Statistical Analysis Plan (SAP)

A phase II placebo controlled, double blind, randomised clinical trial to assess the safety and tolerability of 30 mg/KG daily Ursodeoxycholic Acid (UDCA) in patients with Parkinson's Disease (PD)

Contents

Abbreviations

AE	Adverse Event
AR	Adverse Reaction
ATP	Adenosine triphosphate
CRF	Case Report Form
CT	Computerised Tomography
FAS	Full analysis set
ITT	Intention to treat
LED	Levodopa equivalent dose
MADRS	Montgomery-Asberg Depression Rating Scale
MDS-UPDRS	Movement Disorders Society Unified Parkinson's Disease Rating Scale
MoCA	Montreal Cognitive Assessment
NMS	Non-Motor Symptoms
NMS - Quest	Non Motor Symptoms Questionnaire
PAM	Physical Activity Monitor
PCr	Phosphocreatinine
PD	Parkinson's disease
PDQ39	Parkinson's disease quality of life questionnaire
Pi	Inorganic Phosphate
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAR	Serious Adverse Reaction
UDCA	Ursodeoxycholic Acid
^{31}P-MRS	^{31}Phosphorous Magnetic Resonance Spectroscopy

1. Introduction:

Ursodeoxycholic acid (UCDA) is a licensed drug for the treatment of primary biliary cholangitis which has been in clinical use for more than 30 years, it is well tolerated and safe in this established setting. Standard therapeutic doses of UDCA are 13-15mg/kg/day. In September 2015, the CURE Parkinson's Linked Clinical Trials committee, reviewed ~25 promising drugs already in clinical use for their neuroprotective potential. UDCA was rated as the most promising compound for future neuroprotection trials in Parkinson's disease (PD). The planned dose in this study is 30mg/kg/day to ensure sufficient levels of UDCA in the brain.

The primary aim of this study is to generate clinical data to examine the safety, tolerability and potential effectiveness of 48 weeks exposure to UDCA at a dose of 30 mg/kg compared to placebo in a PD population. A definitive study to determine the neuroprotective effect of UDCA using currently available tools would need to involve several hundred patients and cost several million pounds. This pilot study will determine whether such an expensive study would be safe, feasible and justified. Furthermore, it will determine the usefulness of novel objective readouts (namely the objective sensor based quantification of motor progression and ^{31}P-MRS/imaging-based *in vivo* quantification of ATP) which may allow a reduction of the sample size (and thus cost) of future trials.

2. Study Methods

2.1 Study Design

A double-blind, randomised, placebo-controlled, multi-centre, parallel group trial in patients with PD who have been diagnosed ≤ 3 years ago. This trial will be run on two sites: Sheffield and UCLH.

30 patients will be randomised to UDCA at a dose of 30 mg /kg or matched placebo using a 2:1 split (20 patients on UDCA, 10 on placebo).

The study will include 48 week exposure period and a subsequent 8 week washout period.

Detailed evaluations of all patients will take place at Screening, Baseline, 12, 24, 36, 48 and 56 weeks which are referred to as Screening and then Visits 1 to 6 respectively.

The trial medication will be taken at three equal doses per day, to be taken orally with food. The dose will be increased gradually by 250 mg (one capsule) every three days until patient reaches a dose of 30 mg/kg.

Key exclusion criteria include: Patients diagnosed or suspected to have other cause parkinsonism; abnormality in a CT or MRI scan or unsuitability to ^{31}P-MRS.

2.2 Randomisation

Randomisation will take place at the baseline visit after confirmation of eligibility using a centralised, web-based system hosted by epiGenesys (a wholly owned subsidiary of the University of Sheffield) on behalf of the University of Sheffield Clinical Trials Research Unit (CTRU). All participants will be assigned a unique participant ID number at screening that will link all of the clinical information collected for them on the study database, these will be in the format Sxx/nnnn; where xx is the site number and "nnnn" is a unique number starting at 0001 and incrementing by 1. Once the participant ID number has been entered, the system will supply a randomisation number which will identify the treatment pack to be dispensed. The randomisation system will not reveal the actual treatment; although the system can be used to unblind individual participants in cases of emergency. The randomisation will be 2:1 in favour of UDCA. The randomisation system will stratify by site.

UDCA and placebo will be supplied in identical packaging.

2.3 Sample size:

The primary outcome of interest for this study is the safety and tolerability of UDCA which will be assessed using the rate of SAEs in both the UDCA and placebo groups, alongside the review of adverse treatment reactions, and study completion. As the study is a pilot it is not powered to compare the SAE rate between the groups statistically, but any SAEs in either group will be presented descriptively, the placebo group providing a baseline against which to view any SAEs in the UDCA group.

Should this small study result in no SAEs then it would be of interest to determine how likely it is that a larger study would find an intolerable rate of SAEs. For this purpose, we will consider the rate of SAEs reported by Aviles-Olmos et al. in their Exenatide trial to be tolerable and acceptable (i.e. 20%)[1] It is estimated that, in this study, should no SAEs be found in the group receiving UDCA (n=20) then the likelihood that the true SAE rate is less than 20% is 0.990778 (i.e. there is a less than 1% chance that the true SAE rate is intolerable).

2.4 Timelines

Analysis will take place once all participants have had their final assessment at week 56 or have discontinued. A blind review of key data will take place prior to the database being locked and unblinded.

