University of BRISTOL

CHIEF PD | CHOLINESTERASE INHIBITORS TO PREVENT FALLS IN PARKINSON'S DISEASE

Bristol Trials Centre (BTC)

CHIEF-PD

CHolinesterase **I**nhibitor to pr**E**vent **F**alls in **P**arkinson's **D**isease

Statistical and Health Economic Analysis Plan (SAP/HEAP)

Version 1.0 (23/03/2023)

The following people have reviewed the analysis plan and are in agreement with the contents			
Name	**Role**	**Approval signature**	**Date**
Grace Young	Trial Statistician		24/03/23
Chris Metcalfe	Lead Statistician		24/03/23
Ammar Annaw	Trial Health Economist		24/03/23
Will Hollingworth	Lead Health Economist		27/03/23
Emily Henderson	Chief Investigator		04/04/23
Danielle Pendry-Brazier	Trial Manager		24/03/23
Yoav Ben-Shlomo	Co-Investigator		04/04/23

CONTENTS

ABBREVATIONS

Abbreviation	Phrase
AE	Adverse Event
AR	Adverse Reaction
CI	Chief Investigator
ChEi	Cholinesterase inhibitor
CRF	Case Report Form
CTU	Clinical Trials Unit
DMC	Data Monitoring Committee
HES	Hospital Episode Statistics
IMP	Investigational Medicinal Product
ISRCTN	International Standard Randomised Controlled Trials Number
ITT	Intention to Treat
NICE	National Institute for Health and Care Excellence
PD	Parkinson's disease
PPI	Patient and Public Involvement
QALY	Quality Adjusted Life Years
RCT	Randomised Control Trial
REC	Research Ethics Committee
SAE	Serious Adverse Event
SAP	Statistical analysis plan
SAR	Serious Adverse Reaction
SOP	Standard Operating Procedure
SUSAR	Suspected Unexpected Serious Adverse Reaction
TMG	Trial Management Group
TSC	Trial Steering Committee

1. INTRODUCTION AND PURPOSE

This document details the rules proposed and the presentation that will be followed, as closely as possible, when analysing and reporting the main results from **CHIEF-PD**.

The purpose of the plan is to:

1. Ensure that the analysis is appropriate for the aims of the trial, reflects good statistical practice, and that interpretation of *a priori* and *post hoc* analyses respectively is appropriate.
2. Explain in detail how the data will be handled and analysed to enable others to perform the actual analysis in the event of sickness or other absence.

Additional exploratory or auxiliary analyses of data not specified in the protocol are permitted but fall outside the scope of this analysis plan (although such analyses would be expected to follow Good Statistical Practice).

The analysis strategy will be made available if required by journal editors or referees when the main papers are submitted for publication. Additional analyses suggested by reviewers or editors will, if considered appropriate, be performed in accordance with the Analysis Plan, but if reported the source of such a post-hoc analysis will be declared.

Amendments to the statistical analysis plan will be described and justified in the final report of the trial.

1.1 MOTIVATION

There is a high incidence of falls in patients with PD. A quarter of people with PD fall at least once a month and they are twice as likely to fall on recurrent occasions compared to older people. Falls lead to hospital admissions, hip fracture, fear of further falling, increased dependency and nursing home placement. The propensity to fall results from underlying loss of cholinergic function in cognitive (frontocortical) and gait (mesencephalic locomotor area) critical brain areas. Animal studies have established that the dual loss of dopaminergic and cholinergic networks precipitates gait instability, freezing of gait and falls in PD (1). Amelioration of this underlying cholinergic deficient with cholinesterase inhibitors (ChEis) represents a promising strategy, targeting the aetiology of falls in PD.

Efficacy has been suggested in 3 small, phase 2 trials (2–4). A phase 3 trial is required with a larger sample to definitively estimate the clinical and cost effectiveness of ChEis in this population. This will determine whether the reduction in falls achieved in small single centre studies can be matched in a larger multi–centre trial delivered in standard NHS clinics. Unlike phase 2 trials, the treatment will be delivered as in clinical practice (transdermally, over a longer time, at higher doses).

2. SYNOPSIS OF STUDY DESIGN AND PROCEDURES

This is a summary of the study design as described in the study Protocol (version 6.0, 15th October 2020) with the single purpose of insuring an informed statistical analysis. For all other purposes reference MUST be made to the current version of the protocol.

The CHIEF-PD trial aims to determine whether the cholinesterase inhibitor (ChEi) treatment prevents people with Parkinson's disease (PD) from falling and is cost-effective.

2.1. TRIAL OBJECTIVES AND AIMS

2.1.1. PRIMARY CLINICAL OBJECTIVES

To determine the difference in fall rate over 12 months between people with PD treated for 12 months with a ChEi and those treated with a placebo.

2.1.2. SECONDARY CLINICAL OBJECTIVES

To determine the effect of 12 months of treatment with ChEi versus placebo on:

i. **PD symptom severity**, assessed using the MDS-UPDRS total score and individual subscales

ii. **Freezing of gait**, assessed using the NFOG questionnaire

iii. **Frailty**, assessed using the Survey of Health, Agesing and Retirement in Europe (SHARE) frailty instrument

iv. **Physical performance,** assessed using the short physical performance battery (SPPB)

v. **Presence of orthostatic hypotension**, determined from blood pressure readings

vi. **Dysphagia**, measured by the Swallowing Disturbance Questionnaire (SDQ)

vii. **Wellbeing/capability**, assessed using the ICEpop CAPability measure for Older people (ICECAP-O)

viii. **Cognitive ability**, assessed using the Montreal Cognitive Assessment (MoCA)

ix. **Depressive symptoms**, assessed using the Geriatric Depression Scale (GDS)

x. **Apathy**, using the Starkstein Apathy Scale (SAS)

xi. **Fear of falling**, using the Iconographical Fall Efficacy Scale (ICON-FES)

xii. **Mortality** (all cause and PD-related), based on ONS data

xiii. **Care-related QoL**, measured using the carer experience scale (CES)

2.1.3. HEALTH ECONOMIC OBJECTIVES

i. **Cost effectiveness** of rivastigmine patches

ii. **Health-related quality of life** using the EuroQoL 5D-5L

iii. **Health care use**, collected using participant questionnaires and routine data

2.2. TRIAL DESIGN AND CONFIGURATION

A multicentre, phase III, double blind, randomised control trial (RCT) of the cholinesterase inhibitor, **rivastigmine** versus placebo to prevent falls in Parkinson's disease. CHIEF-PD will provide definitive evidence that can be immediately translated into clinical practice, if appropriate. Powered to detect a treatment effect that is meaningful to clinicians and patients it will also establish the cost-effectiveness of the treatment in preventing falls which are potentially devastating for patients and expensive for the NHS. Positive findings will provide robust evidence to change clinical practice.

Two-arm blinded CTIMP trial:

Intervention: Rivastigmine is a reversible non-competitive inhibitor of acetylcholinesterase. Participants will be asked to wear patches on their arm for 12 months. There are 3 strengths of patch; LOW 4.6mg, MED 9.5mg, HIGH 13.3mg which will be titrated up if clinically tolerated.

Placebo: Placebo patches will match the intervention patches. It is not anticipated that application of placebo represents any risk above that of standard care.

