Official Title: A Multi-Site, Prospective, Longitudinal, Cohort Study Measuring Cerebrospinal Fluid-Mutant Huntingtin Protein in Patients With Huntington's Disease

NCT Number: NCT03664804

Document Date: Statistical Analysis Plan Version 6: 19-Sep-2020

STATISTICAL ANALYSIS PLAN

TITLE: **A MULTI-SITE, PROSPECTIVE, LONGITUDINAL, COHORT STUDY MEASURING CEREBROSPINAL FLUID-MUTANT HUNTINGTIN PROTEIN IN PATIENTS WITH HUNTINGTON'S DISEASE**

PROTOCOL NUMBER: BN40422

STUDY DRUG: None

VERSION NUMBER: 4

IND NUMBER: Not applicable

EUDRACT NUMBER: 2018-000632-98

SPONSOR: F. Hoffmann-La Roche Ltd.

PLAN PREPARED BY: [redacted], M.Sc.
[redacted], Ph.D.

DATE FINAL: 08-Jul-2019

STATISTICAL ANALYSIS PLAN AMENDMENT APPROVAL

Date and Time(UTC)	Reason for Signing	Name
19-Sep-2020 18:37:31	Company Signatory	[redacted]

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Changes from Version 1 to Version 2

Baseline definition (in Section 4.4) was now generally defined as prior to lumbar puncture (successful or attempted) rather than prior to a 'successful' lumbar puncture (similarly also for the reference date in Section 4). This change was implemented because it was not straightforward to have a univocal definition of 'successful' lumbar puncture. This was also reflected in the efficacy population, who previously included patients who had a 'successful' lumbar puncture, but now includes all patients who had at least one lumbar puncture (successful or attempted). The efficacy population (in this version of the SAP renamed as Intent-to-Treat population) became the same as the Safety population (see Section 4.1).

Baseline for clinical endpoints has been made explicit (in Section 4.4).

A typo in Table 3 was corrected.

Changes from Version 2 to Version 3

The changes were mainly done to address identified inconsistencies or to increase clarity of the paragraphs.

Section 4.4 clarified further definition of baseline for clinical assessments with multiple predose measurements and for companion assessments.

Sections 4.5.1.6 and 4.5.1.7 corrected the value of the higher score and, as in section 4.5.1.12, corrected the definition of the direction of worsening /improvement for the change from baseline.

In Table 3 "country" was added as covariate where was previously omitted in error.

Section 4.7.3 and 4.7.3 corrected sentence on the use of local ranges for labs values.

Section 4.7.7 changed to aligned CSSR "Improvement on Suicidal ideation" derivation with the other studies in the program.

Changes from Version 3 to Version 4

Section 4.5.1.5 Additional parameter for normalization of the volumetric measurements for the intracranial volume was added.

Section 4.2.1 Analysis for documenting the impact of the COVID-19 pandemic was added.

Section 4.5.1 Additional sensitivity analysis to account for potential impact of the COVID-19 pandemic was added.

Section 4.5.2.1 and section 4.7 A time window for visit to account for potential impact of the COVID-19 pandemic on the visit schedule was added.

Section 4.5.1.15 Table 2 and Section 4.5.2.1 Table 3 adjusted MRI analysis to account for the for the additional normalization parameter.

Section 4.5.1.15 Table 2 added further exploratory analysis of correlation between CSF biomarker and imaging data.

Section 4.5.1.15 HDSI has been removed from the table. The assessment is conducted weekly and remotely. It will be assessed as part of the digital exploratory analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. BACKGROUND

This Statistical Analysis Plan (SAP) describes all planned methods of summarizing and analyzing data to be collected in Study BN40422.

Study BN4022 is a multi-site, prospective, 15-month longitudinal, cohort study measuring cerebrospinal fluid (CSF) mutant huntingtin protein (mHTT) in patients with early manifest Stage I or Stage II Huntington's disease (HD).

This longitudinal study in patients with early HD will explore the prognostic value of baseline CSF mHTT; the relationship of CSF mHTT to other key biological markers in the putative causal pathway; rate and variability of CSF mHTT increase in untreated patients; and the utility of recording clinical outcome measures via sensors contained in the Roche HD mobile app. Since no study medication will be given to the patients, safety assessments will consist of monitoring and recording adverse events (AEs), including serious adverse events (SAEs), related to study procedures.

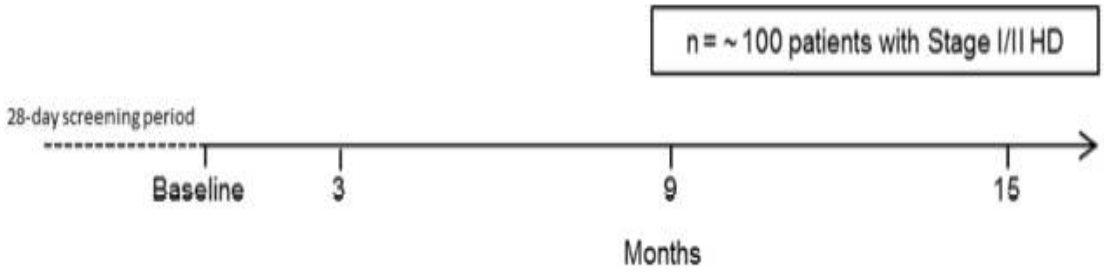
The SAP language supersedes the language in the Protocol and Protocol Synopsis (where specified).

2. STUDY DESIGN

This study will enroll up to approximately 100 patients at approximately 20 sites. After a 28-day screening period, patients will return to the clinic at baseline and at Months 3, 9, and 15. During these visits, patients will receive clinical, magnetic resonance imaging (MRI), and Roche HD mobile app assessments; they will also provide CSF and blood samples. A post-observational follow-up visit will take place 2 weeks after the 15-month longitudinal period and may occur by telephone or in the clinic.

Figure 1 presents an overview of the study design, with 1 month defined as one 28-day period. For further details, see Section 4.2 of the protocol.

Figure 1 Study Schema



HD = Huntington's disease.

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is in Appendix 1. For additional details, see the Schedule of Assessments in Appendix 2.

2.2 ENDPOINTS

Primary and secondary endpoints are provided, additional details are provided below on exploratory endpoints.

2.2.1 Primary Endpoints

The primary objective of this study is to examine the predictive value of baseline cerebrospinal fluid (CSF) mHTT levels on measures of disease progression using the following endpoints:

- Change from baseline in the following clinical endpoints at 3, 9, and 15 months: composite Unified Huntington's Disease Rating Scale (cUHDRS), Total Functional Capacity Scale (TFC), Total Motor Scale (TMS), Symbol Digit Modalities Test (SDMT), Stroop Word Reading (SWR) Test, and Independence Scale (IS)
- Change from baseline in biomarkers of neuronal injury (e.g, CSF neurofilament light chain [NfL] and tau) at 3, 9, and 15 months
- Change from baseline in brain atrophy endpoints (e.g., whole brain volume decline, caudate volume decline) as determined by brain MRI, at 3, 9, and 15 months

2.2.2 Secondary Endpoints

The secondary objective of this study is to investigate the temporal profile of longitudinal CSF mHTT changes within patient, and the association of these changes with corresponding changes in biomarkers of neuronal injury (e.g., CSF NfL and tau) and clinical outcomes using the following endpoints:

- Within-patient change from baseline in CSF mHTT levels at 3, 9, and 15 months
- Association of change from baseline in CSF mHTT levels at 3, 9, and 15 months and:
  - Change from baseline in clinical measures (cUHDRS, TFC, TMS, SDMT, SWR, and IS) at 3, 9, and 15 months.
  - Change from baseline in biomarkers of neuronal injury (e.g., CSF NfL and tau) at 3, 9, and 15 months.
  - Change from baseline in brain atrophy endpoints, as determined by brain MRI at 3, 9, and 15 months.

2.2.3 Exploratory Endpoints

- CGI-S: Clinical Global Impression–Severity
- CGI-C: Clinical Global Impression–Change
- PGI-S: Patient Global Impression–Severity
- PGI-C: Patient Global Impression–Change
- AES: Apathy Evaluation Scale
- EQ-5D-5L: Euro Qol 5 Dimension 5 level Questionnaire
- SMDDS: Symptoms of Mayor Depressive Disorder Scale
- WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0 (12 items scale)
- WPAI: Work Productivity and Activity Impairment
- HD-SDI: Huntington's Disease Speaking Difficulty Item
- MoCA: Montreal cognitive assessment (baseline only)
- Neuro-Qol Cog Func: Neuro Qol Cognitive Function Short form
- Roche HD mobile app digital endpoints:
  - Chorea
  - Speed tapping
  - Symbol digit modality test (SDMT)
  - Stroop Word Reading Test (SWT)
  - 2 minute walking
  - U turn
  - Balance test
  - Draw a shape

  Additional digital endpoints may be considered.

2.2.4 Pharmacokinetic Efficacy Outcome Measures

Not applicable

2.2.5 Safety Outcome Measures

- AEs
- Laboratory parameters (including CSF safety parameters)
- Vital signs
- Columbia-Suicide Severity Rating Scale (C-SSRS)
- ECG

2.3 DETERMINATION OF SAMPLE SIZE

In this study, up to approximately 100 patients are expected to be enrolled across approximately 20 sites. This number is driven by the objective of having a 2:1 match, based on baseline characteristics, to the 46 patients enrolled into the BN40697 open-label extension (OLE) study.

2.4 ANALYSIS TIMING

The final analysis will be conducted once all patients have completed the study as defined in the protocol or withdrawn early from the study and data have been cleaned and verified and the database has been locked.

The sponsor may conduct one or more interim analyses for internal decisions, frontloading activities or in support of clinical development of RO7234292 (see Section 4.9).

3. STUDY CONDUCT

This study will be conducted in full conformance with the ICH E6 guideline for Good Clinical Practice and the principles of the Declaration of Helsinki, or the laws and regulations of the country in which the research is conducted, whichever affords the greater protection to the individual. The study will comply with the requirements of the ICH E2A guideline (Clinical Safety Data Management: Definitions and Standards for Expedited Reporting). Studies conducted in the United States or under a U.S. Investigational New Drug (IND) Application will comply with U.S. FDA regulations and applicable local, state, and federal laws. Studies conducted in the European Union or European Economic Area will comply with the E.U. Clinical Trial Directive (2001/20/EC).