3. Data Collection:

Data, including Adverse Event data, is to be entered onto Prospect, on a secure drive available only to the necessary study staff and with a user log-in system that restricts access to the minimum required by that user. Data will be provided to the SSU as clean SAS data files.

4. Analysis Objectives:

4.1 Primary objective:

The primary objective is safety and tolerability of UDCA at a dose of 30 mg/kg over 48 weeks of exposure which will be assessed by each of the following:

- Serious adverse event rate
- Summary of SAEs
- Summaries of adverse treatment reactions
- Number of patients still taking the study treatment at any dose at the 48 week visit

4.2 Secondary objective:

The secondary objectives are to assess the effect of UDCA compared to placebo on disease progression in PD at 48 weeks (assessed as a change from baseline) by:

- Clinical assessment using Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) part 3 motor subsection in the "OFF" medication state.
- In vivo parameter estimates of high and low energy metabolite levels (ATP, PCr, Pi), derived from cranial ^{31}P-MRS centered on the midbrain and putamen.
- Objective quantification of motor impairment, using motion sensors (Optogait and Opal sensor-based assessment: Sheffield patients only; Dynaport Movemonitor: all patients).

4.3 Exploratory objectives:

The exploratory objectives are to compare changes from week 48 to week 56, and from baseline to week 56, between UDCA and placebo in:

- MDS-UPDRS part III motor subscale in the practically defined "OFF" medication state.
- MDS-UPDRS rating scale parts I, II, III and IV in the "ON" medication state.
- Levodopa equivalent dose (LED)
- Montreal Cognitive assessment (MoCA)
- Montgomery-Ashberg Depression Rating Scale (MADRS)
- Non-Motor Symptoms Scale (NMS-QUEST)
- Parkinson's Disease 39 item quality of life questionnaire (PDQ-39)

Additional exploratory objectives are planned to explore associations between treatment response as quantified by changes of MDS-UPDRS part III motor score at the practically defined "OFF" medication state and:

- Predicted disease progression (Williams-Gray score).
- Energy metabolite levels in brain tissue as quantified by ^{31}P-MRS
- Genetic variants

These analyses will be performed outside the formal analysis of the study as the genetic data will not be available at the time of data base lock.

5. Analysis Sets and Protocol Non Compliances

Full Analysis Set (FAS): All subjects who were randomised into the study and received at least one dose of treatment (UDCA or placebo).
Patients will be analysed according to their randomised treatment unless it can be shown that any misallocation was due to a purely administrative error. This is a modified intention to treat population whereby patients must have taken at least one dose of the trial medication to be included in the analysis.

All analyses will be performed on the FAS.

Protocol Non Compliances will be reviewed prior to the data being unblinded and those which, in the opinion of the study team, may impact the study outcomes will be identified. These will be summarised in the study report. In particular protocol non-compliances due to the coronavirus pandemic will be described including (but not limited to):

- Delayed assessments – for example ^{31}P-MRS, or in clinic motor impairment assessment, length of delay
- Assessments performed virtually rather than in person, for example, MDS-UPDRS etc.
- Missed assessments, either those not possible during virtual visits or whole visits missed.

Following a review of all protocol non compliances prior to unblinding, a sensitivity analysis may be performed for one or more of the primary or secondary objectives, to help assess the impact of the non-compliances.
The subgroup analysis described in section 7.4 will be considered alongside information on protocol non compliances and final analysis decisions made prior to unblinding.

6. Endpoints and Covariates

6.1 Primary endpoints

The primary outcome is the safety and tolerability of UDCA which will be assessed using the measures listed below. Each of these will be assessed from start of treatment to week 56:

- The number of SAEs
- The number of adverse treatment reactions
- The number of patients completing the study

Treatment compliance will provide supportive information to these endpoints.

6.2 Secondary endpoints

The secondary outcomes listed below will be assessed as a change from baseline to week 48.

- MDS-UPDRS part III motor subsection “OFF” medication score (week 48 – baseline)
- In vivo parameter estimates of high and low energy metabolite levels (ATP, PCr, Pi) derived from cranial ^{31}P-MRS centred on the midbrain and putamen (week 48 – baseline)
- Objective quantification of motor impairment, using motion sensors (Optogait and Opal sensor-based assessment, Dynaport Movemonitor) (week 48-baseline).

6.3 Exploratory endpoints

The exploratory outcomes listed below will be assessed as the change from week 48 to week 56 (week 56-week48) and also the change from baseline to week 56 (week 56-baseline):

- MDS-UPDRS part III motor subsection "OFF" medication score
- MDS-UPDRS part I, II, III and IV motor subsection "ON" medication score
- LED
- MoCA
- MADRS
- NMS-QUEST
- PDQ-39

In addition the following endpoints will be included in the exploratory analyses of associations:

- MDS-UPDRS part III motor subsection "OFF" medication score (week 56 – week 48)
- MDS-UPDRS part III motor subsection "OFF" medication score (week 48 – baseline)
- Probability of disease progression calculated at baseline (Williams-Gray score) developed by Velesboer et al[2]
- Changes in energy metabolite levels in brain tissue as quantified by ^{31}P-MRS (week 48-baseline)
- Presence/absence of specific genetic variants at baseline

This analysis of associations may not form part of the formal analysis of the study as the genetic variant data will not be available at the time of database lock.