2.3 TRIAL MILESTONES

Stop/go criteria were planned for the CHIEF-PD study. Participant and site recruitment were meant to be reviewed at 9 months and assessed on an on-going basis. However, as Covid-19 halted the trial for almost a year, this was amended to a target number of participants: 115 by the end of October 2021. This target was met and the trial was given a new recruitment end date (August 2022). At the time of writing (Feb 2023) an extension had been granted, for 8 months, to allow the trial to recruit the full 600 participants by the end of April 2023, with follow up concluding in April 2024.

2.4. ELIGIBILITY CRITERIA

Inclusion criteria

a. Diagnosis of idiopathic Parkinson's disease.
b. Modified Hoehn and Yahr stage 1-4 disease as determined at baseline.
c. Have experienced a fall in the previous year.
d. Able to walk ≥10m without aids or assistance.
e. 18 years of age or above.

Exclusion criteria

a. Previous ChEi use in 12 months prior to enrolment.
b. Hypersensitivity to rivastigmine
c. Dementia diagnosed according to MDS criteria (6).
d. Inability to attend or comply with treatment or follow-up scheduling.
e. Non-English-speaking patients (cognitive tests performed in English).
f. Falling ≥4x per day.

g. Unwillingness to use an acceptable method of contraception for the duration of the trial if they are of childbearing potential.
h. Pregnancy and/or breastfeeding

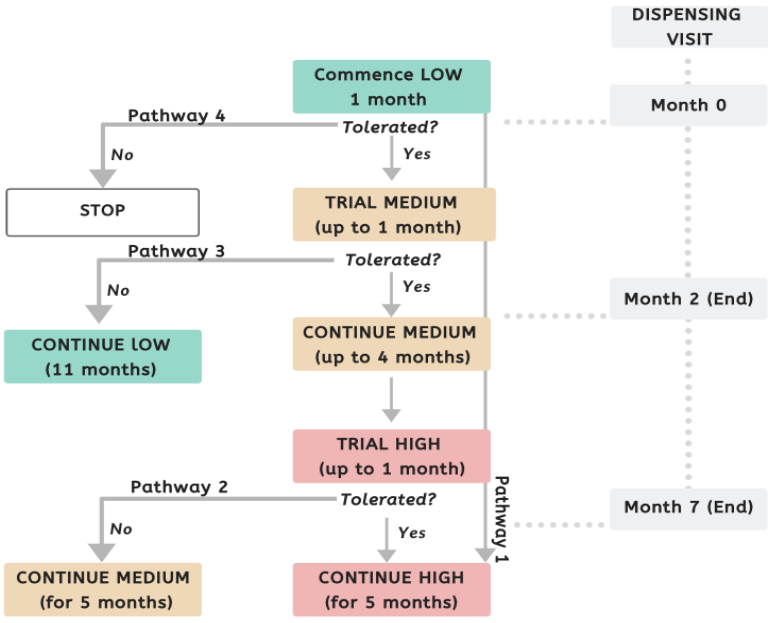
2.5. DESCRIPTION OF INTERVENTION

Trial participants will be allocated to one of two treatment groups (active Rivastigmine patch or matched placebo patch). The active and placebo patches will be supplied by Luye Pharma AG and subsequently packaged/labelled by RenaClinical.

Rivastigmine is a reversible non-competitive inhibitor of acetylcholinesterase. It is currently licensed for use in mild to moderate dementia in Parkinson's disease and in mild to moderate dementia in Alzheimer's disease. Up-titration to (now licensed) 13.3mg is routine in clinical practice. It is not anticipated that application of placebo represents any risk above that of standard care.

There are 3 strengths of patch. All participants will commence LOW dose (4.6mg / 24 hours) patches with the aim to escalate up through MEDIUM (9.5mg / 24 hours) dose to HIGH dose (13.3mg / 24 hours), according to tolerability (Figure 1).

Figure 1. Titration of IMP, according to tolerability



Medication packs will be dispensed at month 0 and end of months 2 and 7. Each pack allows for a one-month trial of up titration to the next dose prior to the next pack being dispensed. Additional dispensing may occur when patches reach their expiry date and replacements are required.

Month 0 - 1 LOW 4.6mg/24hrs active/ placebo IMP

Month 2 - 7 MEDIUM 9.5mg/24hrs active/ placebo IMP

Month 8 - 12 HIGH 13.3mg/24hrs active/ placebo IMP

Titration for side effects

In the event of a participant experiencing side effects the central and local research team will refer to the relevant trial specific instructions. This will provide advice whether down titration or ceasing the treatment is necessary. This decision should be made in liaison with the normal treating clinician.

The same dosing schedule will be applied to participants in both groups with dummy dose titration being applied to the placebo group. Compliance will be monitored through the monthly telephone calls as well as stickers in the falls diary. Participants will also be asked to return any unused patches to the research team.

2.6. RANDOMISATION PROCEDURES

The randomisation sequence will be generated by Sealed Envelope[i] using their online randomisation system. Participants will only be randomised after eligibility and consent have been confirmed. Randomisation will be stratified by site and by the following three minimisation criteria (number of falls in previous year (1-4 falls low, 5+ high); age (18-64 low, 65 + high) and cognitive impairment measured using MoCA (1-25 low, 26-30 high)). Participants will be allocated to each treatment group using minimisation method with a probability of 0.8.

The PI (or authorised delegate) will log onto the randomisation site, enter the minimisation variables and then receive the code that allocates the participant to a treatment pack. They and the participant will remain blinded as to which treatment group this code refers to. The unblinded randomisation code will be held by the trial pharmacy, central pharmacy and BTC trial statistician. The remainder of the trial team, including the PI and trial manager, will stay blinded to treatment allocation. Trial participants will be allocated to one of two treatment groups (active rivastigmine patch or matched placebo patch).

2.7. BLINDING AND BREAKING OF BLIND

The central research team, investigator site staff and participants will be blinded to the allocation of treatment group. The central trials team will remain blinded, including the CI and trial manager. The BTC trial statistician will be unblinded in-order to report to the Data Monitoring Committee (DMC).

Treatment codes will only be released to the investigative team once written confirmation has been received that the trial database has been locked. Participants will be informed of their allocation by the central trial team after the trial has concluded. In the event of a medical emergency the participant's treating physician will contact the central pharmacy at the University Hospitals Bristol NHS Foundation Trust who will hold the unblinding codes. Participants that have been unblinded, as part of this process, will be removed in a sensitivity analysis.

2.8. SAMPLE SIZE CALCULATION

The mean falls per month on the log-transformed scale in the placebo group from the feasibility data was 0.3 (SD 1.2) and a 25% reduction in falls on the log-transformed scale is -0.2877, so the mean in

the Rivastigmine group is 0.0123 on the log-transformed scale. The correlation is 0.589 (lower bound of the 95% confidence interval for correlation found in feasibility trial) between baseline and follow-up measurements of log-transformed fall rate. Using these values in the ANCOVA sample size calculation, we will have 480 (240 per group) participants to detect a 25% difference in mean fall rate with 90% power. We will recruit 600 people assuming a 20% drop out (n=120).

The sample size calculation is based on an Analysis of Covariance (ANCOVA) approach where any variance between individuals in post-treatment falls rate, which is correlated to the corresponding measure taken at baseline, is removed from the error variance resulting in increased statistical power. With one baseline assessment and one post-treatment assessment of outcome, the standard sample size target is reduced by a factor of (1 - r2), where r2 is the squared Pearson's correlation coefficient. This is a standard approach to sample size calculation for quantitative outcome measures, as detailed in Machin et al 1997[ii].

