Official Title: Phase 2, Randomized, Double-Blind, Placebo-Controlled Study to Assess the Effects of RO6889450 (Ralmitaront) in Patients With Schizophrenia or Schizoaffective Disorder and Negative Symptoms

NCT Number: NCT03669640

Document Dates: SAP Version 1: 04-March-2020

TECHNICAL DOCUMENT FOR STATISTICAL ANALYSIS – PART A

TITLE:	**PHASE 2, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY TO ASSESS THE EFFECTS OF RO6889450 IN PATIENTS WITH SCHIZOPHRENIA OR SCHIZOAFFECTIVE DISORDER AND NEGATIVE SYMPTOMS**
PROTOCOL NUMBER:	BP40283
STUDY DRUG:	RO6889450
VERSION NUMBER:	1.0
IND NUMBER:	[redacted]
EUDRACT NUMBER:	To be determined
SPONSOR:	F. Hoffmann-La Roche Ltd
PLAN PREPARED BY:	[redacted]
DATE FINAL:	04 March 2020

Signatures

[redacted]

Date: [redacted] Statistician [redacted]

Date: [redacted]

TECHNICAL DOCUMENT FOR STATISTICAL ANALYSIS – PART A

TITLE: **PHASE 2, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY TO ASSESS THE EFFECTS OF RO6889450 IN PATIENTS WITH SCHIZOPHRENIA OR SCHIZOAFFECTIVE DISORDER AND NEGATIVE SYMPTOMS**

PROTOCOL NUMBER: BP40283

STUDY DRUG: RO6889450

VERSION NUMBER: 1.0

IND NUMBER: [redacted]

EUDRACT NUMBER: To be determined

SPONSOR: F. Hoffmann-La Roche Ltd

PLAN PREPARED BY: [redacted]

DATE FINAL: 04 March 2020

Signatures

[redacted]

Date, [redacted] Statistician

Date, [redacted]

CONFIDENTIAL

TABLE OF CONTENTS

LIST OF TABLES

Es konnten keine Einträge für ein Abbildungsverzeichnis gefunden werden.

LIST OF APPENDICES

1. BACKGROUND

This Technical Document for Statistical Analysis (TDSA) provides the data-handling rules, derivation rules, and statistical methods of summarizing and analyzing the efficacy and safety data that will be used for the data analysis and the reporting of the results from **Part A** of study BP40283.

Recruitment rate in **Part A** has remained substantially below the original estimates despite the implementation of a protocol amendment and the addition of more sites. Therefore the Roche TAAR1 Project Team and the Clinical Team have decided to stop study Part A (monotherapy) and start Part B (add-on treatment). In total 27 patients have been randomized into **Part A**.

There will be a separate TDSA for **Part B** of the study.

2. STUDY DESIGN

This is a Phase 2, multicenter, randomized, double-blind, two-part, placebo-controlled, parallel-group study of 12 weeks of treatment in patients with schizophrenia or schizoaffective disorder. An overview of the study design is provided in Figure 1. The study is designed to collect potential efficacy data as well as safety and tolerability data in a patient population.

This study is a proof of concept study (**Part A**) followed by a dose-ranging study (**Part B**): RO6889450 will first be evaluated in a monotherapy setting (Part A), i.e. the participants will withdraw from their antipsychotic treatment(s), then potential efficacy of two dose levels of RO6889450 in combination with usual antipsychotic treatments(s) will be evaluated.

For both Parts A and B randomization occurs on Day 1 (Baseline), after screening period including placebo compliance period (inclusive of one-week-wash-out period for Part A). Treatment will be allocated by stratified block randomization on the bases of the three (Part A) or four (Part B) factors: baseline Brief Negative Symptoms Scale (BNSS) avolition/apathy subscores (≤6 vs. >6), age (18-35 years vs. 36-55 years), sex (Male vs. Female), and only for Part B antipsychotic treatment (full D2/5HT2A antagonist vs. partial D2 receptor agonist).

Part A (monotherapy)

Participants will be washed out (one week) from their antipsychotic therapy before being randomized 1:1 to either [redacted] QD of RO6889450 or placebo for 12 weeks.

Part B (add-on therapy)

Participants will be randomized 1:1:1 to [redacted] QD of RO6889450, [redacted] QD of RO6889450, or placebo for 12 weeks.

Part B of the study is designed to evaluate potential efficacy of two dose levels of RO6889450 as adjunct to usual antipsychotic treatment(s). Part B is planned to start after Part A. Participants who are enrolled in Part A cannot be screened for Part B.

Length of Study

Part A: The total duration will be approximately 147 days (from screening through study completion) for each enrolled participant as follows:

- Screening up to 35 days (including the placebo compliance and a 7-day washout period)
- Treatment period: 84 days (12 weeks)
- Follow up: 28 days

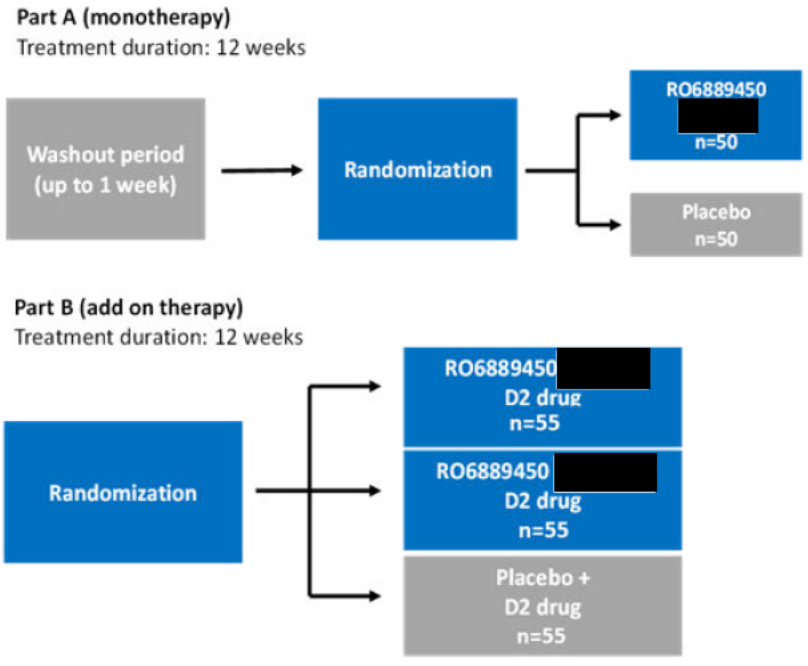
Part B: The total duration will be up to approximately 147 days (from screening through study completion) for each enrolled participant as follows:

- Screening: up to 35 days (including the placebo compliance period)
- Treatment period: 84 days (12 weeks)
- Follow up: 28 days

End of Study

The end of study is defined as the date when the last participant last observation (LPLO) occurs. LPLO is expected to occur approximately 28 days after the last participant's last dose.

Figure 1 Overview of Study Design



D2 = D2/5HT2A antagonists or D2 partial agonist; n = number of completers

2.1 PROTOCOL SYNOPSIS

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

2.2 OBJECTIVES AND ENDPOINTS

The objectives and corresponding endpoints are provided in Table 1.

Table 1 Objectives and Endpoints

Objectives	Endpoints
Primary	
Part A: To compare the efficacy of [redacted] QD of RO6889450 as monotherapy with placebo on negative symptoms in patients with schizophrenia or schizoaffective disorder. **Part B:** To compare the efficacy of [redacted] QD of RO6889450 as add-on therapy with placebo on negative symptoms in patients with schizophrenia or schizoaffective disorder.	Part A & B: Change from baseline at Week 12 in the Brief Negative Symptoms Scale (BNSS) avolition/apathy subscore (sum of items "behavior" and "internal experience").
Secondary	
To compare the effect of RO6889450 with placebo on: • Clinical Global Impression Severity (CGI-S) and Improvement (CGI-I) (overall and negative symptoms)	• Change from baseline in CGI-S and CGI-S negative symptoms (CGI-S NS) scores • CGI-I and CGI-I negative symptoms (CGI-I NS) scores
To compare the effect of RO6889450 with placebo on: • Symptoms of schizophrenia or schizoaffective disorder as assessed with the Positive and Negative Syndrome Scale (PANSS), BNSS, and Defeatist Performance Attitude Scale (DPAS)	• Change from baseline in PANSS total and symptom factor scores • Change from baseline in BNSS total and subscores • Change from baseline in DPAS scores

Objectives	Endpoints
To compare the safety and tolerability of 12 weeks of treatment with RO6889450 as monotherapy (Part A) or add-on therapy with placebo (Part B).	• Incidence, nature, and severity of AEs • Incidence, nature, and severity of SAEs • Incidence, nature, and severity of treatment discontinuations due to AEs • Change from baseline in orthostatic vital signs recordings • Change from baseline in ECG intervals: heart rate, PQ (PR), QRS, QT, RR and QTcF along with information on T- and U-waves • Incidence of laboratory abnormalities, based on hematology, clinical chemistry, and urinalysis test results • Change from baseline in C-SSRS and Extrapyramidal symptom rating scale, abbreviated (ESRS-A)
To evaluate the PK of RO6889450 and RO6889450-derived metabolite(s).	• Concentration per time point • AUC_{ss} of RO6889450 and RO6889450-derived metabolite(s) • C_{max} of and RO6889450-derived metabolite(s) • Other PK parameters as appropriate
Tertiary/Exploratory	
To explore the effects of RO6889450 on: • Reward learning tasks (Effort-Choice Benefit (ECB) task; Reinforcement Learning Working Memory (RLWM) task)	• ECB: Percentage of high effort choices under deterministic reward conditions for high reward magnitudes • RLWM Learning Phase: Proportion of correct choices in late trials and for large blocks of sizes • RLWM Testing Phase: Interaction 'value difference modulated by value mean' (assessing the 'choose A avoid B' paradigm).
To explore the effects of RO6889450 on: • Overall functioning/cognition	• Change from baseline in Brief Assessment of Cognition in Schizophrenia (BACS) • Change from baseline in adjusted total time as measured by the Virtual Reality Functional Capacity Assessment Tool (VRFCAT)
To explore the effects of RO6889450 on: • Changes in smoking dependence	• Change from baseline in Fagerstrom Test for Nicotine Dependence (FTND) total score
To explore the effects of RO6889450 on: • Levels and patterns of social and general activity	• Phone usage behavior data
• To assess CYP2C19 status on PK parameters	• Genotyping of CYP2C19

Objectives	Endpoints
• Part A only: To assess effects on relapse prevention	• Compare rate and timing of relapses and study withdrawal between placebo and RO6889450

2.3 DETERMINATION OF SAMPLE SIZE

In **Part A,** 50 evaluable participants per arm were planned to be randomized to obtain 80% power at the one-sided type I error of 0.05, assuming the true effect size of 0.5.

Based on data from previously conducted studies, the standard deviation for the primary endpoint, the change from baseline in the BNSS Avolition/Apathy Subscale at week 12 (consisting of the sum of the two items 'Avolition/Apathy: Behavior' and Avolition/Apathy: Internal Experience) is about 2.8. Hence, the effect size of 0.5 translates to a reduction (vs. placebo) of 1.4 points for the primary endpoint.

With the anticipated sample size of 50 participants per arm, at the one-sided type I error of 0.05, any observed reduction (vs. placebo) by approximately 0.92 or more points is expected to be statistically significant.

Assumign 20% of the participants randomized will not complete 12 weeks of treatment, in **Part A**, it was planned to randomize approximately 125 participants to one of two treatment arms to ensure that approximately 100 participants complete the 12 weeks of treatment period.

[redacted]

For sample size determination of **Part B**, pease refer to the protocol.

2.4 ANALYSIS TIMING AND UNBLINDING

2.4.1 Interim Efficacy Analysis

No interim analysis will be performed for study **Part A**. Study **Part B** starts without interim analysis of **Part A** [redacted]

The Sponsor may choose to conduct an interim analysis of **Part B** to determine the chance of final success. The results of such an analysis may support an early termination of Part B. Interim analyses of **Part B** will be described in the TDSA of **Part B**.

2.4.2 Interim Safety Reviews

The independent Data Monitoring Committee (iDMC, see Section 3.3) will review unblinded safety and primary/secondary efficacy endpoint data, especially focusing on **Part A** which involves a one-week-wash-out period and only placebo arm.