6.4 Variables

Adverse Events and Adverse Treatment Reactions

Adverse events(AEs); Adverse Reactions (AR) which are AEs suspected to have a causal relationship (relationship to study treatment is recorded as recorded as definite, probable or possible) with the study treatment and Serious Adverse Events (SAEs) be recorded at each study visit.

Serious Adverse Event rate will be defined as the percentage of patients with at least one SAE. All patients with at least 28 days exposure to study treatment or an occurrence of a SAE prior to 28 days exposure will be included in this analysis.

Objective quantification of motor impairment, using motion sensors (Optogait and Opal sensor-based assessment; Dynaport Movemonitor).

The motor impairment is measured in two ways: at home and in-clinic:

At-home real-life monitoring of gait performance (prior to baseline and week 48 visits)
The Dynaport movement monitor is worn by the patients for approximately 7 days. For each patient, data is recorded, on the amount of time spent and movement intensity associated with:

- inactive periods (lying, sitting),
- static periods (standing, shuffling), and
- active periods (walking, stair walking, and cycling).

Duration of worn period (inactive total time + static total time + moving total time) is summarised as a measure of data captured. Data will be cross referenced against patient activity diaries and data captured between 6am and 12am (18hour) will be summarised for 5 days that are considered valid (percentage worn time >=60%) (perc_wornmeasured).

This analysis will concentrate on

- total average movement intensity (MI_worn_mean)
- movement intensity in active period (MI_active_mean)
- active period as a proportion of total time worn (PERC_activeworn) and
- moving total time as a proportion of total time worn (PERC_movingworn)

In-clinic assessment of gait capacity (at baseline and 48 weeks)

There are 3 instrumented gait analysis tasks for Sheffield participants:

1 - 3m 'timed up and go' (TUG) – the time to complete this is recorded on the CRF by the site
2 - walk test at preferred speed
3 - walk test at fast pace

For tests 2 and 3, a combination of OPALS and OPTOgait were used and the variables in table 1 will be recorded. The following variables from the preferred-speed walking trial will be considered as the primary variables within this assessment:

- Step Time variability (mean, "Step_time_var", measures variability)
- Jerk Ratio - Anterior-Posterior/Vertical - Forehead sensor (mean, "Jerk_ratio_AP_Head", measures stability)
- Stride Regularity - Anterior-Posterior axis - Lumbar sensor (mean, "Stride_AP_Pelvis", measures regularity)
- Autosymmetry - Anterior-Posterior axis - Lumbar sensor (mean, "autosym_AP_Pelvis", measures symmetry)

Note that data is not available for all patients and summaries will be calculated for all patients in the FAS with valid observations, in addition some visits may have been delayed due to the coronavirus pandemic, see section 5.

Table 1: Parameters derived from in-clinic assessment of objective quantification of motor impairment

Domain	FieldName	Description	Units
Ambulatory activity	Gait_speed	Gait speed	m/s
	Cadence	Cadence	steps/min
Intensity	RMS_AP_Head	RMS Accelerations - Anterior-Posterior axis - Forehead sensor (mean)	m/s2
	RMS_AP_Pelvis	RMS Accelerations - Anterior-Posterior axis - Lumbar sensor (mean)	m/s2
Pace/rhythm	Step_time	Step time	s
	Stride_time	Stride time	s
	Step_length	Step Length	cm
	Stride_length	Stride Length	cm
	Step_width	Step width	cm
Variability	Step_time_var	**Step time variability**	ms
	Stride_time_var	Stride time variability	ms
	Stride_length_var	Stride length variability	mm
	Step_length_var	Step length variability	mm
Asymmetry	Step_time_asym	Step time asymmetry	ms
	Step_length_asym	Step length asymmetry	ms
Balance (Head)	Jerk_ratio_AP_Head	**Jerk Ratio - Anterior-Posterior/Vertical - Forehead sensor (mean)**	dB
Regularity (Head)	Stride_AP_Head	Stride Regularity - Anterior-Posterior axis - Forehead sensor (mean)	N/A
	Stride_ML_Head	Stride Regularity - Mediolateral axis - Forehead sensor (mean)	N/A
Symmetry (Head)	Autosymm_AP_Head	Auto-Symmetry - Anterior-Posterior axis - Forehead sensor (mean)	N/A
	Autosymm_ML_Head	Auto-Symmetry - Mediolateral axis - Forehead sensor (mean)	N/A
	HR_AP_Head	Harmonic Ratio - Anterior-Posterior axis - Forehead sensor (mean)	N/A
	HR_ML_Head	Harmonic Ratio - Mediolateral axis - Forehead sensor (mean)	N/A
Balance (Pelvis)	Jerk_ratio_AP_Pelvis	Jerk Ratio - Anterior-Posterior/Vertical - Lumbar sensor (mean)	dB
Regularity (Pelvis)	Stride_AP_Pelvis	**Stride Regularity - Anterior-Posterior axis - Lumbar sensor (mean)**	N/A
	Stride_ML_Pelvis	Stride Regularity - Mediolateral axis - Lumbar sensor (mean)	N/A
Symmetry (Pelvis)	Autosymm_AP_Pelvis	**Auto-Symmetry - Anterior-Posterior axis - Lumbar sensor (mean)**	N/A
	Autosymm_ML_Pelvis	Auto-Symmetry - Mediolateral axis - Lumbar sensor (mean)	N/A
	HR_AP_Pelvis	Harmonic Ratio - Anterior-Posterior axis - Lumbar sensor (mean)	N/A
	HR_ML_Pelvis	Harmonic Ratio - Mediolateral axis - Lumbar sensor (mean)	N/A

MDS-UPDRS

The MDS-UPDRS is the standard validated tool for the assessment of patients with PD. This scale includes subsections collecting data regarding the impact of PD on a patient's mood and mental state, (UPDRS part I), their activities of daily living (UPDRS part II) an examination of the motor features of PD (UPDRS part III), and complications arising from the use of dopamine replacement (part IV).

