Official Title: A Phase II Multicenter, Randomized, Double-Blind, 12-Week Treatment, 3-Arm, Parallel-Group, Placebo-Controlled Study to Investigate the Efficacy, Safety and Tolerability of RO7017773 in Participants Aged 15 to 45 Years With Autism Spectrum Disorder (ASD)

NCT Number: NCT04299464

Document Date: Statistical Analysis Plan Version 2: 28-June-2024

STATISTICAL ANALYSIS PLAN	
STUDY TITLE:	**A PHASE II MULTICENTER, RANDOMIZED, DOUBLE-BLIND, 12-WEEK TREATMENT, 3-ARM, PARALLEL-GROUP, PLACEBO-CONTROLLED STUDY TO INVESTIGATE THE EFFICACY, SAFETY AND TOLERABILITY OF ALOGABAT IN PARTICIPANTS AGED 15 TO 45 YEARS WITH AUTISM SPECTRUM DISORDER (ASD)**

STUDY NUMBER:	BP41316
STUDY NAME:	**Aurora**
VERSION NUMBER:	v.2
ROCHE COMPOUND: **COMPOUND NAME:**	RO7017773 Alogabat
EUDRACT NUMBER:	2019-003524-20
IND NUMBER:	141893
PLAN PREPARED BY:	[redacted]

STATISTICAL ANALYSIS PLAN APPROVAL	
28-Jun-2024	DocuSigned by: [redacted] Signer Name: [redacted] Signing Reason: I approve this document Signing Time: 28-Jun-2024 | 1:55:28 PM CEST 10EC3CF5F664481FBAE0AD933DC0A009
SPONSOR: **LEGAL REGISTERED ADDRESS:**	F. Hoffmann-La Roche Ltd Grenzacherstrasse 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp above

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This SAP was developed based on Roche SAP model document V 2.0.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
2	|see electronic date stamp on title page|	US V5, 02 JUNE 2022

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

No key changes to the SAP occurred since version 1. Minor changes were made to detail further the derivation of endpoints, and the specification of subgroup analyses. Further edits were made to the list of outputs, and the entire document to improve clarity and consistency.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
AASP	Adolescent / Adult Sensory Profile
ADaM	Analysis Data Model
ADOS	Autism Diagnostic Observation Schedule
AE	Adverse Event
ANCOVA	Analysis of Covariance
ASD	Autism Spectrum Disorder
BAI	Beck Anxiety Inventory
BL	Baseline
CDISC	Clinical Data Interchange Standard Consortium
CGI	Clinical Global Impression
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide-Severity Rating Scale
CY-BOCS-ASD	Children's Yale-Brown Obsessive Compulsive Scale Modified for ASD
ECG	Electrocardiogram
eCOA	electronic Clinical Outcome Assessments
eCRF	electronic Case Report Form
EEG	Electroencephalography, Electroencephalogram
ESS	Epworth Sleepiness Scale
ESS-CHAD	Epworth Sleepiness Scale for Children and Adolescents
GSV	Growth Scale Value
HF	High Functioning
IA	Interim Analysis
IMC	Internal Monitoring Committee
IRF	Independent Review Facility
IxRS	Interactive voice/web-based Response System
KSS	Karolinska Sleepiness Scale
LF	Low Functioning
MDD	Minimally Detectable Difference
MedDRA	Medical Dictionary for Regulatory Activities
PedsQL™	Pediatric Quality of Life Inventory TM

PD	Pharmacodynamic
PK	Pharmacokinetic
PRAS-ASD	Parent-rated Anxiety Scale for ASD
PROMIS	Patient-Reported Outcomes Measurements Information System
RBS-R	Repetitive Behavior Scale-Revised
RRB	Restricted and Repetitive Behavior
SAE	Serious Adverse Events
SAP	Statistical Analysis Plan
SD	Standard deviation
SoA	Schedule of Assessments
SOC	Scientific Oversight Committee
SSP	Short Sensory Profile
TFL	Table Figure Listing
WASI	Wechsler Abbreviated Scale of Intelligence