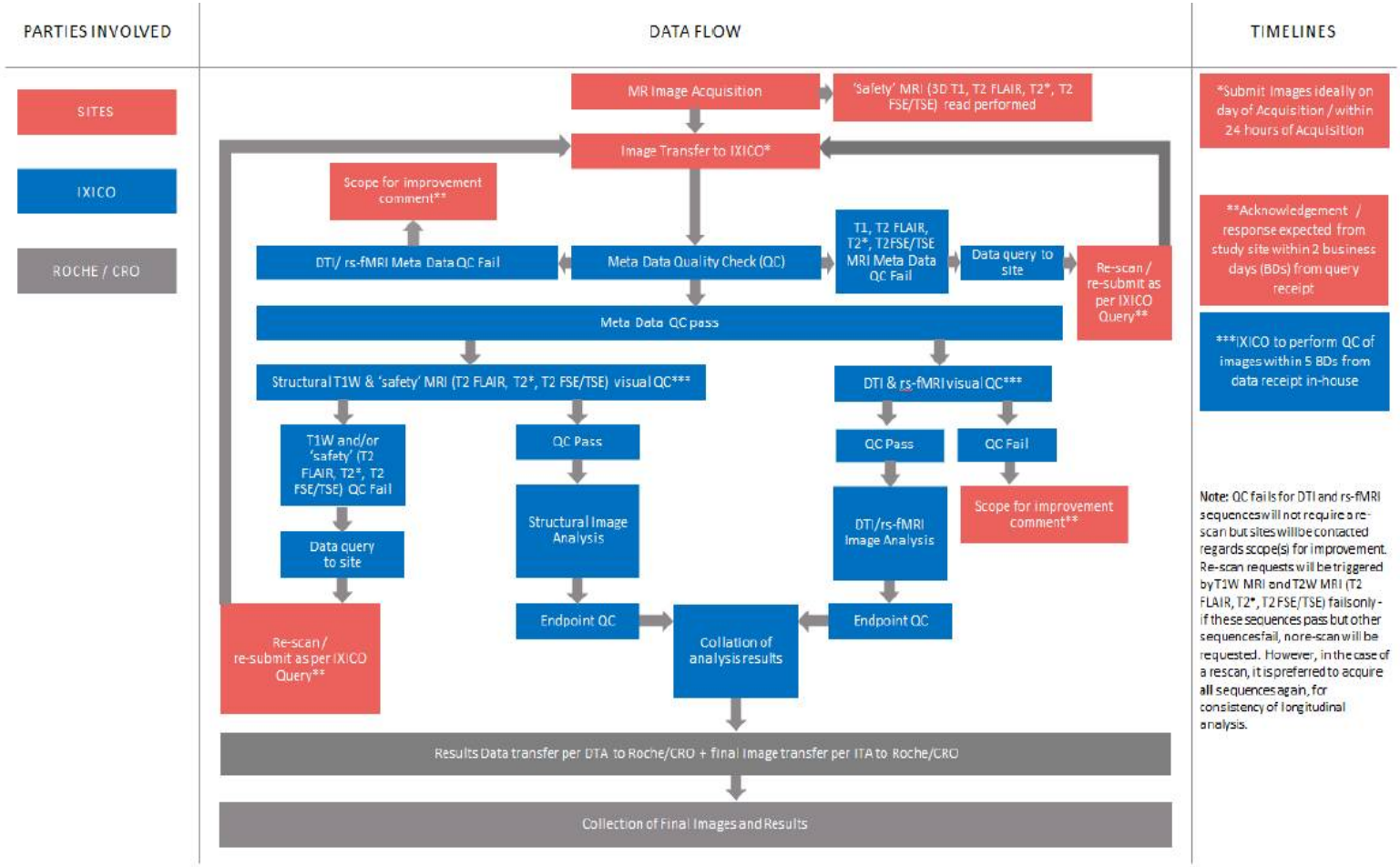
3.1 RANDOMIZATION

There was no randomization in this study.

3.2 INDEPENDENT REVIEW FACILITY

MRI scans will be managed by a central laboratory to monitor and ensure the integrity and quality of the acquired data. Local neuro-radiologists will assess MRI-related study eligibility criteria at screening and for safety monitoring. The sites will carry out MRI acquisition, and data will be transferred to the central laboratory (IXICO). IXICO will then perform QC and quantitative assessment of MRI data which includes volumetric analysis. Details of the roles, processes and analysis are described in a separate charter. A simplified summary of parties involved and data flow diagram is shown in Figure 2.

Figure 2 MRI Data Flow Diagram – A Simplified Summary of Imaging Data Management and Subsequent Reporting of Results



3.3 DATA MONITORING

Not applicable.

4. STATISTICAL METHODS

This is longitudinal prospective cohort study and it is exploratory in nature. As such, there is no predefined hypothesis testing, and all analyses will be regarded as exploratory. Where applicable, parameter estimates will be reported together with 95% CI. p-values may still be reported. Basic assumptions of the proposed analyses will be checked, and data transformation (e.g., logarithm transformation) or other analyses could be carried out, if appropriate.

Any patient data listings will be presented for all patients enrolled into the study.

Descriptive summary statistics including but not limited to n, mean, median, SD, range (minimum, maximum) for continuous variables, and counts and percentages for categorical variables will be used to summarize most data by visit.

Study day will be calculated from a reference date, there is no Day 0, as follows:

- Study Day=assessment date – reference date+1,

  if the assessment date is on or after the reference date; or
- Study Day=assessment date – reference date,

  if the assessment date is earlier than the reference date

The reference date for patients who had at least one lumbar puncture (attempted or successful) is the date of the last lumbar puncture for visit Day 1. For patients who did not have a lumbar punctures (LP) the reference date is the date of enrollment in the study (via the interactive voice or web-based response system [IxVRS] system).

For all analysis, where applicable, the value for age, gender and disease stage as reported in the clinical database as well as the value of CAG repeats as measured by the central lab will be used instead of the values coming from IxVRS system.

For all analysis the clinical cutoff date will be applied for all data, i.e. data occurring after the clinical cutoff date will not be included in the analysis.

The fact that patients in this study are matched to the BN40697 study may offer the future opportunity to compare the current observational study against the OLE study. Details of this analysis are not in the scope of this SAP (neither will be reported in the Clinical Study Report [CSR] of Study BN40422) and may be described in a separate document.

4.1 ANALYSIS POPULATIONS

4.1.1 Safety Population

All patients who have been enrolled in the study and had at least one lumbar puncture (either attempted or successful) will be included in the analysis.

Safety population will be the population for all safety analysis unless stated otherwise.

4.1.2 Intent-to-Treat (ITT)

It will be the same as the Safety Population.

4.2 ANALYSIS OF STUDY CONDUCT

Population for analysis of study conduct will be all enrolled patients unless stated otherwise. All major protocol deviations will be listed and summarized. Number of patients enrolled will be summarized per center and country level. Reasons for discontinuation from study will be listed and summarized.

Time on study (in days) will be calculated as end of study visit (or withdrawal) date – reference date+1 (for definition of reference date see Section 4). Summary tables for time on study will be presented.

For each of the outcome assessments (patient-, observer- and clinician-reported, and performance) the number and proportion of patients (or observer where applicable) who completed the assessment (among those expected to complete the assessment) will be summarized at each time point. For this analysis the ITT population will be used but for safety assessments (e.i. C-SSRS), where the safety population will be used instead. For the evaluation of completeness of each outcome assessment please refer to each subsections within Section 4.5.1 and Section 4.7.7.

4.2.1 Analysis for documenting the impact of the COVID-19 pandemic

Additional analysis may be required to quantify the impact on study conduct/ compliance due to COVID -19 pandemic and to assess their potential impact on the analysis of study endpoints.

Missed and partial assessments are captured as major protocol deviations. Major protocol deviation related to COVID-19 will be listed and summarized.

The number of patients who, due to COVID-19, did not have the assessment done (e.i. LP, MRI and cUHDRS assessments), or had the assessment done outside the specified protocol time window from the planned visit, will be summarized by visit and assessment where applicable (ITT population). The delay in the timing of the scheduled visit may be summarized for selected endpoint of interest.

The number of subjects who discontinued due to COVID-19 will be listed.

Additional TLG may be produced to assess the impact of different modality of collection of eCOA or different MRI centers if applicable.

4.3 ANALYSIS OF BASELINE CHARACTERISTICS

Demographic and baseline characteristics will be listed and summarized overall using means, SD, medians and ranges for continuous variables and proportions for categorical variables, as appropriate. ITT population will be used for summary analysis.

The following baseline variables will be included: age, sex, disease stage, race/ethnicity number of CAG repeats, derived CAP score, educational level, TFC, TMS and overall cognitive status measured by the Montreal cognitive assessment (MoCA). If a study companion was available and consented to take part in the study outcome assessments, then demographic and social status data for the companion will be listed and summarized.

CAP score will be derived as:

- CAP=(CAG-33.66)* AGE_BASELINE

Medical history will be listed and summarized.

4.3.1 Montreal Cognitive Assessment

The Montreal Cognitive Assessment is an 11-item clinician-reported assessment that is used to detect cognitive impairment. It contains a series of basic assessments, including attention and visuospatial tasks. The total score ranges from 0–30, where lower scores indicate greater impairment. A scoring ≥ 26/30 is considered normal. This assessment is only conducted at screening. Individual items contributing to the total score are listed below:

- Visuospatial /executive (0 to 5)
- Naming (0 to 3)
- Attention (0 to 6)
- Language (0 to 3)
- Abstraction (0 to 2)
- Delayed recall (0 to 5)
- Orientation (0 to 6)

The total score is given by the sum of all sub scores; one point will be added by the assessor for subject who has 12 years or fewer of formal education, for a possible maximum of 30 points.

Missing values will not be imputed and total score derived only if all items are present. The assessment is considered completed if there are no missing items.

Descriptive summaries of the baseline values will be provided for the overall score (as continuous variable).

4.4 BASELINE DEFINITIONS

For vital signs (blood pressure [BP], heart rate [HR], respiration rate by position and temperature), baseline will be defined as the average of the available values collected prior to first LP (attempt or successful lumbar puncture) (Screening, Study Day −1 and Study Day 1).

For CSF baseline will be defined as the last available CSF sample collected (Study Day −1 or Study Day 1).

For clinical endpoints (patient-, observer-, clinician-reported and performance outcomes) measured multiple times before first lumbar puncture (successful or attempt), baseline will be defined as the first available measure prior to the first lumbar puncture (successful or attempt) (Screening, Study Day −1 and Study Day 1). Otherwise the only measurement before the first LP will be used. For the observer reported outcomes any measurement occurring on the same day of the first LP, irrespective if it was before or after the LP, can be used for the evaluation of the baseline.

For the Columbia Suicide Severity Scale questionnaire see Section 4.7.7 .

For all other measures and parameters, baseline will be defined as the last non-missing measure prior to last lumbar puncture (successful or attempted) (Screening, Study Day -1 and Study Day 1), unless specified otherwise.

4.5 EFFICACY ANALYSIS

Clinical endpoints, biomarkers and pharmacodynamic parameters considered for analysis are outlined in the sub-sections below. The population to be used for the analysis is the ITT evaluable population unless stated otherwise.

Given the exploratory nature of HD app digital endpoints, analyses similar to the one described below will be performed for digital endpoints as listed in Section 2.2.3 (additional digital endpoints may be considered), with the objective of assessing the longitudinal changes over time and the relationship with clinical endpoint and biomarkers. Details of those analyses will be described in a separate SAP.

4.5.1 Longitudinal Changes over Time

The objective of this section is to describe the longitudinal changes over time, to understand disease progression, for all different clinical, digital and PD endpoints, while adjusting for known prognostic factors (such as CAG repeats). Data outside a time window of +- 56 days from the expected scheduled visit will not be used for analysis of summary tables and mean plots by visit.

For the primary and secondary endpoints descriptive tables will be used to summarize the different variables on absolute value, change from baseline and percentage change (where appropriate) from baseline by visit.

