STATISTICAL ANALYSIS PLAN (SAP)

Protocol Title	A randomized, double-blind, placebo-controlled trial of urate-elevating inosine treatment to slow clinical decline in early Parkinson's disease (SURE-PD3)
Protocol Number	INO-PD-P3-2014
Protocol Version	4.0, 28 Nov 2017
Regulatory Sponsor	Michael A. Schwarzschild, MD, PhD IND # 100,896
Supported by	National Institute of Neurological Disorders and Stroke (NINDS), National Institutes of Health (NIH)
SAP Version	1.0, 05 Sep 2019

SAP APPROVAL SIGNATURES

Michael Alan Schwarzschild

Digitally signed by Michael Alan Schwarzschild
DN: cn=Michael Alan Schwarzschild, o=Massachusetts General Hospital, ou=Neurology, email=michaels@helix.mgh.harvard.edu, c=US
Date: 2019.09.05 09:00:12 -04'00'

Michael A. Schwarzschild, MD, PhD — Date
Sponsor, CCC Principal Investigator, Study Chair

Alberto Ascherio

Digitally signed by Alberto Ascherio
DN: cn=Alberto Ascherio, o=Harvard University, ou=Harvard TH Chan School of Public Health, email=aascheri@hsph.harvard.edu, c=US
Date: 2019.09.05 05:39:36 -04'00'

Alberto Ascherio, MD, DrPH — Date
Study Co-Chair

David Oakes 9/5/2019

David Oakes, PhD — Date
DCC Principal Investigator and Biostatistician

Eric A. Macklin

Digitally signed by Eric A. Macklin
DN: cn=Eric A. Macklin, o=Massachusetts General Hospital, ou=MGH Biostatistics Center, email=emacklin@mgh.harvard.edu, c=US
Date: 2019.09.05 00:39:39 -04'00'

Eric A. Macklin, PhD — Date
CCC Biostatistician

SAP REVISION HISTORY

Version	Date	Description of Changes
1.0	05 Sep 2019	Initial version

Table of Contents

1. Introduction

This statistical analysis plan (SAP) defines the outcome measures and analysis samples and specifies the planned analyses of data for the SURE-PD3 trial. The SAP supplements the clinical protocol. Please refer to the clinical protocol for details on the rationale for the intervention, eligibility criteria, conduct of the trial, clinical assessments and the timing of their use in the trial, definitions and reporting of adverse events, data management conventions, and regulatory oversight and compliance procedures. In case of discrepancies between the SAP and the clinical protocol concerning matters of data analysis, the SAP is authoritative. On all other matters, the clinical protocol is authoritative.

This SAP specifies data and planned analyses for the main trial. Specification of data and analyses for ancillary studies will be detailed in ancillary SAPs if not covered here.

2. Study Design

2.1 Overview

This is a phase 3, randomized, two-arm, parallel-group, placebo-controlled, triple-blind, two-period, multicenter clinical trial of oral inosine titrated to elevate trough serum urate to 7.1 to 8.0 mg/dL over 24 months with a 3-month wash-out among early Parkinson disease (PD) patients exclusive of those with scans without evidence of dopaminergic deficit (SWEDD). The primary aim is to test the effectiveness of oral inosine dosed to elevate serum urate in slowing or delaying PD progression over 24 months based on rate of change in the Movement Disorders Society Uniform Parkinson's Disease Rating Scale (MDS-UPDRS) Parts I-III total score while participants are not receiving dopaminergic therapy (except possibly a stable dosage of a monoamine oxidase-B [MAO-B] inhibitor at baseline). Randomizations are stratified by site to avoid chance confounding between site characteristics and treatment. The trial is registered at Clinicaltrials.gov as study NCT02642393 (see https://clinicaltrials.gov/ct2/show/NCT02642393).

The randomized, two-arm, parallel-group design provides an unbiased estimate of effectiveness of oral inosine dosed to elevate serum urate in slowing or delaying PD progression over 24 months during period 1. Tracking symptoms during randomized wash-in and during the non-randomized 3-month wash-out of period 2 will allow estimation of symptomatic effects of serum urate elevation and thereby better evaluate whether any observed effects on PD progression reflect a disease-modifying effect. Note that although suspension of study drug during the wash-out will be unmasked, subjects and staff will remain blinded as to whether this represents a transition off of active treatment or of placebo treatment, and thus treatment comparison of changes during the 3-month wash-out should remain unbiased. The difference in the proportion of subjects requiring dopaminergic therapy after the 3-month wash-out among all those randomized will provide an estimate of disease-modification by serum urate elevation free of any symptomatic effects of study drug.

2.2 Study Objectives

The primary objective of the trial is to determine whether oral inosine dosed to elevate serum urate is effective in slowing or delaying PD progression over 24 months. Secondary objectives include evaluating the safety and tolerability of serum urate elevation, determining whether serum urate elevation causes short-term, reversible effects on clinical symptoms, and whether

serum urate elevation is effective in delaying time to need for or initiation of dopaminergic therapy or in slowing worsening of cognitive function, mood, autonomic function, quality of life, or functional disability.

2.3 Study Population

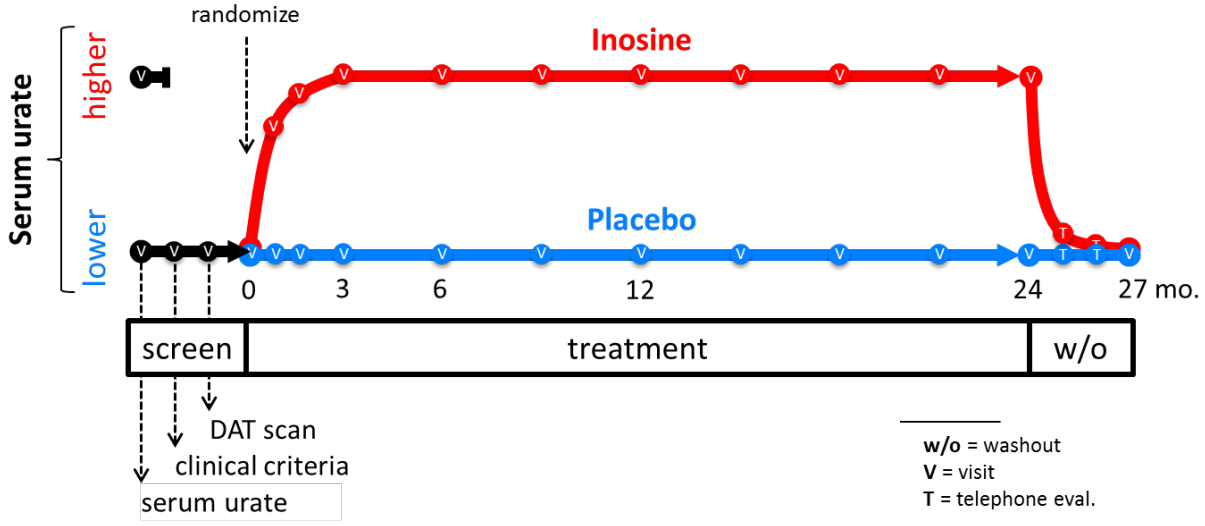
Individuals eligible for trial participation are men or women age 30 years or older with early idiopathic PD, serum urate ≤5.7 mg/dL, with evidence of dopamine deficit by dopamine transporter brain scan, and free of risk factors for potential urate-related adverse events (AEs). Early PD requires a modified Hoehn and Yahr Scale stage less than 3, diagnosis within 3 years of screening, and without current or imminent disability requiring dopaminergic therapy other than a stable dosage of an MAO-B inhibitor. Detailed inclusion and exclusion criteria are specified in the clinical protocol. Participants will be recruited from approximately 60 clinical sites located through the US.

2.4 Participant Flow

After providing informed consent, participants will complete a sequence of three screening visits. At Screening Visit 1 (SC1), participants will be screening for serum urate levels and urine pH to evaluate the highest yield eligibility criteria (serum urate ≤5.7 mg/dL) and evaluate risk of serum urate elevation (urine pH ≤5.0 or history of gout,). For those subjects who pass SC1, most of the remaining screening procedures are conducted on a longer second screening visit (SC2). If participants are still eligible, then a dopamine transporter (DAT) ligand binding brain scan will be acquired at a local neuroimaging center and read centrally to exclude participants classified as having a scan without evidence of dopaminergic deficit (SWEDD).

Eligible participants will be randomized, complete a baseline visit, and initiate study drug. Participants will be seen in-clinic at 3, 6, and 12 weeks and then quarterly to 24 months for study visits and collection of blood samples to measure trough serum urate. Following study drug discontinuation, participants will be followed during a 3-month wash-out period with monthly telephone calls and a final in-clinic study visit 27 months after baseline.

Detailed descriptions of study procedures and timing are specified in the clinical protocol. A schematic of participant flow is given below.



2.5 Treatment Allocation

Prior to the baseline visit, eligible participants will be randomly allocated in equal proportions to one of two treatment groups, oral inosine titrated to achieve trough serum urate in the range 7.1 to 8.0 mg/dL or placebo, according to a permuted-block randomization schedule, stratified by site. The randomization schedule was prepared by computer program by the unblinded study statisticians.

2.6 Treatment Administration

Study drug will be orally self-administered in capsules containing 500 mg of inosine (active drug) or lactose (placebo). Participants will take up to two capsules three times per day (i.e., up to 3.0 gm/day) based on individualized titration that aims to achieve trough serum urate in the range 7.1 to 8.0 mg/dL in the active arm and equivalent dosage adjustments in the placebo arm. Details of the titration algorithm are specified in the clinical protocol. Note, the maximum study drug dosage was reduced from 3.0 to 2.0 gm/day beginning 24 Jan 2018 based on safety concerns over a higher than expected rate of kidney stones with greatest incidence among participants on study drug dosages greater than 2.0 gm/day.

2.7 Allocation Concealment

The randomization schedule is known only by the unblinded study statisticians who generated the schedule and implement the titration algorithm and a calibrated urine alkalinization process and by the study drug distributor. Concealment of the true treatment allocation of specific participants is achieved by use of matched active and placebo capsules, matched titration schedules, and matched urine alkalinization rates. Clinical members of the Steering Committee, site investigators and other site staff, clinical coordination and data management staff, the medical monitor, and all participants are blinded to participant treatment allocations. The Data and Safety Monitoring Board (DSMB) members are provided treatment-specific information in order to monitor the trial but such information is masked by use of coded values to identify the treatment groups. The DSMB may request the true treatment identities.