2.9. TRIAL COMMITTEES

Trial Management Group (TMG)

The TMG will have responsibility for the day-to-day management of the trial and will report to the TSC. The TMG will meet on a regular basis with a core working group of staff having frequent progress meetings.

Trial Steering Committee (TSC)

The TSC will make recommendations/key decisions during the trial to the TMG and minutes will be sent to the funder.

Data Monitoring Committee (DMC)

The Data Monitoring Committee will meet once prior to the first participant, first visit and then convene prior to the TSC meeting to review the adverse event data and any other ethical aspects that arise and report to the TSC. The DMC will comprise Dr T Quinn (University of Glasgow) as Chairperson, Dr A McConnachie and Dr J Burns as independent members. In addition, Dr E Henderson (CI), Dr Metcalfe (Lead statistician, open session only) and Ms G Young (BTC trial statistician, attending both open and closed sessions).

2.10. DATA COLLECTION

Participants in the trial will undergo a face-to-face assessment at baseline (0 months) and follow-up (12 months). Monthly phone calls to participants will be coordinated centrally (from Bristol) by research nurse/assistants, blinded to treatment status. These phone calls remind participants to return their fall diaries, corroborate and / or collect missing fall data and screen for adverse events and collect data on health care use and quality of life.

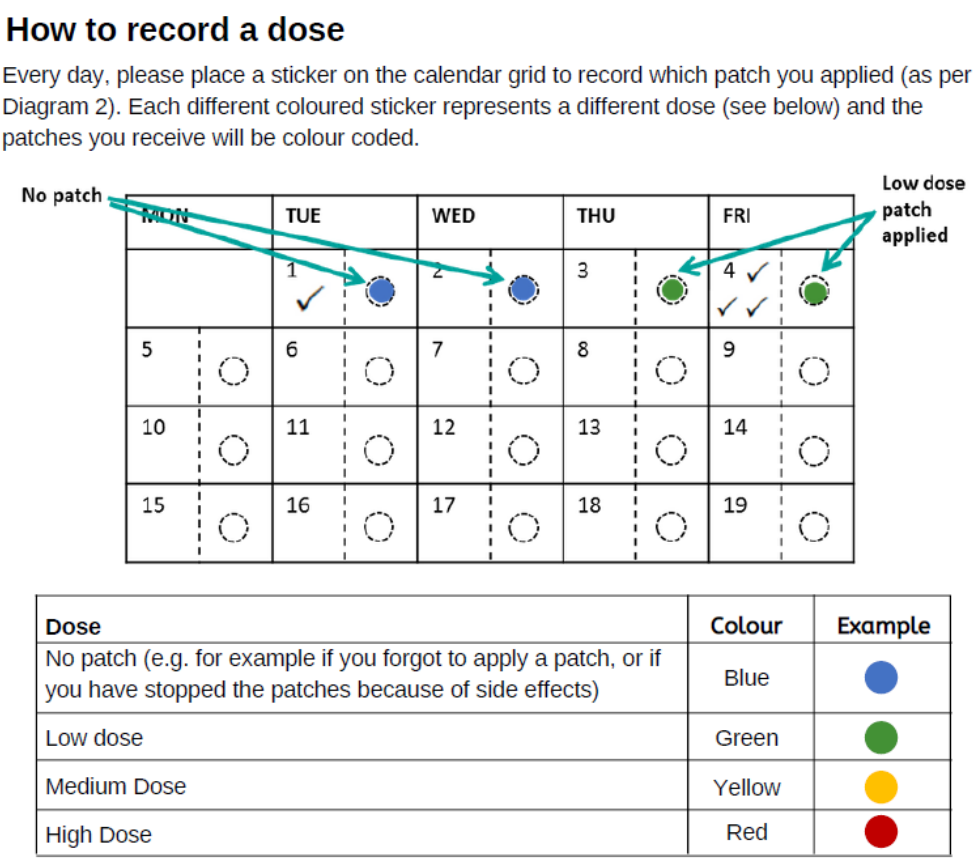
Table 1. Schedule of assessment visits and outcome measurements

Month	0	1	2	3	4	5	6	7	8	9	10	11	12±
Activity													
Procedures													
Eligibility criteria review	●												
Informed consent	●												
Sociodemographics	●												
Medical history	●												
Drug history	●												●
Examination (HR, BP, height, weight, MDS-UPDRS III, frailty, gait, SPPB)	●												●
Falls	●	●	●	●	●	●	●	●	●	●	●	●	●
MoCA	●												●
GDS	●												●
SAS	●												●
MDS-UPDRS I, II, IV	●												●
NFOGQ	●												●
ICON-FES	●												●
SDQ	●												●
ICECAP-O	●												●
EQ5D-5L	●	●		●			●			●			●
CES**	○												○
Medication													
IMP dispensing	●		●					●					
IMP return			●					●					●
Safety													
Adverse events		●	●	●	●	●	●	●	●	●	●	●	●
ECG*	○		○				○						
Formal & informal care use		●		●			●			●			●
Hospital care and mortality	**(HES via NHS England and ONS data linkage)**												

*IMP: Investigational Medicinal Product, SPPB: Short Performance Physical Battery MDS-UPDRS: Movement Disorder Society-Unified Parkinson's Disease Rating Scale NFOGQ: New Freezing of Gait Questionnaire ICON-FES: Iconographical Falls Efficacy Scale GDS: Geriatric Depression Scale SAS: Starkstein Apathy Scale MoCA: Montreal Cognitive Assessment CES: Carer Experience Scale. SDQ: Swallowing Disturbance Questionnaire *ECG as per arrhythmia safety protocol **Completed by carer ±Completed within* +/-10 days of the annual follow up date

The falls diary will be the primary source of data for the primary outcome and consists of monthly calendar views, which allow the participant to record the number of falls each day, along with a sticker, to represent the patch dose used that day (Figure 2). Participants are asked to include the number of falls per day as 'ticks' and also include a blue, green, yellow or red sticker to indicate 'no patch', 'low dose patch', 'medium dose patch' or 'high dose patch', respectively. The falls diary will be utilised in the first instance but the number of falls recorded in the calls log may be used if the diary entry has been left blank, to avoid missing data. Details of this imputation, along with justification will be included in the final manuscript. Participants will be called on a monthly basis to confirm what dosage they are receiving as well as how many times they have fallen in the last month. This will be checked against their falls diary.

Figure 2. Falls diary calendar view



2.11. OUTCOME MEASURES

All outcome data will be obtained from participant diaries and research nurse/assistant phone calls. In all analyses we will present regression coefficients/rate ratios, with 95% confidence intervals and p values. If the data does not conform to the assumptions of parametric tests, suitable transformations or non-parametric methods will be utilised and justification given.

2.11.1. CLINICAL PRIMARY OUTCOMES

Primary Outcome	**Type**	**Outcome detail**
Fall rate	Rate	The primary outcome is fall rate measured using monthly diaries and telephone calls prospectively over 12 months from the day the IMP is commenced. The falls in the diary will be used, in the first instance, and supplemented by the telephone calls if necessary (e.g. missing day/month in diary). All available diaries will be used regardless of whether the patient is compliant with medication and has or has not completed the full 12 months of treatment as long as there is a baseline and one follow-up diary on treatment.