1. INTRODUCTION

The objectives of this Phase 2 proof-of-concept study (BP41316) is to investigate the efficacy, safety, and tolerability of Alogabat in adolescents and adults (15 to 45 years of age) with autism spectrum disorder (ASD), with a target on the core symptoms of ASD of social communication deficits, and restricted and repetitive behaviors (RRBs) and interests. In addition, a set of biomarkers linked to the mechanism of action of the study drug or to ASD-related alterations is investigated in a representative patient population to pilot the identification of responder subgroups. To date, very few studies have investigated the effect of pharmacological treatments on the core symptoms of ASD and consequently no pharmacological agents are approved that offer a treatment for the core symptoms in this population of high unmet clinical need.

A key challenge in the development of new therapies to treat the core symptoms in ASD is the change of symptoms across different stages of development in longitudinal studies in both social communication deficits as well as in RRBs (Simonoff et al. 2020; Richler et al. 2010; Pender et al. 2020). The presentation of RRBs is known to be associated with age and the level of intellectual functioning (Lam and Aman 2007, Esbensen et al 2009, Richler et al 2010, Kim and Lord 2010, Uljarevic et al 2022). Thus, interventions may have different effects, depending on age and intelligence quotient (IQ), on individuals with ASD and treatment effects in older adolescents and adults may not be extrapolated to effects in the pediatric population, given the potential difference in manifestations, both quantitatively and qualitatively, of social communication deficits and RRBs with age.

Adolescents and adults aged 15 to 45 years with ASD are enrolled in this study to allow for the generation of data on the safety and tolerability, and the potential efficacy in these age groups.

Analyses of the primary and secondary endpoints as described in the protocol will be performed as planned.

1.1 OBJECTIVES AND ENDPOINTS

Table 1 Primary and Secondary Objectives

Primary Objective	Endpoint
• To evaluate the efficacy of 12-week treatment with alogabat compared with placebo in treating social communication deficits in participants with Autism Spectrum Disorder (ASD)	• Change from baseline to Week 12 in the Adaptive Behavior Composite score of the Vineland™ Adaptive Behavior Scales, Third Edition (Vineland™-3)
Secondary Objectives	**Endpoints**
• To evaluate the safety and tolerability of a 12-week treatment with alogabat in 15- to 45-year-old participants with ASD	• Incidence and severity of adverse events (AEs) and serious adverse events (SAEs)

	• Incidence of treatment discontinuations due to AEs • Change from baseline over time and incidence of clinically relevant abnormalities in vital signs including orthostatic changes, electrocardiogram (ECG) parameters and safety laboratory values including the incidence of marked laboratory abnormalities • Change from baseline over time in suicide risk (using the Columbia-Suicide-Severity Rating Scale [C-SSRS]) • Incidence of daytime sleepiness assessed with the Karolinska Sleepiness Scale (KSS), the Epworth Sleepiness Scale (ESS) or the ESS for children and adolescents (ESS-CHAD), and the Sudden Onset of Sleep Questionnaire
• To evaluate the efficacy of a 12-week treatment with alogabat compared with placebo on restricted and repetitive behaviors (RRBs)	• Change from baseline to Week 12 in behavior/symptoms as measured by all domains of the Repetitive Behavior Scale-Revised (RBS-R) scale
• To evaluate the efficacy of a 12-week treatment with alogabat compared with placebo on social behaviors	• Change from baseline to Week 12 on the Vineland-3 Socialization domain
• To evaluate the efficacy of a 12-week treatment with alogabat compared with placebo on communication skills	• Change from baseline to Week 12 on the Vineland-3 Communication domain

1.2 STUDY DESIGN

Study BP41316 is a multicenter, randomized, double-blind, parallel-group, 3-arm, placebo-controlled, 12-week treatment Phase II study to investigate the efficacy, safety tolerability, and pharmacokinetics of Alogabat in participants aged 15 to 45 years with a diagnosis of ASD and a Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II) score of ≥ 50. Approximately 105 participants with ASD (in order to have at least 84 evaluable participants at study end) aged 15 to 45 years are enrolled into the study. Participants are randomized [redacted] in a 1:1:1 ratio (approximately 35 participants randomized to each treatment arm).