Two safety reviews are scheduled for the review of the safety data accrued in **Part A** of the study:

1. Once approximately 10 participants per treatment arm reach the Week 2 visit, i.e., after the one-week-wash-out period. – This safety review was done at 14 November 2019, the recommendation was to continue the study without modification.
2. Once approximately 30 participants per treatment arm reach the Week 12 visit, i.e., during and after the Treatment Period with the main focus being on placebo arm. – This safety review will not be conducted as the Part A of the study was stopped before approximately 30 patients per treatment arm reached the Week 12 visit.

Additional safety reviews for the safety data may be conducted for study **Part B** at the request of the iDMC or the Study Team, *e.g.* if safety concerns arise.

The Sponsor's study statistician as well as statistical programmer will be unblinded to provide unblinded safety information to the iDMC. The data will be handled and cleaned in a secure area that is not accessible by any blinded SMT member.

2.4.3 Final Analyses

The final analysis of data from the trial will be performed separately for **Part A** and **Part B** when all patients of the respective study part have either completed the study or discontinued early from the study, all data of the trial are in the database, have been cleaned and verified, and the database has been locked.

3. STUDY CONDUCT

3.1 RANDOMIZATION

An Interactive Voice/Web Response System (IxRS) is used to randomize the participants in equal proportion to the treatment arms.

In **Part A** treatment will be allocated by **stratified block randomization** on the basis of the following factors:

- Baseline BNSS avolition/apathy subscore (≤6 or >6)
- Age (18-35 years vs. 36-55 years)
- Sex (Male vs. Female)

In **Part A**, on Day 1 (baseline) participants within each block stratum are randomly assigned in 1:1 ratio to either [redacted] QD of RO6889450 or placebo. Approximately 125 participants will be randomized to ensure approximately 100 number of completers.

3.2 INDEPENDENT REVIEW FACILITY

PANSS and BNSS will be assessed live via secure videoconferece by an approved remote centralized rater. If a remote administration session cannot occur for an unforeseen reason, it will be rescheduled.

3.3 DATA MONITORING

3.3.1 Independent Data Monitoring Committee (iDMC)

An iDMC will regularly review unblinded safety and efficacy (as appropriate) data. The iDMC will be composed of at least two psychiatrists external to the Sponsor and a statistician external to the Sponsor and will be responsible for monitoring safety data to help ensure that continuation of the trial in its current design does not pose unacceptable safety risks to participants. The iDMC will also assess safety as part of the planned interim safety analysis.

A separate iDMC charter documents the roles, responsibilities, membership as well as scope of activities, time of meetings, and communication plan. Planned and ad hoc safety reviews will be performed as described in the iDMC charter.

4. STATISTICAL METHODS

The final analysis of data from the trial will be performed separately for **Part A** and **Part B**. This TDSA describes the analyses for **Part A**.

[redacted] Data for study **Part A** will be listed and summarized descriptively. No formal hypothesis test will be performed.

4.1 ANALYSIS POPULATIONS

For purposes of analysis, the following populations are defined in Table 2.

Table 2 Analysis Populations

Population	Description
Modified intent-to-treat (ITT)	All randomized participants who have received at least one dose of study drug, including participants who were withdrawn prematurely, will be included in the efficacy population. The modified ITT analysis population will be the primary analysis population for all efficacy data analyses.
Per-protocol (PP)	No PP analysis population is defined for Part A
Safety evaluable (SE)	Same as modified ITT analysis population.

Pharmacokinetic (PK)	All participants who have received at least one dose of study drug and who have data from at least one post-dose sample will be included in the PK analysis population. Participants will be excluded from the PK analysis population if they significantly violate the inclusion or exclusion criteria, deviate significantly from the protocol, or if data are unavailable or incomplete that may influence the PK analysis. Excluded cases will be documented together with the reason for exclusion. All decisions on exclusions from the analysis will be made prior to database closure.

4.2 ANALYSIS OF STUDY CONDUCT

4.2.1 Disposition of Participants

For all patients in the modified ITT analysis population enrollment in the study, discontinuation of study drug and from the study, and reason for discontinuation will be summarized per treatment arm to which participants are randomized. In addition, protocol violations will be summarized by treatment arm.

4.2.2 Overview of Analysis Population

The number (percentage) of participants excluded from the analysis populations will be summarized by reason for exclusion per treatment arm

4.2.3 Protocol Violations

The major protocol violations will be identified before database lock. The number and percentage of participants with major protocol violations will be summarized per protocol violation criterion and treatment arm to which participants are randomized.

4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

The evaluation of treatment arm comparability between the two treatment arms in study **Part A** will include summaries of demographics, baseline disease characteristics, and participant medical history. Data will be summarized by treatment arm to which participants were randomized.

Descriptive statistics (mean, standard deviation, median, and minimum-maximum) will be presented by treatment arm for continuous variables. Frequency counts and proportions will be presented by treatment arm for categorical variables.

4.3.1 Demographic Characteristics

Summary tables of demographics will be produced for the modified ITT analysis population by treatment arm to which participants were randomized. The following variables will be summarized:

Categorical variables:

- Sex (Male vs. Female)
- Age group (18-35 years vs. 36-55 years) at baseline
- Race (Black or African American vs. White) or Ethnicity (Hispanic or Latino vs. non-Hispanic or Latino)

Continuous variables:

- Weight [kg], Height [cm], Body Mass Index (BMI) [kg/m^2] at baseline
- Age (years) at baseline

and others if appropriate.

4.3.2 Baseline Disease Characteristics

Baseline data for BNSS total- subscores, PANSS total- and symptom factor scores, CGI-S and CGI-S NS scores will be included in the respective efficacy outputs. No separate outputs will be presented for baseline.

4.3.3 History of Schizophrenia

Summary statistics will be generated for each treatment group and will include each patient's history of schizophrenia. The key variables from the electronic Case Report Form (eCRF) listed in the psychiatric history page will be summarized. The time (in months) since the diagnosis of schizophrenia and the time (in months) since the most recent hospitalization for worsening of Schizophrenia will be calculated relative to the randomization date.

4.3.4 Previous and Current Diseases at Baseline

For all diseases at baseline (other than schizophrenia), the term entered by the investigator describing the disease (the "verbatim term") will be assigned to a standardized term (the "preferred term") and System Organ Class from MedDRA. All analyses will be performed using these preferred terms and body systems. Diagnoses will be categorized by condition at baseline (ongoing with treatment/ongoing without treatment/resolved) with use of the responses from the medical history page of the eCRF.

The number and percentage of patients with previous (resolved) diseases will be summarized by treatment group. Diseases concurrent (ongoing with or without treatment) at baseline will be summarized similarly in a separate table. Multiple occurrences of a disease (i.e., same coded term) will be counted only once.

4.3.5 Previous and Concomitant Treatments (Other than Study Medication)

For all medications, the term entered by the investigator describing the medication (the "verbatim term") will be assigned to a standardized term (the "preferred term") and drug class from the WHO drug dictionary. All analyses will use these preferred terms and medication classes.

The number and percentage of patients who received previous or concomitant treatment (medication, medical procedure, or non-medical therapy) will be summarized by treatment group.

Multiple occurrences of a medication in an individual (i.e., same coded term) will be counted only once. All summary tables will be sorted by medication class (in decreasing order of overall incidence), then by preferred term (in decreasing order of overall incidence).

4.4 DEFINITION OF BASELINE

Baseline for all efficacy and safety analyses is define das the last non-missing value recorded prior to or on the first day of randomized double-blind study drug administration.

4.5 EFFICACY ANALYSIS

Endpoints will be summarized descriptively for the modified ITT analysis population, no statistical modelling will be perfomed, no statistical hypothesis will be tested.

4.5.1 Primary Efficacy Endpoint

4.5.1.1 Definition of Primary Efficacy Endpoint

The primary efficacy endpoint of the study is the change from baseline in the BNSS Avolition/Apathy subscore at Week 12.

The BNSS (Strauss et al 2012) is a 13-item instrument designed for clinical trials and other studies that measure the severity of negative symptoms in six domains (subscales), inclusive of one single-item domain of 'Distress' (see Appendix 3):

- **Blunted Affect** (three items): facial expression, vocal expression, expressive gestures;
- **Alogia** (two items): quantity of speech, spontaneous elaboration;

- **Asociality** (two items): asociality - behavior, asociality - internal experience;
- **Anhedonia** (three items): intensity of pleasure during activities, frequency of pleasure during activities, intensity of expected pleasure from future activities;
- **Avolition/Apathy** (two items): avolition - behavior, avolition - internal experience;
- **Distress** (single item): distress.

All ratings are based on a semi-structured interview with prompts and queries. Items are rated on a 7-point scale, ranging from the symptom being no impairement/normal (0), very slight deficit/questionable (1), mild deficit/mild (2), moderate deficit/moderate (3), moderately severe deficit/moderately severe (4), marked deficit/severe (5) to severe deficit/extremely severe (6). A total BNSS score is calculated by summing the 13 individual items so that the possible total score ranges from 0 to 78. Subscores are calculated by summing the individual items within each domain (subscale). The primary efficacy endpoint, the change in avolition/apathy subscale ranges from 0 to 12.

Higher scores indicate more negative symptoms. Hence negative values for change from baseline indicate improvement.

In study **Part A**, the BNSS is assessed by an approved remote centralized rater at Week -4, Day 1 (Baseline), Day 14, Day 28, Day 42, Day 56, Day 70, Day 84/EOT and Day 112 (FU). All assessments will be assigned to analysis study days according to the following table.

Assessment Days	Assigned Analysis Study Day
>88 or >4 days after last study medication	FU, will not be included in statistical analysis
78-88	84
64-77	70
50-63	56
36-49	42
22-35	28
2-21	14
-2-1	1
<-2	-35, will not be included in statistical analysis

Descriptive summary statistics and listings will be presented for the BNSS **Avolition/Apathy subscore** by visit and treatment arm for the modified ITT analysis population by treatment arm to which participants were randomized. The same descriptive summary statistics and will be presented for the change from baseline in BNSS Avolition/Apathy subscore.

4.5.2 Secondary Efficacy Endpoints

Descriptive summary statistics will be presented for all secondary efficacy endpoints by visit and treatment arm for the modified ITT analysis population by treatment arm to

which participants were randomized. The same descriptive summary statistics and will be presented for the change from baseline.

Descriptive summary statistics will include the number of patients contributing to the summary statistic as well as the mean, median, lower and upper quartile, standard deviation, minimum and maximum.

In general, if at least one item score is missing or not ratable, the total/factor/subscale scores that include the item will be set to missing.

4.5.2.1 CGI-S and CGI-I Overall and Negative Symptoms

The CGI rating scales are tools used to evaluate both the severity of illness and change from baseline. It provides an overall clinician-determined summary measure that takes into account all available information, including a knowledge of the participants' history, psychological circumstances, symptoms, behavior, and the impact of the symptoms on the participant's ability to function.

CGI-S and CGI-S NS

The CGI-S is a 7-point scale ranging from 1 (normal, not at all ill) to 7 (among the most extremely ill patients) (see Appendix 5). On the scale, the value 0 corresponds to 'Not assessed', which will be excluded from the analyses. The values 1 to 7 will be transformed into the 0 to 6 range prior to statistical analysis. The CGI-S is an absolute measure assessing how mentally ill the participant is at the time of the assessment.

Smaller values indicate less severity of illness. Hence negative values for change from baseline indicate improvement.

The CGI-S NS is very similar to the CGI-S and will be analyzed in the similar manner.

In study **Part A** CGI-S along with CGI-S NS will be evaluated at Day 1 (baseline), Day 14, Day 28, Day 42, Day 56, Day 70, Day 84 /EOT, Day 112 (FU) by an experienced clinician. All assessments will be assigned to analysis study days as described for the primary endpoint.

The CGI-S and CGI-S NS scores and change from baseline will be summarized.

CGI-I and CGI-I NS

The CGI-I, a 7-point scale rating, is used to assess the clinical changes as compared to symptoms at baseline. The scale requests an assessment by how much, compared to baseline, the condition of the participant has changed (see Appendix 5). Values range from 1 (very much improved) to 7 (very much worse) where the value 0 corresponds to 'Not assessed', which will be excluded from any analyses. The values 1 to 7 will be centered prior to analysis so that 0 becomes 'no change' and the values -3 and +3 correspond to 'very much improved' and 'very much worse', respectively.

The CGI-I NS is very similar to the CGI-I and will be analyzed in the similar manner.

In study **Part A** CGI-I along with CGI-I NS will be evaluated at Day 14, Day 28, Day 42, Day 56, Day 70, Day 84 /EOT, Day 112 (FU) by an experienced clinician. All assessments will be assigned to analysis study days as described for the primary endpoint.