2.8 Schedule of Assessments

ACTIVITY	SCREEN			TREATMENT (Period 1)											WASHOUT (Period 2)				UNPLANNED VISITS			
	Wash-in			Maintenance																		
Day Interval	- 60 to - 13	- 58 to - 11	- 52 to - 5	00	Week 3 21 ± 3	Week 6 42 ± 5	Week 12 84 ± 7	Month 6 180 ± 7	Month 9 270 ± 7	Month 12 360 ± 7	Month 15 450 ± 7	Month 18 540 ± 7	Month 21 630 ± 7	Month 24 720 ± 7	Month 25 V10 + 30 ± 3	Month 25-26 V10 + 45 ± 15	Month 26 V10 + 60 ± 3	Month 27 V10 + 90 ± 3	n/a	n/a	n/a	n/a
Visit#/code	SC1	SC2	DS1[6]	BL[6]	V01	V02	V03	V04	V05	V06	V07	V08	V09	V10	TE1	DS2[7]	TE2	SV	U01, U02, etc.	DF	T(vst #) e.g.,T06	PW
Written Informed Consent	x																					
Consent for Optional Blood Sample Collection/Storage/Future Use[1]	±																					
Smart SURE Consent, Registration, & Training[1]		±	±	±	±	±	±	±	±	±	±	±	±									
Consent for 2nd DAT Scan[1]										±	±	±	±	±								
Assign Subject ID Number	x																					
Screening/Demographics	x																					
Socio-Economics				x																		
PD Risk Factors				x																		
Smoking, Alcohol and Caffeine Status				x																		
Inclusion/Exclusion	x	x		x																		
Medical History General		x																				
Vital Signs (orthostatics & weight)		x		x			x			x				x				x	±	x		
Height				x																		
Electrocardiogram		x								x				x						x		
General Neuro & Medical Exams		x																				
Blood Draw ---> Serum Urate[2]	x	x		x	x	x	x	x	x	x	x	x	x	x				x	±	x		
Blood Draw ---> Chem 7 Panel[2]		x		x	x	x	x	x	x	x	x	x	x	x				x	±	x		
Blood Draw ---> LFTs, TSH and Lipids[2]				x			x							x					±	x		
Blood Draw ---> CBC[2]				x	±	±	x	±	±	±	±	±	±	x					±	x		
Urine Collection ---> Analysis/Sediment[2]	x			x	x	x	x	x	x	x	x	x	x	x				x	±	x		
Blood Draw ---> Biomarker Study[1]				±										±				±	±			
Most Revent Food Intake, Study Drug and PD Medication				x	x	x	x	x	x	x	x	x	x	x								
1st DAT Scan[1]			±																			
2nd DAT Scan[1]																±						
24 Hr Urine Collection: Process/Ship				x						x												
Pregnancy Form[3]		x																				
Mini-Mental State Exam		x																				
PD Features		x																				
Diagnostic Features - PD																		x		x		
Primary Diagnosis																		x		x		
MDS-UPDRS Parts I-III		x		x	x	x	x	x	x	x	x	x	x	x				x		x		
MDS-UPDRS Part IV (as warranted)						±	±	±	±	±	±	±	±	±				x		x		
Modified Hoehn & Yahr		x		x			x			x				x				x		x		
Modified Schwab & England ADL and PDQ-39				x			x			x				x				x		x		
NeuroQOL		x																x				
Depression module of NeuroQOL				x			x			x				x								
MoCA (cognition test)				x			x			x				x				x		x		
RBD Question	x			x			x			x				x				x		x		
Assess Need for Dopaminergic Therapy[4]				x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	
Concomitant Medication Log		x		x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	
Adverse Event Log[5]		x		x	x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	
Randomization to Med Assign				x																		
Dose Management Log				x	x	x	x	x	x	x	x	x	x	x					x	x		
Study Drug Dispensing/Return Log				x	x	x	x	x	x	x	x	x	x	x						x		
PD-EQ				x																		
Blindedness Questionnaire						x								x					±	x		
Post Visit Telephone Call Regarding Eligibility and DAT Scan	x	x	x																			
Post Visit Telephone Call to Adjust Dose					x	x	x	x	x	x	x	x	x									
Conclusion of Study Participation																		x				x

[1] Optional (opt-in) study

[2] Safety lab samples sent to central laboratory

[3] For women of child-bearing potential; repeated at other times if clinically indicated

[4] Once need for dopaminergic therapy has been determined, no longer complete the form at future visits

[5] For subjects that discontinue study participant prior to Month 27, AEs active at the time the subject discontinues participation should be followed for 30 days post discontinuation or until resolved or stablized (whichever comes first) as long as the subject agrees to the follow-up.

[6] DAT scan may not be conducted until SC lab results are reviewed; and Baseline visit may not be conducted until SC lab results and DAT scan results are reviewed

[7] End of study DAT scan will only be conducted with subjects who have consented to the additional DAT scan.

± = Optional as specified in protocol

3. Statistical Design

3.1 Primary Outcome

Our choice of rate of change in MDS-UPDRS I-III total score as our primary efficacy measure is based on an interest in an outcome that can be evaluated within 2 years and an interest in a patient-reported outcome. Moreover, our preliminary data from the SURE-PD trial (NCT NCT00833690) showed dosage-dependent efficacy of serum urate elevation for 24-month change in UPDRS I-III total score. Anticipating that recruitment will require 18 months, evaluations of the primary outcome needs to be completed in 2 to 2.5 years in order to launch, implement, and publish results from the trial within a 5-year grant period.

Assessing change in MDS-UPDRS or other measures of motor symptoms among initially de novo PD patients becomes more complex as subjects begin initiating dopaminergic therapy, with roughly two-thirds expected to have initiated dopaminergic therapy by 2 years in the placebo group. We feel that assessing motor symptoms when not on dopaminergic therapy best reflects the efficacy of the intervention under study, whereas evaluation in the OFF condition after dopaminergic therapy has already been initiated is burdensome to subjects and unreliable due to incomplete and variable wash-out. Evaluating motor symptoms in the ON condition is prone to bias due to adjustment of dopaminergic therapy to achieve a pre-targeted level of symptoms. We propose instead to analyze MDS-UPDRS measurements made prior to initiating dopaminergic therapy, considering use of such treatments a censoring event that precludes observation of future untreated MDS-UPDRS scores. While delaying progression over only 2 years provides less benefit to patients than we would ultimately wish to offer, we feel that a demonstration of delay over 2 years in the absence of symptomatic effects is an achievable result, one that could provide evidence of disease modification early in the clinical course, and may motivate a long-term trial focused on functional and quality of life endpoints.

3.2 Efficacy Outcomes

Additional efficacy outcomes will include time to disability warranting dopaminergic therapy, the prescribed levodopa equivalent dosage (LEDD), motor function as measured by MDS-UPDRS part III, ambulation subset, and patient-reported sections (IB and II), striatal DAT binding by DaTscan™, cognition as measured by MoCA, mood as measured by Neuro-QOL Depression, quality of life as measured by PDQ-39, and functional disability as measured by modified Schwab and England. Measures of parkinsonian symptoms assessed by the MDS-UPDRS will be analyzed in two ways: (a) excluding assessments completed after initiation of dopaminergic medication, and (b) adjusting for time-dependent LEDD.

3.3 Safety Outcomes

Safety of oral inosine titrated to achieve trough serum urate in the range 7.1 to 8.0 mg/dL will be evaluated by comparing active vs. placebo treatment groups with respect to overall adverse event (AE) and serious adverse event (SAE) rate, time to first SAE, and proportions of subjects experiencing (a) each type of AE, classified by Medical Dictionary for Regulatory Activities (MedDRA) preferred term and system organ class, (b) clinically-significant abnormal labs, and (c) clinically significant abnormal vital signs, including orthostatic hypotension defined as positional dizziness or other clinical symptoms of orthostatic hypotension or a drop in systolic

and diastolic blood pressures of 20 mm Hg and 10 mg Hg or more, respectively, when moving from a supine to a standing position.

3.4 Tolerability

Tolerance of treatment by an individual participant will be defined as remaining on-study and on his/her assigned treatment without one or more AE-associated dosage reductions lasting more than 4 weeks cumulative. Tolerability of oral inosine titrated to achieve trough serum urate in the range 7.1 to 8.0 mg/dL will be defined as the proportion of all participants in the active arm who are tolerant of the treatment at 12 weeks (short-term), 12 months (medium-term), and 24 months (long-term). Oral inosine titrated to achieve trough serum urate in the range 7.1 to 8.0 mg/dL will be declared tolerable for a given duration of treatment if the proportion who are tolerant is significantly greater than 50%.

3.5 Symptomatic Effects

Symptomatic effects will be estimated by changes in motor symptoms (a) during the first 3 months of wash-in at the start of period 1, and (b) during the 3-month wash-out during period 2 or at the end of study drug exposure, if earlier.

3.6 Two-period Evaluations

Analysis of the persistence of benefit after the 3-month wash-out of period 2 among subjects randomized to active treatment during period 1 permits an evaluation of disease-modifying effect of serum urate elevation. Not all measures are amenable to two-period evaluation because of the censoring effect of initiation of dopaminergic therapy. In particular, with only one-third of subjects expected to have not initiated dopaminergic therapy at the end of period 1, few data on uncensored MDS-UPDRS trajectories during wash-out will be available and subjects contributing those data are a non-random subset of slow progressors. For LEDD-adjusted analysis of MDS-UPDRS and other measures that are equally evaluable at the end of both periods, a three-part test of significantly slower worsening during period 1, non-inferior rates of worsening during period 2, and a significant net benefit at the end of period 2 would provide evidence of disease modification if all three evaluations were favorable.

3.7 Effect Size for Primary Outcome

The effect size for determining power is based on estimates of the minimum clinically important difference (MCID) for changes in MDS-UPDRS I-III total scores and on the previously observed association between baseline serum urate levels and the rate of change in UPDRS I-III total scores in the DATATOP, PRECEPT, and SURE-PD trials.

Hauser et al.[1] report that placebo-treated subjects who reported being minimally worse over 26 weeks on a global impression of change assessment experienced an average increase in UPDRS I-III total scores of 4.9 units. Given our expectation that elevated serum urate will delay progression, not improve symptoms, a MCID related to minimal worsening among placebo subjects seems most appropriate in this context. Based on estimates by Goetz and colleagues[2-5] for PD patients with mild symptoms (Hoehn and Yahr stage I/II), MDS-UPDRS I-III total scores are roughly 30% larger than UPDRS I-III total scores (2.5x for 16 questions in Part 1, 1.1x for 52 questions in Part II, and 1.2x for 108 questions in Part III, yielding a weighted conversion of 1.29x), implying a MCID of 6.3 MDS-UPDRS I-III total score units.

We expect to enroll subjects with mean baseline serum urate of approximately 4.5 mg/dL based on data from SURE-PD and then raise levels in the active arm 3 mg/dL on average at trough sampling. A 3 mg/dL difference in baseline serum urate predicted 8.1 and 2.1 unit per year differences in UPDRS I-III total score slopes in the DATATOP[6] and PRECEPT[7] trials. In the SURE-PD trial[8], subjects randomized to moderate serum urate elevation (7 to 8 mg/dL at random sampling) progressed 1.1 unit / year slower than placebo arm subjects with wide confidence bounds (95% CI 4.8 units / year slower to 2.8 units / year faster). Note that this estimate from SURE-PD (see protocol Fig. 7D) is conservative relative to an alternative model for the effect of serum urate elevation (see protocol Fig. 7C). A random-effects meta-analysis of results from the three studies yields a weighted estimate of 4.5 units slower progression in UPDRS I-III total score over 2 years among patients with higher serum urate levels.