In addition, randomization was staggered with the enrollment of high-functioning adults (defined as WASI-II ≥ 70; [HFA]) first, followed by adolescents and low-functioning adults defined as WASI-II < 70; [LFA]). After approximately 18 high-functioning adults (aged 18 to 45 years) had completed 6 weeks of treatment [redacted] the IMC, together with the SOC conducted a PK, safety, and tolerability review (1st Interim Analysis). Adequate PK, safety, and tolerability in high-functioning adults was established and PK-model predictions from healthy adult volunteers were confirmed to be comparable with the observed PK results in high-functioning adults. Low-functioning adults and adolescents aged 15 to 17 years (both high- and low-functioning) were included in the study. A 2nd Interim Analysis was performed by the IMC + SOC to evaluate PK, safety, and tolerability data from all available participants once 12 adolescents had completed 2 weeks of treatment. Hereby, adequate PK, safety, and tolerability was confirmed in adults and established in adolescents.

Each participant was evaluated using several scales and tests at each visit at the clinic. A total of 9 clinic visits (screening period included) and 1 phone contact, to assess for any clinically significant symptoms and/or new medication, were planned. In addition, digital biomarker data was collected. Each participant needed a reliable caregiver, parent, or support person who oversaw the participant's adherence with protocol-specified procedures and provided feedback on all informant-based assessments throughout the study. The same support person attended all on-site visits. For support persons of high-functioning adult participants, no on-site visits were necessary after the screening visit if the support persons opt to complete all electronic clinical outcome assessments (eCOA), interview, and forms remotely.

The primary analysis will be performed at database lock after Last Participant Last Visit.

1.2.1 Treatment Assignment and Blinding

Randomization took place on Day 1. Randomization to the different treatment arms (placebo, [redacted] alogabat or [redacted] alogabat [redacted] with a ratio of 1:1:1) was stratified by age [adolescents, adults] and IQ [LF, HF], and capping rules were applied as follows:

Table 2 Capping Rules

Clinical Factor	Capping Rule
IQ (as assessed by WASI-II) ≥ 50 to 69 vs ≥ 70	The proportion of low-functioning participants (IQ ≥ 50 to 69) is limited to 25% overall.
Sex female vs male	The proportion of female participants is limited to a maximum of 25%.

All participants were centrally assigned to randomized study treatment using an IxRS. Before the study was initiated, the telephone number and call-in directions for the IxRS and/or the login information and directions for the IxRS was provided to each site.

The randomization list was computer generated centrally by a third party and made available only to those individuals responsible for PK sample bioanalysis and to statisticians or programmers at the Sponsor directly involved in the production of outputs to support IMC and SOC meetings. The data was analyzed in a secure area which is not accessible to any blinded personnel.

Blinding

The IxRS was programmed with blind-breaking instructions: The blind may be broken if, in the opinion of the Investigator, it is in the participant's best interest to know the study treatment assignment. The Sponsor must be notified before the blind is broken unless identification of the study treatment is required for a medical emergency in which the knowledge of the specific blinded study treatment will affect the immediate management of the participant's condition (e.g., antidote available). In this case, the Sponsor must be notified within 24 hours after breaking the blind. The date and reason that the blind was broken must be recorded in the source documentation and Clinical Study Report (CSR), as applicable.

Unblinding should not result in the withdrawal of the participants from the study. Every effort should be made to retain unblinded participants. As per Health Authority reporting requirements, the Sponsor will break the treatment code for all unexpected SAEs that are considered by the Investigator to be related to study treatment. Whenever disclosure of the identity of the test medication is necessary, adequate procedures will be in place to ensure integrity of the data.

1.2.2 Independent Review Facility

No Independent Review Facility (IRF) was planned for this study.

1.2.3 Data Monitoring

An IMC consisting of a selected subset of Roche representatives as identified in the "IMC + SOC Agreement" was responsible to perform interim analyses of accumulated safety, tolerability, PK, and efficacy (as required) data. The IMC members, the clinical pharmacologist, and the pharmacometrician were unblinded to individual study treatment allocation.

