Statistical Analysis Plan (SAP)

Sponsor:	SOM Biotech
Protocol No	SOMCT03 NCT number: NCT05475483
Study Title:	Phase IIb, randomized, double-blind, placebo-controlled study in parallel groups assessing the efficacy and safety of two doses of SOM3355 in patients suffering from Huntington's Disease with choreic movements.
Protocol Version/Date:	V1.1, Spain 2022-02-24 V1.1, UK 2022-05-19 V1.1, Italy 2022-05-17 V1.1, France 2022-05-20 V1.1, Germany 2022-07-25 V1.1 Switzerland 2022-08-03
SAP Version/Date:	Version 1.0, 2024-07-12
Supersedes SAP Version:	Version 0.5
Appendices (external documents):	1. List of Tables, Listing, Figures (TLFs) 2. Additional details on statistical methods

Revision history

SAP Version	Version date	Reason(s) for change
0.1	2023-03-31	1st draft
0.2	2023-05-15	2nd draft for sponsor review
0.3	2023-08-04	3rd draft for sponsor review
0.4	2024-06-05	4th draft for sponsor review
0.5	2024-07-09	5th draft for sponsor review
1.0	2024-07-12	Final version

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Description
AE	Adverse Event
ANOVA	Analysis of Variance
AUC	Area under the Curve
ATC	Anatomic, Therapeutic, Chemical (Classification System for Drugs)
BARS	Barnes Akathisia Rating Scale
BDI	Beck Depression Inventory
BID	Twice daily
BMI	Body Mass Index
BP	Blood Pressure
bpm	Beats per minute
CAG	Cytosine-Adenine-Guanine
CGI	Clinical Global Impression
CI	Confidence Interval
C_{max}	Maximum observed plasma concentration
CPMP	Committee for Propriety Medicinal Products
CRF	Case Report Form
CS	Clinically Significant
CSP	Clinical Study Protocol
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide Severity Rating Scale
CTC	Common Toxicity Criteria
DBP	Diastolic Blood Pressure
DD	Drug Dictionary (WHO Coding Thesaurus)
deuTBZ	Deutetrabenazine
DMP	Data Management Plan
DRM	Data Review Meeting
DVP	Data Validation Plan
ECG	Electrocardiogram
ESS	Epworth Sleepiness Scale
FAS	Full Analysis Set
GCP	Good Clinical Practice
HD	Huntington's Disease
HPLC	High Performance Liquid Chromatography
HR	Heart rate
ICH	International Council for Harmonisation
IMP	Investigational Medicinal Product
ITT	Intention-to-treat
LOD	Level of Detection
LOQ	Level of Quantification
LSmeans	Least Square Means
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
MoCA	Montreal Cognitive Assessment
NCS	Non-Clinically Significant

Abbreviation	Description
PBA-s	Problem Behavior Assessment-short form
PD	Pharmacodynamics
PGI	Patient Global Impression
PK	Pharmacokinetics
PP	Per-protocol
PPA	Per-protocol Analysis
PPS	Per-protocol Set
PT	Preferred Term
REML	Restricted Maximum Likelihood
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SAS®	Statistical Analysis Software Package
SBP	Systolic Blood Pressure
SD	Standard Deviation
SEM	Standard Error of the Mean
SOC	System Organ Class
SOM3355	Bevantolol hydrochloride (racemate)
SOP	Standard Operating Procedure
SP	Statistical Programmer
SUSAR	Suspected Unexpected Serious Adverse Reaction
TBZ	Tetrabenazine
TEAE	Treatment Emergent Adverse Event
TFC	Total Functional Capacity
TLFs	Tables, Listings, Figures
T_{max}	Time corresponding to occurrence of C_{max}
TMC	Total Maximal Chorea
TMS	Total Motor Score
TS	Trial Statistician
UHDRS®	Unified Huntington's Disease Rating Scale
VMAT2	Vesicular monoamine transporter type 2
WHO	World Health Organization

1 STUDY INFORMATION

Huntington's disease (HD) is an inherited progressive neurodegenerative disease. Chorea, an involuntary jerky movement that is purposeless and abrupt, is one of the most prominent symptoms in HD. No effective disease-modifying therapies exist for HD, and treatment is symptomatic relief. Chorea is treated by tetrabenazine (TBZ), approved in Europe and the United States (US), and the long-acting derivative deutetrabenazine (deuTBZ), approved in the US only. Both drugs are vesicular monoamine transporter type 2 (VMAT2) inhibitors depleting monoamines in presynaptic neurons. It is thought that the consequent modulation of dopamine signalling across neural synapses reduces the dyskinetic movements, such as chorea.

SOM3355 provided functional inhibition of the VMAT2 with similar potency to that of TBZ but different binding properties that may provide less risk of severe side effects. Relying on its VMAT2 inhibition activity, and the good tolerability shown in previous trials at doses from 100 to 600 mg/day administered once or twice daily (OD or BID), SOM Biotech is now developing SOM3355 for the treatment of the dyskinetic movement disorders related to HD.

1.1 Primary objective

The primary objective of the study is to assess the efficacy of two doses of SOM3355 (200 mg taken twice daily [BID] i.e. 400 mg/day and 300 mg BID i.e. 600 mg/day over at least 8 weeks at maintenance dose) compared to placebo to reduce chorea in HD patients measured by the change from baseline in total maximal chorea (TMC) score (primary efficacy endpoint).

1.2 Secondary objective

The secondary objectives of the study are:

- To evaluate the safety and tolerability of two doses of SOM3355 (400 mg/day and 600 mg/day) compared with placebo in HD patients, with particular attention to depression and suicidality and to the cardiovascular hemodynamic parameters.
- To evaluate the efficacy of the two doses of SOM3355 (400 mg/day and 600 mg/day) compared with placebo on:
    - the change of Clinical Global Impression (CGI-C),
    - the change of Patient Global Impression (PGI-C),
    - the percentage of responders based on the improvement ≥2 in TMC score,
    - the percentage of change of TMC score.
- A pharmacokinetic (PK) sub-study in a subset of 24 patients will allow to determine the PK profile of the two doses of SOM3355 (200 mg BID and 300 mg BID) at steady state and to model PK-PD assessments of hemodynamic and cardiac parameters. Analysis of these data will be described in a separate document and the results will be also presented in a separate report.

1.3 Study design

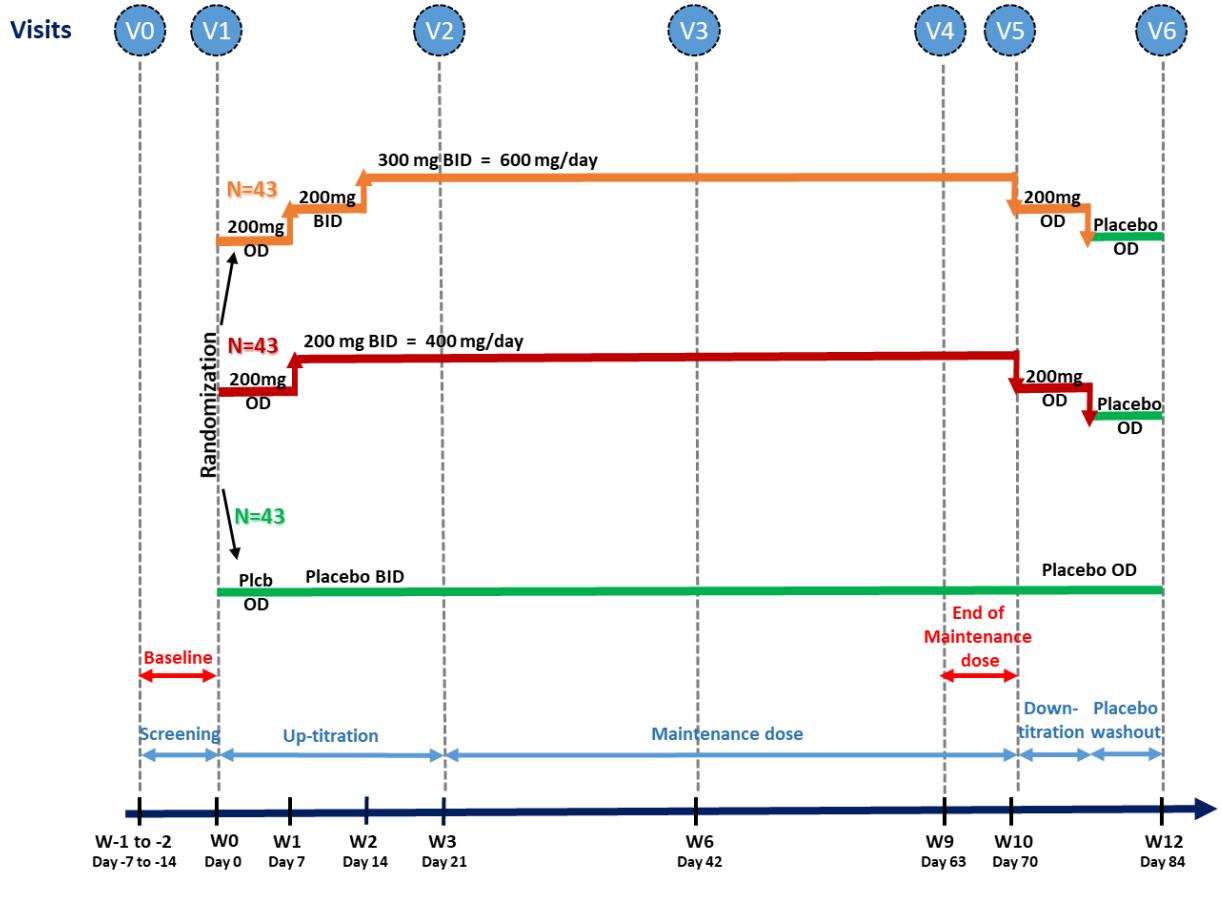
This is an international, multicenter, double-blind, randomized, placebo-controlled, parallel-group Phase IIb study assessing two doses of SOM3355 (400 or 600 mg/day), taken as 200 mg BID or 300 mg BID, in patients presenting choreic symptoms related to HD.