We propose that 6.3 units over two years or a difference in slopes of 3.15 units per year is a reasonable minimum clinically important difference (MCID) in MDS-UPDRS I-III total scores, equal to a MCID of 4.9 units on the scale of UPDRS I-III total scores. This would correspond to a reduction of 20% of the expected placebo rate based on a random effects meta-analysis (Fig. 1 [from protocol Fig. 13]) of the mean rates of decline in UPDRS I-III total scores among de novo PD patients with baseline serum urate levels below 5.9 mg/dL in the DATATOP[9,10], PRECEPT[11], and SURE-PD[8] cohorts. Changes in UPDRS I-III total scores were analyzed using random-slope mixed models censoring follow-up when dopaminergic therapy was initiated. The weighted mean estimate was 0.98 points / month or 23.5 points over 2 years with substantial heterogeneity in the estimate among studies (see Fig. 1).

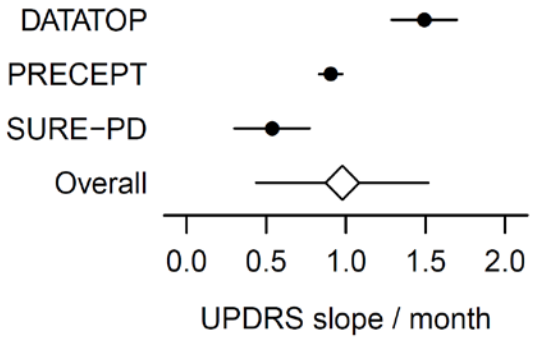


Figure 1. Meta-analysis of the rate of UPDRS change in early PD subjects with lower serum urate. Early, largely untreated PD subjects with a serum urate of <5.9 mg/dL at baseline from DATATOP (n=155; double-placebo only), PRECEPT (n=446) and SURE-PD (n=25; placebo only) were followed for up to two years. The overall UPDRS rate was obtained by weighting each estimate by the inverse of its standard error.

3.8 Sample Size

Power for the primary outcome of rate of change in MDS-UPDRS I-III total score is based on a random slopes model with shared baseline. The model will include fixed effects of time, treatment x time, sex, sex x time, an indicator of baseline MAO-B inhibitor use, and baseline MAO-B inhibitor use x time and random site- and subject-specific intercepts and slopes, each with unstructured covariance. MDS-UPDRS assessments completed after a subject has initiated dopaminergic therapy will be censored.

Based on applying the same primary analysis model to data from SURE-PD, the following variance components were estimated for UPDRS I-III total scores: site-level variance (intercept = 9.75, slope = 0.0123 / month, covariance = -0.346), subject-level variance (intercept = 77.4, slope = 0.230 / month, covariance =2.87), and residual variance = 13.9. We assume that 70% of subjects will initiate dopaminergic therapy based on experience in SURE-PD plus up to 8% additional lost to follow-up prior to initiating dopaminergic therapy. With the planned schedule for MDS-UPDRS assessments (screening, baseline, 3 weeks, 6 weeks, 12 weeks, and then

quarterly through 24 months), the variance and censoring estimates above imply an effective standard deviation for UPDRS I-III total score slopes of 0.587 units / month or 0.758 units / month on the scale of MDS-UPDRS I-III total scores.[12]

Given a standard deviation of 0.76 units / month, a final two-sided test at alpha = 0.046 allowing conservatively for two interim analyses at alpha = 0.001 each, the study would have 80% power with n = 270 subjects randomized 1:1 to placebo or urate elevation if the true effect of treatment were to reduce the rate of increase in MDS-UPDRS I-III total score by 6.3 points over 2 years.

This estimate of power is robust to variable sex-specific enrollment rates and treatment efficacy as long as the average effect of treatment across sexes in the ratio enrolled is 6.3 points over 2 years (Fig. 2A). If urate elevation reduces the average rate of progression by 6.3 points over 2 years only among male or female participants and the other sex experiences less benefit, power would be lower (Fig. 2B). This is unavoidable if a large proportion of the enrolled population accrues less benefit from the intervention. Conversely, power would be greater than 80% if one sex experiences a benefit greater than 6.3 points over 2 years and the other sex experiences at least that large a benefit.

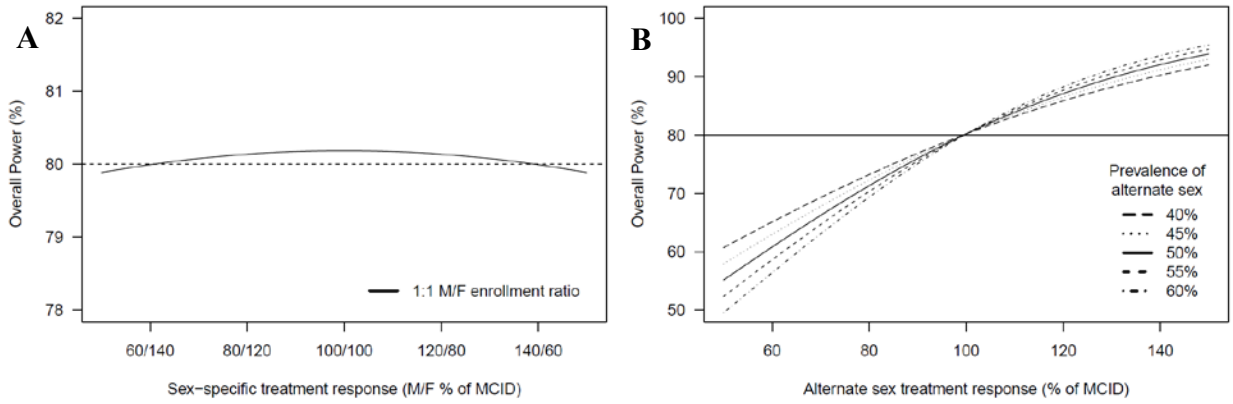


Figure 2. (A) Power for the primary aim across a range of sex-specific treatment responses, all cases with the average treatment response across sexes equal to the MCID of 6.3 points over 2 years. Note that equal prevalence of male (M) and female (F) participants is plotted, but the power curve for other prevalence ratios are so similar as to be indistinguishable at this scale. (B) Power for the primary aim when the effect of treatment on one sex is equal to the MCID and the effect on the alternate sex varies from 50% to 150% of MCID for a range of different prevalence ratios.

3.9 Power for Secondary Outcomes

The trial will have an 80% probability of observing at least one instance of any class of adverse event expected to occur in at least 1.2% of individuals receiving inosine. The trial will have 80% power to detect increased risk of any class of adverse event among subjects receiving inosine if the true risk is two-fold higher and the expected proportion among placebo subjects is at least 14% or if the true risk if four-fold higher and the expected proportion among placebo subjects is at least 3.4%. The trial has 80% probability of declaring serum urate elevation tolerable if the true proportion tolerant under the definition given in Sec. 3.4 at 12 weeks, 12 months, and 24 months is at least 62%.

The estimated SE for symptomatic effects on UPDRS I-III total scores during wash-in between the moderate elevation and placebo arms in the SURE-PD trial was 0.53 units / month. With n = 25 subjects per treatment group in SURE-PD, the expected SE for estimating symptomatic effects in this trial will be roughly sqrt (25/135) * 0.53 = 0.23 units / month. With that SE, the trial would have 80% power to detect symptomatic effects on the order of a difference in slopes during wash-in or wash-out of 0.65 units/month.

The Kaplan-Meier product-limit estimate from the SURE-PD trial of the proportion of placebo subjects requiring dopaminergic therapy by 24 months was 62%. Assuming constant hazard at the same rate among placebo subjects in this trial during period 1 (0 to 24 months) and 8% loss to follow-up prior to determining need for dopaminergic therapy, the trial will have 80% power to detect a hazard ratio of 0.62. Extrapolating to 27 months, we expect a total of 67% of placebo arm subjects to have disability warranting dopaminergic therapy at the end of period 2 (24 to 27 months). The study would have 80% power to detect a difference in the proportion of subjects with disability warranting dopaminergic therapy if 50% or fewer of active arm subjects have progressed by the end of the trial.

Given preliminary data from SURE-PD on the variance components from random-slope models for MoCA after Rasch score conversion, GDS-15, and S&E ADL, the effective standard deviations (in points per month) for these measures given our planned follow-up schedule are 0.038, 0.097, and 0.31, respectively. Weaver et al.[13] report an among-person standard deviation for PDQ-39 at 24 months after deep brain stimulation of 14.2 to 15.3. Assuming conservatively no within-person covariance, that would imply a standard deviation for 24-month change of approximately 30. Given these estimates, the trial will have 80% power to detect 2-year treatment differences on these outcomes as small as 0.31, 0.80, 2.5, and 10.3, respectively.

4. Baseline Characteristics, Study Endpoints, and Final Disposition

4.1 Baseline Characteristics

Each analysis sample will be summarized overall and by treatment group for the following characteristics: randomization site; age, sex, race, ethnicity, time since symptom onset, time since PD diagnosis, modified Hoehn and Yahr score, resting tremor, use of an MAO-B inhibitor, serum urate, MDS-UPDRS scores, NeuroQoL module scores, years of education, MMSE, MOCA score, PDQ-39 score, modified Schwab and England ADL score, smoking history, caffeine consumption, sense of smell, RBD history, vigorous physical activity, serum urate, DAT ligand uptake, BMI, supine systolic and diastolic blood pressure, and orthostatic hypotension.

4.2 Expectancy and Treatment Preference

Measures of expectancy will be obtained once at baseline by self-report. Participants will be asked their expectations of benefit from placebo and oral inosine treatment and their relative preference for placebo vs. oral inosine.

4.3 Efficacy Endpoints

The primary outcome of the trial is rate of change in MDS-UPDRS I-III total score over 24 months estimated in the ITT sample from a shared-baseline, random-slopes mixed model, censoring follow-up of subjects after initiation of dopaminergic therapy.

Secondary efficacy endpoints include the following:

- MDS-UPDRS scores: Parts I, II, III, IV, ambulation, and patient-reported sections (Parts IB and II),
- Striatal DAT binding by DaTscan™,
- Neuro-QOL module scores: Lower Extremity Function, Upper Extremity Function, Anxiety, Depression, Positive Affect and Well-being, Cognitive Function, Fatigue, Sleep Disturbance, Emotional and Behavioral Dyscontrol, Communication, Ability to Participate in Social Roles and Activities, Stigma, and Satisfaction with Social Roles and Activities,
- Montreal Cognitive Assessment (MoCA) score to assess cognition,
- Parkinson's Disease Questionnaire - 39 item version (PDQ-39) score to assess quality of life,
- Modified Schwab and England ADL score to assess functional disability,
- Orthostatic vital signs: supine to standing change in systolic blood pressure, diastolic blood pressure, and heart rate; orthostatic hypotension, and
- Medical intervention: levodopa equivalent daily dosage, disability warranting initiation of dopaminergic therapy.