An SOC consisting of one external expert supported the IMC in making their recommendations. The SOC member was likewise unblinded to individual study treatment allocation. Access to treatment assignment information followed the Sponsor's standard procedures. Based on accumulated study data available at their review, the IMC and SOC could have recommended for example adjustment(s) of the dose levels, dosing regimen, or the timing of dosing. Further details can be found in the "IMC + SOC Agreement".

2. STATISTICAL HYPOTHESES AND SAMPLE SIZE DETERMINATION

2.1 STATISTICAL HYPOTHESES

No formal hypothesis testing will be done. However, to quantify gating criteria it was useful to specify a standardized minimal effect size (MDD) in a way that if the observed drug effect would be above the MDD, a hypothesis test based on p-values of the

ANCOVA models as described in section 3.3 would come out significant, assuming at least 28 subjects are evaluable per arm. Gating criteria will be based on p-values of the ANCOVA models.

2.2 SAMPLE SIZE DETERMINATION

Approximately 105 participants with ASD randomized in a 1:1:1 ratio between two active dosing regimens and placebo need to be recruited in order to obtain evaluable data from approximately 84 participants after 12 weeks of treatment. Evaluable data from 84 ASD participants provides at least 80% power to detect a mean difference between baseline and 12-week follow-up of 4.5 points on a Vineland™ -3 composite between any active treatment arm and the placebo arm with $\alpha = 0.10$ error control (1-sided). The minimally detectable difference in this cohort is 2.7 points. It is assumed that the standard deviation of the differences between baseline and 12 weeks follow up is 8.0 points, that 20% of participants will not provide evaluable data, and that the participants recruited will consist of at most 25% low-functioning participants, at most 25% female participants. Because the purpose of this study is exploratory, no adjustment for comparison of multiple active doses to a single placebo arm is performed.

3. ANALYSIS SETS

The analysis data sets for the purposes of analyses are presented in Table 3.

Efficacy analyses based on the Efficacy Analysis Population will be restricted in a way that assessments at week 12 will be excluded from the efficacy analyses if the participant missed more than 7 consecutive study medications immediately preceding the day of the assessment (if performed), or if the overall IMP compliance, i.e. the ratio of administered study medication to planned study medication until the assessment, was < 60%.

Post treatment EEG assessments will be excluded from the PD analysis if the last dose of study medication was taken more than 6 hours before the EEG. Additionally, Week 2 and Week 12 EEG assessments will be excluded from the PD analysis if the study participant missed more than 3 doses in the week preceding the EEG.

Table 3 Analysis Sets

Population	Description
Safety Population	All participants randomized to study treatment and who received at least one dose of the study treatment, whether prematurely withdrawn from the study or not, will be included in the safety analysis.

Pharmacokinetic Population	All participants providing pharmacokinetic data will be included in the PK 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 which may influence the PK analysis. Excluded cases will be documented together with the reason for exclusion.
Efficacy Analysis Population	All participants who give informed consent, are randomized, and receive at least one dose of double-blind study medication.

4. STATISTICAL ANALYSES

4.1 GENERAL CONSIDERATIONS

Table 4 Analysis Time Windows

Assessments will be assigned to analysis study days according to visit windows defined below:

Analysis Window Label	**Baseline**	**Day 1**	**Day 7**	**Day 14**	**Day 42**	**Day 63 (c)**	**Day 84**
Planned Study Day	1 (a)	1	-	14	42	63	84
ECGs							
Study Day	≤ 1 (a)	1	-	2 to 28	-	-	≥ 71
Vital Signs, KSS							
Study Day	≤ 1 (a)	1	-	2 to 28	29 to 55	-	≥ 71
Lab (Hematology, Chemistry, Urinalysis)							
Study Day	≤ 1 (a)	-	-	2 to 28 (b)	29 to 55	56 to 70	≥ 71
ESS							