Patients will be randomized at the inclusion visit (Visit 1) to one of 3 treatment arms to receive either SOM3355 400 mg/day, SOM3355 600 mg/day, or placebo, all given BID. Dosing will start with an up titration in the first 3 weeks and will be maintained at a fixed dose for 7 weeks. After 10 weeks, patients will down titrate to 1 capsule 200mg / day of study drug for 1 week, followed by 1 capsule / day of placebo to assess the effect of drug interruption.

The study will enroll 129 patients with a genetically proven diagnosis of Huntington's Disease (confirmed by HTT gene CAG repeats ≥36), of mild to moderate severity (TFC score ≥7) with choreic symptoms (TMC score ≥10). The patients will be recruited in investigating sites in various European countries. Ambulatory patients will be recruited by reference centers for HD in collaboration with patient advocacy groups and the EHDN.

A sub-study will characterize the PK profile of the 2 doses (400 mg/day or 600 mg/day) given twice daily for at least one week in a subset of patients (n = 24) who consent to be hospitalized for 24 hours for blood sampling and cardiac and hemodynamic assessments.

Figure1



The flowchart/schedule of events can be found in Table 1: Schedule of Activities by Visit of the protocol.

An independent data safety monitoring board (DSMB) will be established by the Sponsor. DSMB will assess at intervals the progress of the clinical trial and the safety for participants in reviewing accumulated data on safety, tolerability, and efficacy outcomes.

1.4 Planned sample size

Power calculation is based on results from previous trials that tested VMAT2 inhibitors (TBZ and deuterbenazine) on chorea in HD patients with TMC score as primary endpoint, which showed a standard deviation of 3.5 and 3.9 respectively, and a treatment effect compared to placebo in TMC score of -3.5 and -2.5 respectively.

The sample size will be a total of 129 patients. Eligible patients will be randomized in blinded fashion in a 1:1:1 ratio i.e. 43 patients in each arm (SOM3355 400 mg/day, SOM3355 600 mg/day, or placebo). Assuming a treatment difference of -2.5 points in the change from baseline of TMC score compared to placebo and a standard deviation of 3.5 (to ensure a clinically relevant effect of at least -2 in the study), 39 patients per arm will provide at least 80% power that at least 1 dose of SOM3355 will be significantly superior to placebo at the 2-sided 0.025 significance level. This is based on the use of a two-sided t-test for the mean TMC score change. For the primary endpoint, the overall type I error will be controlled using a Holm procedure. To account for a 10% dropout rate, 43 patients per arm (i.e., a total of 129 patients) will be enrolled.

2 GENERAL INFORMATION

2.1 Background details

All study data will be transferred to a SAS database (version 9.4 or later) for statistical analysis purposes. Data will be imported from a Data Capture System OPVerdi and external data vendors (cf. Data Management Plan) via validated SAS programs.

The SAP will be finalized before database lock after agreement with the Sponsor on patient disposition and coding.

2.2 Deviations from the trial protocol with regard to statistical analyses

The following deviations from the protocol were identified as being important for the statistical analysis:

For the key secondary efficacy endpoint (Change from baseline in the Clinical Global Impression (CGI-C)) multiplicity adjustment will be applied.

> Original text in protocol: The type I error will be set at 0.05 (two-sided) and the coverage level of confidence intervals to 95% for all secondary efficacy analyses. No multiplicity adjustment will be applied for the statistical inferences of the secondary efficacy endpoints.

The revised definition in SAP only for key secondary endpoint: To preserve the family-wise error rate at the 0.05 level when comparing the two doses of SOM3355 versus placebo and comparing a primary endpoint and key secondary endpoint, statistical testing will be carried out using the Holm procedure for CGI-C if the primary endpoint is significant for both doses. The smallest p-value of the two comparisons will be first compared at the two-sided 0.025 level and if statistically significant, the other p-value will be compared at the two-sided 0.05 level.

2.3 Individual protocol deviations

Apart from violations of inclusion or exclusion criteria the table below lists conditions that may affect the evaluation of patient data and the proposed actions:

Condition	Decision Rule
Informed consent not given	Exclude from all data analyses.
Violation of any inclusion or exclusion criteria	Potentially exclude from Per Protocol Population (PP)
Patients not exposed to IMP	Exclude from Safety Analysis Population (SAF), Modified Intent-to-Treat Population (mITT) and Per Protocol Population (PP)
Premature discontinuation during the treatment period	Potentially exclude from Per Protocol Population (PP)
Less than 80% or more than 120% compliance with IMP intake	Potentially exclude from Per Protocol Population (PP)
No post-baseline assessment of the TMC score	Exclude from Modified Intent-to-Treat Population (mITT) and Per Protocol Population (PP)
Patients receiving prohibited treatments	Potentially exclude from Per Protocol Population (PP)
Patient received a treatment which the patients was not randomized for	Exclude from Per Protocol Population (PP)

A detailed review of all documented and derived deviations from protocol will be part of the Blinded Data Review Meeting (BDRM) before database lock. During this BDRM the impact of protocol deviations on the analysis will be assessed and the conclusions recorded. Also, additional deviations could be identified during the meeting.

A complete listing of documented and derived protocol deviations and the judgment for assessment of patient disposition will be signed before database lock. A description of protocol violations that led to exclusion from any analysis sets will be included in the CSR.

3 ANALYSIS POPULATIONS

In general, the disposition of patients will be displayed for following populations, if applicable:

- All screened population (having signed ICF)
- Screening failures

- Intention-to-Treat (ITT) population
- Safety (SAF) population
- Modified Intention-to-Treat (mITT) population
- Per-Protocol (PP) population

Population assignment of patients will be decided upon in a treatment-blinded DRM with the Sponsor before database lock. The proper flags for analysis sets exclusion (e.g., exclusion from PP set), will be included in the analysis datasets. The protocol deviation list should be finalized before database lock.

3.1 Screen failures

Screen failures are patients who have consented to participate in the clinical trial by signing the Informed Consent. Form before performing the screening assessments but who did not meet all the selection criteria to be randomized at the inclusion visit V1, by consequence they are not eligible to be subsequently randomly assigned to the study intervention or entered in the study.

3.2 Intention-to-Treat (ITT) population

The ITT Population will include all randomized patients whether they received the study treatment or not. Patients will be analysed in the treatment group to which they were randomized, regardless of the treatment that was actually received or if they withdrew prior to first dosing.

3.3 Safety (SAF) population

The Safety Analysis Set (SAF) includes all patients who were administered at least one dose of any study drug. Patients who are randomized and assigned a treatment but withdrew prior to dosing will not be included in the safety population. Patients will be included in the treatment group based on the treatment actually received. The analysis of all safety variables will be based on the SAF.