Exploratory efficacy endpoints include: REM sleep behavior disorder (RBD) symptoms and symptom severity, quantitative measures derived from tasks completed using the Smart*4*SURE implementation of the mPower smartphone app.

4.4 Intervention and Pharmacodynamic Endpoints

Intervention intensity will be assessed by capsule count. Pharmacodynamic effect of oral inosine on serum urate will be assessed by direct measurement. Effects of oral inosine on urate excretion will be assessed by comparing measurements of serum urate and 24-hour urine urate excretion.

4.5 Safety Endpoints

The following safety endpoints will be evaluated:

- Overall TEAE and serious TEAE incidence rate,
- Time to first serious TEAE,
- Proportion of participants experiencing and number of unique events of each type of TEAE and serious TEAE classified by MedDRA system organ class and preferred term,
- Proportion of participants experiencing and number of unique events of each type of TEAE within the group of events specified as of special interest,
- Proportion of participants experiencing and number of unique events of TEAEs classified by seriousness, severity, relatedness to study drug, action taken with study drug, and outcome, summarized across all MedDRA terms,
- Mean change from baseline in safety labs: complete blood count, leukocyte differential, electrolytes, renal panel, liver function tests, lipid profile, thyroid stimulating hormone, and urinalysis of spot and 24-hr samples,

- Mean change from baseline in blood pressure (supine and standing), heart rate (supine and standing), orthostasis (change from supine to standing), and body weight,
- Mean change from baseline in ECG parameters,

Reported proportions will use as their denominator all participants in the Safety and Tolerability sample (see Section 5.2 below).

4.6 Tolerability Endpoint

Tolerance of treatment by an individual participant will be defined as remaining on-study and on his/her assigned treatment without one or more AE-associated dosage reductions lasting more than 4 weeks cumulative. Tolerance will only be assessed through 24 Oct 2018, the day before clinical sites were notified of early study closeout, as discontinuation of study drug after that date may reflect a change in participants' or investigators' equipoise. Tolerability of oral inosine titrated to achieve trough serum urate in the range 7.1 to 8.0 mg/dL will be defined as the proportion of all participants in the active arm who are tolerant of the treatment at 12 weeks (short-term), 12 months (medium-term), and 24 months (long-term). Oral inosine titrated to achieve trough serum urate in the range 7.1 to 8.0 mg/dL will be declared tolerable for a given duration of treatment if the proportion who are tolerant is significantly greater than 50%.

4.7 Blinding

The presumed treatment group to which a participant was assigned, a respondent's confidence in that presumption, and the reason for that presumption will be collected from participants, site investigators, and coordinators at the 6 week visit and at the end of treatment. The proportion guessing correctly or incorrectly and the odds ratio of guessing the true treatment vs. the wrong treatment will be evaluated for each class of respondent, both overall and among the subset of respondents who assert being at least somewhat sure of their guess.

4.8 Final Disposition

The proportion of participants who permanently discontinue study drug, discontinue clinic visits, fully withdraw consent, or are lost to follow-up prior to planned study completionwill be summarized overall and by treatment group for each period, both overall and stratified by occurrence before vs. on or after 25 Oct 2018, when clinical sites were notified of early study closeout. The probability that each participant has idiopathic PD based on the site investigator's primary diagnostic assessment at the participant's final visit will be summarized overall and by treatment group.

5. Measurement Definitions

5.1 Movement Disorders Society Unified PD Rating Scale (MDS-UPDRS)

The MDS-UPDRS[138] will serve as the primary outcome variable of the study and will be conducted at all standard visits beginning with SC2. The MDS-UPDRS was designed by movement disorders experts to address weaknesses of the original UPDRS (e.g., by adding questions on constipation and sialorrhea) while preserving its overall format.

The instrument is divided into four parts:

- **Part I (non-motor experiences of daily living)**, comprising
    - Part IA concerning behaviors that are assessed by the Site Investigator with all pertinent information from participants and caregivers
    - Part IB that is completed by the participant with or without the aid of the caregiver, but independently of the Site Investigator.
- **Part II (motor experiences of daily living)**, designed to be a self-administered questionnaire like Part IB, but similarly can be reviewed by the Site Investigator to ensure completeness and clarity.
- **Part III (motor examination)** has instructions for the rater to give or demonstrate to the participant; it is completed by the clinician rater.
- **Part IV (motor complications)** with instructions for the rater and also instructions to be read to the participant. This part integrates participant-derived information with the rater's clinical observations and judgments and is completed by the rater.

The full MDS-UPDRS has sixty-five items, each assessed on a 5-point Likert scale ranging from 0 to 4 with 0=none, 1=slight, 2=mild, 3=moderate, 4=severe. Total scores for Parts I, II, III, and IV and for Parts I through III collectively are calculated as simple sums of component items with mean imputation by Part if no more than 1, 2, 7, or 0 items is missing for Parts I through IV, respectively.[16] Two additional summary scores will also be constructed: ambulatory capacity (sum of 5 MDS-UPDRS questions: walking and balance [question 2.12], freezing [q. 2.13], gait [q. 3.10], freezing of gait [q. 3.11], and postural stability [q. 3.12]) and patient-reported symptoms (sum of Parts IB and II). Higher scores imply worse symptoms.

Participants will self-administer Parts IB and II and will review responses for accuracy and clarity with the Site Investigator or Coordinator. Parts IA, III and IV will be conducted by the Site Investigator. Parts I, II, and III will be conducted at study visits as indicated on the Schedule of Activities (Sec. 2.8). Part IV will be conducted at visits where MDS-UPDRS Parts I-III are conducted but only for participants who have started on symptomatic therapy after the Baseline visit.

Use of MDS-UPDRS is responsive to core instrument recommendations for the Quality of Life subdomain of the NINDS CDEs for PD and to FDA guidance encouraging use of patient-reported outcomes (PROs) as a substantial portion of the responses are patient-reported. The same Site Investigator should assess all subjects on parts IA and III of the MDS-UPDRS at all study visits.

5.2 Modified Schwab & England Activities of Daily Living Scale

The Schwab & England scale[181,182] is a Site Investigator and subject assessment of the subject's level of independence. The subject will be scored on a percentage scale reflective of his/her ability to perform acts of daily living. Printed scores with associated descriptors range from 0% to 100% in increments of 5%, with higher percentages associated with more independence. A score of 0% implies "vegetative functions such as swallowing, bladder and bowel functions are not functioning; bedridden". A score of 100% implies "subject has full ability and is completely independent; essentially normal". This joint participant/Site Investigator assessment will be conducted periodically at the visits indicated in Sec. 2.8.

5.3 Parkinson's Disease Questionnaire - 39 item version (PDQ-39) scale

The PDQ-39 asks 39 questions organized over eight domains (scales): mobility (10 items), activities of daily living (6 items), emotional well-being (6 items), stigma (4 items), social support (3 items), cognition (4 items), communication (3 items), and bodily discomfort (3 items). Each item has five possible ordinal responses, from never to always, depending on frequency of the symptom over the preceding month. The eight scales' scores are generated by Likert's method of summated ratings and then transformed to a single figure that ranges from 0 to 100. Higher scores are associated with more symptoms. The PDQ-39 is the most widely used health related-QoL instrument in PD, and is considered to have generally good psychometric properties and content validity.[183,184] Use of PDQ-39 is responsive to core instrument recommendations for the Quality of Life subdomain of the NINDS CDEs for PD, and to FDA guidance encouraging use of PROs. This assessment will be collected from subjects periodically at the visits indicated in Sec. 2.8.

5.4 Modified Hoehn and Yahr Scale

The Modified Hoehn and Yahr Scale[181, 185] is a 6-level PD staging instrument of motor manifestations and disability. It is an ordinal scale, scored by the Site Investigator. Scores range from 0 to 5 with higher scores associated with more motor symptoms and disability. Stage 0 is "no signs of disease", stage 1 is "unilateral disease', stage 1.5 is "unilateral disease with axial involvement", stage 2 is "bilateral disease, without balance impairment", stage 2.5 is "bilateral disease, with recovery on the pull test", stage 3 is "mild to moderate bilateral disease; needs assistance to prevent falling on pull test", stage 4 is "severe disability, but still able to walk or stand unassisted" and stage 5 is "wheelchair bound or bedridden unless aided." This Site Investigator assessment will be conducted periodically at the visits indicated in Sec. 2.8.

5.5 Assess Need for Dopaminergic Therapy

At each visit beginning with Baseline Visit, the Site Investigator will assess the subject's need for dopaminergic therapy. (See Sec. 5.3.3.1 for definition of dopaminergic therapy.) A questionnaire will be used to facilitate the Site Investigator's decision. As in the DATATOP[9,10] and PRECEPT[11] trials, this will be based on PD disability posing a threat to the subject's current occupational status, current abilities (potential capacities) related to occupational matters, to handle routine personal finances and domestic responsibilities, and activities of daily living.

Subjects who are judged to require dopaminergic therapy at Baseline or are thought likely to need therapy within the 3 months after Baseline, will be excluded from participation in this study. Subjects who are judged to require dopaminergic therapy after starting study drug will continue in the study after anti-parkinsonian therapy is instituted.

5.6 Mini-Mental State Examination (MMSE)

The Mini-Mental State Examination[186] is a 30-point ordinal scale that is widely used for the evaluation of degenerative dementia in patients with a variety of neurologic and psychiatric disorders, and is designated an NINDS CDE for PD. The MMSE includes 11 questions which evaluate orientation (10 points), immediate recall (3 points), attention (5 points), delayed recall (3 points), naming (2 points), repetition (1 point), 3-stage command (3 points), reading (1 point), copying (1 point) and writing (1 point). The test is referred to as "mini" because it focuses only on the cognitive aspects of mental functions and does not include questions related to mood,

abnormal mental experiences, or the form of thinking. The total score ranges from 0 to 30, with higher scores signifying better cognition. A total score of 23 or less is associated with varying severities of cognitive impairment.

Subjects will complete the MMSE at the Screening Visit 2. Participants with an MMSE score less than 25 will be excluded from participation in the study.