The The CGI-I and CGI-I NS scores will be summarized. There is not baseline value for CGI-I and CGI-I NS, therefore no change from baseline can be derived.

4.5.2.2 PANSS Total and Symptom Factor Scores

PANSS assessments are performed by an approved remote centralized rater.

The PANSS is a 30-item clinician-rated instrument for assessing positive, negative and other symptoms in partients with schizophrenia (see Appendix 4). The symptoms are rated on a 7-point scale capturing absent to extreme psychopathology, and the tool demonstrated sensitivity to effects seen with medication. The assessment for each of the 30 items is to allow calculation of a total score describing overall symptomatology as well as the factor scores (the classification of symptoms into five factors: negative symptoms; positive symptoms; disorganized thought; uncontrolled hostility/excitement; and anxiety/depression) which are calculated for each factor. Additonally, negative and positive symptom factore scores according to Marder (PANSS-NSFS, PANSS-PSFS, see Table 5 in protocol V2) are used.

Each item of the PANSS is rated on a 7-point scale on the basis of the following anchors: 1 = Absent, 2 = Minimal, 3 = Mild, 4 = Moderate, 5 = Moderately Severe, 6 = Severe, and 7 = Extreme. For the analyses of PANSS data, the scores will be transformed into 0 to 6 points with “Absent” expressed as 0.

Higher scores indicate more positive and negative symptoms. Hence negative values for change from baseline indicate improvement in symptoms.

In study **Part A**, the PANSS is assessed at Week -4, Day 1 (Baseline), Day 14, Day 28, Day 42, Day 56, Day 70, Day 84/EOT, Day 112 (FU). All assessments will be assigned to analysis study days as described for the primary endpoint.

The PANSS total factor and symptom factor scores and change from baseline will be summarized.

4.5.2.3 BNSS Total and Subscores

See section 4.4.1.1.

The BNSS total and subscores (other than Avolition/Apathy) and change from baseline will be summarized.

4.5.2.4 DPAS Scores

The DPAS is a 15-item, patient-reported assessment that evaluates expectations of failures or self-defeating beliefs related to prior failed experiences as well as illness on a 7-point Likert scale ranging from disagree totally (1), disagree very much (2), disagree slightly (3), neutral (4), agree slightly (5), agree very much (6), to agree totally (7) (for items see Appendix 6). For the analyses of DPAS data, the scores will be transformed into 0 to 6 points with "disagree totally" expressed as 0.

Higher scores indicate higher expectation of failure or self-defeating beliefs. Hence negative values for change from baseline indicate improvement.

In study **Part A** DPAS is assessed at Day 1 (baseline), Day 42, and Day 84/EOT. All assessments will be assigned to analysis study days according to the following table:

Assessment Day	Assigned Study Day
>88 or >4 days after last study medication	FU, will not be included in statistical analysis
64-88	84
2-63	42
-2-1	1
<-2	-35, will not be included in statistical analysis

The DPAS sum score and change from baseline will be summarized.

4.5.3 Exploratory Efficacy Endpoints

4.5.3.1 Effort-Choice Benefit Task (ECBT)

The ECBT assesses the degree to which participants are willing to engage in effortful responding as a function of the probability and magnitude of potential rewards. Participants with negative symptoms show reduced willingness to engage in effortful behavior. The ECBT task is assigned to the participants to complete.

In the experiment, participant has the choice between a low effort task and a high effort task. When making a choice, the participant will be aware of the number of **pumps ('effort')** required for the low (always 20 pumps) and the high (100, 120, or 150 pumps) effort task. Also, the participant is informed about the maximum possible **reward** (1 point for the low effort task and 3, 5, 7 points for the high effort task). Each set of cumulated 20 points covert to a e.g. $1 bonus: i) if the low effort task is selected, the participant needs to complete all 20 pumps to obtain his reward; ii) if the high effort task is selected, the participant will obtain the reward after 25 seconds, and the reward will be in proportion to the required amount of effort. For example, if the participant selects the task for 5 points and 120 pumps, and end up pumping 90 times, he receives a reward of $5 \times 90/120$ points. The participant can earn more than 5 points if pumping more than 120 times. The participant will be told the **certainty** with which the reward will be paid (100% or 50%). This certainty applies equally to the low and high effort task. Furthermore, the high effort task (in balloon) may appear on the left or the right **side** of

the screen. In total this leads to 3 × 3 × 2 × 2 = 36 different combinations of 'number of pumps for high effort task (# of pumps required for maximum reward in high effort task, '100', '120', '150')', 'reward of high effort task (3, 5, 7 points)', 'reward certainty (100% or 50%)', and 'side of display of high effort task ('left' or 'right')'. Each combination is presented twice (variable 'trialno', 1 or 2), leading to the total of 72 experiments per session.

In study **Part A** the ECB task will be performed at Day 1 (baseline), Day 14, Day 2, Day 42, and Day 84/EOT.

Descriptive Summaries:
Overall and separately for all levels of the factors 'reward' (levels 3, 5, or 7) and within each patient, the percentage how often the high effort task was selected will be derived. These percentages obtained for each patient will then be summarized by treatment per assessment. The summaries will show the mean, median, lower and upper quartile as well as minimum and maximum.

Exclusion of data from analysis:
In all analyses, data from a session will only be included if the first two of the following three QC criteria applied to the data collected from a patient in a session are met:

1. The proportion of failed trials must be below or equal to 20%;
2. The proportion of same side choices must be below or equal to 85%; and
3. The proportion of same balloon type choices must be below or equal to 85%.

4.5.3.2 Reinforcement Learning Working Memory (RLWM) task

The RLWM task assesses the ability to learn from positive reward and establish corresponding mental value representations, functions that are deficient in patients with negative symptoms. The RLWM task is assigned to the participants to complete.

Task Description:

Learning Phase (LP): To complete the RLWM task, at the learning phase, participants have to learn to select one out of three button presses for each **stimulus** (=**image**): on each presentation (trial), a single stimulus is presented to which participants can respond with one of three responses (button presses). Participants have to learn over time which of those responses is the correct one for each stimulus, based on the binary deterministic feedback they receive.

The stimuli are presented in **blocks**. The number of stimuli in the block to be learned at a time ranges from 2, 3, 4, 5, 6. There are two blocks of size n_s =2 (i.e., two blocks containing two stimuli each), three blocks of size n_s= 3, two blocks or size n_s= 4, three blocks of size n_s =5. The total number of blocks is 10. The stimuli in each block correspond to a different category of images (e.g., sports, fruits, places, etc.) leading to a

total of 2×2 + 3×3 + 2×4 + 3×5= 36 different stimuli. Within each block, stimuli are presented in a pseudo-randomly intermixed order, with 10 presentations of each stimulus within a block, i.e., a total of 360 presentations within a session. When participants select the incorrect action for a stimulus, they receive feedback indicating they have won 0 points, but if they select the correct action, they obtain a reward of 1 or 2 points. Each stimulus has a **fixed** probability *q* of resulting in 2 (vs. 1) points of reward for correct actions (*q* is 0.25, 0.50, and 0.75); the probabilities are pre-assigned to counterbalance across set sizes and other factors[1]. The correct action is always correct and the incorrect one is always incorrect. The participant would not know the probability q of receiving 2 points (rather than 1 point) in case of a correct answer. Although the participants have no control over the reward values, their striatal reinforcement learning system is expected to learn to represent higher expected reward values for those stimuli associated with higher probability of two points, which can be assessed in the test phase.

Test Phase (TP): After the learning phase, participants face a test phase, in which they are presented with **pairs of stimuli** they previously encountered. They are asked to choose the stimulus that they perceive to have given them most points in the learning phase. Instead of selecting the response that had been 'correct' for each of the individual stimuli, in this phase participants are asked to select among pairs of stimuli and to indicate their preferences (more likely to choose ones they think had led to more points). They receive no reward. The test phase includes 156 trials, i.e., 156 pairs of stimuli. These 156 pairs are selected among all 1176 (49 choose 2) possible pairs. The selection is based on the actual responses of the participant in the learning phase and done in a way to ensure that sufficient pairs with a range of 'Value Differences' but also 'Value Means' (see below) are included. Participants are more likely to select the more rewarding stimuli, and they do so more reliably as the difference between their reward values grows. This test also allows us to assess whether participants perform better either **at choosing the most rewarding stimuli or at avoiding the least rewarding stimuli, when each of these are paired with neutral stimuli.** Such difference in performance in these two situations is related to striatal dopaminergic measures in various studies.

With lower set sizes, participants can rely on (fast but subject to decay in time and capacity-limited) working memory (WM). As set size exceeds their WM capacity, they have to rely on incremental (slow, cumulative learning) reinforcement learning (RL). The paradigm can thus isolate WM from RL processes that are striatal dependent.

[1] As an example, there are 2 blocks of size 2 each, leading to a total of 4 different images across these two blocks. Among these 4, two each may be assigned probability 0.25, 0.50, and 0.75; the remaining two may be assigned either of the three. Similarly, there are 3 blocks of size 3. Across the 9 images, each probability is assigned to three images, but not all images from the same block need to have distinct probabilities.

Assessment schedule are at Weeks 1 (baseline), 2, 6, 12 for both Parts A and B. No assessment is planned for Screening or Follow-up.

Statistical Analyses

At the RLWM Learning Phase the **proportion of correct choices** in **late trials** and for **large blocks** of stimuli is investigated while at the RLWM Testing Phase the interaction '**value difference modulated by value mean**' is investigated to assess the 'choose A avoid B' paradigm.

Repeated presentations of a stimulus within a block are referred as **'iteration**': iteration 1 refers to the first presentation of a stimulus, iteration 2 to the second presentation of a stimulus etc. There are 10 iterations per stimulus. Iterations 2 and 3 are referred to as 'early trials', whereas the iterations 8 to 10 are referred to as **'late trials'**. All stimuli are presented 10 times, but there may be on occasion less than 10 responses by a participant e.g., due to the participant not responding within the required time. Such records are retained in the data set with a missing value for participant response. The records will be excluded from any analysis.

RLWM LP: The percentages of correct choices will be calculated for each iteration within each block size. These will be derived in a sequential manner: first, across all blocks of the same size, i.e., separately for blocks of size 2, 3, 4, and 5, the percentage of correct choices per each iteration will be derived for each participant. Then those percentages per iteration within each block size per participant will be averaged across all participants by treatment arm. Descriptive statistics for the **percentages** in **late trials** (iteration 8, 9, 10) and for **large blocks** (block size 4, 5) will be displayed by treatment arm.

RLWM TP: For the RLWM Testing Phase the following factors will be considered in the analysis:

- **Value Difference (F1).** The **'mean value' of a stimulus** is derived based on the average number of points the participant gained in the LP for this stimulus: the sum of all points gained for the stimulus divided by the total number of responses made (typically 13 but may be less). The **value difference** is derived by ('mean value' of right stimulus) – ('mean value' of left stimulus) for the given pair of stimuli. 'Value-difference' ranges between -2 (the left is more rewarding) and 2 (the right is more rewarding). Value difference is an indicator of **RL** value learning.

- **Value Mean (F2).** The **'value mean' of a pair** of stimuli is the average of the two mean values of the pair in the presentation. Values range from 0 to 2.

- **Value Difference Modulated by Value Mean (F12).** Factor F12 will assess the interaction between value difference (F1) and value mean (F2). F12 is an indicator of value learning bias, corresponding to the classic Probabilistic Selection Task (PST)

"choose A vs. avoid B" measure (assessing whether participants perform better either at choosing the most rewarding stimuli or at avoiding the least rewarding stimuli, when each of these are paired with neutral stimuli).

- **Block Size Mean (F3)** is the mean of the sizes of the blocks the two stimuli originated from in the learning phase. For example, if one stimulus (image) in a pair stems from a block of size 2 and the other stimulus stems from a block of size 4, the set size mean is 3. Possible values are from 2 to 5 in steps of 0.5.

- **Value Difference Modulated by Block Size Mean (F13).** Factor F13 will assess the interaction between value difference (F1) and block size mean (F3). High values for block size mean (e.g., 5 or 6) imply that both stimuli in a pair are from rather large blocks (e.g., of sizes 5 or 6). In such cases the effect of value difference may be less pronounced compared to what might be observed if both stimuli in a pair are from smaller blocks. Smaller block sizes might enable more robust learning leading to a more pronounced effect of value difference for smaller compared to large block size means. Alternatively, larger block sizes might require more use of RL over WM and thus lead to a more pronounced effect of value difference.