3.4 Modified Intention-to-Treat (mITT) population

The mITT Population will include all patients from the ITT Population who were randomized to treatment, received at least one dose of study drug and had at least one post-baseline assessment of the TMC score. The primary efficacy analysis will be conducted on the mITT Population. Patients will be included in the treatment group to which they were randomized, regardless of the treatment that was received.

3.5 Per-protocol (PP) Population

The Per-Protocol Population (PP) is a subset of patients in the mITT.

It includes all randomized patients who completed the study without important deviations from the protocol. The decision of the membership in the PP Population will be made during a blinded Data Review Meeting prior to database lock, reviewing all deviations occurred.

3.6 Subgroup analyses

Subgroup analyses of the primary efficacy (change in Total Maximal Chorea (TMC) score from baseline to the end of maintenance therapy at V5) are planned for the following sub-groups:

- patients not taking neuroleptics during the study vs patients taking neuroleptics as concomitant medication during the study
- gender: male, female
- age: < median, ≥ median
- country: France, Spain, UK, Poland, Italy, Germany
- baseline TMC: < median, ≥ median
- baseline TMS: < median, ≥ median

This analysis will be performed on the mITT population, except the neuroleptics subgroup, which will also be repeated on the PP population.

Additional subgroup, correct versus incorrect down titration phase between Visit 5 and Visit 6, is defined as:

- patients who have performed a correct down titration phase i.e. who correctly took the IMP at half dose (1 capsule/ day) for a week and then capsules from placebo bottle for at least 4 days.
- patients who did incorrect down titration
    - who took 2 capsules/day of IMP so remaining at full dose with no down titration, and then placebo with abrupt IMP discontinuation
    - or who took the IMP at half dose (1 capsule/ day) for 2 weeks with no placebo wash out or less than 4 days under placebo

For this subgroup, only descriptive statistics on visits for the TMC score will be repeated on the mITT population to show the evolution of the TMC score between V5 and V6 after IMP interruption. In addition, time profiles of the mean TMC scores with standard error of the mean (SEM) will be presented.

4 STATISTICAL ANALYSES

All statistical analyses will be performed using the SAS® software (Version 9.4 or later). It will be based on data according to CDISC standards, SDTM (SDTM IG v3.3) and ADaM (ADaM IG v1.3).

Descriptive statistics will be given by treatment group and overall, where appropriate.

If not stated otherwise, the following standard types of descriptive analyses will be presented:

– Descriptive statistics for continuous data

  N, mean, SD, min, median and max will be presented. These descriptive statistics will be determined for measured values and for changes from baseline.

– Descriptive statistics for categorical data

  Absolute frequencies and percentages will be presented. Percentage bases (denominators) will be identified in the table title or footnote (i.e. all patients at risk, all non-missing cases, all cases). For changes from baseline, shift tables may be generated.

– Inferential statistics

  Unless otherwise stated, all statistical tests will be performed two-sided and at a type I error probability of α=0.05. The p-values will be printed consistently with 4 decimals ($p<0.0001$ will be displayed, if the p-values are less than 0.0001 and $p>0.9999$ will be displayed, if the p-values is more than 0.9999).

  Unless otherwise stated, all confidence intervals (CIs) will be derived two-sided and at a confidence probability of 1-α=0.95.

– Listings

  All recorded data will be listed by patient (sorted by treatment group Placebo, 200 mg BID, 300 mg BID). Identification variables are centre number, patient number and treatment group.

Derived data will be stored in special analysis data sets and will be calculated as outlined in section 6.1.

4.1 Conventions

4.1.1 Baseline definition

Baseline is defined according to the endpoint as follows:

- For the Total Maximal Chorea score (TMC), and motor part of Unified Huntington's Disease Rating Scale (UHDRS-TMS) the baseline is defined as the average of the values for Visit 0 and Visit 1.
- The severity of choreic symptoms is assessed at Visit 1 with Clinical Global Impression-severity (CGI-S) scale and Patient Global Impression-severity (PGI-S) scales. Score of these 2 scales is assessed at baseline but cannot be used as baseline value when calculating changes to baseline, since change is already assessed at each visit (Visit 2 to Visit 6) with Clinical Global Impression of Change (CGI-C) and Patient Global Impression of Change (PGI-C).
- For EQ-5D, Barnes Akathisia Rating Scale (BARS), Montreal Cognitive Assessment (MoCA) scale and Problem Behavior Assessment short form (PBA-s) baseline is the value at Visit 1.

- For the Total Functional Capacity score (TFC), Columbia Suicide Severity Rating Scale (C-SSRS), Beck Depression Inventory (BDI), Epworth Sleepiness Scale (ESS), ECG and vital signs, baseline is defined as the value at Visit 1.
- For Functional Assessment (FA) of the UHDRS baseline is at Visit 0.
- For laboratory data (haematology and clinical chemistry), baseline is at Visit 0.
- For any other parameter, baseline is defined as the last available measurement before the first administration of study drug, i.e., the value of the parameter at Visit 0 or at Visit 1.

4.1.2 Missing data

In case of premature discontinuation, values collected at the end of study visit (EOS) or Unscheduled Visit (if it is the last visit) will be imputed to the first missing visit of the patients. For example, a patient withdrew between V3 and V4 (i.e. V4 is missing), then the value recorded at EOS will be imputed to V4. If a patient discontinues exactly at the time of a visit, but the value is recorded in the EOS page, this value will be imputed to the visit when the discontinuation occurred.

Partial dates are allowed on the eCRF for adverse event (AE) onset and resolution dates, concomitant medication start and stop dates, and concomitant procedure dates. An entry for the year is required in the eCRF system for each of these dates. Only the month and day may be entered as unknown. Dates from these forms will be reported in listings as collected. Every effort will be made to query missing dates particularly to clarify the onset day and time of AE occurring at visit 1 to determine if the AE is treatment emergent or not.

For records with missing AE onset date, the following procedure will be employed for use in determining whether the AE is treatment emergent:

- AE onset dates with missing day and non-missing month will be assumed to occur on the first day of the non-missing month, except for AEs occurring in the first month of dosing, in which case the date will be set to the first day of dosing.
- AE onset dates with missing month will be assumed to occur on the first day of the non-missing year (i.e., January 1), except for AEs occurring in the first year of dosing, in which case the date will be the first day of dosing.

For records with a missing medication start and/or stop date, the following procedure will be employed for use in determining whether the medication is prior or concomitant:

- Medication start dates with a missing day and non-missing month will be assumed to occur on the first day of the non-missing month, except for medications occurring in the first month of dosing, in which case the date will be the first day of dosing.
- Medication start dates with missing month will be assumed to occur on the first day of the non-missing year (i.e., January 1), except for medications occurring in the first year of dosing, in which case the date will be the first day of dosing.

- Medications that are not ongoing and have a medication stop date with a missing day and non-missing month will be assumed to occur on the last day of the non-missing month.
- Medications that are not ongoing and have a medication stop date with a missing month will be assumed to occur on the last day of the non-missing year (i.e., December 31).

For records with a missing procedure date, the following procedure will be employed for use in determining whether the procedure is prior or concomitant:

- Procedure dates with a missing day and non-missing month will be assumed to occur on the first day of the non-missing month, except for procedures occurring in the first month of dosing, in which case the date will be the first day of dosing.
- Procedure dates with missing month will be assumed to occur on the first day of the non-missing year (i.e., January 1), except for procedures occurring in the first year of dosing, in which case the date will be the first day of dosing.

4.1.3 Pooling of centres

No pooling of centres will be performed.

4.1.4 End of Study

A patient is considered to have completed the study if he/she has completed all visits of the study (until visit 6).

The end of the study is defined as the date of the last visit of the last patient in the study.

4.2 Patient disposition

The disposition of patients will be summarized using the following tabulations:

- the number of all screened subjects (included in Visit 0 - ICF signed)
- the number of screening failures with reasons of exclusion
- the number of randomised subjects (at Visit 1)
- the number of subjects who took at least one dose of IMP
- the number of subjects who prematurely discontinued the study i.e. withdrawals with reasons (Adverse Events, non-compliance to IMP, Protocol selection criteria violation, Consent withdrawn, others)
- the number of patients who completed V5 (primary endpoint assessment at the end of maintenance dose)
- the number of subjects having completed the entire study until Visit 6

Then this will define the number of subjects in the ITT, Safety, mITT and PP population.

Frequencies or different reasons for exclusions from analysis sets (as defined in section 3) will be tabulated and listed.