For continuous variables, spaghetti plots showing patient profiles and boxplots across time will be presented.

Where the numeric variable indicates an ordinal, rather than continuous, outcome, the analysis will focus on frequency counts and percentages by visit; no other summary statistics will be provided.

In order to model the change from baseline on each individual continuous parameter across time, a mixed-model repeated measures (MMRM) analysis will be applied. This model will include as covariates the baseline parameter, number of CAG repeats, baseline age, CAP, country (as categorical, only for outcome assessments [patient-, observer- and clinician-reported, and performance]) and time (as categorical). Data outside a time window of +- 56 days from the expected scheduled visit will not be used for this analysis. Estimates for the change from baseline based on the MMRM model, at each time-point and computed for the baseline average value of the other baseline covariate, along with corresponding standard error, 95% CI and unadjusted p-value will be presented. Technical details of the MMRM model (e.g. methods for the denominator degree of freedom, variance-covariance matrix, etc.) are given in Appendix 3.

Although confidence intervals and p-values will be reported, no formal hypothesis testing will be considered.

A detailed list of outputs produced for each parameter for describing the longitudinal change over time is described in Table 2.

The procedure that will be used to fit mixed-models in SAS is as follows (with Day 421 as example):

```
proc mixed data=dataname plots=all;
class subjid time Country;
model y-variable=Baseline_y-variable + CAG + CAP + Baseline_Age + Country + time;
repeated time / subject= subjid type=UN;
Estimate "Day 421"
        Intercept 1
        Baseline_y-variable &BASEY
        CAG &CAGBL
        CAP &CAPBL
        Baseline_Age &AAGE
        time 0 0 1
/ e cl alpha=0.05

run;
```

where &BASEY, &CAGBL &CAPBL and &AAGE will be fixed at the population average at baseline. Time contrast coefficients will be verified with the actual data and

thus the final version of the SAS code may differ from the example above depending on the ordering of the different levels.

Notice that the country variable will only be included when modelling outcome assessments [patient-, observer- and clinician-reported, and performance], pooling country across region (North America -USA plus Canada- vs EU) may be considered in case of convergence problems.

In case a significant number of visits will be delayed beyond the +-56 days, the following additional sensitivity analysis will be conducted, using all data available (irrespective of visit delay).

In order to model the change from baseline on each individual continuous parameter across time, a linear mixed effect model analysis will be applied. The model will include a random intercept and a random slope for the time of assessment (as continuous) . Additionally it will include as fixed effect the time of assessment (as continuous) , the baseline parameter, number of CAG repeats, baseline age, CAP, country (as categorical, only for outcome assessments [patient-, observer- and clinician-reported, and performance]). An unstructured variance covariance matrix will be used for the two random effects (in case of convergence issues, the two random effects will be assumed to be independent).

Least squares means (LSM) estimates for the change from baseline based on the model, at the time of interest, along with corresponding standard error, 95% CI and unadjusted p-value will be presented. The estimate for the fixed effect of time of assessment (population level slope) will be reported along with corresponding 95% CI. For each parameter, all available time-points will be used.

No formal hypothesis testing will be considered and results are presented descriptively with confidence intervals

A detailed list of outputs produced for each parameter for describing the longitudinal change over time is described in Table 2.

The procedure that will be used to fit mixed-models in SAS is as follows with day 421 as example:

```
proc sort data=dataname;
by subjid time;
run;

proc mixed data=dataname plots=all;
class subjid Country;
model y-variable= Baseline_y-variable + Baseline_CAG + Baseline_CAP +
Baseline_Age + Country + time/Solution ;
random Intercept time/ sub= subjid type=un s;
Estimate "Day 421"
        Intercept 1
```

```
        Baseline_y-variable &BASEY
        CAG &CAGBL
        CAP &CAPBL
        Baseline_Age &AAGE
        time 421
/ e cl alpha=0.05
run;
```

where &BASEY, &CAGBL &CAPBL and &AAGE will be fixed at the population average at baseline.

Notice that the country variable will only be included when modelling outcome assessments [patient-, observer- and clinician-reported, and performance]; pooling country across region (North America -USA plus Canada- vs EU) may be considered in case of convergence problems.

Linearity over time will be assess and the linear mixed effect model will only be applied to endpoint with a monotonic pattern of change over time.

4.5.1.1 Mutant Huntingtin protein in CSF

For CSF mHTT, all statistical modeling will be based on the natural log-transformed parameter. Results on the modelling for the change from baseline will be back-transformed to derive the corresponding percent change in the geometric mean as explained in Appendix 3.

Values below the lower limit of quantitation (BLQ) will be imputed as [The BLQ value]*0.5. In case a sample has been assessed on more than one date then the value corresponding to the last measurement date will be used.

Descriptive tables and plots as well results from the MMRM model will be provided as described in Section 4.5.1 and Table 2.

4.5.1.2 CSF/Plasma Biomarkers

Assessments of CSF biomarkers will include the following parameters:

- NfL
- Tau
- YKL-40

Additional parameters may be explored.

NFL is the only of the above biomarkers which is also assessed in plasma.

For CSF/plasma biomarkers (e.g. NfL), all statistical modeling will be based on the natural log-transformed parameter. Results on the modelling for the change from

baseline will be back-transformed to derive the corresponding percent change in the geometric mean as explained in Appendix 3.

Values below the lower limit of quantitation (BLQ) will be imputed as [The BLQ value]*0.5. In case a sample has been assessed with more than once then the value with the last measurement date will be used.

Descriptive tables and plots as well results from the MMRM model will be provided as described in Section 4.5.1 and Table 2.

4.5.1.3 Huntington's Disease Rating Scale

The Composite Unified Huntington's Disease Rating Scale (cUHDRS) (Schobel et al. 2017) is defined as a combination of the following first 4 individual assessments:

4.5.1.3.1 Total Functional Capacity

The Total Functional Capacity (TFC) represents the Investigator's assessment of the patient's capacity to perform a wide range of activities of daily living including working, chores, managing finances, eating, dressing and bathing. It is based on a brief interview with the patient and, potentially, the trial partner. Individual item scores range from 0 to 3 (by 1), total score ranges from 0 to 13 and higher scores represent better functioning. The total score is the sum of the individual item scores.

The scale is considered completed if there are no missing values. There will be no imputation for missing values. The total score will be derived only if each of the individual item scores is not missing.

The individual items contributing to the total score are listed below:

- Occupation (0 to 3)
- Finances (0 to 3)
- Domestic chores (0 to 2)
- ADL (0 to 3)
- Care level (0 to 2)

4.5.1.3.2 Total Motor Scale

The Total Motor Scale (TMS) is the sum of the individual motor ratings obtained during administration of the motor assessment portion of the UHDRS. Individual item scores range from 0 to 4 (by 1), total score ranges from 0 to 124 and higher scores represent more severe impairment. Individual items contributing to the total score are listed below:

- Ocular Pursuit (two items: horizontal and vertical)
- Saccade Initiation (two items: horizontal and vertical)
- Saccade Velocity (two items: horizontal and vertical)

- Dysarthria
- Tongue Protrusion
- Finger Taps (two items: left and right)
- Pronate/Supinate-Hands (two items: left and right)
- Luria – Fist-Hand-Palm Sequencing
- Rigidity – Arms (two items: left and right)
- Bradykinesia - Body
- Maximal Dystonia (five items: trunk, RUE, LUE, RLE, LLE)*
- Maximal Chorea (seven items: face, bol, trunk, RUE, LUE, RLE, LLE)*
- Gait
- Tandem Walking
- Retropulsion Pull Test

(*: RUE= right upper limb, LUE= left upper limb, RLE= right lower limb, LLE= left lower limb)

The scale is considered completed if there are no missing values.

For each of the scores (total or domains) if any of items is missing then there will be no imputation for missing data and the score will not be computed.

In addition to TMS total score, the ‘maximal chorea in upper limb’ item (left and right; from TMS) and the ‘maximal chorea score’ will also be considered. The maximal chorea score is the sum of the seven maximal chorea items (from TMS).

4.5.1.3.3 Symbol Digit Modalities Test

The Symbol Digit Modalities Test (SDMT) is used to assess attention, visuoperceptual processing, working memory and psychomotor speed. The patient must pair abstract symbols with specific numbers according to a translation key. The test measures the number of items correctly paired (maximum of 110) in 90 seconds.

The scale is not considered completed if there is no result available. There will be no imputation for missing data even if the patient is not able or refuses to complete the task.

4.5.1.3.4 Stroop Word Reading Test

The Stroop Word Reading Test (SWRT) is a measure of processing and psychomotor speed. Patients are presented with a page of color names printed in black ink and are asked to read aloud as many words as possible within a given amount of time (45 seconds in Study BN40422). The number of words read correctly is counted and higher scores indicate better cognitive performance. The raw score will be used for the analysis.

The scale is not considered completed if there is no result available. There will be no imputation for missing data even if the patient is not able or refuses to complete the task.

4.5.1.3.5 cUHDRS

Based on the aforementioned assessments, the following formula is used to derive the cUHDRS where TFC, TMS, SDMT and SWRT are the raw total scores of each assessment:

$$\mathbf{cUHDRS} = \left(\frac{\mathrm{TFC} - 8.8}{2.8}\right) - \left(\frac{\mathrm{TMS} - 34.4}{17.4}\right) + \left(\frac{\mathrm{SDMT} - 25.2}{12.4}\right) + \left(\frac{\mathrm{SWRT} - 58}{21.2}\right) + 10.0$$

The cUHDRS will be derived only if each of the individual components is not missing. No imputation for missing values will be conducted. The minimum possible score is -3.06. The maximum is not defined.

Descriptive tables and plots as well results from the MMRM model will be provided as described in Section 4.5.1 and Table 2 for cUHDRS, its individual components total raw scores, the 'maximal chorea in upper limb' item (left and right; from TMS) and the 'maximal chorea score'.

4.5.1.4 Independence Scale

The patient's Independence Scale (IS) is the Investigator's assessment of the patient's degree of independence. The scale consists of 19 discrete levels ranging from 10 to 100 (by 5) where no special care needed corresponds to a scale of 100 and tube fed and total bed care corresponds to a scale of 10.

The scale is not considered completed if there is no result available. There will be no imputation for missing data.

Descriptive tables and plots (as continuous score) as well results from the MMRM model will be provided as described in Section 4.5.1 and Table 2.

4.5.1.5 MRI Volumes and Boundary Shift Integrals

A 3D T1-weighted structural MRI scan will be used to quantitate whole brain, caudate and intraventricular volumes. The only absolute measures are the structural volumes (baseline), while boundary shift integral (BSI) are used to assess changes in volumes over time. Three boundary shift integrals will be considered: K-means ventricular boundary shift integral (VBSI), K-means caudate boundary shift integral (CBSI) and K-means brain boundary shift integral (KNBBSI). To be noticed is that BSI values represent already a measure of change from baseline.