- **Block Size Difference (F4)** is the difference in the sizes of the blocks the two stimuli originated from in the LP. For example, if one stimulus in a pair stems from a block of size 2 and the other image stems from a block of size 4, the block size difference is 2. Block size differences will also be calculated as 'block size right stimulus originated from' minus 'block size left stimulus originated from'; hence results can range from -4 to 4 in steps of 1.

- **Learning Block Difference (F5).** 13 blocks are used in the LP. If e.g. the left image was learned in the 2nd block and the right image in the 6th block, the learning block difference will be 4 (=6-2). Differences will be derived as 'right' minus 'left'. Values can thus range from -12 to 12 in steps of 1.

- **Previous Trial Choice (F6).** The factor is coded as 'L' or 'R' and indicates the side of the previously chosen stimulus. Significant effects on this factor may be interpreted as an increase in choice stickiness, and thus possibly a decrease in involvement in the task.

For each pair of stimuli, denote by *z* the binary response taking values 1 (if the participant chose the right stimulus in the pair) and 0 (if the participant chose the left stimulus in a pair).

Logistic regression analysis modelling the likelihood the participant **chooses the right stimulus of a pair (variable z)** will be performed for each session with model

$$\log(p/(1-p)) = \beta_0 + \beta_1 \cdot F1 + \beta_{12} \cdot F12 + \beta_{13} \cdot F13 + \beta_4 \cdot F4 + \beta_5 \cdot F5 + \beta_6 \cdot F6 + \varepsilon.$$

Factor F2 will not included because there is no reason to believe why value mean should be systematically associated with the likelihood to choose the right stimulus of the pair. Only the three parameters $\boldsymbol{\beta_0}$, $\boldsymbol{\beta_1}$, and $\boldsymbol{\beta_{12}}$ are of primary interest. The other factors have been added to the model to improve the model fit but (even if significant) are not of scientific interest. Any records where any of the six covariates is missing or where information on the side chosen is not available will be excluded from analysis. Based on the resulting data set,

a) Each variable of F1 to F6 will be standardized within each participant and visit by subtracting the within participant and visit mean and dividing by the within participant and visit standard deviation.
b) The two interaction terms F12 and F13 will then be derived as the product of the corresponding standardized variables F1 and F2 as well as F1 and F3.
c) Subsequently, the two interaction terms F12 and F13 will again be standardized using the same approach as described in a).

The regression coefficients will allow the following interpretation:

Factor F1 (Value Difference): Positive values for β_1 will mean that the likelihood to choose the right side will increase with the (right minus left) value difference F1. Positive value differences imply that the right side is the correct choice; hence positive values for β_1 imply increasing likelihood to make a correct choice with increasing difference between stimuli.

F1 and F2 (Value Mean) Interaction (F12, Value Difference modulated by Value Mean): Effect of F1 (value difference) on *p* depends on the level of F2 (value mean): F2 being close to 0 or close to 2 imply that there can only be very small value differences in such pairs whereas F2 being close to 1 allows F1 (right minus left) to range from -2 to 2. Hence, a positive association between *p* is expected for F2 around 1 but less so for F2 close to 0 or 2 and should manifest itself in the β_{12} interaction term associated with F12. Positive values of β_{12} are assiociated with “choose A (choose one with better reward)” being improved over "avoid B (avoid one with less reward)", while negative values would correspond to “choose A” being worse than “avoid B”. An increase in β_{12} is expected for an effective treatment.

F1 and F3 (Block Size Mean) Interaction (F13): Like F12 smaller block size means are expected to lead to a stronger association of value difference (F1) with the likelihood to choose the right stimulus.

Note that main effects for F2 and F3 are not included. The value mean (F2) of a pair should not impact the likelihood to choose R or L. Same applies to F3, which also as such is not expected to have an effect.

Factor F4 (Block Size Difference): Significant effects of F4 (i.e., positive coefficients for β_4) could be interpreted as a change in effort avoidance, as it indicates a tendency to prefer those stimuli encountered from easier blocks independently of learned value. In other words, F4 measures to what extent the choice of the participant is driven by block size beyond value they have gained.

Factor F5 (Learning Block Difference): i) In case a participant would always correctly choose the higher value stimulus and ii) assuming the higher value stimulus is equally often on the right or left side, the coefficient β_5 associated with factor F5 in the model above is expected to zero. Hence, significant effect of F5 could be interpreted as a change in long term memory use, as it indicates a tendency to be sensitive to recency. F5 is intended to capture effects on learning that are independent of objective value. Decision making literature shows that many other factors may influence choice, beyond strict economic preference. A preference based on how recently the stimulus was seen is a possible factor, which F5 is attempting to capture.

The analysis will focus on the evaluation of the coefficient β_{12} associated with the interaction 'value difference modulated by value mean' in the model above. This is considered to capture the "choose A/avoid B" paradigm. Model

$$Y_{ijk} = \mu + \tau_i + \pi_j + \varepsilon_{ijk},$$

where μ is the general mean, τ_i are the effects of treatment, π_j are the effects of visit, and ε_{ijk} are random errors assumed to be correlated across visits within each participant, i.e., $Cov(\varepsilon_{i1k,}\, \varepsilon_{i'2k}) = Cov(\varepsilon_{i2k,}\, \varepsilon_{i'3k}) = \sigma_1$ and $Cov(\varepsilon_{i1k,}\, \varepsilon_{i'3k}) = \sigma_2$ and $Var(\varepsilon_{ijk}) = \sigma^2$, will be applied to the coefficients β_{12}. Results will be reported as least squares means by treatment arm as well as differences between treatment arms together with 90% confidence intervals.

4.5.3.3 Brief Assessment of Cognition in Schizophrenia (BACS)

To explore the effects of RO6889450 on overall functioning/cognition, the Brief Assessment of Cognition in Schizophrenia (BACS) is employed.

The BACS is designed to measure cognition in schizophrenia. The domains of cognitive function measured by the BACS include:

- verbal memory (measure: number of items recalled total, range: 0–75, ZDSCAT= VERBAL MEMORY; VERBAL MEMORY TSCORE),
- working memory (measure: digit sequencing, number of correct sequences, range: 0–28, ZDSCAT=DIGIT SEQUENCING; DIGIT SEQUENCING TSCORE),
- motor speed (measures: the number of tokens correctly placed into the container, range: 0–xx, ZDSCAT=TOKEN MOTOR TASK; TOKEN MOTOR TASK TSCORE),

- attention and speed of processing (measure: symSbol coding - number of correct numerals, range: 0–110, ZDSCAT=SYMBOL CODING; SYMBOL CODING TSCORE),
- executive functions (measure: tower of london, number of correct responses, range: 0–22, ZDSCAT=TOWER OF LONDON; TOWER OF LONDON TSCORE), and
- verbal fluency (measure: number of words generated = verbal fluency total, range: 0-xx, ZDSCAT=VERBAL FLUENCY; VERBAL FLUENCY TSCORE)

The BACS domains found to be impaired in schizophrenia. The BACS is found to be a valid outcome measure for the effects of medication, and therapy on cognitive function, and is an indicator of functional relevance with respect to independent living skills, performance-based assessments of everyday living skills, and interview-based assessments of cognition in participants with schizophrenia. The BACS is administered by qualified site raters using the tablet-based BACS App.

In study **Part A**, the assessments for BACS are performed at Day 1 (baseline), Day 42, and Day 84/EOT. All assessments will be assigned to analysis study days according to the following table:

Assessment Day	Assigned Study Day
>88 or >4 days after last study medication	FU, will not be included in statistical analysis
64-88	84
2-63	42
-2-1	1
<-2	-35, will not be included in statistical analysis

In general, if at least one item score is missing or not ratable, the total/factor/subscale scores that include the item will be set to missing.

The BACS measure from each of the 6 test (number of items recalled total, number of correct sequences, number of tokens correctly placed, cat) will be provided by absolut values and standardized t-scores (verbal memory, digit sequencing, token motor, verbal fluency, symbol coding, tower of London, BACS composite score) will be provided.

BACS t-scores and change from baseline in t-scores will be summarized by assessment visit.

4.5.3.4 Virtual Reality Functional Capacity Assessment Tool (VRFCAT)

To explore the effects of RO6889450 on overall functioning/cognition, the Virtual Reality Functional Capacity Assessment Tool (VRFCAT) is employed.

The VRFCAT is a computerized measure of functional capacity, with the potential to demonstrate real-world functional improvements associated with cognitive change. It presents participants with a realistic simulated environment to recreate routine activities

of daily living (Ruse et al 2014). It consists of four mini-scenarios that follow a story, including: (1) exploring a kitchen, (2) catching a bus to a grocery store, (3) finding/purchasing food in a grocery store, and (4) returning home on a bus.

In study **Part A**, the assessments for VRFCAT are performed at Day 1 (baseline), Day 42, and Day 84/EOT. All assessments will be assigned to analysis study days according to the following table:

Assessment Day	Assigned Study Day
>88 or >4 days after last study medication	FU, will not be included in statistical analysis
64-88	84
2-63	42
-2-1	1
<-2	-35, will not be included in statistical analysis

VRFCAT adjusted total time, errors and forced progressions t-scores and change from baseline in t-scores will be summarized.

4.5.3.5 Smoking Dependence

To explore the effects of RO6889450 on changes in smoking dependence Fagerstrom Test for Nicotine Dependence (FTND) total score is employed.

The Fagerstrom Tolerance Scale is used to assess nicotine consumption during the study. The Fagerstrom test is a standard instrument for assessing the intensity of physical addiction to nicotine providing an ordinal measure of nicotine dependence related to cigarette smoking. It contains six items that evaluate the quantity of cigarette consumption, the compulsion to use, and dependence. Yes/No items are scored from 0 to 1 and multiple choices are scored from 0 to 3. The items are summed to yield a total score of 0 to 10, see Appendix 8). The higher the total Fagerstrom score, the more intense is the participant's physical dependence on nicotine.

In study **Part A**, FTND is assessed at Day 1 (baseline), Day 42, Day 84/EOT and Day 112 (FU)..

In general, if at least one item score is missing or not ratable, the total/factor/subscale scores that include the item will be set to missing.

The FTND total score and the change from baseline in total score will be summarized by assessment visit.

4.5.3.6 Levels and Patterns of Social and General Activity (optional)

'Phone Usage Monitoring' will be carried out as an exploratory and optional tool to help understand a potential therapeutic effect of RO6889450, regarding participant's functioning and behavior and to refine such outcome measures for future studies.

Participants who have a personal Android-based smartphone will be asked to download a Phone Usage Monitoring App.Through the Smartphone App the phone usage of the participants will be monitored and evaluated, from which levels and patterns of social and general activity of the participants can be inferred. The App will record:

- **Time spent on the phone and on different Apps**: the power state of the phone (screen on/off, power connected/disconnected, and shutdown/restart/boot) and all screen touch events will be recorded. Given the wide variety of apps, each app is grouped into one of the three broad categories: social, engagement, and entertainment. Every 15 mins the total number of apps that belong to each of these three categories will be computed from the process stack. Then the increases in the number of apps for each of categories will be computed.

- **Number and timing of phone calls, and their duration**: the number and duration of incoming and outgoing calls per day will be recorded

- **Number and timing of messages**: the number of incoming and outgoing text messages per day will be recorded

- **Number of contacts and level of interaction with contacts**

- **Ambient noise**: the mean audio amplitude will be computed to determine the acoustic conditions raging from quiet to loud environments. Also the standard deviation of the audio amplitude will be computed to gauge if the audio environment is ambient (e.g., loud air conditioning) or active (e.g., people talking).

- **Location**: the GPS chip will be used to record the spatial location of the phone over time. Total distance traveled, maximum displacement from the home, standard deviation of distances, location entropy, and duration of time spent at primary location and at secondary location will be recorded on a daily basis. A locational routine index over a period of time will be computed to quantify the degree of repetition in terms of places visited. The number of new locations visited in a day that have not been seen previously can be also monitored. Sampled location readings/coordinates will be clustered into primary, secondary or other location over the entirety of a participant's data. The first and second largest clusters are labeled as the primary and secondary locations, respectively.

Collected data will be immediately encrypted and buffered on the device until a Wi-Fi connection becomes available, at which point the data are uploaded to a database on the study server. The data can be analyzed using an evolving suite of software.

The decision to install the App is optional but encouraged. Participants can at any time point decide to remove the App from the smartphone.