5.7 Montreal Cognitive Assessment (MoCA)

The MoCA (Nasreddine et al. 2005) consists of 8 clinician-administered cognitive tasks designed to screen for mild cognitive impairment. The MoCA assesses attention and concentration, executive functions, memory, language, visuoconstructional skills, conceptual thinking, calculations, and orientation. The MoCA was developed to be more sensitive than the MMSE to patients presenting with mild cognitive complaint and may be less prone to a ceiling effect (Zadikoff et al. 2008). One point is awarded for correct completion of each item of the visuospatial/executive function task (5 items), naming task (3 items), digit vigilance and tapping items of the attention task (3 items), the sentence repetition items of the language task (2 items), abstraction task (2 items), delayed recall task (5 items), and orientation task (6 items). One point is awarded for naming 11 or more words during the fluency item of the language task. Zero (none correct) to 3 (4 or more correct) points are awarded based on the number of correct subtractions by 7 starting at 100 in the attention task. One point is awarded if the participant has 12 years or less of education unless the score is already 30. Scores for each task are summed for a total score (range 0 to 30) with higher scores indicating greater cognitive capacity. The MoCA will be administered by Site Investigators at Baseline, Week 12, Month 12, Month 24 and Month 27 or the Discontinuation of In-person Follow-up Visit for subjects unwilling to continue in-person visits.

5.8 Quality of Life in Neurological Disorders (Neuro-QOL)

Neuro-QOL is a set of patient-reported outcome (PRO) measures that assess health-related quality of life (HRQoL) of people with neurological disorders.[190,191] It facilitates comparisons between diseases and within individual patients over time. Developed through a collaborative NINDS-sponsored research initiative, Neuro-QOL is an NINDS CDE and has been validated in multiple patient populations including PD.[190,192] It comprises 17 domains of HRQL covering physical, psychological and social health. It has 13 item banks, 3 item pools and 1 stand-alone scale. There are short forms for each item bank which each form containing 8 to 9 items. Domains tested include anxiety, cognitive function, communication, depression, emotional and behavioral dyscontrol, fatigue, lower extremity function- mobility, positive affect and well-being, stigma, upper extremity function- fine motor and ADL, sleep disturbance, satisfaction with social roles and activities, and ability to participate in social roles and activities. Higher raw scores are associated with more of the concept being measured. In a PD population, Neuro-QoL measures including its short forms have demonstrated high internal consistency, with acceptable test-retest reliability and support for convergent validity with PD specific measures including PDQ-39 and MDS-UPDRS.[190] An instrument comprising multiple short form domains will be employed before and after the study drug treatment period (at SC2 and SV) while the depression domain will be employed on its own at additional visits during the treatment period, as indicated in Sec. 6.3.

5.9 Initiation of Dopaminergic Therapy

For the purposes of this trial, the date at which a participant initiates dopaminergic therapy will be calculated as the first day that the participant either (a) initiates use of one of the following medications as labeled in the WHO Drug Dictionary: amantadine, pramipexole, ropinirole, rotigotine, sinemet, trihexyphenidyl, benztropine, rasagiline, selegiline, safinamide, or levodopa at a LEDD greater than 2 mg/day, or (b) increases the LEDD of one of the following medications over the dosage taken at baseline by more than 2 mg/day: rasagiline, selegiline, safinamide.

5.10 Primary Diagnosis Assessment

The Primary Diagnosis form captures, in the Site Investigator's opinion, a current percentile probability the subject has idiopathic Parkinson disease based on available information. Ranges include: 90-100%; 50-89%, 10-49% and 0-9%. In addition the Site Investigator selects the most likely primary diagnosis from a listing that includes idiopathic PD, many other neurological disorders, and the option of no neurological disorder. To correlate with the MDS-UPDRS, this percentile probability and most likely diagnosis will be captured at Month 27 or the Discontinuation of In-person Follow-up Visit for subjects unwilling to continue in-person visits.

5.11 Diagnostic Features Assessment

The Diagnostic Features form is a companion to the Primary Diagnosis form. It is a review by the Site Investigator of factors that do and do not suggest a diagnosis of Parkinson disease. This assessment is completed at Month 27 or the Discontinuation of In-person Follow-up Visit for subjects unwilling to continue in-person visits.

5.12 Dopamine Transporter (DAT) Neuroimaging

A radionuclide-labeled dopamine transporter (DAT) ligand, and specifically DaTscan™ (^{123}I-ioflupane injection), is approved by the FDA for striatal DAT visualization using single photon emission computed tomography (SPECT) brain imaging to assist in the evaluation of adult patients with suspected parkinsonian syndromes. DaTscan™ is approved to help differentiate essential tremor from tremor due to parkinsonian syndromes (including idiopathic PD, multiple system atrophy and progressive supranuclear palsy) and not for the diagnosis of PD among parkinsonian syndromes. Nevertheless, DAT brain scans (using any of several radioligands) in clinical research have consistently identified a small but substantial (~10%) portion of subjects who are enrolled in clinical trials based on an expert clinician diagnosis of probable early PD but who turn out to be unlikely to have PD.[30,140,143,145-147]

DAT scans will be performed prior to the baseline visit, and will generally be conducted as the final screening evaluation (DS1) at a certified neuroimaging center at or near the clinical site. A determination of whether DaTscan™ imaging supports a diagnosis of PD and therefore study eligibility will be made by the study imaging core. Its experienced nuclear medicine specialists (trained in the visual read method appropriate for DaTscan™) will perform the qualitative eligibility assessment. Each scan will be assessed independently by at least two readers as described in the Imaging Core Charter.

An additional follow-up DaTscan™ imaging study (DS2) will be performed one to two months following the final study clinic visit (V10, or its equivalent), and before the final safety visit (SV), for active SURE-PD3 subjects who have consented to participate in a serial DAT scan substudy, who have been on study drug through at least the 1-year visit (V06), and who are not

using or expected to use any of the following medications within 90 days prior to DS2: modafinil, armodafinil, metoclopramide, alpha-methyldopa, methylphenidate, reserpine, or amphetamine derivative. The substudy will quantify changes in DAT binding between the pre-study drug exposure and post-study drug exposure timepoints.

5.13 Orthostasis

Orthostatic hypotension will be defined as a maximum decrease of 20 mm Hg or greater in systolic blood pressure (SBP) or a maximum decrease of 10 mm Hg or greater in diastolic blood pressure (DBP) when moving from a supine to a standing position, with measurements completed 1 min and 3 min after standing.

5.14 Levodopa Equivalent Daily Dosage

The levodopa equivalent daily dosage (LEDD) will be calculated using data from the concomitant medications log. The conversion from dopamergic drugs other than carbidopa-levodopa will follow the recommendations by Tomlinson et al. (2010) with the following additions for anti-cholinergic medications and more recently approved drugs.

- Trihexyphenidyl (Artane™) will be converted at 25 mg levodopa equivalent per mg trihexyphenidyl based on equivalent symptomatic effect of 8 mg/day trihexyphenidyl and 200 mg/day amantadine reported in Parkes et al. (1974);
- Benztropine (Cogentin™) will be converted at 50 mg levodopa equivalent per mg benztropine based on the ratio of the recommended range of total daily dosages for treating Parkinson disease patients between benztropine and trihexyphenidyl;
- Extended release formations of carbidopa-levodopa (Rytary™) will be converted at 60% of their levodopa content based on mean post-baseline daily dosages reported in Hauser et al. (2013);
- Extended release formations of ropinorole (RequipXL™) will be converted at 20 mg levodopa equivalent per mg ropinirole based on equivalent maximum recommended daily dosage;
- Safinamide (Xadago™) will be converted at 2 mg levodopa equivalent per mg safinamide based on a retrospective study by Mancini et al. (2018).

The full calculation is as follows:

1. At each time any oral levodopa is taken per day, calculate the immediate release levodopa (e.g., Sinemet, Parcopa) dosage (mg) based on the number of tablets taken per dose and the dosage of levodopa in each tablet
2. At each time any oral levodopa is taken per day, calculate the controlled release levodopa (e.g., Sinemet CR) dosage (mg) and multiply by 0.75 to account for loss of bioavailability
3. At each time any oral levodopa is taken per day, calculate the extended release levodopa (e.g., Rytary) dosage (mg) and multiply by 0.60 to account for loss of bioavailability
4. At each time any oral levodopa is taken per day, calculate the total daily oral levodopa dosage by summing together the oral levodopa dosage from immediate release, controlled release, and extended release formulations (quantities 1 through 3)

5. At each time any oral levodopa is taken per day, if entacapone but not tolcapone is taken at the same time, either separately or as a combination drug (e.g., Stalevo), then multiply the total levodopa dosage at that time (quantity 4) by 0.33
6. At each time any oral levodopa is taken per day, if tolcapone (e.g., Tasmar) is taken at the same time, then multiply the total levodopa dosage at that time (quantity 4) by 0.50
7. Calculated the COMT-adjusted total daily levodopa equivalent dosage associated with oral levodopa as the sum over all time points of quantities 4 through 6
8. If enteral levodopa is taken without any COMT inhibitor, then multiply the total enteral levodopa (e.g., Duopa) daily dosage (mg) by 1.11
9. If enteral levodopa is taken with entacapone but not tolcapone, then multiply the total daily enteral levodopa dosage (mg) by 1.48
10. If enteral levodopa is taken with tolcapone, then multiply the total daily enteral levodopa dosage (mg) by 1.67
11. If immediate release or modified release pramipexole (e.g., Mirapex) is taken, multiply the daily dosage (mg/day of the dihydrochloride monohydrate salt) by 100
12. If immediate release or extended release ropinirole (e.g., Requip, Requip XL) is taken, multiply the dosage (mg/day) by 20
13. If a rotigotine patch (e.g., Neupro) is used, multiply the dosage (mg/day) by 30
14. If oral selegiline (e.g., Eldepryl) is taken, multiply the dosage (mg/day) by 10
15. If sublingual selegiline (e.g., Zelapar) is taken, multiply the dosage (mg/day) by 80
16. If oral rasagiline (e.g., Azilect) is taken, multiply the dosage (mg/day) by 100
17. If oral safinamide (e.g., Xadago) is taken, multiply the dosage (mg/day) by 2
18. If oral amantadine (e.g., Symmetrel) is taken, multiply the dosage (mg/day) by 1
19. If an injection or infusion of apomorphine (e.g., Apokyn) is taken, multiply the dosage (mg/day) by 10
20. If oral trihexyphenidyl (e.g., Artane) is taken, multiply the dosage (mg/day) by 25
21. If oral benztropine (e.g., Cogentin) is taken, multiply the dosage (mg/day) by 50
22. Calculate total LEDD as the sum of quantities 7 through 21.

5.15 Blindedness Evaluation

At week 6 and either month 24 (V10), V10-equivalent visit due early study closure, an Unscheduled Visit due to study drug discontinuation, or a Discontinuation of In-person Follow-up Visit (whichever of the 4 visits comes first), the Site Investigator, Coordinator, and subject will complete a blindedness evaluation in which each is asked to give his/her independent impression of the subject’s treatment assignment and the primary and secondary reasons for this opinion. Participants' responses will not be available to the site Investigator or Coordinator when they make their assessments.