The affine scaling factor will be used for normalizing volumetric imaging values (both absolute volumes and BSI), in order to remove the confounding influences of the head size of the patients. Normalized versions are derived as follow:

VOLn = VOL / ASF

BSIn = BSI / ASF

where ASF= Affine Scaling Factor and VOL= absolute structural volumes

Normalization is not required for BSI percent change from baseline (or annualized percent change) given that:

BSI_percent-change = BSI / VOL * 100 = BSIn / VOLn * 100

Only data with a QC flag equals to 'Pass' will be used in the analyses.

Descriptive tables and plots as well results from the MMRM model will be provided as described in Section 4.5.1 and Table 2.

4.5.1.6 Clinical Global Impression-Severity and Clinical Global Impression-Change

The CGI-S is a single-item measure of current global severity. The CGI-S is assessed using an 11-point numeric rating scale (NRS) where higher scores indicate greater severity (0=“not at all severe” to 10=“Extremely severe”, by 1).

The CGI-C is a single-item measure of change in global status (from baseline). The CGI-C has seven response options: "very much worse," "much worse," "minimally worse," "no change," "minimally improved," "much improved," and "very much improved". To assess the relevance of this change, a follow-up question with dichotomous response options (“yes” or “no”) asks if the change has had a meaningful impact on the patient’s wellbeing.

There will be no imputation for missing data

Descriptive tables and plots for CGI-S (as continuous) and CGI-C (as categorical) will be provided as described in Section 4.5.1 and Table 2.

For CGI-C descriptive statistics of the number and percentage of patients with each score will be summarized by visit.

For CGI-S change from baseline in will be summarized at each visit, where change from baseline is calculated as the post-baseline score minus the score assigned at the baseline assessment time point; in addition, negative values will be mapped to a category of “Improvement ”, zero values will be mapped to “No Change”, and positive values will be mapped to “Worsening”. The number and percentage of patients within each of these categories will be summarized by visit.

4.5.1.7 Patient Global Impression-Severity and Patient Global Impression Change

The PGI-S is a single-item measure of current global severity. The PGI-S is assessed using an 11-point NRS, where higher scores indicate greater severity (0=“not at all severe” to 10=“Extremely severe” by 1).

The PGI-C is a single-item measure of change in global status (from baseline). The PGI-C has seven response options: "very much worse," "much worse," "minimally worse," "no change," "minimally improved," "much improved," and "very much improved". To assess the relevance of this change, a follow-up question with dichotomous response options (“yes” or “no”) asks if the change has had a meaningful impact on the patient's wellbeing.

Descriptive tables and plots for PGI-S (as continuous) and PGI-C (as categorical) will be provided as described in Section 4.5.1 and Table 2.

For PGI-C descriptive statistics of the number and percentage of patients with each score will be summarized by visit.

For PGI-S change from baseline will be summarized at each visit, where change from baseline is calculated as the post-baseline score minus the score assigned at the baseline assessment time point; in addition, negative values will be mapped to a category of “Improvement”, zero values will be mapped to “No Change”, and positive values will be mapped to “Worsening”. The number and percentage of patients within each of these categories will be summarized by visit.

4.5.1.8 World Health Organization Disability Assessment Schedule (2.0)

The WHODAS 2.0 (12-item version) assesses disability and aspects of health related quality of life (HRQoL). It is intended for use in any disease population, and it is linked to the concepts in the International Classification of Functioning, Disability, and Health.

The 12-items version contains 12 items scored on a 0-4 Likert type scale, where higher scores indicate greater difficulty. The total raw score is the sum of the 12 scored items, which is then converted onto a 0–100% point scale (by dividing by 48 and multiplying by 100; Manual for WHO Disability Assessment Schedule, WHODAS 2.0).

There are 3 additional questions that are not part of the total raw score and are administered to assess overall difficult and activity impairment measured as number of days of inability over the past 30 days (values from 0 to 30).

The WHODAS 2.0 will be completed by both the patient and study companion (if available, about the patient) at baseline and post-baseline clinic visits on an electronic device.

Missing values will not be replaced and total score derived only if all items are present. The assessment is considered completed if none of the core 12 items are missing.

Descriptive tables and plots will be provided as described in Section 4.5.1 and Table 2 for the patients and study companion versions for the overall converted score (as continuous variable).

4.5.1.9 Work Productivity and Activity Impairment

The WPAI contains 6 items assessing the impact of disease on employment status (yes/no), hours missed due to disease, hours missed due to other reasons, hours worked, and impact on productivity and on daily activities (both using an 11-point NRS, from 0 to 10, where higher scores indicate greater impact). The following sub-scores will be computed (WPAI Scores):

- Percent work time missed due to problem: Q2/(Q2+Q4)
- Percent impairment while working due to problem: Q5/10
- Percent overall work impairment due to problem: Q2/(Q2+Q4)+[(1−(Q2/(Q2+Q4)))x(Q5/10)]
- Percent activity impairment due to problem: Q6/10

The WPAI will be completed by the patient and study companion (if available, about him/herself). It will be completed monthly by the patient remotely on the Roche HD mobile app and by the study companion remotely on an electronic device.

Missing values will not be replaced and scores derived only if the items used to compute the score are present. The assessment is considered completed if there are no missing items. The assessment will be used for informing pharmaco-economic evaluations and will not be reported in the CSR.

4.5.1.10 Apathy Evaluation Scale

The AES is an 18-item assessment of apathy, including overt behavior, cognitive aspects of motivation, and emotional responsivity. Each item is scored on a 4-point Likert scale, from 1 ("Not at all") to 4 ("A lot").

A total score is created by summing the 18 items (scores range from 18 to 72; 3 of the items [tem 6 “little effort”, item 10 “Tell me what to do” and item 11 “Less concern than should be”] are reverse scored), with higher scores indicating greater apathy.

Missing values will not be replaced and total score derived only if all items are present. The assessment is considered completed if there are no missing items.

Descriptive tables will be provided as described in Section 4.5.1 and Table 2 for the patients and study companion versions for the overall score (as continuous variable) at baseline.

4.5.1.11 Neuro-Qol Cognition Function Short Form, Version 2

The Neuro-QoL Cognition Function Short Form contains 8 items (including "trouble concentrating" and difficulty "learning new tasks or instructions"), each assessed using a 5-point Likert scale, from 1 to 5, where lower scores indicate greater difficulty (4 items) or greater frequency (4 items). The raw sum score is the sum of each individual item (range 8 to 40). The raw sum score is then converted to a T-score distribution (mean = 50, SD = 10). The conversion table from raw score to T score is provided in Table 1.

For short forms with at least 5 items, if a patient skips one or more questions, a prorate score can be approximated if 4 (50%) of items have been answered. If enough responses are provided, the sum of the response scores from the items that were answered is multiplied by the total number of items in the short form and then divided by the number of items that were answered. If the result is a fraction, it should be rounded up to the nearest whole number. This is a pro-rated raw score. However, this prorated score should be used with caution as the advantages of IRT calibrations and their contribution to precision is lost in the process.

The pro-rated raw score is then converted to a normalized T score using the table below (Table 1).

Table 1 Conversion table from raw score to T score

Cognitive Function v2.0 8-item Short Form (Adult)					
Raw Score	T-Score	SE	Raw Score	T-Score	SE
8	17.3	4.3	25	39.9	2.6
9	20.4	3.8	26	40.9	2.6
10	22.6	3.5	27	41.9	2.6
11	24.4	3.3	28	42.9	2.6
12	25.9	3.1	29	43.9	2.7
13	27.3	3	30	44.9	2.7
14	28.6	2.9	31	46	2.7
15	29.8	2.8	32	47.1	2.7
16	30.9	2.7	33	48.3	2.8
17	32	2.7	34	49.6	2.8
18	33	2.6	35	50.9	2.9
19	34	2.6	36	52.4	3.1
20	35	2.6	37	54.2	3.3
21	36	2.6	38	56.3	3.7
22	37	2.6	39	59	4.2
23	37.9	2.6	40	64.2	5.7
24	38.9	2.6			

Source: http://www.healthmeasures.net/score-and-interpret/calculate-scores/scoring-instructions , http://www.healthmeasures.net/images/neuro_qol/Neuro_QOL_Scoring_Manual_Mar2015.pdf).

Descriptive tables and plots for the T score (as continuous) will be provided as described in Section 4.5.1 and Table 2.

4.5.1.12 Huntington's Disease—Speaking Difficulty Item

The HD-SDI is a single question assessing difficulty speaking over the past 7 days. It is assessed using a 5-point Likert scale, where higher scores indicate a greater frequency of difficulty.

The scale is not considered completed if there is no result available. There will be no imputation for missing data.

4.5.1.13 Symptoms of Major Depressive Disorder Scale

The SMDDS is a newly developed, valid, and reliable self-report assessment of depression (McCarrier et al. 2016). It contains 16 items, measuring concepts such as sadness, irritability, worry, and sleep disturbance. Each item is assessed on a 5-point Likert scale, from "Not at all" (0), to "Extremely" (4) (9 items) and from "Never" to "Always" (7 items). Item scores from 15 of the items (the least severe of the two eating items is dropped) are summed to create a 0–60 score, where higher scores indicate more severe depressive symptomatology.

Missing values will not be replaced and total score derived only if all items are present. The assessment is considered completed if there are no missing items.

Descriptive tables and plots will be provided as described in Section 4.5.1 and Table 2 for the overall score (as continuous variable).

4.5.1.14 EuroQol 5-Dimension, 5-Level Questionnaire

The EQ-5D-5L is a validated self-report health status questionnaire used to calculate a health status utility score for use in health economic analyses (EuroQol Group 1990; Brooks 1996; Herdman et al. 2011; Janssen et al. 2013). There are two components to the EQ-5D-5L: a 5-item health state profile that assesses mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, as well as a visual analogue scale (VAS) (range from 0 worst - to 100 best) that measures health state. Published weighting systems allow for creation of a single composite score of the patient's health status (Index score) from the 5-item scores (i.e., does not include the VAS).

Missing values will not be imputed, the composite score will only be computed if all questions have been answered.

The Index score will be used for informing pharmaco-economic evaluations. The VAS score will be used to assess HRQoL. The EQ-5D-5L index score will not be reported in the study CSR, VAS score will be summarized (as continuous variable) by study visit as described in Section 4.5.1 and Table 2.

The EQ-5D-5L will be completed both by the patient (about him/herself) and by the study companion (if available, about him/herself) at baseline and post-baseline clinic visits on an electronic device. In addition, it will be completed weekly by the patient remotely on the Roche HD mobile app and weekly by the study companion remotely on an electronic device.

4.5.1.15 Summary of Outputs for Describing the Longitudinal Changes Over Time

Table 2 Overview of Planned Outputs to Describe Parameter Longitudinal Changes Over Time

Parameter	Output Presentation			Data Type			Time points	MMRM Performed
	Plot	Table	Listing	Absolute	Change	%Change		
mHTT	Y	Y		Y	Y	Y	All available	Yes
CSF Biomarkers§	Y	Y		Y	Y	Y	All available	Yes
Plasma NFL	Y	Y		Y	Y	Y	All available	Yes
BSI	Y	Y			Y*	Y‡	All available	Yes
normalized BSI	Y	Y			Y*	Y‡	All available	Yes
BSL MRI Volume		Y		Y			BSL	No
BSL normalized MRI Volume		Y		Y			BSL	No
UHDRS+	Y	Y		Y	Y	Y	All available	Yes, except for IS
CGI-S		Y		Y	Y	Y	All available	No
CGI-C		Y			Y*		All available	No
PGI-S		Y		Y	Y	Y	All available	No
PGI-C		Y			Y*		All available	No
WHODAS 2.0		Y		Y	Y		All available	No
AES		Y		Y	Y		All available	No
Neuro-Qol Cog. Funct.		Y		Y	Y		All available	No
SMDDS		Y		Y	Y		All available	No
EQ-5D-5L (VAS only)		Y		Y	Y	Y	All available	No

Table 2 Overview of Planned Outputs to Describe Parameter Longitudinal Changes Over Time (cont.)