A CONSORT flow diagram on patient disposition, will be created in the CSR.

Also, the disposition by country and centre will be presented.

A by-patient listing of the randomisation scheme and codes will be provided.

4.3 Demographic and other background data

4.3.1 Basic description

Demographic and baseline characteristics will be summarized in frequency or sampling statistics tables for Safety population (SAF) and mITT, if not identical.

Screen failure demographic data will be also listed.

Demographic variables:

- Age
- Sex
- BMI
- Ethnic origin
- Country and centre
- Females only: childbearing potential (Yes/No)

Disease specific patient characteristics measured at screening/baseline:

- Number of Cytosine-Adenine-Guanine (CAG) repeats
- Total Functional Capacity (TFC) score at baseline
- Level of chorea at baseline (TMC score)
- Total Motor score (TMS) at baseline
- Age at diagnosis: *Year of HD diagnosis - Year of Birth*, where year of birth will be calculated as: *Year of inform consent - Age*
- Age at first symptoms: *Year of first symptoms - Year of Birth*, where year of birth will be calculated as: *Year of inform consent - Age*
- Documented evidence on any other HD-related symptom(s) available such
    - Presence of depression
    - Presence of Behaviour disorders or aggressiveness
    - Presence of Cognitive disturbance
- Neuroleptic treatment in co-medications to treat behaviour disorder (accepted by Steering Committee)
- Antidepressant treatment in co-medications

Antidepressant and neuroleptic treatments will be identified by the Sponsor and Medical Monitor.

4.4 IMP exposure and compliance

Exposure:

- Total number of patients with at least one dose of study treatment: overall and per treatment periods (W1 (200mg once daily (OD) or Placebo), W2 (200mg BID or Placebo), W3 (200 or 300mg BID or Placebo), W4-W10 (200 or 300mg BID or Placebo), W11 (200mg OD or Placebo), W12 (Placebo))
- Total number of days with at least one dose of study treatment: overall and per treatment periods (W1 (200mg OD or Placebo), W2(200mg BID or Placebo), W3, W4-W10 (200 or 300mg BID or Placebo), W11(200mg OD or Placebo), W12 (Placebo)).

The earliest date of study withdrawal/completion or product withdrawn (AE) will be selected and the day before this date will be defined as the last treatment intake. The first date of administration is collected by visit in the eCRF. Within visits, the intake of study treatment will be defined according to the study protocol and is theoretical.

Compliance:

Compliance will be evaluated by calculating at each visit the number of capsules used (capsules dispensed minus capsules returned) divided by the expected number of capsules to be used x 100. A patient will be deemed compliant if the patient has taken ≥80% but ≤120% of the expected tablets of study drug.

Total dose taken per patient and per bottle can be calculated as:

Total dose (mg) for each bottle = (total number of capsules in the bottle – number of capsules returned) x *dose for the given bottle*

Compliance (%) will be assessed during the whole study and by study period (up-titration period W1-W3 (from V1 to V2), maintenance period W4-W10 (from V2 to V5), and down-titration period W11-W12 (from V5 to V6).

A summary of the total dose by study period and IMP compliance (%) by study period will be presented per treatment group for the SAF.

A by patient listing of IMP compliance by period will be provided.

4.5 Medical history, physical examination

All medical history and physical examination data will be listed.

Analysis of physical examination is described in Section 4.10.5.

4.6 Prior and concomitant medication

Prior and concomitant medications will be coded using the WHO Drug Global thesaurus version Q1 2022. Coding will be performed by Ergomed and agreed upon with the sponsor before data base lock (cf. DMP).

Prior medications are defined as medications which were taken before the study and stopped before the start of study treatment. For some drugs they should be stopped 1 or 3 months before the randomization as stated in protocol §5.3.1.

Concomitant medications are defined as medications which were taken during the study with the investigational treatment (regardless of the start time).

Prior medications and Concomitant medications will be listed separately. Summaries will show the frequencies of patients by WHO Drug Global INN.

All details of concomitant medications will also be listed including the route, dose, frequency, start and stop date and indication.

Analysis of concomitant medications is described in Section 4.10.6.

4.7 Efficacy

The primary efficacy analysis will be conducted on the mITT population including all patients who were randomized to treatment, received at least one dose of study drug and had at least one post-baseline assessment of the TMC score. Patients will be included in the treatment group to which they were randomized, regardless of the treatment that was received.

Additional analyses for the primary endpoint will also be performed for the ITT population and PP population. The latter analyses serve as sensitivity analysis to explore the robustness (cf. Section 2.1).

4.7.1 Primary endpoint

The primary endpoint is the change in Total Maximal Chorea (TMC) score from baseline (defined for each patient as the average of values at the screening visit (Visit 0) and the inclusion visit (Visit 1)) to the end of maintenance therapy (defined for each patient as the average of values at the end of Week 9 (Visit 4) and the end of Week 10 (Visit 5)).

As a result, the different factors included in the MMRM model (below) will be defined as following:

- Baseline TMC: calculated as (TMC V0 + TMC V1)/2,

- TMC at each following Visits: V2 (week 3), V3 (week 6), V4 (week 9), V5 (week 10)),

- Change from baseline TMC: [(TMC V4 + TMC V5)/2 – Baseline TMC]

If the value at V5 is missing and V4 is not missing, then the value at V4 will be used for the calculation. If the value at V4 is missing and V5 is not missing, then the value at V5 will be used for the calculation. If both V4 and V5 values are missing, change from baseline TMC will be treated as missing.

The estimand of the primary efficacy analysis is defined according to ICH-E9-R1. The estimand reflects a mixed strategy for the inter-current events as presented in the table below.

Estimand attribute	Primary definition	Rationale (as needed)
Population	Patients with Huntington's disease with choreic movements	

Estimand attribute	Primary definition		Rationale (as needed)
Endpoint	Change from baseline (Visit 0 and Visit 1) to end of maintenance dose (Visit 4 and Visit 5) in TMC score		
Inter-current events	**Event**	**Strategy**	**Rationale (as needed)**
	Receipt of assigned study treatment	Treatment policy	This strategy applied to align the estimand with the application of the ITT principle.
	Prohibited concomitant neuroleptic treatment during the treatment period	While on treatmen	This strategy applied up to the time of the prohibited concomitant treatment.
	Discontinuation from study	Hypothetical policy	This strategy applied for randomized and treated patients using a MMRM approach under MAR assumptions.
Summary measure	Difference in the change of TMC score from baseline to end of maintenance visit adjusted LS means between active arms and placebo		

The primary analysis of the primary efficacy endpoint will be based on a mixed-effect linear model for repeated measures (MMRM) to compare each active treatment group to placebo. The model will include the following factors: treatment group, visit, and treatment-by-visit interaction as fixed effects, and baseline TMC score as covariate. An unstructured (UN) variance-covariance matrix will be used to model the within-patient errors. If no reliable results under the UN assumption can be obtained, the Toeplitz (TP) covariance structure, followed by the autoregressive (AR1) and Compound Symmetry (CS) covariance structure will be used. The Kenward-Roger method will be used to estimate the denominator degree of freedom.

The SAS code for the primary analysis is as follows:

```
PROC MIXED data=dataset method=REML;
CLASS subject treatment visit;
MODEL changeTMC = treatment visit treatment*visit baselineTMC / ddfm=KR;
REPEATED visit / subject = subject type = UN;
LSMESTIMATE treatment*visit “TMC change 200 mg vs Placebo” 0 0 -1  0 0 1  0 0 0 / cl elsm;
LSMESTIMATE treatment*visit “TMC change 300 mg vs Placebo” 0 0 -1  0 0 0  0 0 1/ cl elsm;
RUN;
```

The least squares mean (LSM) estimate for each treatment group and difference in LSM estimates at end of maintenance dose between active treatments and placebo group with their confidence intervals will be provided. For the lowest p-value 97.5% CI will be provided and for the highest p-value 95% CI will be provided.

To preserve the family-wise error rate at the 0.05 level when comparing the two doses of SOM3355 versus placebo, statistical testing will be carried out using the Holm procedure. The smallest p-value of the two comparisons will be first compared at the two-sided 0.025 level and if statistically significant, the other p-value will be compared at the two-sided 0.05 level.

Descriptive statistics will be presented for the TMC score and changes from baseline by treatment group and by visit (including V6), on imputed data (as described in 4.1.). In addition, time profiles of the mean TMC scores with standard error of the mean (SEM) will be presented.