Each part of the MDS-UPRDS consists of a number of items each requiring a response from 0 to 4 (0=Normal, 1=slight, 2=Mild, 3=Moderate, 4=Severe). These are summed to give a total score for each part.

Part I consists of 13 items; part II similarly consists of 13 items; part III has 33 items and part IV has 6 items. Each part is scored individually by adding together the score from each item.

The MDS-UPDRS part III score has to be ascertained in the practically defined "OFF" at baseline (visit 1), visit 5 (48 weeks) and visit 6 (56 weeks) to ensure that the obtained motor score reflects the severity of the underlying neurodegenerative disease process which can be partially or completely masked by the effect of the symptomatic Parkinson's disease medication. The practically defined "OFF" medication state refers to the patient assessment conducted in the absence of their regular medication with the aim of exposing the severity of the underlying PD.

The full MDS-UPDRS assessment will also be carried out in the practically defined "ON" state at baseline (visit 1), visit 5 (48 weeks) and visit 6 (56 weeks). Part III only will be assessed in the "ON" state at visit 3 (24 weeks). The practically defined "ON" stage refers to patient assessments conducted after patient has taken their regular medication (typically 30-60 min after patient has taken PD symptomatic medication).

In vivo parameter estimates of high and low energy metabolite levels (ATP, PCr, Pi) derived from cranial ^{31}P-MRS centred on the midbrain and putamen

The key parameters to be explored from the ^{31}P-MRS scans are:

- Total ATP
- Total PCr
- Total Inorganic Phosphate

There are two voxels placed in the midbrain: left and right. There are 4 voxels placed in the putamen: left posterior, left anterior, right posterior and right anterior putamen. Measures can be derived from a single voxel or as means of a combination of voxels. Each voxel may have varying amounts of brain tissue and/or CSF within, this is the partial volume effect which is anticipated to be small upon any measured ^{31}P-MRS variables (details of variables below).

The above parameters (ATP, PCr, Pi) will be examined at each of the following locations as part of the **secondary analyses**:

- Mean midbrain (variables MSN_TATP, MSN_PCR, MSN_Pi)
- Mean posterior putamen (variables: MPBG_TATP, MPBG_PCR, MPBG_Pi).

The key parameters (ATP, PCr, Pi) will be explored further in the following locations in the putamen as a supportive measure, this analysis will be viewed as **exploratory** within the formal study report:

- Midbrain voxel contralateral to the worst clinically affected side. That is for participants with worse symptoms on the left the variable required will be coded RSN, for those with worse symptoms on the right it will be LSN. (Variables: RSN_TATP, RSN_PCR, RSN_PI)
- Posterior putamen contralateral to the worst clinically affected side. That is for participants with worse symptoms on the left the variable required will be coded RPBG, for those with worse symptoms on the right it will be LPBG. (variables: RPBG_TATP, RPBG_PCR, RPBG_Pi, LPBG_TATP, LPBG_PCR, LPBG_Pi)
- Mean putamen value contralateral to the worst clinically affected side. That is for participants with worse symptoms on the left the variable required will be coded MRBG, for those with worse symptoms on the right it will be MLBG. (variables: MRBG_TATP, MRBG_PCR, MRBG_Pi, MLBG_TATP, MLBG_PCR, MLBG_Pi)

Note that the worst clinically affected side is recorded in the medical history CRF.

The following parameters may also be explored as supportive measures in the above described locations, this analysis will be **exploratory and may be performed separately from the formal analysis** and reporting described in this SAP:

- PCr:ATP ratio
- Phosphate:ATP ratio
- Total high energy phosphates
- Phopshate:High energy Phosphates ratio
- Gibbs free energy

In the midbrain the partial volume will be summarised by the proportion of brain tissue within the voxel and the ratio of grey matter volume to white matter volume, in the putamen it will the grey/white matter ratio only (variables: RSN_non_CSF_prop, LSN_non_CSF_prop, MSN_non_CSF_prop, RSN_GM_WM_ratio, LSN_GM_WM_ratio, MSN_GM_WM_ratio, RPBG_GM_WM_ratio, LPBG_GM_WM_ratio, RABG_GM_WM_ratio, LABG_GM_WM_ratio, MPBG_GM_WM_ratio, MRBG_GM_WM_ratio, MLBG_GM_WM_ratio)

Levodopa Equivalent Dose (LED)

Parkinson's-related dopaminergic medications will be collected on the concomitant medications form, with dose (mg), frequency and start/end dates. The medications will be converted to a LED using the conversion factor listed in the table below, the conversion takes the form:

Doseage x frequency x conversion factor.

Note that is a patient is taking entacapone or opicapone OD then the LED of standard release levodopa drugs is increased by 1.33 after the conversion. See further details in Table 3 below.