Study Day	≤ 1	-	-	2 to 28	29 to 70	-	≥ 71
ESS-CHAD							
Study Day	≤ 1	-	-	2 to 28	29 to 55	56 to 70	≥ 71
Coagulation							
Study Day	≤ 1 (a)	-	-	-	-	56 to 70	≥ 71
Vineland-3, RBS-R, CRRI, BAI, PRAS-ASD, SSP, AASP, PedsQL, PROMIS, Anthropometric Measurements							
Study Day	≤ 1 (a)	-	-	-	-	-	≥ 71
CGI-S, CGI-I (d)							
Study Day	≤ 1 (a)	-	-	2 to 28 (b)	29 to 70	-	≥ 71
Neurocognitive Battery							
Study Day	≤ 1 (a)	1	-	2 to 28 (b)	29 to 70	-	≥ 71
Sudden Onset of Sleep							
Study Day	≤ 1	-	2 to 10	11 to 28	29 to 55	56 to 70	≥ 71
(a) Latest assessment prior to intake of first study medication. Use date and time, if available. (b) May include assessments on day 1 if after the time of first study medication. (c) Planned for adolescent participants only. (d) No baseline for CGI-I.							

If more than one assessment is in the same time window, the assessment that is closest to the planned study day will be selected to enter the summary tables.

Table 5 General Descriptive Statistics

Continuous variables will be summarized with the following descriptive statistics and nomenclature:

n	number of observations or participants
mean	arithmetic mean
SD	standard deviation

min	minimum value
q1	first quartile value
median	median value
q3	third quartile value
max	maximum value

Categorical variables will be summarized and presented with the following nomenclature:

n	frequency
%	percentage

Percentage by category will be based on the number of participants exposed within a category. Categories of NOT REPORTED, MISSING, and/or UNKNOWN will be excluded from summaries if null or n/a.

Besides descriptive summaries, listings, and plots will be provided where deemed necessary. Listings will also include multiple assessments within the same time window.

Table 6 ADaM Datasets

The analysis datasets will be generated in accordance with Clinical Data Interchange Standard Consortium (CDISC) Analysis Data Model (ADaM) Version 2.1. ADaM compliant datasets will be created by the sponsor.

ADaM Dataset Name	Description	Double Programmed ?
ADSL	Subject Level Analysis Dataset	Yes
ADAE	Adverse Events Analysis Dataset	No
ADCM	Concomitant Medications Analysis Dataset	No
ADMH	Medical History Analysis Dataset	No
ADEX	Exposure Analysis Dataset	No

ADVS	Vital Signs Analysis Dataset	Yes
ADEG	ECG Analysis Dataset	Yes
ADLB	Laboratory Analysis Dataset	No
ADDV	Protocol Deviations Analysis Dataset	No
ADSUB	Subcategory Analysis Dataset	Yes
ADCSSRS	C-SSRS Analysis Dataset	No
ADKSS	KSS Analysis Dataset	No
ADVABS3	VABS-III Questionnaire Analysis Dataset	Yes
ADRBSR	RBS-R Questionnaire Analysis Dataset	Yes
ADESS	ESS Questionnaire Analysis Dataset	No
ADSOSQ	Sudden Onset of Sleep Questionnaire Analysis Dataset	No
ADNEURO	Neurocognitive Battery Analysis Dataset	No
ADCRRI	CRRI Questionnaire Analysis Dataset	Yes
ADCGII	CGI-I Analysis Dataset	Yes
ADCGIS	CGI-S Analysis Dataset	No

ADBAI	BAI Analysis Dataset	Yes
ADPRAS	PRAS-ASD Analysis Dataset	Yes
ADSSP	SSP Analysis Dataset	No
ADAASP	AASP Analysis Dataset	No
ADPROMIS	PROMIS Analysis Dataset	No
ADPEDSQL	PedsQL Analysis Dataset	No

4.1.1 Missing Data

Generally, no data imputations will be made.

4.1.2 Baseline Definitions

Unless otherwise stated, baseline is defined as the last observation recorded before the first study drug administration. The last observation can be an unscheduled / repeated measurement. If a pre-treatment observation is missing, then the pre-baseline (preferred) or screening value may be used.

As the age at screening was determining the choice of questionnaires, and for reasons of consistency, the age at screening was chosen to be the analysis age.