2.11.2. CLINICAL SECONDARY OUTCOMES

Parkinson's disease symptom severity	Total score (0-260) and sub scores	MDS-UPDRS[iii] questionnaire total score range 0-260 with higher values indicating greater disease severity and each individual subscale: 1. Part 1 - Non-Motor Aspects of Experiences of Daily Livings: Consists of 6 items completed by the examiner (*found in CRF*) and 7 items completed by the participant/carer (*found in questionnaire*). All items are on a scale of 0 (Normal) to 4 (Severe), with a combined score range of 0-52. 2. Part 2 - Motor Aspects of Experiences of Daily Living: Consists of 13 items completed by the participant/carer (*found in questionnaire*). All items are on a scale of 0 (Normal) to 4 (Severe), with a combined score range of 0-52. 3. Part 3 - Motor Examination: Consists of 33 scores across 18 items completed by the examiner (*found in CRF*). All items are on a scale of 0 (Normal) to 4 (Severe), with an additional option of "UR: Unable to rate", which will be treated as missing. Some questions have separate sections (e.g. left and right hands) so the total combined score ranges from 0-132. 4. Part 4 - Motor Complications: Consists of 6 items completed by the examiner (*found in CRF*). All items are on a scale of 0 (Normal) to 4 (Severe) with a combined score of 0-24.
	Ordinal	Hoehn & Yahr stage
Freezing of Gait*	Score (0-28)	NFOG[iv] questionnaire. Consists of two parts – a Y/N question for whether the participant experiences "freezing episodes" then, for those that answer Y there are 9 further items on freezing of gait; 8 of these questions are on a scale of 0-3 and 1 is on a scale of 0-4. Therefore, the combined score has a range of 0-28.
Frailty	Score	Survey of Health, Ageing and Retirement in Europe (SHARE[v]) frailty instrument with 4 questions with an overall range of 1-7 with a maximum grip strength (continuous).
Physical performance		Short physical performance battery (SPPB[vi]) and gait speed
Presence of orthostatic hypotension	Binary	The presence of orthostatic hypotension[vii] will be determined from blood pressure readings taken upon immediate, 1 min and 3 min of standing following 10 min of supine rest
Dysphagia	Score	Swallowing Disturbance Questionnaire (SDQ[viii]), consisting of 14 questions on scale 0 (never) – 3 (very frequently) and 1 Y/N question (2.5 for Y, 0.5 for N). The overall score is on a scale of 0.5-44.5.
Wellbeing/capability	Score (0-1)	ICEpop CAPability measure for Older people (ICECAP-O[ix])
Health related quality of life	Score	Euroqol five dimension five level (EQ-5D-5L). The summary score is anchored at 0 (as bad as death) and 1 (best health).

Cognitive ability	Score (0-31)	Montreal Cognitive Assessment (MoCA[x]) with a total score of 0-30, with higher scores indicating less impairment. One additional question asks whether the participants has had less than 12 years of education since kindergarten (pre-school). Where this is the case, an additional 1 point is added to the MoCA score.
Depression	Score (0-15)	Geriatric Depression Scale (GDS[xi]): 15 items with yes or no answers (range 0-15). A score above 5 indicates depression and higher scores indicate more severe depression.
Apathy (lack of enthusiasm)	Score (0-42)	Starkstein Apathy Scale (SAS[xii]). Apathy is considered more severe as the total score (0-42) increases.
Fear of falling	Score (10-40)	Iconographical Fall Efficacy Scale (ICON-FES[xiii]) consisting of 10-items on a scale of 1 (not at all concerned about falling) to 4 (very concerned about falling). Total score range 10-40, with higher scores indicating greater fear of falling.
Mortality	Binary	Death from all causes (Y/N) and PD-specific mortality, based on ONS data.

*During the COVID-19 pandemic, some patients were assessed at home over a shorter distance. An indicator variable will be included in this analysis in an attempt to account for any variation caused by this.

Note: The carer experience and quality of life will explored in a subsequent secondary analysis.