Absolute=measured data at a given time point; BSL=Baseline

Change (%Change)=change from baseline from a given post-baseline measurement (expressed as a percentage)

* BSI, CGI-C and PGI-C are already a measure of change thus the change is not derived as for the other variables)

ǂ %-Change for BSI is derived as $\left(\frac{BSI}{baseline\ volume}\right) \cdot 100$

§ CSF Biomarkers will include NFL, YKL-40 and Tau

\+ The parameters considered are cUHDRS, its individual total raw score components as well the Independence Scale(IS), Maximal chorea score and Maximal chorea of upper limb

4.5.2 Relationship Between Selected Parameters

The objective of this section is to investigate the relationship between selected pairs of parameters, e.g. CSF mHTT and cUHDRS, to assess:

- the prognostic value of baseline CSF mHTT for the change in cUHDRS at pre-specified time-points
- the relationship between change in CSF mHTT and change in cUHDRS at pre-specified time-points

Both descriptive statistics (scatter plots) and statistical modelling will be applied to the data. All analyses in this section will be based on the ITT analysis population.

4.5.2.1 Descriptive Analyses

Scatter plots between selected pairs of continuous variables will be presented based on data types shown in Table 3 for X (main predictor, e.g. CSF mHTT) and Y (response, e.g. cUHDRS) variables at each time point separately. At each time point only those data from visit occurring ± 56 days from the scheduled visit will be used.

4.5.2.2 Model-based Analyses

In order to model the relationship between a continuous response variable (e.g. the change in cUHDRS) and a main predictor (e.g. change in CSF mHTT) at a given time-point, a linear regression model (LRM) will be used. At each time point only those data from visit occurring ± 56 days from the scheduled visit will be used.

The full list of analyses which will be considered, together with specifications about additional covariates to account for, is provided in Table 3. Additional analyses, not listed in this SAP, may be conducted to better understand any safety concern which may arise.

Least squares means (LSM) estimates, at each time-point, for the slope of the main predictor (X-variable in Table 3) along with corresponding standard error, 95% CI and unadjusted p-value will be presented. The estimated effect represents the effect of the main predictor on the response adjusted for all other covariates in the model. Distributional assumptions underlying the model used for analysis will be examined to gain confidence that the model assumptions are reasonable and where needed, sensitivity analyses will be performed.

The procedure that will be used to fit linear regression in SAS is as follows:

```
proc mixed data=dataname plots=all;
class
model y-variable= x-variable +  {additional covariates in Table 3}/ddfm=kr cl;
ods output solutionf=outputname;
run;
```

In addition to Table 3, scatter plots and Spearman correlations will be calculated among all baseline fluid biomarkers at baseline. Similar analyses will be performed for CAG and CAP with baseline fluid biomarkers, baseline CSF mHTT and baseline cUHDRS (including individual components). CAP is defined by the product of (CAG – 33.66) by baseline age.

Table 3 Overview of Planned Outputs to Describe the Association Between Selected Pairs of Parameters at Final analysis.

Paired Variables	Data Type				Time points	Additional covariates in statistical models
	X-variable		Y-variable			
	Absolute	Change	Absolute	Change		
X=**mHTT**, Y=**Clinical Endpoints**§	Y		Y		X & Y= BSL	BSL-Age, CAG, CAP, Country
	Y			Y	X=BSL, Y=15M	BSL- Clin.endp, BSL-Age, CAG, CAP, Country
		Y		Y	X =3M, 9M, 15M, Y=15M	BSL- Clin.endp, BSL-mHTT, BSL-Age, CAG, CAP, Country
X= **CSF Biomarkers**, Y=**Clinical Endpoints**	Y		Y		X & Y=BSL	BSL-Age, CAG, CAP, Country
	Y			Y	X=BSL, Y=15M	BSL- Clin.endp, BSL-Age, CAG, CAP, Country
		Y		Y	X=3M, 9M, 15M, Y=15M	BSL- Clin.endp, BSL-CSF.Biom, BSL-Age, CAG, CAP, Country
X=**mHTT**, Y=**CSF Biomarkers**+	Y		Y		X & Y= BSL	BSL-Age, CAG, CAP
	Y			Y	X=BSL, Y=15M	BSL- CSF.Biom, BSL-Age, CAG, CAP
		Y		Y	X & Y=3M, 9M,15M	BSL-CSF.Biom, BSL-mHTT, BSL-Age, CAG, CAP
X=**CSF-Nfl**, Y=**Plasma-NfL**	Y		Y		X & Y= BSL	None
		Y		Y	X & Y=3M, 9M, 15M	None
X=**mHTT**, Y=normalised **MRI**	Y		Y		X & Y= BSL	BSL-Age, CAG, CAP
X=**CSF-Biom** Y=normalised **MRI**	Y		Y		X & Y= BSL	BSL-Age, CAG, CAP
X=**mHTT**, Y=**BSI***	Y			Y (%-change)	X=BSL, Y=3M, 9M, 15M	BSL-Age, CAG, CAP
		Y		Y (%-change)	X=3M, 9M, 15M, Y=15M	BSL-mHTT, BSL-Age, CAG, CAP
X=**CSF-Biom**, Y=**BSI***		Y		Y (%-change)	X=3M, 9M, 15M, Y=15M	BSL-Biom, BSL-Age, CAG, CAP

Table 3 Overview of Planned Outputs to Describe the Association Between Selected Pairs of Parameters at Final analysis (cont.)

BSL= baseline; BSI= boundary shift integral; mHTT=mutant huntingtin protein; CAP= =(CAG-33.66)* AGE_BASELINE; CSF= cerebrospinal fluid; NfL= neurofilament light chain; MRI= magnetic resonance imaging.

§ Clinical Endpoints will include cUHDRS, TFC, TMS, SDMT and SWR

+ CSF Biomarkers will include NFL, YKL-40 and Tau

‡ MRI Volume

* BSI is already a measure of change thus the change is not derived.

In case of an interim analysis (see Section 4.9) the analyses will be based on the available data (e.g. up to 9 months) by the clinical cutoff date.

4.5.3 Exploratory Analysis

Additional exploratory analyses may be performed as appropriate. Data from this study may be combined with data from other studies to inform disease and causal models.

4.6 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

No study medication will be given to the patients therefore no pharmacokinetic or pharmacodynamic analysis will be conducted.

4.7 SAFETY ANALYSES

Safety information will be collected, reviewed, and evaluated by the Sponsor or designee in accordance with the Safety Management Plan throughout the conduct of the clinical trial. Data will be summarized using descriptive statistics. Only scheduled visits will be presented in summary tables and both scheduled and unscheduled visits will be included in listings. All enrolled patients will be used for listings; safety population will be used for summary statistics.

Adverse event reports will not be derived from patient-reported outcome (PRO) or observer-reported outcome (ObsRO) data by the Sponsor, and safety analyses will not be performed using PRO or ObsRO data. Sites are not expected to review the PRO or ObsRO data for adverse events.

Additional analysis to the one outlined below may be conducted depending on emerging safety data.).

Where applicable, in order to account for delays in the visit due to covid-19 a time window of +- 56 days from the expected scheduled visit will be applied to summary tables and plots (details in the list of planned output (LOPO)).

4.7.1 Exposure of Study Medication

No study drug will be given during this study.

4.7.2 Adverse Events

Since no study medication will be given to the patients, safety assessments will consist of monitoring and recording adverse events, including serious adverse events, related to study procedures.

Adverse events and SAEs related to a study procedure reported from date of inform consent until the patient's last visit or withdrawal, will be analyzed.

All AE will be listed, any AE which starts after the reference time point (see Section 4) will be summarized.

The number of unique patients who experience and report an adverse event, each adverse event, and the number of adverse events themselves will be tabulated by medical dictionary for regulatory activities (MedDRA) System Organ Class and Preferred Term. Adverse events will also be summarized by severity.

Serious AEs, AEs that lead to study discontinuation and deaths will also be analysed separately. Selected adverse events may be summarized and listed separately.

4.7.3 Laboratory Data

Laboratory data including chemistry, hematology, CSF safety labs (cell counts, protein, glucose) and urinalysis, will be collected throughout the study. A full list of collected laboratory analytes is contained in the protocol (Refer Section 4.5.7 of Protocol v2).

All clinical laboratory data will be stored on the database in the units in which they were reported. Patients' listings and summary statistics at each assessment time will be presented using the International System of Units (SI units; Système International d'Unités). Laboratory data not reported in SI units will be converted to SI units before processing. For most parameters, the measured laboratory test result will be assessed directly using the Roche standard reference range. Certain laboratory parameters will be transformed to Roche's standard reference ranges. A transformation will be performed on certain laboratory tests that lack sufficiently common procedures and have a wide range of Investigator ranges (e.g., enzyme tests that include AST, ALT, and alkaline phosphatase and total bilirubin). Since the standard reference ranges for these parameters have a lower limit of zero, only the upper limits of the ranges will be used in transforming the data. For all laboratory parameters where Roche standard reference range are not applicable, the local laboratory reference ranges will be used.

In the event that the local laboratory does not have laboratory reference ranges, the Roche predefined standard reference range will be used. Laboratory values falling

outside this standard reference range will be labeled “H” for high or “L” for low in patient listings of laboratory data.

In addition to the standard reference range, a marked reference range has been predefined by Roche for each laboratory parameter. The marked reference range is broader than the standard reference range. Values falling outside the marked reference range that also represent a defined change from baseline will be considered marked laboratory abnormalities (i.e., potentially clinically relevant). If a baseline value is not available for a patient, the midpoint of the standard reference range will be used as the patient’s baseline value for the purposes of determining marked laboratory abnormalities. Marked laboratory abnormalities will be labeled in the patient listings as “HH” for very high or “LL” for very low.

All abnormal values will be listed.

For selected laboratory parameters descriptive tables will be used to summarize absolute change and percentage change from baseline by visit. For the measurement which are reported as ‘<x’ or ‘>x’, the value will be replaced by ‘x’. For the listing, however, original values as collected in the database will be presented. Values outside the normal ranges will be flagged. Shifts in toxicity grade (where grading is based on numeric ranges within NCI-CTCAE v5) from baseline to the worst grade observed during treatment will be presented.

Spaghetti plots and boxplots for the absolute time course as well the change and percentage change from baseline will be presented for selected parameters.

Additional figures/tables/listings may be produced as deemed appropriate.

For all CSF safety laboratory parameters, in case results have been recorded in duplicate or triplicate, the mean value of the visit’s results recorded will be used to produce summary tables and summary plots, while all available values will be shown in the listing and the spaghetti plots.