Sensitivity analyses on the primary efficacy endpoint will be conducted as follows:

Sensitivity #1

A sensitivity analysis will be performed to assess the robustness of the primary analysis based on MMRM. Intermittent missing values will first be imputed using the Markov Chain Monte Carlo (MCMC) method which assumes a joint multivariate normal model for variables in the imputation model. Then a control-based pattern imputation is applied to the monotone missing data under the assumption of data missing not at random (MNAR). PROC MI using the MCMC statement will be used to implement the first step. Then PROC MI using the MONOTONE and MNAR statements will be used to create the complete data sets. MNAR will be imputed using Copy Reference method and number of imputations will be 30. The same MMRM analysis as defined for the primary analysis will be performed on the complete data sets, and the results will be combined using PROC MIANALYZE. This analysis will be performed on the mITT population.

Sensitivity #2

Primary endpoint will be analysed using the observed cases (no imputations) using an analysis of covariance (ANCOVA) model including treatment as fixed factor and baseline score as covariate. Results for the outcome of the primary endpoint will be presented for the least square means estimates and 95% CIs. This analysis will be performed on the mITT population.

Sensitivity #3

Primary efficacy analysis will be repeated for the PP Population.

Sensitivity #4

The primary efficacy analysis will be repeated using data without imputation for the visits of discontinued patients for the mITT population.

Subgroup analysis as defined in section 3.6 will be performed for the primary endpoint across all predefined groups, with a particular focus on comparing patients taking neuroleptics as concomitant medication during the study versus patients not taking neuroleptics during the study. For this subgroup, descriptive statistics for the TMC score and changes

from baseline by treatment group and by visit (including V6) will be presented, as well as time profiles of the mean TMC scores with standard error of the mean (SEM) will be provided.

For the subgroup analysis assessing patients with correct versus incorrect down titration phase between V5 and V6, only descriptive statistics for TMC score on visits will be repeated on the mITT to show the evolution of the TMC score between V5 and V6 after IMP interruption.

4.7.2 Secondary endpoints

Key secondary endpoint:

- CGI change (CGI-C) at V5 (end of maintenance)

The key secondary efficacy endpoint will be analysed for the mITT and PP Population.

To preserve the family-wise error rate at the 0.05 level when comparing the two doses of SOM3355 versus placebo and comparing a primary endpoint and key secondary endpoint, statistical testing will be carried out using the Holm procedure for CGI-C if the primary endpoint is significant for both doses. The smallest p-value of the two comparisons will be first compared at the two-sided 0.025 level and if statistically significant, the other p-value will be compared at the two-sided 0.05 level.

Other secondary endpoints:

They will be assessed in the following order:

- TMC-response defined as improvement ≥ 2 in TMC score between the baseline and the end of maintenance dose
- The percentage of change in TMC score
- The PGI change
- Total Motor Score (TMS) of the motor part of UHDRS®
- Gait sub-score of the motor part of UHDRS®
- Dystonia sub-score of the motor part of UHDRS®
- EQ-5D-5L measuring quality of life

Secondary efficacy endpoints will be analysed for the mITT Population.

The type I error will be set at 0.05 (two-sided) and the coverage level of confidence intervals to 95% for all secondary efficacy analyses. No multiplicity adjustment will be applied for the statistical inferences of the secondary efficacy endpoints.

4.7.2.1 Clinical Global Impression

The CGI-Severity (CGI-S) assessed at baseline is a 7-point Likert Scale that asks the clinicians to rate, considering their total experience with this type of patients, the level of illness due to chorea on the illness severity ranging from normal (1) to extremely ill (7) (CGI-Severity).

Then, after therapy initiation, the CGI-Change (CGI-C) asks the clinician to assess the overall global change in the choreic symptoms since the start of study drug dosing, in a 7-point Likert Scale ranging from very much worse (7) to very much improved (1).

CGI-S will be assessed at the inclusion (Visit 1 at Day 0), and CGI-C at each following visit (from V2 to V6). For the CGI-C, patients with a score of 1 (very much improved), 2 (much improved) or 3 (minimally improved) will be defined as ”Improved” and patients with a score of 4 (No change), 5 (Minimally worse), 6 (Much worse) and 7 (Very much worse) will be defined as “Not improved”.

CGI-S (baseline characteristic) and CGI-C results will be analysed descriptively at each visit.

A logistic regression model will be performed for CGI-C at V5 (end of maintenance dose) including the treatment group as the main factor in the model. Estimates along with 95% confidence intervals for the difference in response proportion and the odds ratio of response will be derived from this model.

In addition, histogram will be presented.

4.7.2.2 Percentage of responders on TMC score

TMC score is assessed at each visit from V0 to V6.

V0 and V1 values are used to calculate the baseline (mean of the 2 values), and V4 V5 values for the end of maintenance dose.

Assessment at V6 occurs after the drug interruption with one week at 200 mg OD and the last week under placebo.

TMC score evolution will be assessed at each visit.

The number and percentage of responders with an improvement of at least 2 points in the TMC score will be calculated in each treatment group at each visit.

The TMC response status (“Responder” vs. “Not responder”) will be described at each visit.

Logistic regression model will be performed for responders on TMC score at V5 (end of maintenance dose) including the treatment group as the main factor in the model. Estimates along with 95% confidence intervals for the difference in response proportion and the odds ratio of response will be derived from this model.

In addition, histogram will be presented.

This analysis will be repeated on the PP population.

4.7.2.3 The percentage of change in TMC score

The percentage of change in TMC score is defined as:

100 * (TMC score at Visit x – TMC score at Baseline) / TMC score at Baseline.

"Baseline" is the average score between V0 and V1, "Visits x" are V2, V3 and the average score between V4 and V5 or V4, V5 in case of missing visit.

For the analysis of the percentage of change in TMC score, descriptive statistics and the same MMRM model, as described for the primary endpoint will be used.

This analysis will be repeated on the PP population.

4.7.2.4 Patient Global Impression

The PGI-Severity (PGI-S) is a 7-point Likert Scale that asks the patients to assess the severity of their choreic symptoms at baseline ranging from not present (1) to extremely severe (7).

Then, after therapy initiation, the PGI-Change (PGI-C), asks the patients to assess the overall global change in their choreic symptoms and their quality of life since the start of study drug dosing, in a 7-point Likert Scale ranging from very much worse (7) to very much improved (1).

PGI-S will be assessed at the inclusion (Visit 1 at Day 0), and PGI-C at each following visit (from V2 to V6). For the PGI-C, patients with a score of 1 (very much improved), 2 (much improved) or 3 (minimally improved) will be defined as "Improved" and patients with a score of 4 (No change), 5 (Minimally worse), 6 (Much worse) and 7 (Very much worse) will be defined as "Not improved".

PGI-S and PGI-C results will be analysed descriptively at baseline and post-baseline visit respectively.

Logistic regression models will be performed for PGI-C at V5 (end of maintenance dose) including the treatment group as the main factor in the model. Estimates along with 95% confidence intervals for the difference in response proportion and the odds ratio of response will be derived from this model. In addition, histogram will be presented.

This analysis will be repeated on the PP population.

4.7.2.5 Motor part of UHDRS® (Total Motor Score, Gait and Dystonia sub-scores)

The motor part of the UHDRS® is assessed at each visit from V0 to V6, including the TMC score and the following other scores:

The **Total Motor Score (TMS)** is made of all 31 items The items are rated from 0 to 4, with 0 indicating normal findings and 4 indicating severe abnormalities. The range of the Total Motor Score (TMS) is 0 to 124, with higher scores indicating more severe motor impairment. If 7 or more item scores are missing for a given assessment, the TMS will not be calculated. If 6 or fewer item scores are missing, the last non missing score for any missing item will be used in calculating the TMS.

The **gait sub-score of the UHDRS** questionnaire measures the severity of gait disturbances in individuals with Huntington's disease and it is made from 3 items: gait, tandem walking and retropulsion pull test. To obtain the gait sub-score all 3 items need to be

summed up. The gait sub-score ranges from 0 to 12, with higher scores indicating greater impairment.

The **dystonia sub-score of the UHDRS** questionnaire is used to evaluate the severity of dystonia in individuals with Huntington's disease and it consists of four items: trunk (involuntary movements of the trunk), right upper extremity (involuntary movements of the right arm and hand), left upper extremity (involuntary movements of the left arm and hand), right lower extremities (involuntary movements of the right leg and feet), left lower extremities (involuntary movements of the left leg and feet). The severity of dystonia is rated on a scale of 0 to 4 for each item, where 0 represents no dystonia and 4 represents severe dystonia. The scores for each item are summed to obtain the dystonia sub-score, which ranges from 0 to 20.