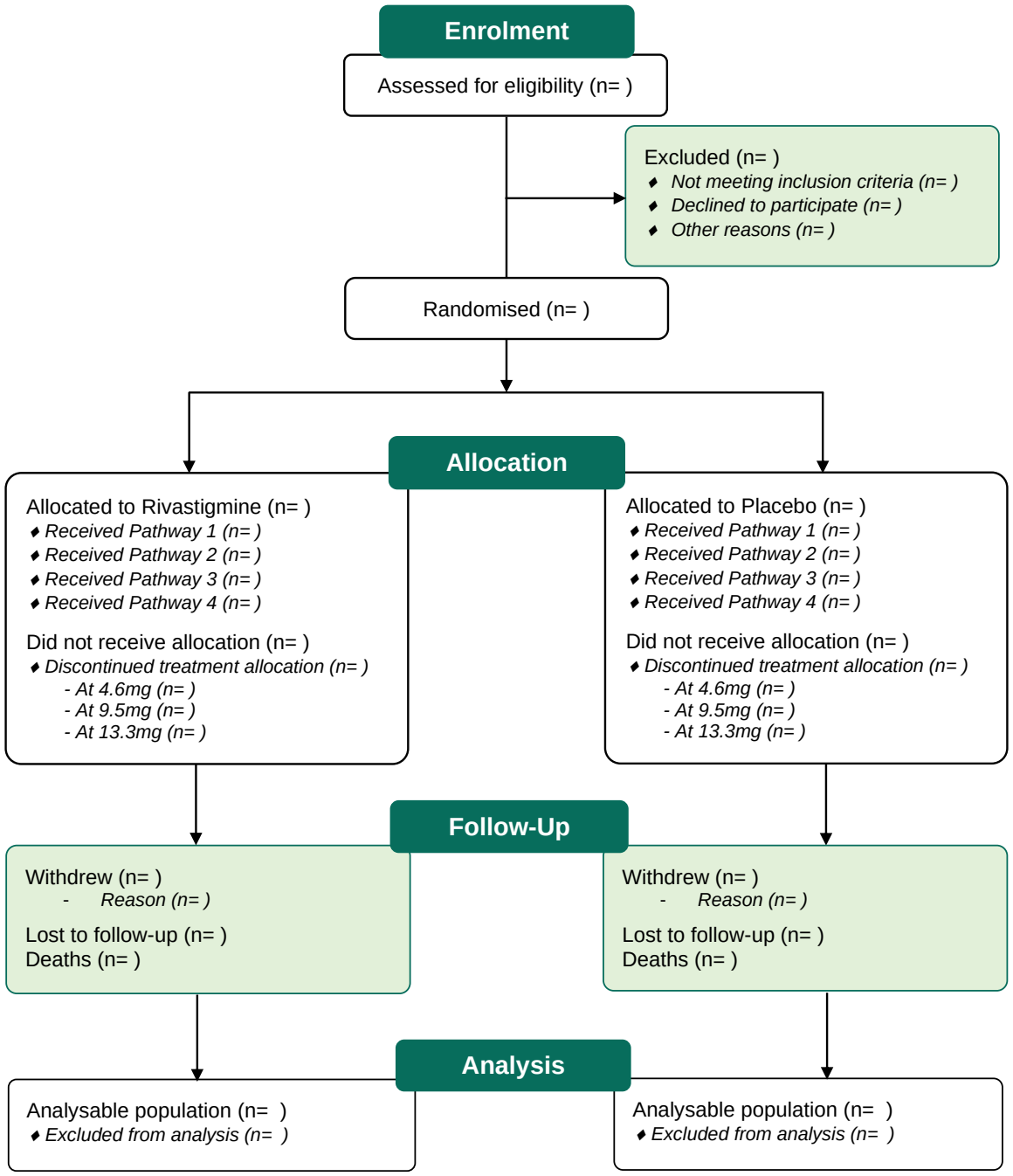
2.11.3. RESOURCE USE

Resource use	Item	Measurement	Unit cost source	Unit cost calculations/ assumptions
Secondary healthcare	Number of A&E visits	NHS England secondary care.	National Cost Collection[xiv]	Accident and Emergency outpatient visit
	Number of hospital outpatient visits	NHS England secondary care.	National Cost Collection[xiv]	Weighted average of outpatient visits
	Number of hospital inpatient admissions; Number of nights in hospital	NHS England secondary care.	National Cost Collection[xiv]	Estimated cost per night in hospital.
Primary healthcare	Number of: GP consultations; Practice Nurse/ Nurse Practitioner contacts	Brief question (based on the Client Service Receipt Inventory (CSRI))	Personal Social Services Research Unit (PSSRU)[xv]	Different assumptions: for example the consultation time, type (e.g. face to face or virtual) and different unit costs e.g. GP vs nurse, home visit vs at surgery
Community/ primary care	Number of visits to walk in centre	Brief question (based on the Client Service Receipt Inventory (CSRI))	National Cost Collection[xiv]	Weighted average
	Number of telephone calls to 111	Brief question (based on the Client Service Receipt Inventory (CSRI))	Pope et al. 2017[xvi]	£12.26 per call (2017) to be adjusted for price year.
Primary care Prescribed	Prescribed medication	Brief question (based on the Client Service Receipt Inventory (CSRI))	British National Formulary (BNF)[xvii]	Cost per quantity.
Informal care costs	Number of hours	CRF_ Baseline & 12 months	The proxy good method[xviii]	Valuation at a shadow price of a market substitute (e.g., a nurse for nursing services and a housekeeper for housekeeping services).

3. DESCRIPTION OF PARTICIPANT CHARACTERISTICS

3.1. DISPOSITION

The flow of participants through the trial will be summarised in a CONSORT diagram (subject to change) that will include eligibility, numbers randomised, withdrawals, deaths, losses to follow up and the numbers analysed.



Note: Pathway 1 is for patients that have tolerated each dose as it's been provided: 1 month LOW, 5 months MED and 6 months HIGH. Pathway 2 is for patients that have not tolerated the HIGH dose and have continued on MED. Pathway 3 is for patients who have not tolerated the MED dose and have continued on LOW. Pathway 4 is for patients who have not tolerated any dose and have discontinued treatment.

4. GENERAL ANALYSIS CONSIDERATIONS

4.1. ANALYSIS POPULATIONS

Full analysis set: All randomised participants: analyses will be based on the intention to treat principle, analysing participants in the groups to which they were randomised regardless of the total duration of follow-up as long as there is a baseline and at least one follow-up fall diary.

Per protocol analysis set 1: All participants in the full analysis set who have applied patches for at least 20 days of the first month.

Per protocol analysis set 2: All participants in the full analysis set who have taken the IMP for 6 months. This can include those that continued on LOW dose for those 6 months.

The per protocol sets will be used to calculate compliance, in the intervention arm, for a compliance average causal effect (CACE) sensitivity analysis, which aims to calculate an unbiased estimate of the treatment effect in compliers in the intervention arm and "potential" compliers in the control arm.

4.2. ADVERSE EVENTS

Pharmacovigilance will be carried out in accordance with the requirements set out by the European Commission Detailed Guidance CT-3 2011 including the terminology of adverse events and reactions and the assessment of seriousness, causality and expectedness of an event. Each event will be categorised in terms of severity (mild, moderate or severe) and relatedness to the IMP (not related, unlikely to be related, possibly related, probably related or definitely related). These categories will be used to categories events into adverse events (AEs), adverse reactions (ARs), serious adverse events (SAEs), serious adverse reactions (SARs) and suspected unexpected serious adverse reactions (SUSARs). Non-serious adverse events (with the exception of falls) that are unrelated to the IMP will not be reported. Data will be presented in terms of the number of events overall as well as the number of events per participant.

5. STATISTICAL ANALYSES

5.1. STATISTICAL ANALYSIS

The latest version of STATA will be used for all statistical analyses; the version used will stated in study reports. A pre-specified random seed of 350938 has been chosen to ensure all analyses can be replicated.

5.2. SUMMARY OF THE PRIMARY AND SECONDARY ANALYSIS

All primary and secondary analyses will be conducted on complete data, using the ITT principle and appropriate regression model. All analyses are testing for superiority therefore emphasis will be on the estimate, 95% confidence interval and p value. Assumptions for each regression will be checked to make sure the correct method of analysis is being used.

The primary analysis will explore the fall rates by arm using all available daily follow-up, regardless of whether or not the participants have completed 12 months of treatment. This will be analysed using a mixed effects Poisson regression model, incorporating individual variation in rate of falls as a random effect term. The analysis will be adjusted for the variables used in randomisation; centre, age, baseline number of falls and MoCA score. An indicator variable for month will also be included, from 1-12, indicating the month of follow up that day of data falls into. There may be delays in starting treatment, e.g. delays collecting prescriptions. Therefore, the origin (day 1) will be the first day that participants start taking the IMP, rather than randomisation, to allow for 360 days of follow up (30 days on LOW dose, 30*5 days on MEDIUM dose and 30*6 days on HIGH dose).

```
mepoisson n_falls arm i.centre age b_falls moca month || pid:, irr
```

The primary analysis will be adapted, by the choice of a suitable regression model, for the analysis of secondary outcomes.

Timing of the origin and outcome assessment

Due to logistical issues, it can be several weeks between random allocation and some participants receiving the first batch of their allocated patches. The participants are blind to their allocation, and so we are assuming that allocation itself will not affect the participant's outcome until they receive their patches. Hence we are defining the origin of data collection (participant completed diaries) as the day the participant applies their first daily intervention patch, with follow-up then continuing for 360 days.

All randomised participants who are lost to follow-up will be recorded in the CONSORT flowchart, and this will include any participants who do not start their allocated intervention. Participants who do not want to start their allocated intervention but who agree to completing the diaries will start follow-up on receipt of the diary. We will report the mean and standard deviation of the number of days between randomisation and treatment initiation for the intervention and comparison groups. We will conduct a sensitivity analysis, repeating the primary analysis including follow-up until 360 days post-randomisation.

5.3. SENSITIVITY ANALYSES

1. A complier average causal effect (CACE) analysis will be carried out which utilises the “per protocol” analysis sets set out in section 4.
2. Imputation model: Due to difficulties in pre-specifying a sensitivity analysis looking at the impact of missing data, we will conduct an exploratory investigation into the potential impact on our conclusions of primary outcome data being missing due to participants becoming too ill to fill out diaries ("missing not at random").
3. Imbalance at baseline: Any baseline characteristics that differ between the arms by 10% or 0.5 standard deviations will be adjusted for in a sensitivity analysis.
4. To investigate variations in the treatment effect over the 12 months of follow-up, time by treatment interaction terms will be added to the primary analysis model. We will additionally consider simplifications to the model for treatment effect variations over time to secure an analysis with greater statistical power (e.g. treating time as continuous). We will also explore the difference in falls rate, across the dose levels, descriptively (see Supplementary Table 4).
5. The primary analysis will be repeated, removing any participants that were unblinded during the trial, for any reason.
6. Randomisation time zero: The primary analysis will be repeated using the day of randomisation as time zero and using diary days that are within 360 days of this date.