In addition, the shifts (relative to the normal range) from baseline to the minimum and maximum post-baseline values will be presented. If a subject is missing a baseline value but had a post-baseline value, then the baseline assessment is labeled as "unknown". Likewise, if a subject had a baseline value but had no post-baseline values, then the minimum and maximum are labeled as "unknown". For each parameter, the incidence of shift to low (below the normal range) will be summarized using the minimum post baseline values (across all samples and time points); the incidence of shift to high (above the normal range) will be summarized using the maximum post-baseline values (across all samples and time points).

In case results have been recorded in duplicate or triplicate, for the shift to maximum, the baseline is defined as the highest of the triplicate; similarly, for the shift to minimum, the baseline is defined as the lowest of the triplicate.

Finally, the number of patients with abnormal and marked abnormal absolute values for the CSF safety labs, according to the parameter ranges presented in Table 4, will be summarized by visit.

Table 4 CSF Safety Laboratory Ranges

Parameter	Low	Normal	High	Marked High
Glucose (mmol/L)	<2.2	2.2-3.9	>3.9	>4.4
Protein (g/L)	<0.15	0.15-0.5	>0.5	–
Erythrocytes (cells/μL)	–	0	>0	>5
Leucocytes (cells/μL)	–	0-5	>5	–

4.7.4 Vital Signs

Four primary vital signs (pulse/HR, systolic and diastolic BP, body temperature (BT), respiratory rate (RR); all taken in seated position and body weight will be measured at protocol specified visits.

Absolute values and changes from baseline will be summarized with descriptive statistics by treatment arm and study day. Individual vital signs data will be also listed. Spaghetti plots and boxplots for the absolute time course as well the change from baseline will be presented for all parameters.

If the indication of position is missing or measurements are conducted in a different position than seated, the values will be reported in the listing but not used for summary statistics or computation of change.

A listing of subjects meeting any of the abnormality criteria (according to Roche GDSR) outlined in Table 5 will be provided.

Table 5 Vital Signs Ranges (Seated Position)

	Low	Normal	High
Pulse/HR (bpm)	<60	60-100	>100
SBP (mmHg)	<90	90-130	>130
DBP (mmHg)	<60	60-80	>80
Temperature (C)	<36.5	36.5-37.5	>37.5
RR (bpm)	<8	8-20	>20

DBP=diastolic blood pressure; HR=heart rate; RR= respiratory rate; SBP= systolic blood pressure.

4.7.5 Electrocardiogram

Abnormal findings and clinical significance for the finding for ECG parameters will be listed.

4.7.6 Physical Physical/Neurological Examination

Any abnormality identified at baseline in physical or neurological examination are included in the medical history listing and summary. New or worsened clinically significant abnormalities in physical or neurological examination associated with a study procedure are recorded as AE and therefore reported in AE listings or summaries.

4.7.7 Columbia-Suicide Severity Scale

The Columbia-Suicide Severity Rating Scale (C-SSRS) is a structured tool to assess suicidal ideation and behavior. Four constructs will be assessed: severity of ideation, intensity of ideation, behavior and lethality of actual suicide attempts. At Screening the "Baseline/Screening version" of the scale is completed. At all subsequent visits (including baseline) the "Since Last Visit version" of the scale is completed.

The C-SSRS collects binary responses to 11 categories: five subtypes of suicidal ideation, five subtypes of suicidal behavior, and self-injurious behavior without suicidal intent. Specifically, the following outcomes are C-SSRS categories and have binary (Yes/No) responses. The categories have been re-ordered from the actual scale to facilitate the definitions of the composite endpoints and to enable clarity in the presentation of the results.

Suicidal Ideation:

- **Category 1** - Wish to Be Dead
- **Category 2** - Non-specific Active Suicidal Thoughts
- **Category 3** - Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
- **Category 4** - Active Suicidal Ideation with Some Intent to Act, without Specific Plan
- **Category 5** - Active Suicidal Ideation with Specific Plan and Intent

Suicidal Behavior:

- **Category 6** - Preparatory Acts or Behavior
- **Category 7** - Aborted Attempt
- **Category 8** - Interrupted Attempt
- **Category 9** - Actual Attempt (non-fatal)
- **Category 10** - Completed Suicide

Other:

- **Category 11** - Non-suicidal Self-injurious Behavior

In addition, a numerical score, the Suicidal Ideation Score, will be defined as the highest suicide ideation category (1-5) at which the patient responded "Yes" for the given visit. If the patient did not respond “Yes” to any of these categories, the score will be set to zero.

For each of the aforementioned 11 categories, the number and percent of patients with a “Yes” response at any time post-baseline (regardless of baseline response) will be summarized. Emergent suicidal ideation or behavior will be summarized.

The listed binary categories and the Suicidal Ideation Score will be used to identify a number of composite suicidal endpoints. Note that "recent history" for these composite endpoints is defined as the 12 months prior to Screening and the Screening period. At the Screening visit, CSSRS data is collected for the prior 12 months using the “Baseline/Screening version” of the scale. Therefore, analyses that utilize data from “recent history'” will include the historical 12-month data collected at Screening as well as all on-study C-SSRS data collected after enrollment but prior to LP (attempt or successful) using the “Since Last visit version” of the scale. The endpoints are defined as follows:

- **Suicidal Ideation**: A “Yes" answer at any time post first LP (attempt or successful) to any one of the five suicidal ideation questions (Categories 1-5), regardless of the pre-LP responses
- **Suicidal Behavior**: A “Yes" answer at any time after post first LP to any one of the five suicidal behavior questions (Categories 6-10), regardless of the pre-LP responses.
- **Suicidal Ideation or Behavior**: A "Yes" answer at any time post first LP to any one of the ten suicidal ideation or behavior questions (Categories 1-10), regardless of the pre- LP responses.
- **Treatment-Emergent Suicidal Ideation compared to recent history**: A maximum post first LP suicidal ideation score that is increased from the maximum suicidal ideation score in recent history.
- **Treatment-Emergent Serious Suicidal Ideation compared to recent history**: A maximum post first LP suicidal ideation score of 4 or 5 when the maximum suicidal ideation score during recent history was less than 4 (i.e., scores of 0-3). Only patients with a recent history score of 0-3 will be considered evaluable for this outcome.
- **Emergence of Serious Suicidal Ideation compared to recent history**: A maximum post first LP suicidal ideation score of 4 or 5 when the maximum suicidal ideation score during recent history was 0. Only patients with a recent history score of 0 will be considered evaluable for this outcome.
- **Improvement in Suicidal Ideation compared to baseline**: A decrease in the suicidal ideation score at the patient's visit Study Day 85 (3 months), visit Study Day 253 (9 months) and visit Study Day 421 (15 months) C-SSRS assessment compared to the baseline score, defined as the last measurement taken just prior to first LP . Only patients with a baseline score >0 will be considered evaluable for these outcomes.

- **Emergence of Suicidal Behavior compared to all prior history**: The occurrence of suicidal behavior (a "Yes" response to one or more of Categories 6-10) post-LP from not having suicidal behavior prior to LP (includes the "lifetime" score collected at the Screening Visit as well as all C-SSRS assessments collected from the Screening Visit through pre-LP on Study Day 1).

Each of the composite endpoints will be summarized by study day. For each on study emergent outcome listed, only those patients with the specified screening condition will be considered evaluable. In addition, patients who discontinue from the study with no post-LP C-SSRS assessment will be considered unevaluable for analyses of suicidality. Percentages will be based on the number of evaluable patients for each outcome.

In addition, a shift table will be created to demonstrate the change in suicidal ideation score from recent history to post- LP period. The maximum suicidal ideation score in each period will be used to create the shift table. If a patient's recent history suicidal ideation score is missing but has a post-LP score, then the recent history assessment will be labeled as "unknown". Likewise, if a patient's recent history suicidal ideation score is available but has no post-LP score, then the scores during the on-study and post-LP period will be labeled as "unknown".

4.7.8 Concomitant medications

The original terms recorded on the patients' electronic Case Report Form (eCRF) by the Investigator for concomitant medications will be standardized by the sponsor. For classification purposes, preferred terms will be assigned by the Sponsor to the original terms entered on the eCRF, using the most up-to-date version of WHODrug dictionary for treatments and surgical and medical procedures. Concomitant medications will be presented in listings and summary tables.

4.8 MISSING DATA

Incomplete dates for AEs and laboratory data will be handled as described in the Roche data analysis standards (GDSR). Missing values for laboratory and/or CSF biomarkers will not be imputed unless otherwise specified (e.g. for BLQs). Missing data for outcome assessments will be handled as described in Sections 4.5.1 and 4.7.7

4.9 INTERIM ANALYSES

One or more interim analyses may be performed to make results of this study available as early as possible, to understand the natural course of the disease in this population.

Patients from this study will also serve as external controls for an OLE study with an active compound (RO7234292). For this reason interim analyses may be triggered by the OLE vs NHS comparison and details of this analysis will be described in a separate document.

The decision to conduct an optional interim analysis and the timing of the analysis will be documented in the Sponsor's trial master file prior to the conduct of the interim analysis. The interim analysis will be performed and interpreted by Sponsor study team personnel

Primary and secondary endpoints will be analyzed as describe in this document as well as demography, baseline and safety data.

Given the observational nature of this study and the absence of formal hypothesis testing, adjustment for multiplicity to account for multiple interim analyses will not be considered.

5. REFERENCES

Brooks R. EuroQol: the current state of play. Health Policy 1996;37:53–72.

EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.

Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research. 2011 Dec 1;20(10):1727-36.

International Classification of Functioning, Disability, and Health. (https://www.who.int/classifications/icf/whodasii/en/)

Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Quality of Life Research. 2013 Sep 1;22(7):1717-27.

McCarrier KP, Deal LS, Abraham L, et al. Patient-centered research to support the development of the Symptoms of Major Depressive Disorder Scale (SMDDS): initial qualitative research. The Patient-Patient-Centered Outcomes Research. 2016 Apr 1;9(2):117-34.

Manual for WHO Disability Assessment Schedule, WHODAS 2.0 https://apps.who.int/iris/bitstream/handle/10665/43974/9789241547598_eng.pdf;jsessionid=FF29B710CF8A24B76A70C3D87FCC5126?sequence=1

Schobel SA, Palermo G, Auinger P, et al. Motor, cognitive, and functional declines contribute to a single progressive factor in early HD. Neurology 2017;89:2495–2502.

WPAI scores: reference http://www.reillyassociates.net/WPAI_Scoring.html

Appendix 1 Protocol Synopsis

TITLE:	**A MULTI-SITE, PROSPECTIVE, LONGITUDINAL, COHORT STUDY MEASURING CEREBROSPINAL FLUID-MUTANT HUNTINGTIN PROTEIN IN PATIENTS WITH HUNTINGTON'S DISEASE**
PROTOCOL NUMBER:	BN40422
VERSION NUMBER:	2 (Germany)
TEST PRODUCT:	None
INDICATION:	Huntington's disease
SPONSOR:	F. Hoffmann-La Roche Ltd

Objectives and Endpoints

This longitudinal study in patients with early Huntington's disease (HD) will explore the prognostic value of baseline cerebrospinal fluid (CSF) mutant huntingtin protein (mHTT); the relationship of CSF mHTT to other key biological markers in the putative causal pathway; rate and variability of CSF mHTT increase in untreated patients; and the utility of recording clinical outcome measures via sensors contained in the Roche HD mobile app. Since no study medication will be given to the patients, safety assessments will consist of monitoring and recording adverse events, including serious adverse events, related to study procedures. Specific objectives are outlined below.