ECG baseline is defined as the mean of the triplicate measurement before the first study drug administration on Day 1.

4.1.3 Characteristics of Efficacy Endpoints

Vineland-3

The Vineland-3 (Sparrow et al. 2016) is a semi-structured interview that measures an individual's adaptive behaviour across three domains of communication, socialisation and daily living skills. Additional domains of motor skills (only used in children up to 9 years of age) and maladaptive behaviours were not captured in this study. Each adaptive behaviour domain is composed of three subdomains. Subdomain raw scores are used to derive two types of scores: Standard (V-scale scores; range from 1-24) and Growth Scale Values (GSVs). Standard scores are scores relative to a normative age group and can range from 20-140. A GSV is a person-ability score that is used to track an individual's progress. A higher score indicates better adaptive functioning.

Repetitive Behaviour Scale-Revised (RBS-R)

The RBS-R (Bodfish, 1999) is a 43-item informant-based questionnaire, assessing the variety of restricted and repetitive behaviours observed in individuals with ASD. A total RBS-R score is calculated as the sum of the scores for the 43 items. The total RBS-R score ranges from 0 to 129, with a higher score indicating more severe repetitive and restrictive behaviours. The RBS-R subscale scores will also be calculated as the sum of the scores for each item within each subscale.

Caregiver-Reported Routines Inventory (CRRI)

The CRRI is a 62-item caregiver-reported restricted and repetitive behaviours scale for older adolescents and adults based on the Children Routine Inventory-Revised (CRI-R; Evans et al 2017). It has been developed by the author in collaboration with the Sponsor in order to extend the age-range to older adults and adolescents. The CRRI questionnaire captures a wide range of restrictive and repetitive behaviours, including stereotypies, tics, compulsions, habits, sensory sensitivities, and focused interests, in the context of typical and atypical development in adults and adolescents. A total score is calculated as the sum of the scores for the 62 items.The scores range from 62 to 310, with a higher score indicating more severe repetitive and restricted behaviours. The CRRI subscale scores will also be calculated as the sum of the scores for each item within each subscale.

Clinical Global Impression-Severity (CGI-S) and Improvement (CGI-I) Scales

The CGI rating scales are tools used to evaluate both the severity of illness and change from Baseline (Guy et al 1976). The CGI-S reflects the rater's impression of the participant's current autism severity on a 7-point scale ranging from no symptoms (1) to very severe symptoms (7). The CGI-I is used to assess the clinical change as compared with symptoms at baseline using a 7-point scale, ranging from very much improved (1) to very much worse (7). For this study, modified versions will be used (Busner and Targum 2007).

Beck Anxiety Inventory (BAI)

The BAI is a well-validated 21-item participant-administered inventory probing for common symptoms of anxiety (Beck et al 1988). Each item contains four possible responses, which range in severity from 0=not at all (e.g., unable to relax) to 3=severely (e.g., it bothered me a lot). Participants are asked to provide answers based on the way they have been feeling over the past month, including the assessment day. A total score is calculated as the sum of the score for the 21 items. The scores range from 0 to 63, with a higher score indicating greater levels of anxiety.

Parent-Rated Anxiety Scale-ASD (PRAS-ASD)

The PRAS-ASD is a support person-rated scale with 25 items to assess the severity of anxiety symptoms in children and adolescents with ASD. Support persons of adolescent participants will be asked to describe their child's worries and anxiety-related behaviors over the past 2 weeks on a 0 to 3-point scale from "NONE=not present", "MILD=present sometimes, not a real problem", "MODERATE=often present and a problem", and "SEVERE=very frequent and a major problem". The PRAS-ASD total score is calculated

as the sum of the scores for the 25 items. The scores range from 0 to 75, with a higher score indicating greater levels of anxiety.