5.4. SUBGROUP ANALYSES

These subgroup analyses will use baseline data to assess whether the intervention is more/less effective in certain sub groups of participants. Rather than assessing each group individually, an interaction term will be added to the model, followed by a likelihood ratio test.

Subgroup	Cut off*	Rationale/hypothesis
Age of participant	Median age	The intervention effect may be more/less effective for those who are older.
Falls over the past year	1-4, 5+	The intervention effect may be more/less effective for those who experience more falls.
MoCA score	0-24, 25-30	The intervention effect may be more/less effective for those with worse cognitive function~.
Freezing of gait	Yes, No	The intervention effect may be more/less effective for those with experience freezing.
Presence of orthostatic hypotension	Yes, No	The intervention effect may be more/less effective for those who had orthostatic hypotension

*A cut off will be used to simplify interpretation of the sub group analysis descriptives. However, continuous versions of the sub group variable will be used in the interaction analysis, where possible. ~If this is the case, the degree of executive dysfunction will be explored

6. HEALTH ECONOMIC ANALYSIS

6.1. OVERVIEW OF ECONOMIC ANALYSIS

The health economics analysis will be conducted on an intention-to-treat approach comparing the intervention (rivastigmine patches) and control (placebo) groups as randomised. A cost-utility analysis (CUA) will be the primary economic analysis comparing the difference in costs to the NHS and the difference in quality-adjusted life years (QALYs).

6.2. PERSPECTIVES

The primary economic analysis will be from the NHS and social services perspective which will involve differences in resource use for secondary care, primary and community care.

In secondary analyses we will estimate the cost per fall prevented and expand the perspective of the analysis to include informal care costs, carer quality of life and participant wellbeing measures.

6.3. IDENTIFICATION OF RESOURCES

NHS resource use identified as relevant to the analysis are: (1) secondary care; (2) primary and community care. Informal care costs in terms of loss in productivity will be considered in the secondary economics analysis.

6.4. MEASUREMENT OF RESOURCE USE

6.4.1. HEALTHCARE USE: ROUTINE DATA: SECONDARY CARE

Patient-level data will be purchased from NHS England using the Data Access Request Service (DARS). NHS England admitted patient (day case & inpatient), outpatient and Emergency Care Data Set (ECDS) hospital episode statistics (HES) datasets will be requested from NHS England covering the estimated 3 years from first participant randomised to 12 months after the last participant is randomised. To link NHS England's secondary care dataset with CHIEF-PD trial data, a bespoke data linkage will be done by NHS England. NHS England will be provided with identifiable data for the patient's involved in the CHIEF-PD trial. These identifiers include NHS number and date of birth. Data linkage will only be done for participants who consent for their routine data to be used.

6.4.2. HEALTHCARE USE: ROUTINE DATA: PRIMARY CARE AND INFORMAL CARE

A brief question to assess primary and community care use, medications and informal care is collected via telephone interviews at 1, 3, 6, and 9 months and at the 12-month research clinic visit.

6.5. VALUATION OF RESOURCE USE DATA

Primary and community care will be valued in monetary terms using the most recent unit costs by the Personal Social Services Research Unit (PSSRU)[xv]. Secondary care and use of hospitals will be valued using the latest NHS National Cost Collection[xiv]. Medication costs will be estimated from the

British National Formulary. All unit costs will be taken from the most recent sources or inflated to the most recent available cost year.

6.6. IDENTIFICATION OF OUTCOMES

Quality Adjusted Life Years (QALYs) will be the primary outcome measure for the economic analysis. QALYs will be derived from utility scores, obtained using the EuroQoL health related quality of life questionnaire, EQ-5D-5L. Secondary outcomes relevant for the expanded perspective in the secondary economic analysis will be falls prevented, carer quality of life (CES) and participant wellbeing (ICECAP-O) measures.

6.7. MEASUREMENT OF OUTCOMES

The EQ-5D-5L will be collected via research clinic visits at baseline and at the 12-month and via telephone interviews at 1, 3, 6, and 9 months.

6.8. VALUATIONS OF OUTCOMES

EQ5D-5L responses will be converted into utility scores using the English value set recommended by NICE at the time of analysis[xix]. Utility scores will be combined with ONS mortality data to estimate quality adjusted life years. The area-under-the-curve method will be used to convert utility scores into QALYs for the 12-month time horizon. Regression methods will be applied to generate appropriate estimates of differential mean QALYs and to control for any imbalance in baseline utility between arms[xx].

6.9. ECONOMIC DATA ANALYSIS

6.9.1. DISCOUNT RATES FOR COSTS AND BENEFITS

As costs and benefits will not be evaluated further than 12 months post randomisation for the primary analysis, discounting will not be needed. If an extrapolation model is developed, the NICE recommended discount rate will be used for long-term costs and outcomes.

6.9.2. COST-EFFECTIVENESS THRESHOLD(S)

Adjusted mean costs and QALYs associated with each trial group will be combined through the Net Benefit (NB) framework. NB will be calculated at threshold values (i.e. willingness to pay per QALY) recommended by NICE at the time of analysis (currently £20,000-£30,000 per QALY). A conservative willingness-to-pay threshold of £20,000 per QALY will be used in the primary analysis.

6.9.3. ANALYSIS OF RESOURCE USE AND COSTS

Individual level costs will be estimated and summed up. Mean cost (SD), for each trial arm per cost category (e.g. pharmaceutical, secondary care, etc.), and incremental costs will be calculated. We will estimate adjusted mean costs and the adjusted incremental mean costs (with their corresponding 95% confidence intervals) between the two trial groups by fitting an appropriate regression method (e.g. Ordinary Least Squares (OLS) or Generalised Linear Models (GLM))[xxi]. We

will compare model observed and predicted values and model goodness of fit, to decide on the selection of the regression model.

6.9.4. ANALYSIS OF OUTCOMES

The primary economic outcome in the economic evaluation is Quality-Adjusted Life Years (QALYs). QALYs for each patient over the 12-month period will be estimated from the utility values using the area under the curve approach. Appropriate regression techniques will be used to estimate mean QALYs (adjusted for baseline utility scores) and the adjusted incremental mean QALYs (and their associated 95% confidence intervals) between the trial arms.

6.9.5. DATA CLEANING FOR ANALYSIS

Face validity tests will be conducted on data (e.g. to identify misspelt text, implausible values) and checked against the source documents. Corrections identified will be documented in the Stata code. Data variables not required for the economic analysis and duplicate data entries will be dropped from the dataset.

6.9.6 MISSING DATA

The economic analysis will take an intention-to-treat approach with imputation of missing data. Missing data will be explored to assess the likely mechanism of missingness and to inform the appropriate method of handling missing data (e.g. multiple imputation MI). Accepted guidelines and best practice will be followed to decide on the imputation procedures[xxii].