Primary Objective

The primary objective of this study is to examine the predictive value of baseline CSF mHTT levels on measures of disease progression using the following endpoints:

- Change from baseline in the following clinical endpoints at 3, 9, and 15 months: composite Unified Huntington's Disease Rating Scale (cUHDRS), Total Functional Capacity Scale (TFC), Total Motor Scale (TMS), Symbol Digit Modalities Test (SDMT), Stroop Word Reading (SWR) Test, and Independence Scale (IS)
- Change from baseline in biomarkers of neuronal injury (e.g, CSF NfL and tau) at 3, 9, and 15 months
- Change from baseline in brain atrophy endpoints (e.g., whole brain volume decline, caudate volume decline) as determined by brain MRI, at 3, 9, and 15 months

Secondary Objective

The secondary objective of this study is to investigate the temporal profile of longitudinal CSF mHTT changes within patient, and the association of these changes with corresponding changes in biomarkers of neuronal injury (e.g., CSF NfL and tau) and clinical outcomes using the following endpoints:

- Within-patient change from baseline in CSF mHTT levels at 3, 9, and 15 months
- Association of change from baseline in CSF mHTT levels at 3, 9, and 15 months and:
  - Change from baseline in clinical measures (cUHDRS, TFC, TMS, SDMT, SWR, and IS) at 3, 9, and 15 months
  - Change from baseline in biomarkers of neuronal injury (e.g., CSF NfL and tau) at 3, 9, and 15 months
  - Change from baseline in brain atrophy endpoints, as determined by brain MRI at 3, 9, and 15 months

Exploratory Objectives

The exploratory objectives of this study are:

- To examine the association of the Roche HD mobile app and standard clinical assessments at baseline and at 3, 9, and 15 months
- To examine the association of the Roche HD mobile app and biological markers under study at baseline and at 3, 9, and 15 months
- To examine the utility of Roche HD mobile app and patient-reported outcomes (PROs) in monitoring disease progression in early HD
- To examine the association of potential prognostic biomarkers and genetic modifiers with disease status and disease progression

Study Design

Description of Study

The study is designed as a multi-site, prospective, 15-month longitudinal, cohort study measuring CSF mHTT in patients with early manifest Stage I or Stage II HD. After study completion, participants will be eligible to enroll in an OLE study with active RO7234292 compound, provided the data from the ongoing RO7234292 program support continued development and subject to approval by the relevant competent authorities and Ethics Committees (ECs).

This study will enroll up to approximately 100 patients at approximately 20 sites. After a 28-day screening period, patients will return to the clinic at baseline and at Months 3, 9, and 15. During these visits, patients will receive clinical, MRI, and Roche HD mobile app assessments; they will also provide CSF and blood samples. A post-observational follow-up visit will take place 2 weeks after the 15-month longitudinal period and may occur by telephone or in the clinic. Early termination visits should take place via clinic visit, where possible.

Number of Patients

The study will enroll up to approximately 100 patients with early manifest Stage I or II HD.

Target Population

Inclusion Criteria

Patients must meet the following criteria for study entry:

- Signed Informed Consent Form
- Capacity to consent to participate in the study as assessed using the Evaluation to Sign Consent tool and investigator judgment
- Age 25 to 65 years, inclusive, at the time of signing Informed Consent Form
- Early manifest, Stage I or Stage II HD (defined as TFC of 7−13, inclusive)
- Genetically confirmed disease (CAG repeat length ≥ 36 in huntingtin gene by direct DNA testing)
- Body mass index ≥ 18 and ≤ 32 kg/m^2; total body weight > 50 kg
- Ability to undergo and tolerate MRI scans (e.g., no claustrophobia, no metal implants including MRI-incompatible intrauterine devices or metal dental braces, chorea of a severity that precludes MRI scans, or any condition that renders testing intolerable for the patient)
- Ability to tolerate blood draws and lumbar puncture
- Ability and willingness to comply with all aspects of the protocol, including completion of interviews and questionnaires and carrying/wearing of a digital monitoring device
- Stable medical, psychiatric, and neurological status for at least 12 weeks prior to screening and at the time of enrollment

  Patients may be managed according to clinical judgment during study observational period.

- Signed study companion consent for participation, if a study companion is available

  A study companion should be reliable, competent, and at least 18 years of age; willing to accompany the patient to clinic visits and to be available to the study site by telephone if needed; and (in the opinion of the investigator) a person who is and will remain sufficiently knowledgeable of the patient's ongoing condition to respond to study site enquiries about the patient. The study companion will complete study companion assessments and will provide demographic and social status data (e.g, relationship to patient and employment status).
- For women of childbearing potential: agreement to remain abstinent (refrain from heterosexual intercourse) or use acceptable contraceptive methods during the observational period.

  A woman is considered to be of childbearing potential if she is postmenarcheal, has not reached a postmenopausal state (≥ 12 continuous months of amenorrhea with no identified cause other than menopause), and has not undergone surgical sterilization (removal of ovaries and/or uterus).

  The following are acceptable contraceptive methods: bilateral tubal ligation; male sterilization; hormonal contraceptives; hormone-releasing intrauterine devices; copper intrauterine devices; male or female condom with or without spermicide; and cap, diaphragm, or sponge with spermicide.

  The reliability of sexual abstinence should be evaluated in relation to the duration of the observational study and the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not acceptable methods of contraception.

Exclusion Criteria

Patients who meet any of the following criteria will be excluded from study entry:

- Any condition, including severe chorea, that would prevent either writing or performing pen-and-paper or smartphone-based tasks
- History of attempted suicide or suicidal ideation with plan (i.e., active suicidal ideation) that required hospital visit and/or change in level of care within 12 months prior to screening

  Current active suicidal ideation is demonstrated by the Columbia-Suicide Severity Rating Scale. If suicidal ideation is generally present, a risk assessment should be done by an appropriately qualified mental health professional to assess whether it is safe for the patient to participate in the study. Mild passive suicidal ideation (i.e., occasional thoughts that life is not worth living or is hard) without history of attempts or hospitalization over the past 12 months is generally acceptable for study participation, but final decision on participation should be made carefully and in consultation with appropriately qualified mental health professional per judgment of the investigator.

- Current active psychosis, confusional state, or violent behavior
- Any serious medical condition or clinically significant laboratory, vital sign, or electrocardiogram abnormalities at screening that, in the investigator's judgement, precludes the patient's safe participation in and completion of the study
- Pregnant or breastfeeding, or intending to become pregnant during the study
- Positive for hepatitis C virus antibody or hepatitis B surface antigen at screening
- Known HIV infection
- Current or previous use of an antisense oligonucleotide (including small interfering RNA)
- Current use of antipsychotics prescribed for psychosis, cholinesterase inhibitors, memantine, amantadine, or riluzole including use within 12 weeks of enrollment

  Use of antipsychotics for motor symptoms and/or tetrabenazine or deutetrabenazine is not permitted unless on stable dose for at least 12 weeks prior to screening.

- Treatment with an investigational drug within 30 days prior to screening or 5 half-lives of the investigational drug, whichever is longer
- Antiplatelet or anticoagulant therapy within the 14 days prior to screening or anticipated use during the study, including, but not limited, to aspirin (unless $\leq$81mg/day), clopidogrel, dipyridamole, warfarin, dabigatran, rivaroxaban, and apixaban
- History of bleeding diathesis or coagulopathy; platelet count $<$lower limit of normal unless stable and assessed by the Investigator and Sponsor Medical Monitor to be not clinically significant
- Malignancy within 5 years prior to screening, except basal or squamous cell carcinoma of the skin or carcinoma in situ of the cervix that has been successfully treated
- History of gene therapy or cell transplantation or any other experimental brain surgery
- Concurrent or planned concurrent participation in any clinical study without approval of the Medical Monitor
- Presence of implanted shunt for the drainage of CSF or an implanted CNS catheter
- Prexisting structural brain lesion (e.g., tumor, arterio-venous malformation) as assessed by MRI scan

End of Study

The end of the study will be the date from which the last patient, last visit is recorded in the study database. The end of the study is expected to occur 15.5 months after the last patient is enrolled.

Length of Study

The total length of the study, from screening of the first patient to the end of the study, is expected to be approximately 27 months.

Statistical Methods

Primary Analysis

The primary objective of this study is to investigate whether baseline CSF mHTT is a predictor of the change from baseline to 15 months for the following variables:

- Clinical endpoints (cUHDRS, TFC, TMS, SDMT, SWR, and IS)
- CSF and blood biomarkers
- Brain imaging (MRI)

Initially, a linear model will be used to evaluate the association between each response variable (e.g., change from baseline to 15 months in cUHDRS) and baseline CSF mHTT without adjustment for any baseline characteristics.

The same model will then be considered again, adjusting for baseline level of the response variable (e.g., baseline cUHDRS), baseline CSF mHTT levels, CAG, and baseline CAP (the latter defined as the product of age at baseline by [CAG—33.66]. A sensitivity analysis to evaluate the impact of additional baseline characteristics (e.g., gender, education level) may be performed.

Other approaches may be used in the case of severe multicollinearity, which would results in model instability.

Determination of Sample Size

In this study, up to approximately 100 patients are expected to be enrolled across approximately 20 sites.

Interim Analyses

The Sponsor may choose to conduct one or more interim analyses. The decision to conduct such an optional interim analysis and the timing of the analysis will be documented in the Sponsor's study master file prior to the conduct of the interim analysis.

Appendix 2 Schedule of Assessments

	Screening Period	Observational Period						Follow-Up (or Early Termination) [a]
Study Day	Days −29 to −2	Day −1 to 1 (Baseline)	Day 85 (Month 3)	Day 169 (Month 6)	Day 253 (Month 9)	Day 337 (Month 12)	Day 421 (Month 15)	
Visit Window (days)			±7	±10	±7	±10	±7	
Type of visit	CL	CL	CL	PH	CL	PH	CL	CL or PH
Capacity to consent assessment and informed consent [b]	x							
Demographic data	x							
Medical history and baseline conditions	x	x						
Vital signs [c]	x	x	x		x		x	
Height [d]	x							
Weight [e]	x	x	x		x		x	
Physical examination [f]	x	x	x		x		x	
Neurological examination [g]	x	x	x		x		x	
CAG repeat length [h]	x							
Chemistry, hematology, urinalysis [i]	x	x	x		x		x	
Viral serology [j]	x							
Pregnancy test [k]	x		x		x		x	
Coagulation [l]	x	x	x		x		x	
Thyroid panel	x							
ECG (12-lead)	x							
Lumbar puncture and CSF sample [m]		x	x		x		x	

Appendix 2 Schedule of Assessments (cont.)