During the blind review prior to data base lock, a listing of this data will be provided to the study team.

The stop/end dates will be considered alongside visit dates so that LED at Baseline, week 48 and week 56 can be identified for analysis.

Table 2: Frequency of medication

1=OD	Once daily
2=BD	Twice daily
3=TDS	Three times daily
4=QDS	Four times daily
5=5XD	Five times daily
6=6XD	Six times daily
7=STAT	Once only

Table 3: Medication conversion factor

Medication	conversion factor	Effect of entacaopone/ opicapone
Levodopa (controlled release)	0.75	x 1.33
Levodopa (immediate release)	1	x 1.33
Levodopa (standard release, any preparation, e.g.: Sinemet Plus, Co-Careldopa, Co-Beneldopa)	1	x 1.33
Levodopa with Entacapone	1.33	#
Madopar-Levodopa/ Benserazide	1	x 1.33
Pramipexole (Mirapex)	100	none
Pramipexole (Modified release)	100	none
Rasagiline (Azilect)	100	none
Ropinirole (Requip)	20	none
Ropinirole CR (RequipXL)	20	none
Rotigotine	30	none
Selegiline-Oral	10	none
Selegiline-Sublingual	80	none
Sinemet-Levodopa/Carbidopa	1	x 1.33
Sinemet CR - controlled release Levodopa/Carbidopa	0.75	x 1.33
Stalevo-Levodopa/Carbidopa/Entacapone	1.33	#
Other, specify	N/A	N/A

Note that no further adjustment is required for these medications as the conversion factor accounts for the entacapone.

The Montreal Cognitive Assessment (MoCA)

The Montreal Cognitive Assessment (MoCA) is a screening tool for mild cognitive impairment, taking the form of a series of cognitive tasks and scored out of 30. The total score is calculated by summing all scored from the tasks, one additional point is added for an individual who has 12 or fewer years of formal education, thus the total score could be 31. The recommended cut-off when screening for dementia is 24/25, individuals with a MoCA score < 25 at screening will be excluded.

Montgomery and Asberg Depression Rating Scale (MADRS)

The Montgomery and Asberg Depression Rating Scale (MADRS) is a 10 item physician rated depression severity scale previously used in the assessment of PD. Summing the ratings on each question gives a score from 0 to 60, the appropriate cut-off in PD for screening purposes is 14/15; individuals with a MADRS score of > 16 at screening will be excluded.

NMS-QUEST

The Non-motor symptoms questionnaire (NMS-QUEST) is a series of 30 questions to be answered yes or no. The number of 'yes's is summed to give a score out of 30. A score of under 10 is considered mild, 10-20 moderate and over 20 severe.

PDQ-39

The Parkinson's Disease Quality of Life Questionnaire (PDQ-39) consists of 39 questions, each with possible responses of Never; Occasionally; Sometimes; Often and Always (or cannot do at all); these responses are scored 0, 1, 2, 3 or 4 respectively.

The questions are summed to give scores for eight dimensions:

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

Each dimension score is calculated by summing the scores of each item in the dimension and dividing by the maximum possible score of all the items in the dimension, multiplied by 100. Giving a dimension score ranging from 0 (never have difficulty) to 100 (always have difficulty). Lower scores reflect better QoL

All items are assumed to impact QoL and must be answered to compute scores for each dimension.

For the Social Support score, if a patient does not have a spouse or partner on question 28 then the summary scale can be calculated using only questions 27 and 29:

$$\text{sum of scores}/(4\text{x}2)\text{x}100.$$

The summary index score, PDSI or PDQ-39 SI is calculated as the sum of dimension total scores divided by 8.

Probability of disease progression (Williams-Gray Score) calculated at baseline

Probability of clinical progression as measured by the Williams-Gray score: The probability will be calculated using the model described by Velseboer et al [1]. Patients will be assessed at screening for their

- MDS-UPDRS motor examination (MDS-UPDRS/III) axial score;
- Animal fluency score for which patients are asked to name as many animals as possible in a 1-minute time frame..

The probability will be calculated using the score in these two assessments and the subject's age, using the validated formula

Probability (unfavourable outcome)=

1/[1+exp-{ 0.059 age +0.3794 UPDRS-ME axial score

- 0.0684 animal names x language correction factor -3.1246}]

for English, the language correction factor is 1.267.

Genetic Variants

Patient DNA extracted from blood samples taken at baseline will be analysed using the NeuroChip. The results will be processed by the study team prior to statistical analysis, such that the data will take the form of a small number of binary variables (present/absent) relating to mutations relevant to PD which will be used in the exploratory analysis.

This data will not be available at the time of database lock and will not form part of the formal analysis of the study.

7. Statistical Analyses

Outcome and demographic data will be summarised using appropriate descriptive statistics such as N, mean, standard deviation, min, lower quartile, median, upper quartile, max.

All analyses will be carried out with a two sided 5% significance level.

Summaries of key demographics and baseline characteristics to include:

- Age
- Sex
- Ethnicity
- Blood pressure (systolic and diastolic)
- Pulse rate
- Height (cm)
- Weight (kg)
- BMI

Summaries of disease history to include:

- Time since Parkinson's diagnosis (to baseline)
- Worst affected side
- Any first-degree relatives with PD?
- Age of relative at onset of PD
- Relationship to participant
- Modified Hoehn & Yahr stage
- Predicted disease progression (Williams-Gray score)

Outcome measures will be summarised by timepoint and treatment group, changes over time as detailed in sections 7.1 to 7.3 will also be summarised by treatment group.