The analysis of the Total Motor Score (TMS) endpoint, **sub-scores of gait and sub-scores of dystonia** will be similar to the analysis of the primary endpoint.

Descriptive statistics for each score by visit will be presented.

The analysis of those endpoints will be based on a mixed-effect linear model for repeated measures (MMRM) to compare each active treatment group to placebo. The model will include the following factors: treatment group, visit, and treatment by-visit interaction as fixed effects, and baseline TMS score/ sub-scores of gait/ sub-scores of dystonia as covariate, depending what is analysed. An un-structured (US) variance-covariance matrix will be used to model the within-patient errors. If no reliable results under the US assumption can be obtained, the Toeplitz (TP) covariance structure, followed by the autoregressive (AR1) and Compound Symmetry (CS) covariance structure will be used. The Kenward-Roger method will be used to estimate the denominator degree of freedom.

The least squares mean (LSM) estimate for each treatment group and difference in LSM estimates at end of maintenance dose between active treatments and placebo group with their 95% confidence intervals will be provided.

For TMS, this analysis will be repeated on the PP population.

4.7.2.6 EQ-5D-5L measuring quality of life

EQ-5D is a standardised instrument for use as a measure of health outcome. The EQ-5D measure is designed in two parts: one descriptive part (EQ-5D) containing five dimensions of health; mobility, self-care, usual activities, pain/discomfort and anxiety/depression (5 categories, each) and a visual analogue scale (EQ-5D VAS) ranging from ‘Best imaginable health state’=100 to ‘Worst imaginable health state’=0.

The EQ5D is rated at the inclusion (Visit 1 at Day 0) and at the end of maintenance dose (Visit 5 at end of Week 10).
The EQ5D ratings at baseline will be compared to those after the treatment by presenting descriptive statistics for the entire questionnaire and for the difference between V1 and V5 and also for the VAS.

EQ5D dimensions and VAS will be analysed using an analysis of covariance (ANCOVA) model including treatment as fixed factor and baseline score as covariate. Results for the

outcome of the primary endpoint will be presented for the least square means estimates and accompanying 95% CIs.

4.8 Pharmacokinetics (PK)

For plasma concentration of SOM3355 and Prolactin plasma concentration measured in all patients at each visit from Visit 1 to Visit 6 (before dosing at each visit and 2 hours after dosing at Visit 2 and Visit 4), analysis will be described in a separate Statistical Analysis Plan and results will be reported in a separate report.

PK analysis for the PK sub-study conducted in a sub-set of 24 patients (8/per arm) with 12-hours PK samples performed at Visit 2 (H0 before dosing, H0.5, H1, H1.5, H2, H3, H5, H8, H12) analysis will be described in a separate Statistical Analysis Plan and results will be reported separately.

4.9 PK-PD assessments

PK-PD analysis is described in a separate Statistical Analysis Plan and results will be reported separately.

4.10 Safety

The Safety Analysis Set (SAF) will be used for the analysis of safety data. SAF includes all patients who were administered at least one dose of any study drug. Patients who are randomized and assigned a treatment but withdrew prior to dosing are not included in the safety population. Patients are included in the treatment group based on the treatment actually received.

Safety endpoints include:

- Incidence, frequency, and severity of AEs in each treatment group.
- Values and change from baseline in:
  - Vital signs: blood pressure (BP) (including orthostatic measurement), and heart rate (HR), weight.
  - ECGs parameters (separately analysed by Banook/Phinc Development)
  - Clinical laboratory test results (chemistry, hematology)
  - Physical examination findings,
  - Concomitant medications
  - C-SSRS assessing suicidality
  - BDI scale assessing depression and suicidality (item #9)
  - ESS assessing sleepiness
  - TFC of the UHDRS® assessing functioning

- FA of the UHDRS® assessing disabilities
- BARS assessing akathisia
- MoCA scale to assess cognitive disturbances
- PBA-s to assess behavioural symptoms in HD.

Tolerability and safety data will be listed by patient (CSR section 16.2) and summarized descriptively by treatment group. Also, differences to Baseline will be presented whenever appropriate.

4.10.1 Adverse events

Adverse events (AEs) will be presented by MedDRA System Organ Class (SOC) and Primary Term (PT). Coding of verbatim terms will be performed by Ergomed and all relevant codes are stored in the database and it will be approved by the Sponsor before database lock (cf. DMP).

Only treatment-emergent adverse events (TEAE) will be analysed, i.e. all new and worsened pre-existing adverse events occurring on or after the initiation of IMP. It is assumed that for each change in intensity, relationship, or seriousness of an AE a new entry of the AE was recorded in the data capture database; hence such cases will be analysed like different phases of the same AE.

An overall summary of TEAEs will be presented including both the number and percentage of patients with a given AE category and the number of events:

- All TEAEs
- Serious TEAEs (STEAEs)
- Treatment-related TEAEs
- Treatment-related Serious TEAEs
- Severe TEAEs
- TEAEs leading to study drug discontinuation
- Fatal TEAEs

A descriptive analysis will be performed. Global incidences of primary system organ classes (SOC) and preferred terms (PT) will be calculated for:

- All TEAEs irrespective of the causality assessment
- Serious TEAEs (STEAEs)
- TEAEs by relationship
- STEAEs related to study drug
- TEAEs by maximum severity
- TEAEs leading to discontinuation
- Fatal TEAEs

These tables will be presented by treatment group and also include summary lines for all SOCs and the occurrence of any event of the specified type. SOC terms will be presented according to the Internationally Agreed Sorting Order (MedDRA).

Multiple counts within a PT or SOC (repeated or different included terms or changes in descriptors) will be counted only once for the calculation of incidences. In addition to the incidences, the number of reported AEs (within a PT or SOC) will be included in the summaries.

Detailed listings of all fatal and other serious events will be provided (for section 14.3.2). All adverse events (i.e. TEAEs as well as non-treatment emergent events) will be listed in section 16.2.

4.10.2 Vital signs

The following vital signs parameters will be assessed at each of the visits (from Visit 0 to Visit 6) as per schedule of assessment (except height only at Visit 0).

- Systolic blood pressure (SBP)
- Diastolic blood pressure (DBP)
- Heart rate
- Body weight

Four sequential measures of Blood Pressure should be performed as follows: sitting position for at least 5 minutes, supine position at least 5 minutes, standing position for 1 minute and 3 minutes. Orthostatic BP is calculated as change from supine to standing position after 1 and 3 minutes.

The following summaries will be provided:

• A summary of the observed absolute values and change from baseline in each vital sign parameter by treatment group and time point separately for sitting, supine, 1 min standing and 3 min standing

• A time profiles of the mean values for each vital sign parameter by treatment group separately for sitting, supine, 1 min standing and 3 min standing with SEM

A summary of the incidence of drop in SBP of ≥ 20 mmHg or drop in DBP of ≥ 10mm Hg will be presented, where drop in BP is calculated as orthostatic BP.

4.10.3 ECG

All ECGs will be performed after at least 5 minutes' rest in a supine or semi-supine position. To assess safety 12-lead ECGs will be recorded at Screening visit (Visit 0), inclusion visit (Visit 1), Visit 2, Visit 3 (in some countries Spain, UK & Germany), Visit 4, and Visit 6. At the inclusion, the ECG will be triplicated i.e., 3 successive records at 5 minutes of intervals and the average of all parameters will be calculated for baseline reference. All ECG are tele transmitted to a central reading centre (Banook) and analyzed within 48 hours by a cardiologist.

ECG analysis is described in a separate Statistical Analysis Plan and results will be reported separately.

4.10.4 Laboratory variables

The following laboratory measurements will be taken at the Visit 0 (screening visit), at Visit 3 (interim visit at the end of Week 6), and at Visit 5 (end of maintenance dose at end of Week 10).