The phone usage is continuously monitored from the placebo compliance period up to end of the follow-up, i.e., continuously from Week -2 to Day 112 (FU)

4.5.3.7 Due to the small number of patients no analysis will be done for Part A.CYP2C19 Status on PK Parameters

To assess CYP2C19 status on PK parameters, genotyping of CYP2C19 will be performed.

A mandatory whole blood sample will be taken for DNA extraction from every participant. The DNA may be used for the anaylsis of genetic variants of cytochrome P450, such as CYP2C19, that might affect the metabolism, PK, pharmacodynamics, or safety of RO6889450.

It has been suggested that genetic variation of CYP2C19 associated with decreased CYP2C19 enzyme activity might be related to an altered PK response to the study treatment. Therefore, to assess the affect of genetic variation of CYP2C19 on PK responses to RO6889450 the genotyping for CYP2C19 will be performed. Genotyping of CYP2C19 would allow for assay of the alleles frequencies. The different alleles of CYP2C19 that make up a genotype of a participant would indicate poor or enhanced metabolizer status.

Due to the small number of patients no analysis will be done for Part A.

4.5.4 Other Efficacy Analyses

4.5.4.1 Sleep, Mood, Well-Being and Cognitive Functioning

Sleep, mood, cognition and treatment expectation questions will be presented on the AiCure app at selected time points between Day 1 and Day 112 that will appear as random to the participant. Likert scales assessing sleep quality, mood / well-being, and cognitive functioning will be administered via the smartphone (see Appendix 9). Participant will be instructed to fill in the Likert scales (made of six smiley icons) at defined time points during the trial. In addition, participants will be asked to answer the following question at the end of the study: “Do you think you were taking placebo or drug?”.

The Likert scale is a 6-point scale rating values range from 0 to 5. The higher the values, the better is the subjective well-being and cognitive functioning. Hence positive values for change from baseline indicate improvement.

The Likert scale data will be aggregated in 3-weeks intervals, i.e. ≤ Day 1, (Day 1, Day 21], (Day 21, Day 42], (Day 42, Day 63], (Day 63. Day 84], (Day 84, Day 112]. If more than one value is in an interval, the last value in the interval will be used.

The Likert scale scores and change from baseline will be summarized by time interval.

PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

To evaluate the PK of RO6889450 and RO6889450-derived metabolite(s), concentration per time point, AUC_{ss} and C_{max} of RO6889450 and RO6889450-derived metabolite(s), and other PK parameters as appropriate will be collected.

Blood samples for measurement of plasma concentrations of RO6889450 and derived metabolite(s) will be collected at Day 7, Day 14, Day 42, Day 84/EOT, Day 112 (FU).

Plasma concentrations will be summarized descriptively by treatment arm per time points: the number of non-missing values, arithmetic mean, geometric mean, coefficient of variation, standard deviation, minimum and maximum will be reported.

AUC_{ss}, C_{max} of RO6889450 and, if feasible, of RO6889450-derived metabolite(s) and other PK parameter as appropriate will be analyzed.

Population PK analyses will only be performed after Part B is completed.
Pharmacodynamic effects will not be assessed in this study.

4.6 SAFETY ANALYSES

The safety analysis will be conducted to compare the safety and the tolerability of 12 weeks of treatment with RO6889450 as monotherapy (Part A) with placebo. Safety assessments will consist of monitoring and recording AEs, including SAEs, measurement of protocol-specified safety laboratory assessments, measurement of protocol-specified vital signs, ECGs, and other protocol-specified tests that are deemed critical to the safety evaluation of the study.

All safety analyses will be based on the safety analysis population grouped according to the treatment assigned at randomization.

4.6.1 Exposure of Study Medication

The following extent of exposure to randomized study medication will be summarized by randomized treatment arm:

- Duration of treatment (weeks), which will be calculated from the first day of randomized double-blind study medication intake as recorded with AICURE to the last day of randomized double-blind study treatment intake as recorded with AICURE, i.e. last dosing date minus first dosing date plus 1 day (divided by 7 to get weeks).

- Number of doses over the whole double-blind treatment period derived from the number of days at which randomized double-blind treatment was taken as recorded with AICURE.

- Total cumulative dose (mg) over the whole double-blind treatment period as

recorded with AICURE.

- Extent of compliance with randomized double-blind treatment (%). Treatment compliance for each participant will be calculated as actually taken dose divided by planned dose of double-blind treatment. The actually taken dose will be the total cumulative dose. The planned dose is the cumulative planned dose over 84 days, i.e., [redacted] ×84 days.

4.6.2 Adverse Events

The original terms recorded on the eCRF by the Investigator for AEs will be assigned by the Sponsor to a standardized term (the "preferred term") based on MedDRA. All data displays of AE will be performed using the system organ class and preferred terms. All AEs will be presented as frequency tables by randomized treatment arm with numbers and percentage of participants experiencing AEs.

AEs will be separately summarized for the following analysis periods:

- Screening/Wash-out period for antipsychotics includes the AEs with onset before the first day of study medication in the randomized double-blind treatment period (Baseline Sign and Symptoms).

- Double-blind treatment period including follow-up includes all AEs from day of first onset of study medication in the double-blind treatment period until end of follow-up.

Intensity and relationship to study treatment will be summarized for the double-blind treatment period including follow-up as assigned by the investigator.

In the summary table of AEs by intensity, if a participant has more than one occurrence of an event, the event with the most severe intensity will be counted. If the intensity of an AE is missing, then the AE will be included only in the total number of events column and not in the count of participants with the event by intensity.

In the summary table of AEs by relationship to trial treatment, if a participant has more than one occurrence of an event, the most closely related event will be counted. If the relationship of an AE is missing, then the AE will be included only in the total number of events column and not in the count of participants with the event by relationship.

All SAEs and AEs that led to death, and AEs that led to withdrawal of study treatment will also be summarized by treatment arm.

The following rules will be applied for AEs with missing onset or end dates:

- Events that are missing both, onset and end dates, will be considered as emerging from treatment, given the participant had at least one dose of study medication.

- If the onset date is missing and the end date is on or after the first dosing date, then the event will be considered as emerging from treatment.

- If the end date is missing and the onset date is on or after the first dosing date of a specific period, then the event will be considered as emerging from treatment.

- If the end date is missing and the most extreme intensity is worse than the initial intensity and the onset date is before the first dosing date of a specific period, then the event will be considered as emerging from treatment.

- If the onset or end dates are missing, the duration will be set to missing.

The analysis will be based on the safety analysis population grouped according to the treatment assigned at randomization.

Time to onset of exacerbation of schizophrenia (AE preferred terms Schizophrenia, Psychotic dissorder) will be evaluated by summary statistics.

4.6.3 Clinical Laboratory Tests

The Investigator will review the laboratory report, document this review, and record any clinically relevant changes occurring during the study in the AE section of the eCRF.

In study **Part A**, hematology, chemistry and lipids are assessed at Week -4, Day 1, Day 14, Day 42, Day 84/EOT, Day 112 (FU). Urinalysis is additionally assessed at Day 28 and Day 56, but not at Day 42. Hematology, chemistry and lipids assessments will be assigned to analysis study days according to the following table:

Assessment Day	Assigned Study Day
>88 or >= 1 day after last study medication	FU, will not be included in statistical analysis
64-88	84
28-63	42
2-28	14
-2-1	1
<-2	-35, will not be included in statistical analysis

Urinalysis assessments will be assigned to analysis study days according to the following table:

Assessment Day	Assigned Study Day
>88 or >=1 day after last study medication	FU, will not be included in statistical analysis
71-88	84
43-70	56
22-42	28
2-21	14

-2-1	1
<-2	-35, will not be included in statistical analysis

Clinical laboratory data are stored on the database in the units in which they were reported. For analysis laboratory data will be converted 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 and will be transformed to common reference range using Roche Safety Lab Standardisation.

Summary statistics for converted values and changes from baseline at each assessment time will be presented. Baseline is the patient's last observation prior to initiation of randomized double-blind study treatment.

The number of patients with a laboratory value abnormality in the direction specified (low/high) among patients during randomzed double-blind treatment without this abnormality at baseline will be summarized.

Mean (+/- SEM) plots including single values will be presented by scheduled time point and randomized treatment arm.

Safety laboratory test values will be presented by individual listing with flagging of values outside the normal range.

The analysis will be based on the safety analysis population grouped according to the treatment assigned at randomization.

4.6.4 Electrocardiograms (ECG) data analysis

Triplicate ECGs are required at screening, where the three individual ECG tracings should be obtained as closely as possible in succession, but no more than two minutes apart. The full set of triplicates should be completed in less than five minutes. At all other visits, single ECGs are collected. Any clinically relevant changes or abnormalities occurring during the study should be recorded in the AE section of the eCRF.

In study **Part A**, twelve-lead ECG recording will be obtained at Week -4, Day 1, Day 14, Day 56, Day 84/EOT, Day 112 (FU). ECG assessments will be assigned to analysis study days according to the following table:

Assessment Day	Assigned Study Day
>88 or >=1 day after last study medication	FU, will not be included in statistical analysis
71-88	84
36-70	56
2-35	14
-2-1	1
<-2	-35, will not be included in statistical analysis

Summary statistics for actual heart rate, PQ (PR), QRS, QT, RR, and QTcF values and changes from baseline at each assessment time will be presented. Baseline is the patient's last observation prior to initiation of randomized double-blind study treatment.

The number of patients with a ECG value abnormality in the direction specified (low/high) during randomzed double-blind treatment will be summarized.

QTcF intervals will be classified into categories ≤450 msec, >450-≤480 msec, >480-≤500 msec and >500 msec, change from baseline will be classified into categories ≤30 msec, >30-≤60 msec, >60 msec; both will be summarized by scheduled time point and overall for the maximum post-baseline interval.

ECG data will be presented by individual listings. In addition, tabular descriptive summaries for the change from baseline in ECG intervals: heart rate, PQ (PR), QRS, QT, RR, and QTcF. T- and U-waves abnormalities will be displayed.

ECG data will be presented by individual listing with flagging of values outside the normal range.

Mean (+/- SEM) plots including single values will be presented by scheduled time point and randomized treatment arm.

The analysis will be based on the safety analysis population grouped according to the treatment assigned at randomization.

4.6.5 Concomitant Medications

Concomitant therapy includes any medication, e.g., prescription drugs, over-the-counter [OTC] drugs, approved dietary and herbal supplements, nutritional supplements and any nonmedication interventions (e.g., individual psychotherapy, cognitive behavioural therapy, smoking cessation therapy, and rehabilitative therapy). As a general rule no new concomitant medications are permitted, with the exception of medications to treat AEs.

The original terms (the "verbatim term") recorded on the participants' eCRF by the investigator will be assigned to a standardized term (the "preferred term") and Anatomical Therapeutic Chemical (ATC) classification using WHODrug Global B3 Format dictionary. Analyses will use preferred terms and grouped by ATC2 (pharmacological or therapeutic properties) levels. Note: If a generic is assigned multiple classes ATC Class paths in WHODrug , it may be listed multiple times.

Previous/Concomitant medications (started before first randomized double-blind treatment intake), concomitant medications (started after first randomized double-blind treatment intake) and medications to treat AEs.will be presented by treatment arm in

summary tables and listings. Multiple occurrences of a medication in an individual (i.e., same coded term) will be counted only once in summary tables.

The analysis will be based on the safety analysis population grouped according to the treatment assigned at randomization.

4.6.6 Suicidality Assessment (C-SSRS)

The C-SSRS is a tool used to assess the lifetime suicidality of a participant (C-SSRS baseline) as well as new instances of suicidality (C-SSRS since last visit). It incorporates a structured interview to prompt recollection of suicidal ideation, including the intensity of the ideation, behavior, and attempts with actual/potential lethality.

In study **Part A**, the "baseline" version will be completed at the Screening visit (Week -4) and a "since last visit" version will be completed at subsequent visits at Day -7, Day 1, Day 7, Day 14, Day 28, Day 42, Day 56, Day 70, Day 84/EOT, Day 112 (FU).

The items of the C-SSRS will be listed. The number and percentage of patients with suicidal ideation, suicidal behavior, and self-injurious behavior without suicidal intent based on the C-SSRS during treatment will be summarized (see Appendix 7).

The analysis will be based on the safety analysis population grouped according to the treatment assigned at randomization.

4.6.7 Extrapyramidal Symptom Rating Scale, Abbreviated Version (ESRS-A)

The presence and severity of extrapyramidal symptoms are evaluated using ESRS-A. The reliability and validity of the ESRS have been demonstrated in antipsychotic-induced movement disorders. The scale is organized into two main components: (1) an assessment of specific symptoms of Parkinsonism, dyskinesia, akathisia, and dystonia (2) the Clinician's Global Impression (CGI) of these symptoms. All are evaluated on a scale of 0 (absent) to 5 (extreme).