All analyses will be carried out on the FAS, unless specified otherwise.

7.1 Primary Analysis

The primary analyses of safety and tolerability will summarise, by treatment group, the counts and details of SAEs, ARs and patients continuing on any dose to the end of the study. These measures will be summarised and descriptive comparisons will be made. Note that as the study is not powered for a formal analysis no formal analysis of these outcomes will be carried out.
Summaries of treatment compliance will be provided by treatment group.

7.2 Secondary Analyses

7.2.1 General approach to secondary analysis

For each secondary analysis, summaries for change from baseline to 48 weeks (calculated as week 48-week1) within treatment group will be presented. Comparisons for change from baseline to 48 weeks between treatment groups will be performed using a t-test with change (week 48 – week 1) as the outcome and treatment as the grouping variable. These analyses will be carried out on the FAS population. After the review of protocol non compliances a sensitivity analysis may be added for certain endpoints, excluding data related to those non-compliances, this decision will be made prior to the unblinding of the data (see section 5).

If there is an imbalance in baseline characteristics, such as age or time since diagnosis, between randomised groups further analyses will be explored within the constraints of the small sample size. This may include an analysis of covariance.

In all analyses, if the assumptions of the test are not met, transformations of the raw data or alternative tests will be considered and the most appropriate solution applied. Where possible analysis decisions will be made prior to unblinding of the data.

This study is not powered to show statistically significant differences in the secondary objectives, therefore the interpretation will concentrate on observed trends and confidence intervals for estimated differences. The results will be presented in such a way (effect size estimates, standard errors and confidence intervals) that this data can be used to inform the design of any future study, including the assessment of sample size and power.

7.2.2 Secondary analysis details by endpoint

MDS-UPDRS part 3 motor subsection "OFF" medication score:
The analysis will compare UDCA to placebo in terms of change in MDS-UPDRS from baseline (week 1) to week 48. This will be done as described above (section 7.2.1).

In vivo parameter estimates of high and low energy metabolite levels (ATP, PCr, Pi), derived from cranial ^{31}P-MRS centred on the midbrain and putamen:
The following variables will be compared between UDCA and placebo groups in terms of change from baseline (week 1) to week 48 in the midbrain:

- Mean midbrain total ATP
- Mean midbrain total PCr
- Mean midbrain total inorganic phosphate
- Mean Posterior Putamen total ATP
- Mean Posterior Putamen PCr
- Mean Posterior Putamen total inorganic Phosphate

To explore differences in tissue composition of each voxel, exploratory plots showing the metabolite against the partial volume variable (scatter plots) and the partial volume variable at baseline and week 48 (line diagrams) will be created. These plots will not form part of the formal study outputs but will be used to assess if it is necessary to include a partial volume variable as a covariate in the analysis.

Change from baseline to week 48 will be compared between UDCA and placebo groups in the following parameters in the midbrain and putamen as supportive measures, **this analysis is exploratory** but will be performed as part of the formal study report:

- Midbrain voxel contralateral to the worst clinically affected side total ATP
- Midbrain voxel contralateral to the worst clinically affected side total PCr
- Midbrain voxel contralateral to the worst clinically affected side total inorganic phosphate
- Posterior putamen contralateral to the worst clinically affected side total ATP
- Posterior putamen contralateral to the worst clinically affected side total PCr
- Posterior putamen contralateral to the worst clinically affected side total inorganic phosphate
- Mean putamen value contralateral to the worst clinically affected side total ATP
- Mean putamen value contralateral to the worst clinically affected side total PCr
- Mean putamen value contralateral to the worst clinically affected side total total inorganic phosphate

Change from baseline to week 48 may also be summarised in these parameters as supportive measures in the above described locations **(this analysis is exploratory and may be performed outside of the formal study reporting)**:

- PCr:ATP ratio
- Phosphate:ATP ratio
- Total high energy phosphates
- Phopshate:High energy Phosphates ratio
- Gibbs free energy

Objective quantification of motor impairment.
For all participant:

- **At-home real-life monitoring:**
  The measurements listed below will be used to compare UDCA and placebo in terms of change from baseline to week 48, as described above.
  Change from baseline to week 48 in:
    - total average movement intensity (MI_worn_mean)
    - movement intensity in active period (MI_active_mean)
    - active period as a proportion of total time worn (PERC_activeworn) and
    - moving total time as a proportion of total time worn (PERC_movingworn)

For Sheffield participants:

- **In clinic assessment of 3m timed up and go:**
  The analysis will compare UDCA and the placebo in terms of change from baseline (week 1) to week 48 in the time to complete, as described above.
- **In clinic assessment of 10m walk at preferred pace:**
  The measurements of:
  - Step time variability;
  - Jerk Ratio Anterior-Posterior/Vertical from the head level sensor;
  - Stride Regularity - Anterior-Posterior axis from pelvis level sensor and
  - Autosymmetry - Anterior-Posterior axis from pelvis level sensor

  will be used to compare UDCA and placebo in terms of change from baseline to week 48, as described above (section 7.2.1). The change from baseline to week 48 for the remaining parameters in Table 1 will be summarised and displayed graphically.