Hematology		
Red blood cell count	Mean corpuscular hemoglobin	Reticulocytes
Hematocrit	Mean corpuscular volume	White blood cell count with differential (neutrophils, lymphocytes, monocytes, eosinophils, basophils)
Hemoglobin	Platelet count	
Clinical chemistry		
Alanine aminotransferase	Aspartate aminotransferase	Alkaline phosphatase
Albumin	C-reactive protein	Calcium
Total Protein	Creatine kinase (CPK)	Phosphate
Bicarbonate	Creatinine	Chloride
Bilirubin, direct and total	Urea	Potassium
Gamma glutamyl transpeptidase	Glucose	Sodium
Other		
Pregnancy test in women of childbearing potential		

Continuous laboratory data:

- Continuous data laboratory safety data (original values and change from baseline) will be summarized using descriptive statistics per visit and treatment group.
- Frequencies of high and low values with respect to the normal range will be displayed by visit and treatment group
- Shift tables comparing each post-baseline visit with baseline with respect to normal range will be displayed by treatment group
- In addition, abnormal values outside reference ranges (with assessment by investigator, NCS/CS) will be summarized in a listing per visit and per patient.

Vertical boxplot of laboratory values per parameter and per visit with outliers will be presented.

The lab values may be transformed to common units before analysis and summaries and analyses will only be performed on values converted to common units. However, original values and units can be found in laboratory listings. If repeated determinations at the same

sampling times were performed (e.g., due to sample clotting) the values of the latest sample will be used in the analysis. However, all measurements will be included in laboratory listings.

The incidence of abnormal values out of ranges will be summarized for laboratory tests using descriptive statistics. These summaries will include all post baseline values (including scheduled, unscheduled, and early termination visits). Listing with outliers greater or lower than upper normal range will be presented.

4.10.5 Physical examination

Physical examination will be performed at each visit. Results of physical examination will be listed by patient and summarised in a frequency table displaying the absolute and relative frequency for normal/abnormal findings for each body system by treatment group and visit.

4.10.6 Concomitant Medications

Concomitant medications, including medications that are taken on an as needed basis and occasional therapies, will be monitored during the study. All concomitant medications will be coded using the WHO drug Global thesaurus.

The incidence of concomitant medications will be summarized using descriptive statistics by WHO Drug Global Preferred Term (INN). Patients are counted only once in each PT category. In addition to the incidences, the number of reported CMs (within PT) will be included in the summaries. Concomitant medications will include all medications taken from the first day of study drug administration up to the end of study as defined in the study protocol.

4.10.7 C-SSRS assessing suicidality

The C-SSRS is a questionnaire performed by trained study personnel that should be done at baseline and at each visit during the study to assess the status of the patient regarding suicide ideation and the risk of suicide attempt. The “Baseline/Screening” C-SSRS form provided at Screening collects the history of suicide attempts, while for subsequent visits the C-SSRS form termed “Since the Last Visit” should be used.

C-SSRS will be assessed at each visit.

The Columbia-Suicide Severity Rating Scale Screening/Baseline Version (C-SSRS-BL) is assessed at Visit 0 and Visit 1. The Columbia-Suicide Severity Rating Scale Since Last Visit (C-SSRS-SLV) is assessed at each visit after randomisation (Visit 2, Visit 3, Visit 4, Visit 5 and Visit 6).

Frequency counts and percentages of the C-SSRS outcomes: Suicidal Behavior or Suicidal Ideation at Baseline, Suicidal Behavior or Suicidal Ideation Post Dosing, and shifts from baseline will be summarized as categorical data using descriptive statistics.

4.10.8 BDI scale assessing depression

BDI is 21-question multiple-choice self-report inventory. The questionnaire is composed of items relating to symptoms of depression such as hopelessness and irritability, cognitions such as guilt or feelings of being punished, as well as physical symptoms such as fatigue, weight loss, and lack of interest in sex.

Each item has four possible degree of severity rated from 0 to 3, and the patient is asked to choose the one that is closest to how he/she feels. The 21 items are summed linearly to create a score which ranges from 0 to 63. It can be administered as an interview by the clinician or as a self-report instrument.

The item #9 assessing suicidality includes the following:

0. "I don't have any thoughts of killing myself."

1. "I have thoughts of killing myself, but I would not carry them out."

2. "I would like to kill myself."

3. "I would kill myself if I had the chance."

The Total Score shows the severity of depression as following:

0 to 9: No depression or minimal depression

10 to 18: Mild to moderate depression

19 to 29: Moderate to severe depression

30 to 63: Severe depression

BDI will be assessed at each visit from V0 to V6.

Descriptive statistics for BDI total score as well as changes to baseline and 95 % confidence intervals will be presented at each visit. In addition, time profiles of the mean values with SEM will be created.

For BDI item #9 frequency table and shift table to baseline will be presented.

4.10.9 ESS assessing sleepiness

The ESS is a self-administered questionnaire comprised of eight questions that provides a measure of a patient's general level of daytime sleepiness. The ESS asks respondents to rate, on a 4-point Likert scale (0 – 3), their usual chances of dozing off or falling asleep in different situations or activities that most people engage in as part of their daily lives. The total ESS score is the sum of 8 item scores and can range between 0 and 24 with a higher score indicating a higher level of daytime sleepiness.

ESS will be assessed at each visit.

Descriptive statistics for the ESS as well as changes to baseline and 95% confidence intervals will be presented at each visit. In addition, time profiles of the mean values with SEM will be created.

4.10.10 TFC of the UHDRS® assessing functioning

The UHDRS®-TFC scale is a tool used by clinicians to assess the stage of HD and the level of functional capacity in patients on five items: patients' ability to work, manage finances, run a house, take care of themselves, and live at home independently. The TFC score ranges from 0 (worse) to 13.

TFC will be assessed at screening (Visit 0), at the inclusion (Visit 1), and at the end of maintenance dose (Visit 5 at end of Week 10).

Descriptive statistics for TFC as well as changes to baseline and 95% confidence intervals will be presented at each visit. In addition, time profiles of the mean values with SEM will be created.

4.10.11 Functional Assessment of the UHDRS®

The UHDRS®-FA scale includes 25 yes/no questions about common daily tasks (total range 0 – 25).

FA will be assessed at screening and at the end of the maintenance dose (Visit 5 at end of Week 10).

Descriptive statistics for the FA as well as changes to baseline and 95% confidence intervals will be presented at each visit. In addition, time profiles of the mean values with SEM will be created.

4.10.12 BARS assessing akathisia

The BARS is administered by healthcare providers to assess the severity of drug-induced akathisia and is scored as follows: Objective Akathisia, Subjective Awareness of Restlessness and Subjective Distress Related to Restlessness are rated on a 4-point scale from 0 – 3 and are summed yielding a total score ranging from 0 to 9. Then Global Clinical Assessment of Akathisia uses a 5-point scale ranging from 0 – 4.

BARS will be assessed at inclusion and at end of the maintenance dose (Visit 5 at end of Week 10).

Descriptive statistics for the BARS as well as changes to baseline and 95% confidence intervals will be presented at each visit. In addition, time profiles of the mean values with SEM will be created.

4.10.13 MoCA scale to assess cognitive disturbances

The MoCA scale is a brief screening instrument recommended for the detection of mild cognitive impairment that is sensitive to executive dysfunction in a variety of neurological conditions. When used in HD, it was more sensitive to detect mild to moderate cognitive impairment.

The items of the MoCA examine attention and concentration, executive functions, memory, language, visuo-constructional skills, conceptual thinking, calculations, and orientation. The attention/executive function items include Trail Making Test B, digit span, target detection, verbal fluency, abstraction, and serial seven subtraction. The visuospatial items include clock drawing and cube copying. The language items include object naming

and sentence repetition. The memory items include recall of five previously presented words. The orientation items include six orientation-based questions.

MoCA scores range between 0 and 30.

MoCA will be assessed at inclusion and at end of the maintenance dose (Visit 5 at end of Week 10).

Descriptive statistics for the MoCA as well as changes to baseline and 95% confidence intervals will be presented at each visit. In addition, time profiles of the mean values with SEM will be created.

4.10.14 PBA-s to assess behavioral symptoms in HD

The PBA short form is the recommended outcome measure for behavioral symptoms in HD. It is a semi-structured interview containing 11 items, each designed to measure the severity and frequency of behavioral neuropsychiatric symptoms that are frequent in HD: depression, suicidal ideation, anxiety, irritability, aggressive behavior, apathy, obsessive-compulsive symptoms, perseverative thinking, paranoid thinking, hallucinations, and dis-orientation.

Severity and frequency during the previous 4 weeks are rated separately for each symptom on a 5 point (0–4) scale. Severity and frequency ratings are multiplied to provide an overall score for each symptom.

Worst severity rating will be only listed.

PBA-s will be assessed at inclusion, at the interim visit (Visit 3 at the end of Week 6), and at the end of the maintenance dose (Visit 5 at the end of Week 10).