5.16 Genetics

DNA samples will be sequenced using whole genome sequencing (WGS) supported by the NIH's Accelerating Medicines Partnership for Parkinson's Disease (AMP-PD; https://www.nih.gov/research-training/accelerating-medicines-partnership-amp/parkinsons-disease) based on established methods (e.g, Allen *et al.* 2016), with variants scored under the AMP program (e.g, http://www.type2diabetesgenetics.org/variantSearch/variantSearchWF; Butkiewicz et al. 2018). The following single nucleotide polymorphisms (SNPs) or repeat polymorphisms may be investigated amongst emerging genetic variants contributing to progression of PD across motor, non-motor and holistic domains (Iwaki et al. 2019), with simple additive, recessive, dominant, interactive, haplotypic, polygenic or other models to be determined based on current understanding of genetic risk determinants upon availability of WGS expected by early 2020. Genetic predictors of PD progression that may be specific for urate-targeting inosine treatment, and those that are more likely generalized will be considered:

- Potential predictors of PD progression, specific to urate/inosine intervention; e.g.:
  - rs1109303 (*INPP5K*), stratifying for TT vs G carriers (Nazeri et al. 2015)
  - s6855911, rs7442295, rs16890979 (*SLC2A9*; encoding urate transporter gene *Glut9*), adjusting for and stratifying by a composite score of proportional to the # of their minor alleles) (Simon et al. 2014)
  - rs2231142 (*ABCG2* encoding urate transporter gene), stratifying for CC vs A carriers (Matsuo et al. 2015)
- Potential predictors of PD progression, generalized; e.g.:
  - rs76904798 (near/5' to *LRRK2*) (Iwaki et al. 2019)
  - D4S3481 (*SNCA*-Rep1; complex repeat polymorphism upstream of *SNCA*) (Ritz et al. 2012)
  - rs76763715 (*GBA* N370S; but MAF 0.007) (Davis et al. 2016)
  - rs9298897 and rs17710829, alone and in a modeled interaction (Latourelle et al. 2017)
  - rs2230288 (*GBA*), rs17649553 (*MAPT*), apolipoprotein E4 (*APOE4*), COMT val^{158}met (*COMT*) and BDNF val^{66}met (*BDNF*) alone and in additive model for cognitive progression outcomes (Iwaki et al. 2019; Caspell-Garcia et al. 2017).

5.17 Exploratory Assessments

Three brief, self-administered questionnaires will be included to explore whether readily ascertained historical factors modify or otherwise interact with inosine effects.

- **REM Sleep Behavior Disorder (RBD) Single-Question** Screen (RBD1Q)– A single "yes-no" question on dream-enactment behaviour will be asked at SC1, BL, and 3 mo, 12 mo, 24 mo, and 27 mo visits.
- **PD Risk Factors Questionnaire** – A self-administered questionnaire assessing exposures and experiences linked to the risk of PD will be collected at BL. The questionnaire was developed by the PSG and derived from NINDS CDEs.
- **PD - Expectancy Questionnaire** – A self-administered questionnaire assessing expectations of study drug effect will be collected at BL.

- **Smart*4*SURE assessments** – A set of smartphone metrics of parkinsonian features adapted from the mPower app developed as a sub-study of SURE-PD3.

6. Statistical Methodology

6.1 General Considerations

6.1.1 Statistical Software

All statistical analyses will be performed using SAS (SAS Institute, NC, USA) or R (R Foundation for Statistical Computing, Vienna, Austria).

6.1.2 Summary Statistics

Data will be summarized with respect to disposition, demographics, pre-treatment characteristics, safety outcomes, tolerability, and efficacy outcomes. Summary statistics for continuous variables will include the number of subjects, the mean, median, standard deviation, and range. For categorical data, summaries will include counts and percentages.

6.1.3 Precision

Results will generally be reported to 3 significant figures. Percentages will generally be reported to 0.1 percentage points. P-values will be reported to two digits when greater than or equal to 0.095, to three digits when greater than or equal to 0.00095 and less than 0.095, and as <0.001 for all smaller values.

6.1.4 Administration

A test set of tables and figures specified in the SAP will be produced prior to breaking the blind using a dummy randomization schedule. The SAP will be finalized and must be approved by the Steering Committee prior to the final lock of the trial data and breaking of the blind.

6.2 Analysis Samples

The following analysis samples will be used for testing effectiveness, efficacy, safety, and tolerability endpoints:

- Intent-to-treat (ITT) Sample: Participants who are randomized, classified according to their randomized treatment assignment. Participants determined to have been ineligible prior to randomization, participants who never initiate study drug, and observations made after premature permanent discontinuation of study drug are included in this sample.
- As-treated (AT) Sample: Participants who are eligible, randomized, and take at least one dose of study drug. If a participant permanently discontinues study drug, observations made during period 1 will be censored following the first visit after study drug discontinuation. Separate analyses will classify participants in the AT sample either according to the actual treatment received or according to the average post-baseline serum urate concentration achieved. Participants randomized to inosine who report not taking inosine will be reclassified as placebo participants. Participants randomized to placebo who report taking inosine or whose serum urate during treatment averages 6.5 mg/dL or greater will be reclassified as inosine participants.

- Safety and Tolerability (ST) Sample: Participants who are randomized and take at least one dose of study drug, classified according to the actual treatment received. Participants determined to have been ineligible prior to randomization and observations made after premature permanent discontinuation of study drug are included in this sample.

The primary efficacy analyses will use the ITT sample to best estimate the expected effectiveness of urate-elevating oral inosine supplementation in clinical practice, recognizing that compliance in clinical use may differ from compliance in the clinical trial. Secondary analyses of efficacy outcomes will use the AT sample to best estimate efficacy of inosine supplementation as administered and serum urate elevation as achieved. Analyses of safety and tolerance will use the ST sample.

6.3 Baseline Comparison

While recognizing that any difference observed between treatment groups is axiomatically due to chance, nominal p-values will be calculated as a measure of the magnitude of difference for each baseline characteristic summarized using Fisher's exact tests for nominal variables, exact Cochran-Armitage trend tests for ordinal variables, and two-sample t-tests for approximately continuous variables.

6.4 Sample Size Review

Estimates of variance components and censoring and drop-out rates required for sample size calculation will be reviewed periodically using blinded data. No formal sample size re-estimation is proposed.

6.5 Interim Analysis

Interim analyses for efficacy or futility will be performed by the unblinded statistician using an information-based group-sequential design. Interim analyses will be performed after approximately 2000 and 5000 person-months of period 1 follow-up are completed, roughly one-third and three-quarters of total anticipated period 1 follow-up. Non-binding early stopping for efficacy will be proposed if the active treatment group is superior to placebo for the primary efficacy outcome based on a one-sided p-value of 0.001 or less. Other criteria used in deciding whether to stop early for efficacy will include evaluation of safety and secondary efficacy outcomes. We propose stringent early stopping criteria for efficacy in order to ensure that evidence for efficacy is unambiguous if the study is stopped early. Non-binding early stopping for futility will be proposed based on a beta-spending rule quadratic in information time. If there were no benefit from treatment, then the trial would have a 51% probability of early stopping for futility based on this rule. As with the efficacy evaluation, the decision to stop early for futility will also include evaluation of feasibility, safety, and secondary efficacy outcomes. Because the efficacy and futility stopping rules are non-binding, we conservatively assume that each look is cumulative for alpha and ignore possible findings of futility when calculating the overall type I error.

6.6 Efficacy Analysis

6.6.1 Primary Analysis of the Primary Efficacy Endpoint

The primary analysis of the primary efficacy endpoint will estimate the effectiveness to slow PD disease progression of oral inosine titrated to achieve serum urate in the range 7.1 to 8.0 mg/dL. The analysis will use the ITT sample to estimate of rate of change of MDS-UPDRS I-III total scores during period 1 in a random slopes model with shared baseline, censoring observations made subsequent to initiation of dopaminergic therapy. The model will include fixed terms for time, treatment x time, and sex, baseline MAO-B inhibitor use, baseline Schwab and England ADL score, and their interactions with time and random site- and participant-specific intercepts and slopes, each with unstructured covariance. Use of a shared baseline adjusts for baseline MDS-UPDRS score[14,15] in addition to the adjustment for sex, baseline MAO-B inhibitor use, and baseline modified Schwab and England ADL score. Inference of benefit from serum urate elevation will be made by testing whether the treatment x time interaction term is significantly less than zero (i.e., slower progression among subjects randomized to the active arm) using a two-sided test at $p < 0.046$ for a cumulative two-sided alpha = 0.05.

This estimate of the effectiveness of serum urate elevation on MDS-UPDRS will be unbiased if observed trajectories are predictive of MDS-UPDRS assessments that are missing due to loss to follow-up or are censored due to initiation of dopaminergic therapy. Data from the SURE-PD trial suggest good conformance of UPDRS I-III total score trajectories to the model assumptions. No non-linearity in treatment effects on observed UPDRS scores was found and models that included quadratic random effects fit worse by Akaike and Bayesian information criteria. Empirical Bayes estimates of 24-month UPDRS scores assuming linear trajectories were all within range of the instrument (range 7.4 to 97.4) even with 85% of the sample being censored prior to the final 24-month observation (median 12 months). Conditional residuals were normally distributed and homoscedastic. Similar assessments and influence statistics will be evaluated in judging the adequacy of the proposed primary analysis in this trial.

6.6.2 Secondary Analyses of the Primary Efficacy Endpoint

Several secondary analyses will be investigated to assess sensitivity of our estimates of treatment effect to alternative modeling assumptions.

- Shared-baseline repeated-measures model: The fixed effects of time and its interactions will be replaced by visit-specific fixed effects. The interaction between fixed effects for treatment group and visit will be restricted to post-baseline visits by including a numeric indicator variable (0 pre-treatment, 1 post-treatment) in the interaction. The participant-specific random intercepts and slopes will be replaced with participant-level unstructured covariance among repeated measures.
- AT sample by treatment group: The primary efficacy model and the secondary repeated-measures model will be applied to the AT sample, replacing the treatment group as randomized by the treatment actually received.
- AT sample by serum urate achieved: The primary efficacy model and the secondary repeated-measures model will be applied to the AT sample, replacing the treatment group as randomized by the average post-baseline serum urate level achieved.

- AT sample by change in serum urate: The primary efficacy model and the secondary repeated-measures model will be applied to the AT sample, replacing the treatment group as randomized by the change from average pre-treatment to average post-baseline serum urate level achieved.
- AT sample by average daily inosine capsule count: The primary efficacy model and the secondary repeated-measures model will be applied to the AT sample, replacing the treatment group as randomized by the daily inosine treatment actually received.
- AT sample excluding participants who initiated dopaminergic therapy by their 12-week (V03) visit: The primary efficacy model and the secondary repeated-measures model will be applied to the subset of the AT sample who initiated dopaminergic therapy by their 12-week visit (V03).
- LEDD-adjusted analysis: The primary efficacy model, the secondary repeated-measures model, and the secondary analyses of the AT sample listed above will be augmented by including observations after initiation of dopaminergic medication and adjusting for time-dependent LEDD.