6.9.7. ANALYSIS OF COST-EFFECTIVENESS

Based on the NICE willingness-to-pay thresholds (i.e. £20,000 and £30,000) for a QALY we will use net benefit regression to estimate the incremental net benefit (and 95% confidence intervals) at 12 months (post-randomisation).

6.9.8. SAMPLING UNCERTAINTY

Uncertainty will be explored using cost effectiveness acceptability curves to estimate the probability that rivastigmine is cost-effective at a range of plausible cost-effectiveness thresholds.

6.9.9. SENSITIVITY ANALYSES

We will explore the methodological and parameter uncertainties in this economic evaluation via sensitivity analyses. This will include:

- Making plausible changes to key methodological assumptions, for instance using a different method to handle missing data.
- Identifying plausible alternative parameter values for key unit costs.

6.10. MODELLING

If the intervention has a sustained effect in reducing falls and is potentially, but not definitively, cost-effective at 12 months, we will develop a simple extrapolation model to assess the potential cost-

effectiveness over a lifetime horizon. A long-term Markov decision model with annual cycle length will be used to evaluate effects of intervention on costs, health gains and cost-effectiveness over the patient's lifetime.

6.11. REPORTING/PUBLISHING

6.11.1. REPORTING STANDARDS

The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) guidelines will be followed when reporting the health economic evaluation, in a format appropriate to stakeholders and policy makers.

6.11.2. REPORTING DEVIATIONS FROM THE HEAP

Any deviation in the final analysis from the published HEAP will be documented and justified in the final published report (e.g. HTA monograph).

7. STATISTICAL TABLES AND FIGURES (SUBJECT TO CHANGE)

Table 1a. Baseline participant demographics

		Rivastigmine		Placebo
	n*	***Mean (SD)* or N (%)**	**n***	***Mean (SD)* or N (%)**
Total number of participants				
Demographics				
Median age (IQR)				
Height (cm)				
Weight (kg)				
Index of Multiple Deprivation				
Gender:				
Male				
Female				
Ethnicity:				
White				
Mixed				
Asian				
Black				
Other ethnic group				
Accommodation:				
Lives independently				
Sheltered accommodation				
Residential/Nursing home				
Other				
Number of adults living in household				
Number of children living in household				
Education:				
No formal qualification				
GCSEs, O Levels or equivalent				
A levels or equivalent				
Degree/Higher Degree				

Table 1b. Baseline participant clinical history

	Rivastigmine		Placebo	
	n*	*Mean (SD)* or N (%)	n*	*Mean (SD)* or N (%)
Baseline clinical measures[a]				
Care needs:				
No care or help required				
1-3 care needs				
4+ care needs				
Number of hours per week				
Comorbidities				
0 comorbidities				
1-2 comorbidities				
3+ comorbidities				
Assessments				
Cognitive and mood measures				
Cognition – MoCA score				
PD severity - MDS UPDRS				
Part 1: Non-motor				
Part 2: Motor				
Part 3: Motor examination				
Part 4: Motor complications				
Total score				
PD severity – Hoehn & Yahr				
1. Unilateral involvement only				
2. Bilateral involvement without impairment of balance				
3. Mild to moderate involvement				
4. Severely disabling disease				
Frailty				
SHARE-FI score				
Maximum grip (kg)				
Physical performance				
SPPB balance score				
SPPB gait speed score				
SPPB chair stand score				
Gait speed[c]				
Gait speed whilst dual tasking				

[a]Care needs included: personal care (e.g. dressing), physical help (e.g. walking), practical help (e.g. preparing meals), keeping company (e.g. talking), taking out (e.g. for a drive) or taking medicines, [b]XX% of participants had their pulse rate measures in person (home visit or at the clinic), [c]During the COVID-19 pandemic some participants were assessed at home. Therefore gait distance was performed over a shorter distance than the 10m set out in the protocol. These may be presented in a separate line in the baseline table, or omitted, depending on the quantity of participants that this affects.

Table 2a. Withdrawals

	Rivastigmine (n=)	Placebo (n=)	P value*
Participant withdrawals, based on most recent dose prior to withdrawal			
Before starting IMP			
During 4.6mg/ml			
During 9.5mg/ml			
During 13.3mg/ml			

*Chi Square, per row

Table 3a. Number of serious adverse events per participant

	Rivastigmine (n=)	Placebo (n=)	P value*
Number of SAEs			
0 events			
1-3 events			
4-6 events			
7-9 events			…
10-12 events			
13-15 events			

*Ordinal logistic regression

Table 3b. All reported serious and non-serious adverse events and their relatedness

	Rivastigmine (n=)	Placebo (n=)
Serious adverse events		
Total # of events		
Relatedness to IMP:		
Unrelated		
Unlikely		
Possibly		
Probably		
Definitely		
Adverse events*		
Total # of events		
Relatedness to IMP:		
Possibly		
Probably		
Definitely		

*Not related and unlikely related AEs will not be reported, as specified in the trial protocol

Table 3c. Serious adverse events by system organ class

	Rivastigmine (n=)		Placebo (n=)	
	Events	**Participants**	**Events**	**Participants**
Serious adverse event classification				
Cardiac disorders				
…				
Endocrine disorders				
…				
Eye disorders				
…				
…				

Table 3d. "Possibly", "probably" or "definitely" related adverse events (non-serious) by system organ class

	Rivastigmine (n=)		Placebo (n=)	
	Events	**Participants**	**Events**	**Participants**
Adverse event classification				
Cardiac disorders				
…				
Endocrine disorders				
…				
Eye disorders				
…				
…				

Table 4. Titration pathway (tolerance) with the CHIEF-PD intervention [to be confirmed]

	Rivastigmine (n=)	Placebo (n=)	P value*
IMP treatment pathway			
IMP not started			
Received 4.6mg			
Stopped 4.6 mg			
Received 9.5mg			
Back down to 4.6 mg			
Stopped 9.5 mg			
Received 13.3mg			
Back down to 9.5 mg			
Stopped 13.3 mg			

*Chi Square, per row

Table 5. Primary outcome for the CHIEF-PD trial

	Rivastigmine (n=) Rate	Placebo (n=) Rate	Unadjusted IRR (95% CI)	Adjusted* IRR (95% CI)	P value*
Primary outcome					
Falls over the 12-months					
Sensitivity analyses					
CACE analysis (using per protocol set 1)					
CACE analysis (using per protocol set 2)					
Imputation model (tbc)					
Adjustment for imbalance					
Time varying coefficient					
Removing those unblinded					
Randomisation time zero					

*Adjusted for age (in years), number of falls in the previous year and MoCA score

Table 6. Subgroup analyses for the CHIEF-PD trial

Variable	Rivastigmine Rate*; n*	Placebo *Rate; n*	Subgroup specific IRR (95% C.I)	Interaction IRR (95% C.I)*	P#
Subgroup analyses					
Age					
<Median					
≥Median					
Falls in past year					
1-4					
5+					
MoCA score					
Low (1-24)					
High (25-30)					
Freezing of gait					
Yes					
No					
Presence of orthostatic hypotension					
Yes					
No					

*The interaction coefficient, #Taken from a likelihood ratio test comparing models with/without the interaction term included, treating the subgroup of interest as a continuous variable where possible.