In study **Part A**, the assessments will be carried out at the Screening visit (Week -4), and subsequent visits Day 1, Day 14, Day 28, Day 84/EOT, Day 112 (FU). Assessments will be assigned to analysis study days according to the following table:

Assessment Days	Assigned Analysis Study Day
>88 or >=1 day after last study medication	FU, will not be included in statistical analysis
57-88	84
22-56	28
2-21	14
-2-1	1
<-2	-35, will not be included in statistical analysis

The movement related rating scales (ESRS-A) will be summarized for the total score of each domain (i.e., Parkinsonism, Akathisia, Dystonia, Dyskinesia) and the associate CGIS score by treatment group using descriptive statistics at each scheduled assessment visit. Change from baseline will be summarized by timepoint using descriptive statistics. In addition, patients (n, %) with clinically relevant severity in each of the four symptoms will be summarized.

The analysis will be based on the safety analysis population grouped according to the treatment assigned at randomization.

4.6.8 CGI-S MTS and CGI-I MTS

At screening, clinicians will gather information from the patient's past psychiatric history, caregiver feedback, and clinical judgement to determine two to three prodromal or MTS (most troubling symptoms) historically associated with a signal of the patient's relapse. The patient's MTS (including frequency, intensity, interference with daily function or family life, and other salient consequences) will be recorded and these symptoms evaluated using the CGI-S MTS. At post-baseline visits, the clinician will use the MTS description and re-evaluate the MTS using the CGI-S and CGI-I. This re-assessment of the patient's prodromal symptoms will support with the ongoing clinical evaluation of relapse risk.

In study **Part A** CGI-S MTS will be evaluated at Week -4, Day -7, Day 1 (baseline), Day 7, Day 14, Day 21, Day 28, Day 35, Day 42, Day 49, Day 56, Day 63, Day 70, Day 77, Day 84 /EOT, Day 112 (FU) by an experienced clinician. All assessments will be assigned to analysis study days according to the following table:

Assessment Days	Assigned Analysis Study Day
>88 or >=1 day after last study medication	FU, will not be included in statistical analysis
81-88	84
74-80	77
67-73	70
60-66	63
53-59	56
46-52	49
39-45	42
32-38	35
25-31	28
18-24	21
11-17	14
2-10	07
-2-1	1
-8- -14	-7, will not be included in statistical analysis
<-14	-35, will not be included in statistical analysis

The The CGI-S MTS and and change from baseline will be summarized.

In study **Part A** CGI-I MTS will be evaluated at Day 7, Day 14, Day 21, Day 28, Day 35, Day 42, Day 49, Day 56, Day 63, Day 70, Day 77, Day 84 /EOT, Day 112 (FU) by an experienced clinician. All assessments will be assigned to analysis study days as described for CGI-S MTS.

The CGI-I MTS scores will be summarized.

4.6.9 Vital Signs Body Weight and BMI

Orthostatic vital signs are vital signs taken when the patient moves from supine to standing position. In the Phase I study BP30134 Postural Orthostatic Tachycardia was the most often observed change in vital signs which occurred at multiple timepoints, including pretreatment in all cohorts and placebo in both Part A (SAD) and Part B (MAD). The definition of POT is a heart rate > 120 bpm on standing or an increase in heart rate by 30 bpm from a resting heart rate without clinically significant orthostatic hypotension.

In study **Part A**, systolic blood pressure, diastolic blood pressure and heart rate are assessed in supine as well as standing position at Week -4, Day 1, Day 14, Day 28, Day 56, Day 84/EOT, Day 112 (FU). Weight is collected at at Week -4, Day 1, Day 56, Day 84/EOT, Day 112 (FU). BMI is calculated as weight (kg) / height $(m)^2$ (height as measured at baseline).

Blood pressure and heart rate assessments will be assigned to analysis study days according to the following table:

Assessment Days	Assigned Analysis Study Day
>88 or >=1 day after last study medication	FU, will not be included in statistical analysis
71-88	84
43-70	56
22-42	28
2-21	14
-2-1	1
<-2	-35, will not be included in statistical analysis

Weight assessments will be assigned to analysis study days according to the following table:

Assessment Days	Assigned Analysis Study Day
>88 or >4 day after last study medication	FU, will not be included in statistical analysis
71-88	84
2-70	56
-2-1	1
<-2	-35, will not be included in statistical analysis

Summary statistics for values and changes from baseline at each assessment time will be presented. Baseline is the patient's last observation prior to initiation of randomized

double-blind study treatment. Blood pressure and heart rate will be presented in units as collected (mmHg and beats/min), while body weight will be converted to kg.

The definition of POT is a heart rate > 120 bpm on standing or an increase in heart rate by 30 bpm from a resting heart rate without clinically significant orthostatic hypotension. To support assessment of POT, the difference between standing and supine heart rate will be calculated and summarized. Patients with a heart rate > 120 bpm in standing position or an increase in heart rate by 30 bpm from a supine heart rate will be listed.

Mean (+/- SEM) plots including single values will be presented by scheduled time point and randomized treatment arm.

The analysis will be based on the safety analysis population grouped according to the treatment assigned at randomization.

4.7 MISSING DATA

Generally, in summaries, all data from the respective analysis population regardless of completeness by Week 12 will be included; participants who have taken the study medication or have visited the sites only partially will be included in the summaries. No imputation will be applied for missing data will be applied.

For all rating scales, if any item score contributing to the total/factor/subscale score is missing, then the total/factor/subscale will be set to missing.

4.8 INTERIM ANALYSES

See section 2.4.

Appendix 1 Protocol Synopsis

PROTOCOL TITLE:	**PHASE 2, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY TO ASSESS THE EFFECTS OF RO6889450 IN PATIENTS WITH SCHIZOPHRENIA OR SCHIZOAFFECTIVE DISORDER AND NEGATIVE SYMPTOMS**
PROTOCOL NUMBER:	BP40283
VERSION:	2
TEST PRODUCT:	RO6889450
PHASE:	2

RATIONALE

The trace amine-associated receptor 1 (TAAR1) partial agonist RO6889450, a modulator of dopaminergic transmission, leads to a normalization of behaviors in rodent models of negative symptoms of schizophrenia. Results of a clinical study with RO6889450 in healthy volunteers suggest performance improvements in tasks that probe reward-based learning and motivated behavior, functions that are deficient in patients with negative symptoms (e.g., blunted affect, avolition/apathy). These results indicate a potential utility of RO6889450 in the treatment of negative symptoms associated with schizophrenia or schizoaffective disorder.

OBJECTIVES AND ENDPOINTS

Objectives	Endpoints
Primary	
Part A: To compare the efficacy of [redacted] once daily (QD) of RO6889450 as monotherapy with placebo on negative symptoms in patients with schizophrenia or schizoaffective disorder.	Part A & B: Change from baseline at Week 12 in the Brief Negative Symptoms Scale (BNSS) avolition/apathy subscore (sum of items "behavior" and "internal experience").
Part B: To compare the efficacy of [redacted] QD of RO6889450 as add-on therapy with placebo on negative symptoms in patients with schizophrenia or schizoaffective disorder.	

Objectives	Endpoints
• **Secondary**	
To compare the effect of RO6889450 with placebo on: • Clinical Global Impression Severity (CGI-S) and change (CGI-I) (overall and negative symptoms)	• Change from baseline in CGI-S and CGI-S negative symptoms scores • CGI-I, and CGI-I negative symptoms scores
To compare the effect of RO6889450 with placebo on: • Symptoms of schizophrenia or schizoaffective as assessed with the Positive and Negative Syndrome Scale (PANSS), BNSS, and Defeatist Performance Attitude Scale (DPAS)	• Change from baseline in PANSS total and symptom factor scores • Change from baseline in BNSS total and subscores • Change from baseline in DPAS scores
• To compare the safety and tolerability of 12 weeks of treatment with RO6889450 as monotherapy (Part A) or add-on therapy with placebo (Part B)	• Incidence, nature, and severity of adverse events (AEs) • Incidence, nature, and severity of serious AEs (SAEs) • Incidence, nature, and severity of treatment discontinuations due to AEs • Change from baseline in orthostatic vital signs recordings • Change from baseline in electrocardiogram (ECG) intervals: heart rate, PQ (PR), QRS, QT, RR and QTcF along with information on T- and U-waves • Incidence of laboratory abnormalities, based on hematology, clinical chemistry, coagulation and urinalysis test results • Change from baseline in Columbia-Suicide Severity Rating Scale (C-SSRS) and Extrapyramidal symptom rating scale, abbreviated (ESRS-A)
• To evaluate the pharmacokinetics (PK) of RO6889450 and RO6889450-derived metabolite(s).	• Concentration per time point • AUC_{ss} of RO6889450 and RO6889450-derived metabolite(s) • C_{max} of RO6889450 and RO6889450-derived metabolite(s) • Other PK parameters as appropriate

OVERALL DESIGN

This is a Phase 2, multicenter, randomized, double-blind, two-part, placebo-controlled, parallel-group, study in patients with schizophrenia or schizoaffective disorder.

Study Design

Part A (monotherapy):

Participants will be washed out (one week) from their antipsychotic therapy before being randomized to either [redacted] QD of RO6889450 or placebo for 12 weeks. Stratification will be based on baseline BNSS avolition/apathy subscore (BNSS ≤6 vs >6), age (18-35 years vs. 36-55 years), and sex. The participants will have the option of an inpatient stay or partial hospitalization during the washout period as well as the first week of study drug.

Part B (add-on therapy):

Participants will be randomized 1:1:1 to [redacted] QD of RO6889450, [redacted] QD of RO6889450, or placebo for 12 weeks. Stratification will be based on BNSS avolition/apathy subscore (sum of items "behavior" and "internal experience"; ≤6 vs. >6), age (18-35 years vs. 36-55 years), sex, and antipsychotic treatment (1:1 stratification ratio between dopamine 2 receptor [D2]/serotonin 2A [5HT2A] receptor antagonists and D2 partial agonist).

Participants who are not enrolled in Part A, can be screened for Part B only after approval by the Medical Monitor or delegate. The reason(s) for the decision must be documented.

Part B is planned to start after Part A. If the result of Part A is negative, the Sponsor may choose not to conduct Part B.

Interim analyses:

The Sponsor may choose to conduct an interim efficacy analysis in Part A. 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 members of the Sponsor study team and appropriate senior management personnel who will be unblinded at the treatment group level. Access to treatment assignment information will follow the Sponsor's standard procedures.

The results of this analysis may support initiation of Part B enrollment during Part A if a positive results threshold is met as defined in the statistical analysis plan (SAP). In case of very slow recruitment, Part A may be stopped and/or Part B may be started without the interim analysis.

The Sponsor may choose to conduct an interim analysis of Part B to determine the chance of final success. The results of such an analysis may support an early termination of Part B.

Treatment Groups and Duration

The investigational medicinal product (IMP) is RO6889450 (and matching placebo) and will be given orally QD at a daily dose of [redacted] in Part A and at a daily dose of either [redacted] in Part B.

Length of Study

Part A: The total duration will be approximately 147 days (from screening through study completion) for each enrolled participant as follows:

- Screening: up to 35 days (including the placebo compliance and a 7-day washout period)
- Treatment period: 84 days (12 weeks)
- Follow up: 28 days

Part B: The total duration will be up to approximately 147 days (from screening through study completion) for each enrolled participant as follows:

- Screening: up to 35 days (including the placebo compliance period)
- Treatment period: 84 days (12 weeks)
- Follow up: 28 days

End of Study

The end of the study is defined as the date when the last participant last observation (LPLO) occurs. LPLO is expected to occur approximately 28 days after the last participant's last dose.

Data Monitoring Committee

An independent Data Monitoring Committee (iDMC) will regularly review unblinded safety and efficacy (as appropriate) data. The iDMC will be composed of at least two psychiatrists external to the Sponsor and will be responsible for monitoring safety data to help assure that continuation

of the study in its current design does not pose unacceptable safety risks to participants. Planned and ad hoc safety reviews will be performed as described within the iDMC charter.

PARTICIPANT POPULATION

The participants of this study will be male and female outpatients between 18 to 55 years of age with a Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) diagnosis of schizophrenia or schizoaffective disorder with negative symptoms, who are stable. For Part B, participants also need to be stable on antipsychotic treatment. Participants who are in a residential inpatient setting are eligible for this study.

