per participant. The final LEDD score will be calculated by summing the converted amounts for each drug. All doses will be converted into mg/ml prior to conversion and the total LEDD amounts will be reported in mg. Where the data for any individual drug is incomplete (e.g. missing units), the total LEDD will be set as missing. Participants who have indicated that they are receiving PD medication but input no further data will have their LEDD set as missing. The LEDD will be set at 0mg for participants who report no usage of PD medication.

15.2.13 PCI

Inter-limb coordination will be assessed using the Phase Coordination Index (PCI), which quantifies deviation and variability around the ideal 180° phase relationship between right and left stride durations. PCI will be derived using established methods combining phase error and phase variability. Lower PCI values indicate improved inter-limb coordination.

15.2.14 APA

Whole-body coordination will be assessed using mediolateral anticipatory postural adjustments estimated from centre-of-mass (COM) mediolateral velocity during preparatory stepping. APA magnitude will be defined as the peak mediolateral COM velocity occurring immediately before step initiation. Higher APA magnitudes indicate improved postural control.

15.2.15 MUSCLE STRENGTH

Muscle strength will be assessed for ankle dorsiflexors, ankle evertors, and hip abductors using maximal voluntary isometric contraction. Strength will be recorded as force. Higher values indicate greater muscle strength.

15.2.16 FOOT VELOCITY DURING SWING PHASE (WALKING)

Foot velocity during the swing phase provides a measure of lower-limb bradykinesia during walking. Peak sagittal-plane foot velocity during the swing phase of gait will be extracted from step-sensor or motion-capture data. Higher velocities indicate reduced bradykinesia.

15.3 APPENDIX 3: EXAMPLE TABLES

Table 4. Example participant demographics

Characteristic	Overall N = 100 ¹	Usual care N = 42 ¹	FES + usual care N = 58 ¹
Age (years)			
Mean (SD)	62 (25)	64 (23)	61 (27)
Median (Q1, Q3)	62 (39, 85)	64 (44, 87)	58 (36, 83)
Min, Max	19, 107	26, 103	19, 107
Missing	0	0	0
Sex at birth			
Male	52 (52%)	22 (52%)	30 (52%)
Female	48 (48%)	20 (48%)	28 (48%)
Missing	0	0	0
Ethnicity			
White	14 (14%)	4 (9.5%)	10 (17%)
Mixed	21 (21%)	8 (19%)	13 (22%)
Asian	23 (23%)	8 (19%)	15 (26%)
Black	20 (20%)	13 (31%)	7 (12%)
Other	22 (22%)	9 (21%)	13 (22%)
Missing	0	0	0

¹Mean (SD); Median (IQR); Min, Max or Frequency (%)

This data is simulated