Table 7. Secondary outcomes for the CHIEF-PD trial

	Rivastigmine (n=)	Placebo (n=)	Unadjusted coeff (95% CI)	Adjusted* coeff (95% CI)	P value*
Secondary outcomes					
PD severity - MDS UPDRS					
Part 1: Non-motor					
Part 2: Motor					
Part 3: Motor examination					
Part 4: Motor complications					
Total score					
PD severity – Hoehn & Yahr					
Asymptomatic					
Unilateral involvement only					
Bilateral involvement without impairment of balance					
Mild to moderate involvement					
Severe disability					
Wheelchair bound or bedridden					
Freezing of Gait					
Experiences freezing episodes					
Freezing of gait (NFOG) score					
Frailty					
SHARE-FI score					
Frail					

Pre-frail

Non-frail

Maximum grip (kg)

Physical performance

SPPB balance score

SPPB gait speed score^

SPPB chair stand score

Gait speed (normal walk)

Gait speed (walk plus task)

Orthostatic hypotension

Present, n(%)

Dysphagia

SDQ score

Wellbeing

ICECAP-O score

Cognitive and mood measures

Cognition – MoCA score

Depression – GDS score

Apathy – SAS score

Fear of falling – ICON-FES score

Mortality

All-cause mortality

PD specific mortality

^As a 10 metre walking speed is captured as part of the MDS-UPDRS we will not be asking frail participants to complete the 4 metre walking speed test. The speed for the 10 metre walk will be scaled down to provide a value for the SPPB gait speed score.

7.1. SUPPLEMENTARY TABLES

Supplementary Table 1. Randomisation variables

	Rivastigmine		Placebo	
	n*	***Mean (SD)* or N (%)**	**n***	***Mean (SD)* or N (%)**
Total number of participants				
Randomisation minimisation variables				
Falls in the previous year:				
Low (1-4 falls)				
High (5+ falls)				
Age in years:				
Low (18-64)				
High (65+)				
Montreal Cognitive Assessment (MoCA):				
Low (1-25)				
High (26-30)				

Supplementary Table 2. Withdrawals, by month

		Rivastigmine (n=)	Placebo (n=)
Withdrawals, per follow up month			
T1	Month 1		
	Month 2		
	Month 3		
T2	Month 4		
	Month 5		
	Month 6		
	Month 7		
	Month 8		
T3	Month 9		
	Month 10		
	Month 11		
	Month 12		

T1, T2 and T3 refer to the up-titration periods

Supplementary Table 3. Descriptives: patches and falls by follow up month

		Rivastigmine (n=)			Placebo (n=)		
		# diary days	# falls	# days on IMP	# diary days	# falls	# days on IMP
Month							
T1	Month 1						
	Month 2						
	Month 3						
T2	Month 4						
	Month 5						
	Month 6						
	Month 7						
	Month 8						
T3	Month 9						
	Month 10						
	Month 11						
	Month 12						

T1, T2 and T3 refer to the up-titration periods

Supplementary Table 4. Descriptives: falls by dose

	Rivastigmine (n=)		Placebo (n=)	
	# diary days	# falls	# diary days	# falls
Dose				
Low				
Medium				
High				

REFERENCES

[i]https://www.sealedenvelope.com/

[ii]Machin D, Campbell MJ, Tan SB, Tan SH. Sample Size Tables for Clinical Studies [Internet]. Oxford, UK: Wiley-Blackwell; 2008. Available from: http://doi.wiley.com/10.1002/9781444300710

[iii]Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 2008; 23: 2129–70.

[iv]Nieuwboer A, Rochester L, Herman T, et al. Reliability of the new freezing of gait questionnaire: agreement between patients with Parkinson's disease and their carers. Gait Posture 2009; 30: 459–63.

[v]Romero-Ortuno R. The SHARE operationalized frailty phenotype: a comparison of two approaches. Eur Geriatr Med. 2013;4(4):255–9 Available from: https://linkinghub.elsevier.com/retrieve/pii/S187876491300048X.

[vi]Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol. 1994;49(2):M85–94 Available from: https://doi.org/10.1093/geronj/49.2.M85

[vii]Brignole M, Moya A, de Lange FJ, Deharo J-C, Elliott PM, Fanciulli A, et al. 2018 ESC guidelines for the diagnosis and management of syncope. Eur Heart J. 2018;39(21):1883–948 Available from: https://academic.oup.com/eurheartj/article/39/21/1883/4939241.

[viii]Manor Y, Giladi N, Cohen A, Fliss DM, Cohen JT. Validation of a swallowing disturbance questionnaire for detecting dysphagia in patients with Parkinson's disease. Mov Disord. 2007;22(13):1917–21 Available from: http://doi.wiley.com/10.1002/mds.21625.

[ix]Flynn TN, Chan P, Coast J, Peters TJ. Assessing quality of life among British older people using the ICEPOP CAPability (ICECAP-O) measure. Appl Health Econ Health Policy. 2011;9(5):317–29 Available from: http://link.springer.com/10.2165/11594150-000000000-00000.

[x]Dalrymple-Alford JC, MacAskill MR, Nakas CT, et al. The MoCA: well-suited screen for cognitive impairment in Parkinson disease. Neurology 2010; 75: 1717–25.

[xi]Meara J, Mitchelmore E, Hobson P. GDS-15 geriatric depression scale as a screening instrument for depressive symptomatology in patients with Parkinson's disease and their carers in the community. Age Ageing 1999; 28: 35–8.

[xii]Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R, Robinson RG. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. J Neuropsychiatry Clin Neurosci. 1992;4(2):134–9. Available from: http://psychiatryonline.org/doi/abs/10.1176/jnp.4.2.134.

[xiii]Delbaere K, Smith ST, Lord SR. Development and initial validation of the Iconographical Falls Efficacy Scale. J Gerontol A Biol Sci Med Sci 2011; 66: 674–80

[xiv]National Cost Collection for the NHS. 2018/19 National Cost Collection data: NHS Improvement https://improvement.nhs.uk/resources/national-cost-collection/

[xv]Personal Social Services Research Unit (PSSRU): https://www.pssru.ac.uk/research/1295/

[xvi]Pope C, Turnbull J, Jones J, Prichard J, Rowsell A, Halford S. Has the NHS 111 urgent care telephone service been a success? Case study and secondary data analysis in England. BMJ Open. 2017;7(5):e014815.

[xvii]Joint Formulary Committee. British National Formulary (online) London, UK: BMJ Group and Pharmaceutical Press. Available from: http://www.medicinescomplete.com

[xviii]Landfeldt, E., Zethraeus, N. & Lindgren, P. Standardized Questionnaire for the Measurement, Valuation, and Estimation of Costs of Informal Care Based on the Opportunity Cost and Proxy Good Method. *Appl Health Econ Health Policy* **17**, 15–24 (2019). https://doi.org/10.1007/s40258-018-0418-2

[xix]Devlin N, Shah K, Feng Y, Mulhern B, Hout B van. Valuing health-related quality of life : An EQ-5D-5L value set for England. Off Heal Econ. 2016;16(1):1–22.

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

[xxi]Kearns, B., Ara, R., Wailoo, A. *et al.* Good Practice Guidelines for the use of Statistical Regression Models in Economic Evaluations. *PharmacoEconomics* **31**, 643–652 (2013). https://doi.org/10.1007/s40273-013-0069-y

[xxii]Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics. 2014;32(12):1157-70.