Inclusion/Exclusion Criteria

Unless otherwise stated, inclusion and exclusion criteria refer to evaluations or assessments at screening.

Inclusion Criteria:

Informed Consent

1. Able and willing to provide written informed consent according to International Council for Harmonisation (ICH) and local regulations as assessed using the Evaluation to Sign Consent. If the participant has a legal representative, the informed consent must be signed by this person as well.

Patient and Disease Characteristics

2. Male or female participants aged 18-55 years (inclusive).
3. Patients with a DSM-5 diagnosis of schizophrenia or schizoaffective disorder as confirmed by the Mini International Neuropsychiatric Interview (MINI).
4. **Part B only:** Stable treatment with a D2/5HT2A antagonists or a D2 partial agonist for a minimum of six months and receiving no more than two antipsychotics (if no blood concentration of the prescribed AP medication [or active metabolites] is detected, the participant, should not be enrolled.). Antipsychotic regimen: Patients must be on a "primary" antipsychotic and may be on a secondary antipsychotic. The amount of the secondary antipsychotic has to be equal to or less than the equivalent dose of the primary antipsychotic and the sum of the primary and secondary antipsychotics must be ≤6 mg of risperidone equivalents (see Appendix 7).
5. Medically stable during the prior six months.
6. Outpatient with no hospitalization for worsening of schizophrenia or schizoaffective disorder within the six prior months (hospitalization for social management within this time is acceptable).
7. PANSS negative symptom factor score of 18 or higher.
8. The following rating on items of the PANSS:
    (a) less than 5 on G8 (uncooperativeness), P1 (delusions), P3 (hallucinations), P4 (excitement/hyperactivity), and P6 (suspiciousness/persecution)
    (b) less than 4 on P7 (hostility) and G14 (poor impulse control)
9. Has an informant who is considered reliable by the Investigator to provide support to the participant and to help ensure compliance with study visits and protocol procedures; who preferably is also able to provide input helpful for completing study rating scales and is in regular contact with the participant in order to be able to alert the Investigator of worsening signs and symptoms for the participant.
10. Body mass index (BMI) between 18 and 40 kg/m^2 inclusive.

Reproductive

11. A female is eligible to participate if she is not pregnant, not breastfeeding, and at least one of the following conditions applies:
    (a) Not a woman of childbearing potential (WOCBP).
    (b) WOCBP, who

- agree to remain abstinent (refrain from heterosexual intercourse) or use acceptable contraceptive methods during the treatment period and for at least 28 days after the last dose of study drug.

- The following are acceptable contraceptive methods: bilateral tubal occlusion, male sterilization, established proper use of hormonal contraceptives that inhibit ovulation, hormone-releasing intrauterine devices and copper intrauterine devices, male or female condom with or without spermicide; and cap, diaphragm, or sponge with spermicide.

The contraception and abstinence requirements are intended to prevent exposure of an embryo to the study treatment. The reliability of sexual abstinence for female enrollment eligibility needs to be evaluated in relation to the duration of the clinical study and the preferred and usual lifestyle of the participant. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or post-ovulation methods) and withdrawal are not acceptable methods of preventing drug exposure.

Exclusion Criteria:

Medical Conditions

1. **Part A only:** Confirmed suicidal behavior based on Investigator judgment or violent behavior resulting in injury or property damage in the prior five years. A history prior to the last five years requires approval by the Patient Review Committee (PRC) on a case-by-case basis.
2. **Part A only:** Lifetime history of homicidal behavior
3. Moderate to severe substance use disorder within six months (excluding nicotine or caffeine) as defined by DSM-5.
4. Other current DSM-5 diagnosis (e.g., bipolar disorder, major depressive disorder).
5. PANSS item G6 (depression) greater than or equal to 5.
6. Significant risk of suicide or harming him- or herself or others according to the Investigator's judgment.
7. A prior or current general medical condition that might be impairing cognition or other psychiatric functioning (e.g., migraine headaches requiring prophylaxis treatment, head trauma, dementia, seizure disorder, stroke; or neurodegenerative, inflammatory, infectious, neoplastic, toxic, metabolic, endocrine etc. conditions).
8. Positive result at screening for hepatitis B surface antigen (HBsAg), hepatitis C (HCV, untreated), or human immunodeficiency virus (HIV)-1 and -2. HCV patients who have been successfully treated and who test negative for HCV RNA, may be considered eligible for entry into the study.
9. Tardive dyskinesia that is moderate to severe or requires treatment.
10. History of neuroleptic malignant syndrome.
11. Average triplicate QTcF interval greater than 450 msec for males and 470 msec for females or other clinically significant abnormality on screening electrocardiogram (ECG) based on centralized reading.
12. Clinically significant abnormalities in laboratory safety test results (including hepatic and renal panels, complete blood count, chemistry panel, coagulation and urinalysis), including:
    - Aspartate transaminase (AST), alanine transaminase (ALT) >2 × upper limit of normal (ULN)
    - Total bilirubin >1.5 ULN with the exception of Gilbert syndrome
    - Serum creatinine >1.5 ULN

    In case of uncertain or questionable results, tests performed during screening may be repeated before randomization to confirm eligibility or may be accepted if they are, in the opinion of the Investigator and the Medical Monitor, not clinically significant.

13. Significant or unstable physical condition that in the Investigator's judgment might require a change in medication or hospitalization during the course of the study.

Prior/Concomitant Therapy

14. On more than one antidepressant, or if on one antidepressant, a change in dose within 28 days prior to screening.
15. History of clozapine treatment.
16. Concomitant use of prohibited medications.
17. Positive urine drug screen for amphetamines, methamphetamines, opiates, buprenorphine, methadone, cannabinoids (including cannabidiol), cocaine and barbiturates. In case of uncertain or questionable results, the urine drug screen may be repeated once during the screening period to confirm eligibility.

Other Exclusions

18. Receipt of an investigational drug within 28 days or five times the half-life of the investigational drug (whichever is longer) before the first study drug administration
19. Donation of blood over 400 mL within three months prior to screening.

Confirmation of Eligibility at Baseline

1. **Part A only:** Patient will be excluded if unable to taper off an antipsychotic in the one week prior to baseline (e.g., in the case of symptom exacerbation or antipsychotics with a longer half-life).
2. Female participants who are of childbearing potential must have a negative pregnancy test result at baseline.

NUMBER OF PARTICIPANTS

Overall, it is planned to have approximately 500 participants randomized in this study in order to achieve 400 evaluable participants.

In Part A, approximately 125 participants will be randomized to one of two treatment arms to ensure that approximately 100 participants complete the treatment phase as described in this protocol, with the primary endpoint assessed at Week 12. In Part B, approximately 375 participants will be randomized to one of three treatment arms to ensure that approximately 300 participants complete the treatment phase as described in this protocol with the primary endpoint assessed at Week 12.

CONCOMITANT MEDICATIONS

Rescue Medicine During Part A

If early signs of relapse requiring treatment are noticed/recorded by the Investigator, an antipsychotic should be prescribed and the participant will be immediately withdrawn from the study drug.

Permitted Therapy

The use of stable, scheduled doses of benzodiazepines are permitted. The use of stable, scheduled doses of non-benzodiazepine hypnotics are permitted.

If deemed necessary by the Investigator to treat occasional anxiety, agitation or sleep problems, lorazepam, alprazolam, or oxazepam are allowed as needed (PRN) (up to a maximum daily dose of lorazepam 6 mg, alprazolam 3 mg, or oxazepam 90 mg). If PRN treatment of more than five consecutive days is necessary, this should be discussed with the Medical Monitor. PRN doses of benzodiazepines should not be administered within 12 hours of a clinic visit.

Non-prohibited medications used for the treatment of stable medical conditions other than schizophrenia (e.g., hypertension, diabetes, oral contraceptives, hormone-replacement therapy) are allowed during the study, provided that they have been stable for 28 days prior to screening and are intended to remain stable throughout the study up to the follow-up visits.

Patients in Part B of the study must be on stable treatment (i.e., no medication changes or significant dose adjustments) with D2/5HT2A antagonists or D2 partial agonist for a minimum of six months prior to screening. Co-administration of RO6889450 and P-gp substrates, such as risperidone, amisulpride, aripiprazole, and paliperidone, though permitted may result in slight elevations of the P-gp substrate so continued monitoring is suggested.

Prohibited Therapy

Use of the following therapies (non-topical formulation where applicable) is prohibited during the study unless otherwise specified:

- Clinically relevant substrates of P-gp, including quinidine and loperamide.

In Part A, all antipsychotic treatment as well as medications used to treat extrapyramidal symptoms (EPS) (e.g., anticholinergics) will be discontinued before randomization. Antipsychotic medications that require longer washout periods (i.e., more than two weeks), such as cariprazine, brexpiprazole, or long-acting injectable antipsychotics are exclusionary.

Patients receiving treatment for tardive dyskinesia (e.g., valbenazine or deutetrabenazine) are excluded. As a general rule, no new concomitant medications will be permitted, with the exception of medications to treat AEs, unless the rationale for exception is discussed and accepted by the Sponsor or designee and clearly documented between the Investigator and the Sponsor.

Appendix 2 Schedule of Activities for Part A of the Study

Period	Screening			Treatment Period												Follow-Up Visits		
		Placebo Compliance																
Week	Week -4	Week -2	Week -1	1		2	3	4	5	6	7	8	9	10	11	12	13	14	16
Day	Day -35 to Day -19	Day -14	Day -7	Baseline[1]/ Day 1	Day 7	Day 14	Day 21	Day 28	Day 35	Day 42[1]	Day 49	Day 56	Day 63	Day 70	Day 77	Day 84/ EOT[1, 2]	Day 91	Day 98	Day 112[1]
Visit window (days)		±1	±1	-1	±1	±1	±1	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±4
On-site visit	X	X		X	X	X		X		X		X		X		X			X
Inpatient/partial hospitalization (optional)			←		→														
Phone call[3]			X				X		X		X		X		X		X	X	
Informed consent	X																		
Inclusion/exclusion criteria	X			X[4]															
Eligibility Assessment Form submission	X																		
Latest start of Antipsychotic washout[8]			X																
Randomization				X															
Demography	X																		
Medical history	X																		
Physical examination [5]	X			X								X				X			X
ECG-12 lead[6]	X			X		X						X				X			X
Meal time	X			X		X						X				X			X

Period	Screening			Treatment Period													Follow-Up Visits		
		Placebo Compliance																	
Week	Week -4	Week -2	Week -1	1		2	3	4	5	6	7	8	9	10	11	12	13	14	16
Day	Day -35 to Day -19	Day -14	Day -7	Baseline[1]/ Day 1	Day 7	Day 14	Day 21	Day 28	Day 35	Day 42[1]	Day 49	Day 56	Day 63	Day 70	Day 77	Day 84/ EOT[1, 2]	Day 91	Day 98	Day 112[1]
Visit window (days)		±1	±1	-1	±1	±1	±1	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±4
Orthostatic vital signs	X			X		X		X				X				X			X
Viral serology	X																		
Prolactin sample	X															X			
Alcohol levels	X			X		X		X				X				X			X
Drugs of abuse	X			X		X		X				X				X			X
Urinalysis	X			X		X		X				X				X			X
Pregnancy test [7]	X			X												X			
Blood chemistry and lipids	X			X		X				X						X			X
Hematology	X			X		X				X						X			X
Coagulation parameters	X			X		X													
Antipsychotic blood levels	X				X					X						X			
PK sample					X[9]	X[9,10]				X[9,10]						X[9,10]			X
Genotyping				X[11]															
RBR DNA (optional)				X[11]															
MINI	X																		
WASI-II	X																		

Period	Screening			Treatment Period													Follow-Up Visits		
		Placebo Compliance																	
Week	Week -4	Week -2	Week -1	1		2	3	4	5	6	7	8	9	10	11	12	13	14	16
Day	Day -35 to Day -19	Day -14	Day -7	Baseline[1]/ Day 1	Day 7	Day 14	Day 21	Day 28	Day 35	Day 42[1]	Day 49	Day 56	Day 63	Day 70	Day 77	Day 84/ EOT[1, 2]	Day 91	Day 98	Day 112[1]
Visit window (days)		±1	±1	-1	±1	±1	±1	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±4
WRAT-4[17]	X																		
PANSS and IC-PANSS	X			X		X		X		X		X		X		X			X
BNSS[17]	X			X		X		X		X		X		X		X			X
DPAS				X						X						X			
CGI-S and CGI-S Negative Symptoms				X		X		X		X		X		X		X			X
CGI-I and CGI-I Negative Symptoms						X		X		X		X		X		X			X
CGI-S MTS	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X			X
CGI-I MTS					X	X	X	X	X	X	X	X	X	X	X	X			X
FTND				X						X						X			X
BACS[17]				X						X						X			
VRFCAT[17]				X						X						X			
Reward learning tasks: RLWM and ECB				X		X				X						X			
Phone usage monitoring (optional)[12]		←														→			