7.3 Exploratory Analysis

For the remaining MDS-UPDRS data (specified below), LED, MoCA, MADRS, NMS-QUEST and PDQ39 data, the change from week 48 to week 56 (calculated as week 56- week 48), and from baseline to week 56 (calculated as week 56-week1) will be summarised within treatment groups using standard summary statistics. Further analysis will follow as for the secondary outcomes. These analyses will include all randomised patients (FAS).

MDS-UPDRS part III motor subsection "OFF" medication score:
The exploratory analyses will compare UDCA to placebo in terms of change in MDS-UPDRS part III motor subsection "OFF" medication score from week 48 to week 56, and from baseline to week 56.

MDS-UPDRS "ON" medication score:
The exploratory analyses will compare UDCA to placebo in terms of change from week 48 to week 56, and baseline to week 56, for each of the following parts:

- MDS-UPDRS part I
- MDS-UPDRS part II,
- MDS-UPDRS part III and
- MDS-UPDRS part IV motor subsection "

In addition to the analyses described above an analysis including an adjustment for LED will be considered.

LED:
The exploratory analyses will compare UDCA to placebo in terms of change in LED value from week 48 to week 56, and from baseline to week 56.

MoCA:
The exploratory analyses will compare UDCA to placebo in terms of change in MoCA score from week 48 to week 56, and change from baseline to week 56.

MADRS:
The exploratory analyses will compare UDCA to placebo in terms of change in MADRS value from week 48 to week 56, and change from baseline to week 56.

NMS-QUEST:
The exploratory analyses will compare UDCA to placebo in terms of change in NMS-QUEST value from week 48 to week 56, and change from baseline to week 56.

PDQ-39:
The exploratory analyses will compare UDCA to placebo in terms of change in PDQ-39 score from week 48 to week 56, and change from baseline to week 56.

In vivo parameter estimates of high and low energy metabolite levels (ATP, PCr, Pi), derived from cranial ^{31}P-MRS centred on the midbrain and putamen:

In addition to the analysis described in section 7.2.2 an exploratory analysis will also be performed using principal component analysis. Variables to be considered for inclusion will be those described for inclusion in the secondary analysis from ^{31}P-MRS. The number of principal components formed will be chosen appropriately according to the number of chosen input variables and observations available.

Investigation of possible associations:
This analysis will take place after the formal analysis and may not include the genetic data if it is not available. The final choice of ^{31}P-MRS variables to be included will made considering the results from the secondary analysis.

Exploratory analyses using regression techniques will be undertaken to investigate any associations between treatment response, measured as change to MDS-UPDRS part 3 motor score in the practically defined "OFF" medication state and:

- Probability of disease progression calculated at baseline (Williams-Gray score) developed by Velesboer et al[2]
- Changes in energy metabolite levels in brain tissue as quantified by ^{31}P-MRS at 48 weeks compared to baseline
- Genetic variants at baseline

The dependent variable in this analysis will be the change in MDS-UPDRS part 3 motor subsection, the variables listed above will be included a covariates (as appropriate), treatment will be included as a factor and if there is sufficient data interactions will be considered.
This analysis will be performed separately for:

- MDS-UPDRS part 3 motor subsection "OFF" medication score (week 48 – baseline)
- MDS-UPDRS part 3 motor subsection "OFF" medication score (week 56 – week 48)

7.4 Subgroup Analysis

If a notable number of patents did not remain on the target dose of 30mg/kg for >24 weeks (>50% of total treatment duration of 48 weeks) then subgroup analysis of the primary and key secondary endpoints will be undertaken for the group of patients who did meet threshold.
The analysis will follow the same analysis plan as described above.

7.5 Adjusted Analysis

Due to the relatively small sample size and pilot nature of the study it is unlikely that an adjusted analysis, in addition to the analyses described above (section 7.4), will be appropriate. However, if there appears to be a significant imbalance between the randomised groups, consideration will be given to adjusting for this in the analysis. Where possible, final analysis decisions will be made prior to unblinding of the data.

7.6 Data Monitoring Committee (DMC):

An independent DMC will monitor the safety and ethics of the trial by overseeing recruitment, primary outcomes data completeness and AEs. The DMC will be provided with unblinded summaries of safety data (SAEs, AEs, relevant safety labs). The study team will remain blinded and the DMC should not communicate any unblinded information to the study team.

7.7 Interim Analysis:

No formal interim analyses are planned. Safety data will be reviewed by the DMC who will be able to recommend the premature closing of the trial if safety concerns arise. In addition, if the investigators have any safety concerns they are able to request an unscheduled meeting of the DMC.

7.8 Multiplicity considerations

Secondary analyses are considered to be exploratory and so no adjustment will be made for multiplicity.

7.9 Missing Data

If there are noteworthy amounts of missing patient data the reason for this will be investigated. Where possible, safety data from missed visits will be completed at future visits e.g. adverse events.

Treatment compliance will be summarised and if appropriate described and compared between treatment groups.

Missing clinical assessments will be noted in the interpretation but no attempt will be made to impute any missing data.

MoCA or NMS-QUEST
No Adjustment is made to the MoCA or NMS-QUEST scores for missing data, but the amount of missing data will be summarised to aid interpretation.