	Screening Period	Observational Period						Follow-Up (or Early Termination) [a]
Study Day	Days −29 to −2	Day −1 to 1 (Baseline)	Day 85 (Month 3)	Day 169 (Month 6)	Day 253 (Month 9)	Day 337 (Month 12)	Day 421 (Month 15)	
Visit Window (days)			±7	±10	±7	±10	±7	
Type of visit	CL	CL	CL	PH	CL	PH	CL	CL or PH
MRI [n]	x		x		x		x	
ClinROs, ObsROs, PerfOs, and PROs conducted in clinic [o]	x	x	x		x		x	
C-SSRS [p]	x	x	x		x		x	
Roche HD mobile app equipment training	x							
Roche HD mobile app in-clinic assessment [q]	x	x	x		x		x	
Roche HD mobile app remote data collection [q]	Continuous remote data collection							
Whole blood sample for clinical genotyping		x						
Plasma samples for biomarkers		x	x		x		x	
RBR DNA (optional) [r]		x						
RBR RNA (optional) [r]		x					x	
RBR serum (optional) [r]		x					x	
Change in medical information since last visit [s]			x	x	x	x	x	
Adverse events [t]	x	x	x	x	x	x	x	x

Appendix 2 Schedule of Assessments (cont.)

	Screening Period	Observational Period						
Study Day	Days −29 to −2	Day −1 to 1 (Baseline)	Day 85 (Month 3)	Day 169 (Month 6)	Day 253 (Month 9)	Day 337 (Month 12)	Day 421 (Month 15)	Follow-Up (or Early Termination) [a]
Visit Window (days)			±7	±10	±7	±10	±7	
Type of visit	CL	CL	CL	PH	CL	PH	CL	CL or PH
Concomitant medications [u]	x	x	x	x	x	x	x	x

CL = clinic; ClinROs = clinician-reported outcomes; CSF = cerebrospinal fluid; C-SSRS = Columbia-Suicide Severity Rating Scale; eCRF = electronic Case Report Form; HBsAg = hepatitis B surface antigen; HCV = hepatitis C virus; HD = Huntington disease; MRI = magnetic resonance imaging; ObsROs = observer-reported outcomes; PerfOs = performance outcomes; PH = phone; PROs = patient-reported outcomes; RBR = Research Biosample Repository.

[a] Post-observational follow-up visit will take place 2 weeks after end of observational period and will occur via clinic visit or by telephone call. Early termination visit should occur via clinic visit where possible.

[b] Capacity to consent assessment (patient only) and informed consent (for both the patient and study companion [if available]).must be documented before any study-specific screening procedure is performed and may be obtained more than 28 days before baseline visit.

[c] Includes measurements of respiratory rate, pulse rate, and systolic and diastolic blood pressure while the patient is in a seated position after resting for approximately 5 minutes, and temperature. Record abnormalities observed at baseline on the General Medical History and Baseline Conditions eCRF.

[d] Height in meters.

[e] Weight in kilograms.

[f] Includes evaluation of the head, eyes, ears, nose, and throat, and the cardiovascular, dermatologic, musculoskeletal, respiratory, gastrointestinal, genitourinary, and neurologic systems. Any abnormality identified at baseline should be recorded on the General Medical History and Baseline Conditions eCRF. Changes from baseline abnormalities should be recorded in patient notes. New or worsened clinically significant abnormalities associated with a study procedure should be recorded as adverse events on the Adverse Event eCFR.

[g] Neurological examination includes assessment of mental status, level of consciousness, sensory function, motor function, cranial nerve function and reflexes. Record abnormalities observed at baseline on the General Medical History and Baseline Conditions eCRF. New or worsened clinically significant abnormalities associated with a study procedure should be recorded as adverse events on the Adverse Event eCRF.

[h] At screening, a mandatory whole blood sample will be obtained for DNA extraction for analysis of CAG repeat length.

Appendix 2 Schedule of Assessments (cont.)

[i] Screening laboratory assessments include hematology (WBC count, RBC count, hemoglobin, hematocrit, platelet count, differential count including neutrophils, eosinophils, basophils, monocytes, lymphocytes, other cells), serum chemistry (sodium, potassium, chloride, bicarbonate, glucose, BUN or urea, creatinine, total protein, albumin, phosphorus, calcium, total and direct bilirubin, ALP, ALT, AST, uric acid, GGT, CPK), and urinalysis (dipstick including pH, specific gravity, glucose, protein, ketones, blood and microscopic examination, including sediment, RBCs, WBCs, casts, crystals, epithelial cells, bacteria).

[j] Viral serology: HBsAg and HCV antibody.

[k] All women of childbearing potential will have a serum pregnancy test at screening. Urine pregnancy tests will be performed at specified subsequent visits. If a urine pregnancy test is positive, it must be confirmed by a serum pregnancy test.

[l] Coagulation: INR, aPTT, PT.

[m] Patients will have CSF collected at baseline and at Months 3, 9, and 15. Prior to performing each scheduled lumbar puncture, local laboratory analysis of coagulation factors (INR, aPTT, PT) and platelets must be conducted and results reviewed. Collection for these local labs may occur at any time in the 72 hours prior to the lumbar puncture. The lumbar puncture should be performed at approximately the same time at each visit (ideally in the morning between 08:00 and 12:00 hours) to minimize potential diurnal variation of CSF parameters. CSF fluid (20 mL) is to be collected for analyses using a standard lumbar puncture collection kit. If there are difficulties in collecting 20 mL of CSF fluid, a minimum of 15 mL should be collected. A 24G Whitacre (atraumatic) needle should be used. Depending on institutional guidelines, local anesthesia may be used for the procedure. Sedation may not be used. Spinal ultrasound may be used for the lumbar puncture procedure if deemed necessary, but is not required. Ultrasound guidance may be used if attempts at lumbar puncture without imaging are unsuccessful, if it is local practice to use ultrasound, or if institutional guidelines dictate use of ultrasound with each lumbar puncture.

[n] For MRI scans at Months 3, 9, and 15, the MRI should be scheduled to occur before the lumbar puncture. It can be scheduled in the days prior to the lumbar puncture (provided it occurs within the visit window), or it can be scheduled to occur on the same day as the lumbar puncture (provided it is prior to the lumbar puncture).

[o] Questionnaires will be self-administered or administered by trained rater (as appropriate) at screening and at each clinic visit prior to the lumbar puncture. See Appendix 2 (Refer Protocol v2) and Appendix 3 (Refer Protocol v2) for details.

[p] The C-SSRS must be administered on scheduled timepoints and may be performed at other timepoints per investigator's discretion.

[q] The Roche HD mobile app includes a short, preconfigured schedule of daily tasks (“active test”) that assess motor symptoms (upper and lower body movements, upper limb dexterity, voice) and non-motor symptoms (processing speed/voice). In addition, passive monitoring data is collected. During the in-clinic assessment, the entire suite of active tests should be performed.

[r] Not applicable for a site that has not been granted approval for RBR sampling. Performed only for patients who have provided written informed consent to participate at participating sites.

Appendix 2 Schedule of Assessments (cont.)

[s] At the time of each follow-up physical examination, an interval medical history should be obtained and any changes in medications, any major procedures or hospitalizations, and any physician visits for HD or general medical care should be recorded.

[t] After informed consent has been obtained, adverse events caused by a protocol-mandated intervention should be reported until the patient's last visit or withdrawal. The investigator should follow each adverse event until the event has resolved to baseline grade or better, the event is assessed as stable by the investigator, the patient is lost to follow-up, or the patient withdraws consent. Every effort should be made to follow all serious adverse events considered to be related to protocol-related procedures until a final outcome can be reported.

[u] Medication (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by a patient during the study (from screening to the study completion/discontinuation visit).

Appendix 3
Percentage Change in the Geometric Means for Log-Transformed Variables

The mean for the change from baseline of a log-transformed *y* variable is given below:

$$\frac{1}{\mathrm{n}}\sum_{i=1}^{n}\left(log\left(y_{i,FU}\right)-log\left(y_{i,BSL}\right)\right)=\sum_{i=1}^{n}\left(log\left(\frac{y_{FU}}{y_{BSL}}\right)^{\frac{1}{\mathrm{n}}}\right)=Z$$

where $y_{i,BSL}$ and $y_{i,FU}$ are the values of y for the *i-th* patient respectively at baseline and at a given follow-up time and n is the total number of patients.

By applying the following transformation to Z

$$\left((e^{Z}-1)\cdot 100\right)=\left(\left(e^{\sum_{i=1}^{n}\left(log\left(\frac{y_{FU}}{y_{BSL}}\right)^{\frac{1}{\mathrm{n}}}\right)}-1\right)\cdot 100\right)=\left(\left(\prod_{i=1}^{n}e^{log\left(\frac{y_{FU}}{y_{BSL}}\right)^{\frac{1}{\mathrm{n}}}}-1\right)\cdot 100\right)=$$

$$=\left(\left(\prod_{i=1}^{n}\left(\frac{y_{FU}}{y_{BSL}}\right)^{\frac{1}{\mathrm{n}}}-1\right)\cdot 100\right)=\left(\left(\frac{\prod_{i=1}^{n}y_{FU}^{\frac{1}{\mathrm{n}}}}{\prod_{i=1}^{n}y_{BSL}^{\frac{1}{\mathrm{n}}}}-1\right)\cdot 100\right)=\left(\left(\frac{\sqrt[n]{\prod_{i=1}^{n}y_{FU}}}{\sqrt[n]{\prod_{i=1}^{n}y_{BSL}}}-1\right)\cdot 100\right)=$$

$$=\left(\left(\frac{Geom\_Mean\_Y_{FU}}{Geom\_Mean\_Y_{BSL}}-1\right)\cdot 100\right)$$

it is shown that this corresponds to the percentage change in the geometric means between follow-up and baseline. Thus, the following transformation will be applied to the least squares estimate, β, obtained from modeling the change from baseline of log-transformed variables

$$\left(\mathrm{e}^{\beta}-1\right)\cdot 100$$

Appendix 4
Mixed-Model Repeated Measures Technical Details

The advantage of mixed-model repeated measures (MMRM) analysis is that it takes into account the fact that measurements from the same patients are correlated and also that all available data are used without eliminating individuals without complete information. The following points will be considered during MMRM model fitting;

- The Kenward and Roger method for approximating the denominator degrees of freedom and correcting for bias in the estimated variance-covariance of the fixed effects will be used.

  In the event the model fails to run using the KR method, then the residual method will be used instead.

- For modelling change from baseline with a limited number of post-baseline time points (3M, 9M and 15M) an Unstructured covariance structure for the R matrix will be used by specifying ‘type=UN’ on the REPEATED line.

  In the event that this model fails to converge, alternative correlation structures may be considered. Akaike's Information Criteria (AIC) will be used to assist with the selection of covariance structure.

- For modelling change from baseline with many post-baseline time points (>3), a Heterogeneous Toeplitz covariance structure for the R matrix will be used by specifying ‘type=TOEPH’ on the REPEATED line.

  In the event that this model fails to converge, alternative correlation structures may be considered. Akaike's Information Criteria (AIC) will be used to assist with the selection of covariance structure.

Distributional assumptions underlying the model used for analysis will be examined to gain confidence that the model assumptions are reasonable and where needed, sensitivity analyses will be performed.