6.6.3 Secondary Efficacy Endpoints

Continuous secondary efficacy endpoints will all be analyzed using the ITT and AT samples and the primary and secondary efficacy analysis described above. Variables that are strongly right-skewed will be log-transformed prior to analysis, and estimates will be back-transformed for reporting.

The binary indicator of orthostatic hypotension will be analyzed in an equivalent generalized mixed model treating the indicator variable as a Bernoulli random variable with logit link.

6.6.4 Disability Warranting Dopaminergic Therapy

Treatment differences in time to disability warranting dopaminergic therapy during period 1 will be tested by a Kaplan-Meier logrank test with a two-sided alpha = 0.05. Participants lost to follow-up will be censored. Note that subjects will continue to be followed even if they discontinue study drug.

The following secondary analyses of time to disability warranting dopaminergic therapy will be pursued:

- Subjects lost to follow-up will be classified as initiating dopaminergic therapy at the time of their withdrawal.
- Cox regression will be used to compare treatments with adjustment for baseline modified Schwab and England ADL score.

6.6.5 Combined Function and Treatment

A rank-based test of a single outcome combining function as measured by change in MDS-UPDRS I-III total scores and time to disability warranting dopaminergic therapy will be constructed paralleling the Combined Assessment of Function and Survival (CAFS) score methodology developed for ALS.[17] MDS-UPDRS I-III total scores will substitute for ALSFRS-R total scores and time to disability warranting dopaminergic therapy will substitute for time to mortality. The test consists of calculating a rank-sum score for each individual relative to pair-

wise comparisons with all other subjects. Subjects are ranked according to time to disability warranting dopaminergic therapy when that is observed for both members of a pair or when one is censored after the observed event time for the other. Pairs that cannot be ranked by time to disability warranting dopaminergic therapy are ranked by absolute change from baseline in MDS-UPDRS I-III total score at the maximum follow-up time at which both subjects have an observation. Inference is drawn by calculating a U-statistic from the rank-sum scores.

6.6.6 Subgroup Analyses

The following subgroups will be considered: sex, race (classified as Asian, Black or African American, Caucasian, or other, including multiracial; or as Caucasian vs. non-Caucasian if fewer than 10% of our sample is non-Caucasian), ethnicity (classified as Hispanic or Latino vs. non-Hispanic), use of MAO-B inhibitors at baseline, and age (both categorized as <65 years vs. ≥65 years and continuous). For each subgroup, the potential for differential benefit from serum urate elevation will be tested by including subgroup, subgroup x time, and subgroup x time x treatment interaction terms into the primary random-slopes model. A significant subgroup x time x treatment 3-way interaction in combination with significantly slower progression among members of a subgroup randomized to serum urate elevation vs. members of the same subgroup randomized to placebo will be taken as evidence of differential benefit. For sex, we will also test for sex-specific effects in selected secondary analyses of efficacy, disability warranting dopaminergic therapy, and symptomatic effects and in safety analyses related to kidney stone risk. For safety analyses, we will also summarize and compare eligible vs. the few ineligible participants in the ST sample.

6.6.7 Pharmacogenetics

The above listed and any subsequently designated genetic variants will be tested as potential modifiers of treatment effect. For each genetic variant, the potential for differential benefit from serum urate elevation will be tested by including genotype, genotype x time, and genotype x time x treatment interaction terms into the primary random-slopes model. A significant genotype x time x treatment 3-way interaction in combination with significantly slower progression among members of a genotype randomized to serum urate elevation vs. participants with the same genotype randomized to placebo will be taken as evidence of differential benefit.

6.6.8 Symptomatic Effects

The presence of symptomatic effects will be tested using a change-point model constructed as a partial linear spline over time with knots at 12 weeks and 24 months. Both fixed and subject-specific random terms for intercept, slope from baseline to 12 weeks, slope from 12 weeks to 24 months, and slope from 24 to 27 months (or from V10-equivalent visit on study drug to SV) will be included. Unstructured covariance will be assumed among the random effects (10 terms). The AT sample will be analyzed. Significantly smaller slopes during wash-in or larger slopes during wash-out among subjects treated with inosine using one-sided testing at alpha = 0.025 will be interpreted as symptomatic effects. If symptomatic effects are not found on that basis, the absence of symptomatic effects will be judged based on a non-inferiority test using a non-inferiority bound of 6.3 / 3 = 2.1 points/month in the direction of a symptomatic effect and using one-sided testing at alpha = 0.05.

6.6.9 Disease Modification

A three-part test of time to disability warranting dopaminergic therapy will be used to evaluate whether serum urate elevation is disease-modifying[18-20]. Inference of disease modification would be supported if (a) a Kaplan-Meier logrank test of time to disability warranting dopaminergic therapy during period 1 favors serum urate elevation, (b) the proportion of subjects who developed disability warranting dopaminergic therapy remains significantly lower among those randomized to serum urate elevation vs. placebo at the end of period 2 by Chi-square test, and (c) the time-dependent hazard ratio for time to disability warranting dopaminergic therapy during period 2 is non-inferior to a rate that would lead to equivalence to placebo by 3 months by Cox regression. Inference from each of the three component tests will be evaluated sequentially. Difference in time to disability warranting dopaminergic therapy will be evaluated first. If significant by logrank test with two-tailed alpha = 0.05, then the proportion not yet requiring dopaminergic therapy at the end of period 2 will be tested second. If significant by Chi-square test with two-tailed alpha = 0.05, then non-inferiority of the time-dependent hazard ratio will be tested by Cox regression based on a one-tailed test at alpha = 0.05.

6.6.10 Multiplicity Adjustments

With two interim efficacy analyses each testing for benefit of serum urate elevation at a one-sided $p < 0.001$, the final primary analysis tested using a two-sided $p < 0.046$ ensures that the overall type I error rate is under 5%. Results from analysis of secondary efficacy endpoints and subgroup analyses will report nominal, comparison-wise p-values, recognizing that the totality of results will be evaluated in judging the potential efficacy of serum urate elevation.

6.6.11 Missing Data

Baseline values for efficacy endpoints will be determined from the last non-missing data collected prior to the first dose of study medication. The planned mixed model yields estimates that are unbiased conditional on the observed scores under a missing at random assumption. In addition, a secondary analysis of the primary endpoint will use placebo-based multiple imputation of missing data. Additional sensitivity analyses may be pursued to impute missing values or otherwise construct models for unobserved outcomes if more than 20% of participants are missing follow-up data for any reason.

6.7 Safety and Tolerability Analysis

6.7.1 Treatment-emergent Adverse Events

The incidence of TEAEs will be summarized by the number of events of a given classification experienced by participants in each treatment group and by the number and proportion of participants experiencing such an event in each treatment group in the ST sample. TEAEs will be summarized in aggregate across all MedDRA terms and separately by MedDRA system organ class and preferred term.

Aggregate summaries of TEAE grade will include characteristics of: (a) seriousness, (b) severity, (c) relatedness to study drug, (d) action taken with study drug, (e) action taken with study procedure, and (f) outcome. For each level of a given TEAE characteristic, summaries will include the number of events of a given classification and by the number and proportion of participants for which that level of a characteristic was the worst they experienced (treating any unknown characteristic as not worst).

6.7.2 Safety Labs

The absolute level and the absolute change from baseline for each safety laboratory assay will be summarized as means, standard deviations, medians, and ranges at each visit by treatment group. The proportion of participants with safety lab levels below the lower limit of normal or above the upper limit of normal will be summarized by treatment group by visit and at any post-baseline visit.

6.7.3 Vital Signs

The absolute level and the absolute change from baseline for vital signs will be summarized as means, standard deviations, medians, and ranges at each visit by treatment group.

6.7.4 Additional Continuous Safety Outcomes

The additional continuous safety outcomes of weight, ECG parameters, and their absolute change from baseline will be summarized as means, standard deviations, medians, and ranges at each visit by treatment group and tested for group differences in linear mixed models.

6.7.5 Tolerability

Participants will be classified as tolerant, intolerant, or censored with subjects lost to follow-up or withdrawing consent prior to administrative termination classified as intolerant. Tolerability will then be estimated from Kaplan-Meier product-limit methods with complementary log-log confidence bounds at 12 weeks, 12 months, and 24 months. Serum urate elevation will be declared tolerable at each time point if its one-sided lower 95% confidence bound is greater than 50%. We will explore the relationship between tolerance and use of dopaminergic therapy using Cox regression with use of dopaminergic therapy as a time-dependent covariate.

6.7.6 Alkalinization and Nephrolithiasis

The frequency of persistent acidic urine will be summarized by treatment group and tested by Fisher's exact test in the ST sample. The efficacy of the alkalinization protocol to increase urine pH will be summarized overall in the ST sample. The frequency of nephrolithiasis and the absolute rate of nephrolithiasis per time exposed to study drug will be summarized by treatment group. Risk factors for nephrolithiasis will be tested in a binary generalized linear model with log link and with log time on study drug as an offset. The following risk factors will be tested in a multivariable model:

- age,
- sex,
- treatment,
- uric acid excretion fraction at baseline estimated from serum urate and 24-hr urine urate,
- uric acid crystals/hpf prior to first nephrolithiasis
    - in most recent urinalysis,
    - average in post-randomization urinalyses,
    - maximum in any urinalysis,
- urine specific gravity prior to first nephrolithiasis
    - in most recent urinalysis,
    - average in all urinalyses,
    - maximum in any urinalysis,
- urine pH prior to first nephrolithiasis

- in most recent urinalysis,
- average in all urinalyses, and
- minimum in any urinalysis.

6.8 Other Analyses

6.8.1 Participant Disposition

The number of participants who were screened, randomized, completed scheduled follow up, and prematurely withdrew study participation will be summarized overall and by treatment group. Reasons for screen failure and for withdrawal from study will be presented.

6.8.2 Study Drug Exposure

The number of capsules of study drug prescribed will be summarized by visit and treatment group. Total study drug exposure will be calculated and will be summarized by treatment group. Compliance with study drug will be calculated as the number of doses taken divided by the scheduled number of doses taken prior to permanent discontinuation, expressed as a percentage.

6.8.3 Prior and Concomitant Medication Use

Concomitant medications taken during the study period will be listed for each participant, coded using the World Health Organization Drug Dictionary Enhanced. The percentage of participants taking each class of medication will be summarized overall and by treatment group.

6.8.4 Determinants of change in MDS-UPDRS

Potential determinants of change in MDS-UPDRS will be tested to assess their relationship to PD progression independent of serum urate elevation. This may motivate further adjustment of the primary model as additional secondary analyses and may also assist in the design of future trial. In addition to those noted above in 6.6.1, we will test the following:

- Prodromal PD features, either individually or as an index
    - Positive RBD screen at SC1 or BL
        - RBD characterization: At SC1, do those who are RBD-positive vs negative (2 vs. 1 on Q1 of RBD Sleep Question) have higher age- and sex-adjusted serum urate?
    - Positive hyposmia screen: Response of 1 on Q9a of the PD Risk Factor Questionnaire
        - Hyposmia characterization: Do hyposmics (1 on Q9a vs. 0 on Q9) have lower age- and sex- adjusted BMI at SC2?
    - Positive constipation screen: Response of >2 on Q1.11 of MDS-UPDRS at SC2 or daily (non-prn) laxative use on concomitant medication log at SC2
        - Constipation characterization: Do those who are positive by this moderately stringent criterion have a:
            - Lower current, vigorous physical activity score (Q3 on PD Risk Factor Questionnaire; among those selecting response 1-4)?