PDQ-39
For the PDQ-39 All items are assumed to impact QoL and must be answered to compute scores for each dimension

MDS-UPDRS

For the MDS-UPDRS a small amount of missing data is allowable, the missing data should be assessed descriptively to assess if items are missing randomly across patients or the same item is missing consistently across all patients. The table below (ref: Handling Missing Values in the MDS-UPDRS, Goetz et al, Movement Disorders, Vol. 30, No. 12, 2015) shows the number of missing values which are allowable in the calculation of total scores. If more observations are missing for an individual patient then their total score (for that part) will be missing.

	Same item missing consistently across all patients	Different items randomly missed across patients
Part I	1	1
Part II	1	2
Part III	3	7
Part IV	0	0

Where items are missing and calculation of the score is allowable (as per table above) then the score will be calculated as follows:

(sum of non missing scores) x (total number of items in scale) / (number of non missing scores).

8. References

1. Aviles-Olmos I, Dickson J, Kefalopoulou Z, et al. Exenatide and the treatment of patients with Parkinson's disease. The Journal of clinical investigation 2013;123:2730-2736.
2. Velseboer DC, de Bie RM, Wieske L, et al. Development and external validation of a prognostic model in newly diagnosed Parkinson disease. Neurology 2016;86:986-993.

9. List of Outputs

Table 1.1 Patient disposition by randomised group (after randomisation)– showing number of patients screened, randomised, completing and withdrawing. This table should also show the number of patients with valid data for each of the secondary objectives.*

Table 1.2 Protocol non compliances (summarised by category, event category and treatment group).*

Table 1.3 Baseline characteristics and or demographics, (see list in section 7)*

Table 1.4 Disease characteristics at baseline (see list in section 7)*

Table 2.1.1 Summary of MDS-UPDRS part 3 motor subsection “OFF” medication score over time*

Table 2.1.2 Summary of change in MDS-UPDRS part 3 motor subsection "OFF" medication score. (to show change from baseline to week 48, week 48 to week 56 and baseline to week 56)

Table 2.1.3 Summary of comparison of change in MDS-UPDRS part 3 subsection "OFF" medication between UDCA and placebo groups (to show change from baseline to week 48, week 48 to week 56 and baseline to week 56)

Table 2.1.4 to include any subgroup/sensitivity analyses

In general, for the following tables (2.x) for secondary endpoints the following pattern will be followed:

- Table 2.x.1 Summary of xxxx over time
    - this table shows summary statistics (mean, standard dev. median, minimum, maximum etc) at each time point by actual (unblinded) treatment group.
- Table 2.x.2 Summary of change in xxxxx. (to show change from baseline to week 48, week 48 to week 56 and baseline to week 56)-
    - this table shows summary statistics for each change by treatment group.
- Table 2.x.3 Summary of comparison of change in xxxx between UDCA and placebo groups (to show change from baseline to week 48, week 48 to week 56 and baseline to week 56)
    - this table shows the results of the analysis for each change, eg estimated effect size, confidence interval, p-value.

Tables2.x Summary of ^{31}P-MRS parameters of high and low energy metabolites

Tables2.x Summary of Sensor-based objective quantification of motor impairment

Tables for exploratory endpoints will follow a similar pattern to those for the secondary endpoints for the comparisons as detailed in section 7.3:

Tables2.x Summary of MDS-UPDRS part 1, 2, 3 and 4 "ON" medication score

Tables2.x Summary of LED

Tables2.x Summary of MoCA

Tables2.x Summary of MADRS

Tables2.x Summary of NMS-QUEST

Tables2.x Summary of PDQ-39

Table 2.x Exploratory analysis of association between changes to MDS-UPDRS part 3 motor score in the practically defined "OFF" medication state and disease progression score (Velseboer) at baseline, ^{31}P-MRS parameters and genetic variants.

Table 3.1.1 Summary of exposure to study drug *

Table 3.1.2 Number of patients still taking study drug at the 48 week visit*

Table 3.1.3 Summary of patients remaining on the target dose of study treatment for >24 weeks*

Table 3.2.1 Summary of Adverse Events by System Organ Class and Preferred Term

Table 3.2.2 Summary of Serious Adverse Events by System Organ Class and Preferred Term

Table 3.2.3 Summary of Serious Adverse Event Rate

Table 3.2.4 Summary of Adverse Treatment Reactions by System Organ Class and Preferred Term

Listing 4.1 Listing of patient disposition

Listing 4.2 Listing of protocol non-conformances*

Listing 4.3 Listing of dosing start and stop dates including days on treatment*

Listing 4.4 Listing of baseline characteristics

Listing 4.5 Listing of disease characteristics

Listing 4.6 Listing of MRD-UPDRS part 3 motor subsection "OFF" medication score

Listing 4.7 Listing of ^{31}P-MRS parameters of high and low energy metabolites

Listing 4.8 Listing of sensor-based objective quantification of motor impairment

Listing 4.9 Listing of MDS-UPDRS part 1, 2, 3 and 4 "ON" medication score

Listing 4.10 Listing of LED*

Listing 4.11 Listing of MoCA

Listing 4.12 Listing of MADRS

Listing 4.13 Listing of NMS-QUEST

Listing 4.14 Listing of PDQ-39

Listing 4.15 Listing of Adverse Events to include SOC, PT, start and stop days (relative to first does) intensity, action taken, outcome, relationship, serious (Y/N & reason)

*= key for Blinded Review prior to data base lock

Figures

Figure 2.x.x Figures will be provided to illustrate the results of the analyses of key secondary endpoints.