Period	Screening			Treatment Period													Follow-Up Visits		
		Placebo Compliance																	
Week	Week -4	Week -2	Week -1	1		2	3	4	5	6	7	8	9	10	11	12	13	14	16
Day	Day -35 to Day -19	Day -14	Day -7	Baseline[1]/ Day 1	Day 7	Day 14	Day 21	Day 28	Day 35	Day 42[1]	Day 49	Day 56	Day 63	Day 70	Day 77	Day 84/ EOT[1, 2]	Day 91	Day 98	Day 112[1]
Visit window (days)		±1	±1	-1	±1	±1	±1	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±2	±4
ESRS-A	X			X		X		X								X			X
C-SSRS	X		X	X	X	X		X		X		X		X		X			X
Adverse events[13]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Previous and concomitant treatments	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Administration of open label placebo[14]		←	→																
Administration of study drug[15]				←												→			
AiCure Adherence App		←														→			
AiCure App Questions[16]		←														→			
Study drug dispensing				X		X		X		X		X		X					

Abbreviations: BACS = Brief Assessment of Cognition in Schizophrenia, BNSS = Brief Negative Symptom Scale, CGI-I = Clinical Global Impression of Improvement, CGI-S = Clinical Global Impression of Severity, C-SSRS = Columbia-Suicide Severity Rating Scale, DNA = deoxyribonucleic acid, DPAS = Defeatist Performance Attitude Scale, ECB = Effort-choice benefit 18, ECG = Electrocardiogram, EOT = End of treatment, ESRS-A = Extrapyramidal symptom rating scale, abbreviated, FTND = Fagerstrom Test for Nicotine Dependence, IC-PANSS = informant checklist Positive and Negative Syndrome Scale, MINI = Mini International Neuropsychiatric Interview, MTS = most troubling symptom, PANSS = Positive and Negative Syndrome Scale, PK = pharmacokinetic, RBR = Research Biosample Repository, RLWM = Reinforcement learning working memory, SAE = serious adverse event, VRFCAT = Virtual Reality Functional Capacity Assessment Tool, WASI-II = Wechsler Abbreviated Scale of Intelligence – Second Edition, WRAT-4 = Wide Range Achievement Test

1 To decrease patient burden, some assessments may be split within the visit window days. Baseline assessments should be done pre-dose on Day 1 or within 24 hours prior to the first double-blind dose (i.e., on Day -1).
2 If treatment is discontinued early, the Week 12 visit assessments as well as the follow-up visit assessments (four weeks after last dose) should be performed.
3 Phone calls listed are for scheduled visits. Informal calls by a case manager up to three times per week to follow up with patients, especially during the washout period, are encouraged.
4 Eligibility will be confirmed at baseline (see Section [redacted]).
5 Complete physical exam includes height and weight at screening. Limited physical exams (symptom driven) and weight will be done at subsequent visits including the follow-up visit (see Section [redacted]).
6 Triplicate ECGs will be performed at screening and a single ECG at subsequent visits ECGs should be performed prior to meals or any scheduled vital sign measurements and blood draws.
7 Pregnancy testing will be using a blood sample at screening and urine sample at subsequent visits. If urine test is positive, pregnancy must be confirmed by blood test.
8 Washout may start earlier for antipsychotics with long half-life (2 week washout for aripiprazole, olanzapine, see [redacted]).
9 Sample should be taken prior to the morning dose (pre-dose/trough).
10 One additional sample should be taken two hours (± 0.5) post-dose.
11 Samples can be taken at any time between baseline and follow-up.
12 Phone usage monitoring will be used only with screened patients who have android phones and agree to the assessment.
13 After informed consent, but prior to initiation of study drug, only SAEs caused by a protocol-mandated intervention will be collected. After initiation of study drug, all adverse events, regardless of relationship to study drug, will be reported until 28 days after the last dose of study drug.
14 The placebo compliance period will start with a visit reviewing instructions with the patient.
15 Study drug is three capsules administered once daily, preferably in the morning. The last dose is given at the Day 84/EOT visit.
16 Sleep, mood, cognition and treatment expectation questions will be presented on the AiCure app at select time points that will appear as random to the participant.
17 The exploratory endpoints Brief Assessment of Cognition in Schizophrenia (BACS) (digital version), Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II), Wide Range Achievement Test (WRAT-4) and Virtual Reality Functional Capacity Assessment Tool (VRFCAT) are not available as validated Spanish translation. Therefore, sites will not be required to administer these tests to Spanish only speaking patients. If validated versions become available, they may be used.

Appendix 3 BNSS Items

Item 1: Intensity of pleasure during activities

Item 2: Frequency of pleasure during activities

Item 3: Intensity of expected pleasure from future activities

Item 4: Distress

Item 5: Asociality behavior

Item 6: Asociality inner experience

Item 7: Avolition behavior

Item 8: Avolition inner experience

Item 9: Facial expression

Item 10: Vocal expression

Item 11: Expressive gestures

Item 12: Quantity of speech

Item 13: Spontaneous elaboration

Appendix 4 PANSS Items and Scores

PANSS *Negative Symptoms, Disorganized Thoughts and Positive Symptoms Factors*

Factor	Original PANSS item number	PANSS item name
Negative symptoms	N1	Blunted affect
	N2	Emotional withdrawal
	N3	Poor rapport
	N4	Passive/apathetic social withdrawal
	N6	Lack of spontaneity and flow of conversation
	G7	Motor retardation
	G16	Active social avoidance
Disorganized thought/cognition	P2	Conceptual disorganization
	N5	Difficulty in abstract thinking
	G5	Mannerisms and posturing
	G10	Disorientation
	G11	Poor attention
	G13	Disturbance of volition
	G15	Preoccupation
Positive symptoms	P1	Delusions
	P3	Hallucinatory behavior
	P5	Grandiosity
	P6	Suspiciousness
	N7	Stereotyped thinking
	G1	Somatic concern
	G9	Unusual thought content
	G12	Lack of judgment and insight

PANSS factors based on “Marder” PANNS factor analysis published in J Clin Psych 58:12, December 1997 p538.

Appendix 5 CGI-S and CGI-I Questions

Clinical Global Impression

Clinical Global Impression – Severity of Illness Scale (CGI-S)

Circle the appropriate number for item below.

SEVERITY OF ILLNESS

Considering your total clinical experience with this particular population, how mentally ill is the patient at this time?

0 = Not assessed	4 = Moderately ill
1 = Normal, not at all ill	5 = Markedly ill
2 = Borderline mentally ill	6 = Severely ill
3 = Mildly ill	7 = Among the most extremely ill patients

Clinical Global Impression – Improvement Scale (CGI-I)

Circle the appropriate number for item below.

GLOBAL IMPROVEMENT

Rate total improvement whether or not, in your judgment, it is due entirely to drug treatment. Compared to his condition at baseline (prior to randomization), how much has patient changed?

0 = Not assessed	4 = No change
1 = Very much improved	5 = Minimally worse
2 = Much improved	6 = Much worse
3 = Minimally improved	7 = Very much worse

Adapted from Guy W. ECDEU Assessment manual for Psychopharmacology, US Department of Health, Education, and Welfare publication (ADM) 76-338, Rockville, MD; National Institute of Mental Health, 1976.

Appendix 6 DPAS Questions

1. It is difficult to be happy unless one is good-looking, intelligent, rich and creative.

2. People will probably think less of me if I make a mistake.

3. If I do not do well all the time, people will not respect me.

4. Taking even a small risk is foolish because the loss is likly to be a disaster

5. If a person asks for help, it is a sign of weakness.

6. If I do not do as well as other people, it means I am an inferior human being.

7. If I fail at my work, then I am a failure as a person.

8. If you cannot do something well, there is little point in doing it atall.

9. Making mistakes is fine because I can learn from them.

10. If I fail partly, it is as bad as being a complete failure.

11. People should have a reasonable likelihood of success before undertaking anything.

12. If I don't set the highest standards for myself, I am likey to end up a second-rate person.

13. If I am to be a worthwhile person, I must be truly outstanding in one major respect.

14. People who have good ideas are more worthy than those who do not.

15. If I ask a question, it makes me look inferior.

Appendix 7 Patients with Suicidal Ideation, Suicidal Behavior, and Self-Injurious Behavior without Suicidal Intent Based on the C-SSRS During Treatment

Events during treatment	Drug Name N=xx n (%)	Comparator Name N=xx n (%)
Suicidal Ideation (1-5)	x (%)	x (%)
1) Wish to be dead	x (%)	x (%)
2) Non-specific active suicidal thoughts	x (%)	x (%)
3) Active suicidal ideation with any methods (not plan) without intent to act	x (%)	x (%)
4) Active suicidal ideation with some intent to act, without specific plan	x (%)	x (%)
5) Active suicidal ideation with specific plan and intent	x (%)	x (%)
Suicidal Behavior (6-10)	x (%)	x (%)
6) Preparatory acts or behavior	x (%)	x (%)
7) Aborted attempt	x (%)	x (%)
8) Interrupted attempt	x (%)	x (%)
9) Non-fatal suicide attempt	x (%)	x (%)
10) Completed suicide	x (%)	x (%)
Suicidal Ideation or Behavior (1-10)	x (%)	x (%)
Self-injurious behavior without suicidal intent	x (%)	x (%)

Notes: N = number of enrolled patients with at least one post-baseline C-SSRS assessment. In this table, n and (%) refer to the number and percent of patients who experience the event at least once during treatment. For the composite endpoint of suicidal ideation (1-5), n and (%) refer to the number and percent of patients who experience any one of the five suicidal ideation events at least once during treatment. For the composite endpoint of suicidal behavior (6-10), n and (%) refer to the number and percent of patients who experience any one of the five suicidal behavior events at least once during treatment. For the composite endpoint of suicidal ideation or behavior (1-10), n and (%) refer to the number and percent of patients who experience any one of the ten suicidal ideation or behavior events at least once during treatment.

			Suicidal Ideation					Suicidal Behavior					
Patient	Trt	Visit	1	2	3	4	5	6	7	8	9	10	Self-Inj Beh wo SI
xxxx			Y	Y	Y	N	Y	N	N	N	N	N	N

1=Wish to be Dead, 2=Non-specific Active Suicidal Thoughts, 3=Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act, 4=Active Suicidal Ideation with Some Intent to Act, without Specific Plan, 5=Active Suicidal Ideation with Specific Plan and Intent, 6=Preparatory Acts or Behavior, 7=Aborted Attempt, 8= Interrupted Attempt, 9=Actual Attempt (non-fatal), 10=Completed Suicide.

Only patients with suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent are displayed. For patients with suicidal ideation, suicidal behavior, or self-injurious behavior without suicidal intent at any time, data from all visits are displayed. Self-Inj Beh wo SI = Self-injurious Behavior without Suicidal Intent.

Appendix 8 Fagerstöm Nicotine Dependence Test (FTND) Scoring

How Long After Waking For First	0 = After 60 minutes 1 = 31-60 minutes 2 = 6-30 minutes 3 = Within 5 minutes
Difficult to Refrain	0 = No 1 = Yes
Hardest Cigarette to Give Up	1 = The first in the morning 0 = Any other
No. of Cigarettes per Day	0 = 10 or less 1 = 11-20 2 = 21-30 3 = 31 or more
More Just After Awakening	0 = No 1 = Yes
Smokes When ill and Bedbound?	0 = No 1 = Yes

Appendix 9 Assessment of Sleep, Mood, Well-being and Cognitive Functioning, and Assumed Treatment Assignment

Sleep (treatment period)

"How did you sleep last night?"

Mood & Well-being (treatment period)

"How are you feeling today?"

"How is your energy level today?"

Cognitive Functioning (treatment period)

"How is your concentration and memory today?"

Treatment Expectation (baseline visit only)

"Do you expect that the study drug will help you?"

Treatment Expectation (treatment period)

"Do you think the drug is helping you?"

Treatment Expectation (Day 84 ±2)

"Do you think you were taking placebo or study drug?"

Placebo ☐

Study drug ☐