- Lower caffeine intake (Q3-4 on Smoking, Alcohol and Caffeine Status)

- SURE-PD3 prodromal PD Index (range 0, 1, 2 or 3) with simple sum values for presence (= 1) for each of above three features.

6.8.5 Determinants of Early Disability Warranting Dopaminergic Therapy

An earlier, mid-study effort to identify determinants of need for dopaminergic therapy occurring prematurely (by the 12-week visit) was prompted by a higher than expected rate of these events and a concern that it would reduce study power to meet its primary analysis objective. Modified Schwab and England ADL score at screening was identified as a significant predictor, and site staff were advised to keep in mind that a Modified Schwab and England ADL score below 90% was predictive of premature need for dopaminergic therapy, the expectation of which is an exclusionary criterion. Because this clinical eligibility criterion is common in PD trials of candidate neuroprotectants, a fuller assessment of predictors would be valuable.

The following potential determinants of Early Disability Warranting Dopaminergic Therapy will be tested: Modified Schwab and England ADL score, time between symptom onset or diagnosis and baseline visit, having taken dopaminergic medication (other than an MAO-B inhibitor) prior to screening, being on a stable dosage of an MAO-B inhibitor at baseline, baseline MDS-UPDRS Part I-III total score, DaTscan^{TM} SBR, expectation of improvement on inosine (on Expectancy Questionnaire), baseline Neuro-QOL Depression score, and baseline MoCA score.

6.8.6 Placebo Response

The frequency of a placebo-associated improvement in activities of daily living (ADL) and motor symptoms as measured by the MDS-UPDRS Parts IB and II and Part III, respectively, will be summarized among all placebo arm participants, by visit, and by site. Following Goetz et al. 2000, we will define placebo-associated improvement as a reduction in the ADL or motor score of at least 50% or an improvement from baseline of two or more points on at least two different items of the respective MDS-UPDRS scale. Stability of the rates of placebo-associated improvements will be tested across visits using mixed model logistic regression with fixed effects of visit and random participant-specific intercepts. The following predictors of placebo-associated improvement at any visit will be tested by logistic regression:

- age at baseline (less than 65 years versus greater or equal to 65 years),
- disease duration at baseline (less than 2 years vs. greater or equal to 2 years),
- disease severity at baseline (modified Hoehn and Yahr Stages 1 and 1.5 vs. 2 and 2.5),
- number of years education (Q1 on Socio-Economics form)
- daily caffeine intake: lower vs higher (Q3 + 0.3xQ4 from Smoking, Alcohol and Caffeine form)
- expectation of inosine: “A lot better” vs “Somewhat better” vs “No change” or worse (Q2 on PD-Expectancy questionnaire)
- expectation of inosine-placebo difference: >/=2 vs <2 units more favorable for inosine (Q2 - Q1 on PD-Expectancy questionnaire)
- preference for inosine: “strong” vs other (Q3 on PD-Expectancy questionnaire)
- treatment belief: placebo vs inosine (Q2 on Blindedness Questionnaire).

6.8.7 Blindedness

Descriptive stats will be summarized for responses on the early and late administrations of the Blindedness Questionnaire and the change between them. The following predictors of treatment assignment belief reported on the Blindedness Questionnaire will be tested by logistic regression. For belief of inosine assignment (stratified by correct and incorrect):

- age at baseline (categorized as <65 years vs. ≥65 years)
- sex
- treatment assignment
- MDS-UPDRS I-III censored change at wash-in (per 6.6.8) for response at the 6-week (V02) administration
- MDS-UPDRS I-III censored change to ~24-mo/V10 for response at the ~24-mo/V10 administration
- history of preceding nephro-lithiasis TEAE
- history of preceding gout, gout-like or other arthritic TEAE
- average preceding daily number study drug capsules
- expectation of inosine: “A lot better” vs “Somewhat better” vs “No change” or worse (Q2 on PD-Expectancy questionnaire),
- expectation of inosine-placebo difference: >/=2 vs <2 units more favorable for inosine (Q2 - Q1 on PD-Expectancy questionnaire)
- preference for inosine: “strong” vs other (Q3 on PD-Expectancy questionnaire).

7. References

1. Hauser RA, Auinger P; Parkinson Study Group. Determination of minimal clinically important change in early and advanced Parkinson's disease. Mov Disord. 2011 Apr;26(5):813-8.
2. Lang AE, Eberly S, Goetz CG, Stebbins G, Oakes D, Marek K, Ravina B, Tanner CM, Shoulson I; LABS-PD investigators. Movement Disorder Society Unified Parkinson Disease Rating Scale experiences in daily living: Longitudinal changes and correlation with other assessments. Mov Disord. 2013 Dec;28(14):1980-6.
3. Goetz CG, Stebbins GT, Tilley BC. Calibration of unified Parkinson's disease rating scale scores to Movement Disorder Society-unified Parkinson's disease rating scale scores. Mov Disord. 2012 Sep 1;27(10):1239-42.
4. Gallagher DA, Goetz CG, Stebbins G, Lees AJ, Schrag A. Validation of the MDS-UPDRS Part I for nonmotor symptoms in Parkinson's disease. Mov Disord. 2012 Jan;27(1):79-83.
5. Stebbins GT, Goetz CG, Burn DJ, Jankovic J, Khoo TK, Tilley BC. How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. Mov Disord. 2013 May;28(5):668-70.
6. Ascherio A, LeWitt PA, Xu K, Eberly S, Watts A, Matson WR, Marras C, Kieburtz K, Rudolph, A, Bogdanov M, Schwid SR, Tennis M, Tanner CM, Beal MF, Lang AE, Oakes D, Fahn S, Shoulson I,

7. Schwarzschild MA, Schwid SR, Marek K, Watts A, Lang AE, Oakes D, Shoulson I, & Ascherio A, and the Parkinson Study Group PRECEPT Investigators. Serum urate as a predictor of clinical and radiographic progression in Parkinson's disease. (2008) Arch Neurol. 65:716-723.
8. The Parkinson Study Group SURE-PD Investigators; Schwarzschild MA, Ascherio A, Beal MF, Cudkowicz ME, Curhan GC, Hare JM, Hooper DC, Kieburtz KD, Macklin EA, Oakes D, Rudolph A, Shoulson I, Tennis MK, Espay AJ, Gartner M, Hung A, Bwala G, Lenehan R, Encarnacion E, Ainslie M, Castillo R, Togasaki D, Barles G, Friedman JH, Niles L, Carter JH, Murray M, Goetz CG, Jaglin J, Ahmed A, Russell DS, Cotto C, Goudreau JL, Russell D, Parashos SA, Ede P, Saint-Hilaire MH, Thomas CA, James R, Stacy MA, Johnson J, Gauger L, Antonelle de Marcaida J, Thurlow S, Isaacson SH, Carvajal L, Rao J, Cook M, Hope-Porche C, McClurg L, Grasso DL, Logan R, Orme C, Ross T, Brocht AF, Constantinescu R, Sharma S, Venuto C, Weber J, Eaton K. (2014). Inosine to Increase Serum and Cerebrospinal Fluid Urate in Parkinson Disease: A Randomized Clinical Trial. JAMA Neurol. 2014 Feb 1;71(2):141-50. [Published online 2013 Dec 23.]
9. Parkinson Study Group. Effect of deprenyl on the progression of disability in early parkinson's disease. N Engl J Med 1989;321(20):1364-71.
10. Parkinson Study Group. Effects of tocopherol and deprenyl on the progression of disability in early Parkinson's disease. N Engl J Med 1993;328:176-83.
11. Parkinson Study Group PRECEPT Investigators. (2007) Mixed lineage kinase inhibitor CEP-1347 fails to delay disability in early Parkinson disease. Neurology. 69:1480-1490.
12. Schoenfeld, DA. 2014. A program to calculate the power or sample size for a random slopes model. MGH Biostatistics Center. Available at: http://hedwig.mgh.harvard.edu/biostatistics/node/42 (Accessed online 9/5/15).
13. Weaver et al. Randomized trial of deep brain stimulation for Parkinson disease: Thirty-six-month outcomes. Neurology. 2012;79:55-65.
14. Liang KY, Zeger S. Longitudinal data analysis of continuous and discrete responses for pre–post designs. Sankhyā: The Indian Journal of Statistics (Series B) 2000; 62(1):134–148.
15. Liu GF, Lu K, Mogg R, Mallick M, Mehrotra DV. Should baseline be a covariate or dependent variable in analyses of change from baseline in clinical trials? Stat Med. 2009 Sep 10;28(20):2509-30.
16. Goetz CG, Luo S, Wang L, Tilley BC, LaPelle NR, Stebbins GT. Handling missing values in the MDS-UPDRS. Mov Disord. 2015 Oct;30(12):1632-8.
17. Berry JD, Miller R, Moore DH, Cudkowicz ME, van den Berg LH, Kerr DA, Dong Y, Ingersoll EW, Archibald D. The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials. Amyotroph Lateral Scler Frontotemporal Degener. 2013 Apr;14(3):162-8.
18. Leber P. Slowing the progression of Alzheimer disease: methodologic issues. Alzheimer Dis Assoc Disord. 1997;11 Suppl 5:S10-21; discussion S37-9.
19. McDermott MP, Hall WJ, Oakes D, Eberly S. Design and analysis of two-period studies of potentially disease-modifying treatments. Control Clin Trials. 2002 Dec;23(6):635-49.

20. Rascol O. "Disease-modification" trials in Parkinson disease: target populations, endpoints and study design. Neurology. 2009 Feb 17;72(7 Suppl):S51-8.
21. Zadikoff C, Fox SH, Tang-Wai DF, Thomsen T, de Bie RM, Wadia P, Miyasaki J, Duff-Canning S, Lang AE, Marras C. A comparison of the mini mental state exam to the Montreal cognitive assessment in identifying cognitive deficits in Parkinson's disease. Mov Disord. 2008 Jan;23(2):297-9.
22. Parkes JD, Baxter RC, Marsden CD, Rees JE. Comparative trial of benzhexol, amantadine, and levodopa in the treatment of Parkinson's disease. J Neurol Neurosurg Psychiatry. 1974 Apr;37(4):422-6.
23. Mancini F, Di Fonzo A, Lazzeri G, Borellini L, Silani V, Lacerenza M, Comi C. Real life evaluation of safinamide effectiveness in Parkinson's disease. Neurol Sci. 2018 Apr;39(4):733-739.