Descriptive statistics for the PBA-s as well as changes to baseline and 95% confidence intervals will be presented at each visit.

4.11 Other variables

Not applicable.

4.12 Interim analyses

An unblinded interim analysis for futility (non-binding) will take place when at least half of the total number of planned patients (N = 65) have completed the study or discontinued early. The analysis will be prepared by a separate team that will communicate the results only to designated recipients, e.g., the Steering Committee, to protect the study blind. The primary criterion to assess the futility of the study will be the probability of rejecting the null hypothesis conditional on the accumulated data and the planned sample size. In line with the chosen Holm-Bonferroni multiplicity adjustment, the decision will focus on that hypothesis that yielded the smallest p-value in the interim analysis of the TMC change. A non-binding stop for futility will be recommended if the conditional probability is less than 0.25. Otherwise, the study shall be continued without modification unless otherwise recommended by the SC.

If a decision is made to stop the trial, then a complete, final analysis will be performed including those patients that were enrolled into the study in the meantime.

The analyses will be performed using the SAS® software (Version 9.4 or later). The method and SAS code will be the same as for primary endpoint described in Section 4.6.1 and Appendix 2.

An unblinded statistician/programmer will perform the unblinded futility analysis. The generated documents will be stored on in a separate file area on the file server with restricted access only to the unblinded person(s).

5 QUALITY CONTROL

The SAP was reviewed by the Trial Statistician (TS) before signature. Particularly the TS has checked the consistency of the described methods and outputs with the actual version of the study protocol. In addition, a sponsor representative has reviewed the SAP before final approval.

Log files of all SAS® programs used in the analysis will be checked for errors, warnings and suspicious notes by the statistical programmer. All findings will be either eliminated or commented upon. The final version of each program will be stored along with its log file in the electronic archive.

All programs will be validated by the program author or an independent statistical programmer depending on the requested validation level selected in the List of TLFs form (FRM/BS/001.02) for a particular program.

The agreement of the program outputs with the SAP, their consistency and plausibility will be checked by the TS. Moreover, the TS will review the outputs regarding completeness, readability and comprehensibility.

The described process is associated with the 'normal' level of program validation. Additional levels of quality control can be specified in the List of TLFs (see Appendix, 1) for individual outputs.

6 DERIVATIONS AND TRANSFORMATIONS

6.1 Formulas for derived variables

Variable	Definition / Derivation
BMI	Weight [kg] / (height [m])**2
Year of birth	Year of inform consent - Age
Age at HD diagnosis	Year of HD diagnosis - Year of Birth
Age at first symptoms	Year of first symptoms - Year of Birth
Compliance	(Capsules dispensed - capsules returned) / expected number of capsules to be used
CGI-C improvement	Patients with a score of 1 (very much improved), 2 (much improved) or 3 (minimally improved)

Variable	Definition / Derivation
TMC improvement	change of score between the baseline and the end of maintenance dose is ≥ - 2
PGI-C improvement	patients with a score of 1 (very much improved), 2 (much improved) or 3 (minimally improved)

6.2 Transformations to be applied

Not applicable.

7 REFERENCES

Not applicable.

8 STANDARDS USED IN PREPARATION OF STATISTICAL OUTPUTS

The below conventions will be followed as agreed with the Sponsor.

8.1 Programming

- One SAS program should create only one output.
- One output file can contain different output types (e.g. descriptive and inferential).
- Individual output files will be created in MS Word format (Rich Text Format, RTF).
- Once delivered to the client, numbering of TLFs will not be altered, unless agreed with the client

8.2 Layout

- TLFs will be produced in landscape format
- TLFs will have a minimum 2 cm on every side
- TLFs will be produced using the Courier New font, size 8
- Section numbering of TLFs will follow ICH E3 guideline.
- Numbering of TLFs will follow the convention XXX-YY, where XXX stands for a (sub-)section number of the ICH E3 guideline and YY represents the sequence number of the output within the section. A dash ('-') will always be used to separate section numbers from output sequence numbers
- Titles and footnotes for figures will also be in Courier New font, size 8.
- Tables and listings will be in black and white (no colour), figures may include only colour that can be distinguished when printed on a grey-scale printer
- Text styles, such as bolding, italics, borders, shading, and superscripted and subscripted text, will be used sparingly in the TLFs
- The ANSI character set will be used in the TLFs. Certain subscripts and superscripts (e.g., m^2, AUC_{norm}) will be employed on a case-by-case basis.
- Mixed case will be used for all titles, footnotes, column headers, and programmer-supplied formats, unless they are derived directly from the data

8.3 Headers, Titles and Footnotes

- All output will have the following header at the top left of each page showing the study ID, the date of output generation and an internal pagination, where Y stands for the total number of pages in the pertaining output.

- Also, all TLFs will have the following footer, identifying the generating SAS program (XXX.SAS), a reference to the relevant subject listing and the date of the data snapshot:

```
SAS program: <XXX>.sas          Ref. list X.X.X-YY          Data status: YYYY-MM-DD
```

- Each TLF will bear a title which is repeated on each page of the output.
- The title at the top of the page will be horizontally cantered in bold font.
- A blank line will separate the title from the body of the output.
- The title will consist of an Output number, a descriptive title and a description of the presented analysis set (if applicable).
- The title will have the following general appearance:

```
Table / Figure / List XX.X.X-YY
   Descriptive Title line 1
   Descriptive Title line 2
(All subjects in the FAS, N=nnn)
```

- Each new footnote should start on a new line, where possible.
- Preferably, footnotes should be left justified. When extending beyond a single line, a manual linefeed should be inserted to avoid meaning distortion.
- An automatic footnote '(continued)' will appear at the bottom of TLFs that extend over more than one page.

8.4 General Conventions

- For measured variables column headers should include the unit in their description
- The order of treatment arms in the TLFs will be consistent throughout the entire TLF presentation
- Alphanumeric values are preferably displayed left-justified;
- Dates are presented left-justified
- Integer numbers (e.g., counts) can be centered or right-aligned
- Numbers containing fractional portions will be decimal-aligned.
- Fractional numbers with absolute value less than 1 will carry a leading zero, i.e. 0.123 not .123.
- Units of measured or derived variables will be included where appropriate
- Unless otherwise warranted, the estimated mean, median and quartiles for a set of values will be displayed with 1 more significant digit than the original values, and standard deviations with 2 more significant digits. The minimum and maximum should report the same significant digits as the original values.

- P-values are output in the format: “0.xxxx”, where xxxx is the value rounded to 4 decimal places. P-values less than 0.0001 will be presented as <0.0001.
- Precision of percentages displayed will depend on the total study size. For studies with less than 1000 patients values will be presented with one decimal place. For studies with more than 1000 patients, values will be presented with two decimal places.
- Tabular display of data for medical history, prior/concomitant medications and all tabular displays of adverse event data are generally presented by body system, treatment class, or SOC according to the Internationally Agreed Sorting Order of the MedDRA, unless otherwise agreed.
- The percentage of patients is normally calculated as a proportion of the number of patients assessed in the relevant treatment group (or overall) for the analysis (sub-) population presented.
- For categorical summaries (number and percentage of patients) where a patient can be included in more than one category, an explanatory text will be added to clarify that multiple answers were possible.
- Missing values will be displayed either by a double-dash (“--”) or as “NA” (=’not available/applicable’), whichever is appropriate.
- Dates are displayed in according to ISO date/time format as YYYY-MM-DD, e.g. 2010 03 24. Missing dates may be represented as “NA”, if not available/applicable.
- Clock times are displayed as HH:MM or HH:MM:SS based on 24-hour clock

APPENDICES

1. List of Tables, Listings, Figures

A complete List of tables, listings, figures (TLFs) will be given in a separate document which can be updated without updating the SAP. The List will serve as a reference for both the Sponsor, the TS and the SP and describes the entire set of statistical output to be produced. Therefore, this List will be versioned and approved by both Ergomed and Sponsor before commencing the statistical programming.

Each output page will have an appropriate heading specifying the study ID and abbreviated study title.

Each output page will show a common date and page numbers in the form 'Page [x / y]' where x denotes the current page within an output and y the total number of pages of that output. The output pages will not contain any other sequential page numbering.

All statistical output will identify the underlying analysis set(s) and indicate the number of patients/events in this set (N) and the number of patients/events actually contributing to the particular output (n).

All patient listings will contain in addition to the patient identification the analysis population.