Short Sensory Profile 2 (SSP-2)

The Short Sensory Profile Version 2 (SSP-2; Dunn 2014) is a 34-item parent or support person-reported questionnaire that probes for the effect of sensory processing anomalies on a person's ability to function in daily life. Item responses occur on a five-point Likert-rating scale from 1 (almost never) to 5 (almost always). The SSP-2 was based on the Sensory Processing Framework (Dunn 1999) and provides scores in four quadrants based on the child's neurological threshold to sensory input and their method of self-regulation. Raw score totals are calculated for each of the four quadrants: Seeking ("the degree to which a child obtains sensory input"), Avoiding ("the degree to which a child is bothered by sensory input"), Sensitivity ("the degree to which a child detects sensory input."), and Registration ("the degree to which a child misses sensory input."). A Sensory raw score and Behavioral raw score will also be calculated. The score ranges from 0 to 100 (depending on the subscale), with a higher score indicating more frequent sensory behaviors and symptoms.

Adolescent/Adult Sensory Profile (AASP)

The Adolescent/Adult Sensory Profile (Brown and Dunn 2002) is a self-reported questionnaire consisting of 60 items rated for frequency of the behavior at home or in the community. Raw score totals are calculated for each of the four quadrants: "Low Registration", "Sensation Seeking", "Sensory Sensitivity", and "Sensation Avoiding". The raw scores for each quadrant range from 15 to 75, with higher scores indicating more frequent sensory behaviors and symptoms.

Patient-Reported Outcomes Measurements Information System (PROMIS)

The Patient-Reported Outcomes Measurements Information System (PROMIS) Sleep Disturbance Short Form (Yu Lan et al 2011, Buysse et al 2010) assesses self-reported perceptions of sleep quality, sleep depth, and restoration associated with sleep. This includes perceived difficulties and concerns with getting to sleep or staying asleep, as well as perceptions of the adequacy of and satisfaction with sleep. This questionnaire assesses sleep disturbance over the past seven days. The following versions of the Sleep Disturbance short form (8 items) will be implemented in this study: adult self-report (ages 18+, high-functioning), pediatric self-report (adolescents aged 15 to -17 years, high-functioning). In addition, proxy forms completed by the support person will be used for all adolescents, as well as low-functioning adults.

The sleep disturbance total raw score is calculated as the sum of the scores for the 8 items. The score will range from 8 to 40, with a higher score indicating greater sleep disturbance.

The PROMIS Sleep-Related Impairment Short-Form (Buysse et al 2010) assesses self-reported perceptions of waking alertness, sleepiness and function within the context of overall sleep-wake function. This questionnaire assesses sleep-related impairment over the past seven days. The following versions of the PROMIS Sleep-Related Impairment Short Form (8 items) will be implemented in this study: adult self-report (ages 18+, high-functioning), pediatric self-report (ages 15 to 17, high-functioning) and for support

persons serving as proxy reporters for the participants (adolescents aged 15 to17 years and low-functioning adults).

The sleep-related impairment total raw score is calculated as the sum of the scores for the 8 items. The score will range from 8 to 40, with a higher score indicating greater sleep impairment.

4.1.4 Characteristics of Safety Endpoints

Suicide Severity Rating Scale (C-SSRS)

A baseline assessment of suicidal ideation is performed and treatment-emergent suicidal ideation and behavior during the study is monitored using the Columbia Suicide Severity Rating Scale (C-SSRS) forms “At Baseline” and “Since Last Visit” at the timepoints indicated in the SoA. C-SSRS assessments are completed by a certified clinician.

Karolinska Sleepiness Scale (KSS)

The Karolinska Sleepiness Scale (KSS; Akersted and Gillberg 1990, Baulk et al 2001) will be completed as part of the safety assessments. The KSS measures the subjective level of sleepiness at a particular time during the day. On this scale, participants (or support persons for adolescents aged 15 to 17 years and for low-functioning participants) indicate which level best reflects the psycho-physical state experienced in the last 5 minutes. The KSS is a 9-point scale (1 = extremely alert, 9 = very sleepy, great effort to keep awake, fighting sleep). This assessment takes approximately 5 minutes to perform.

Epworth Sleepiness Scale (ESS)

The Epworth Sleepiness Scale (ESS; Johns 1991, Johns 1992) will be completed as part of the safety assessments. The ESS is a brief, self-administered eight-item questionnaire that measures daytime sleepiness in adults. Participant will be asked to rate on a scale of 0-3 the chances that, “over the past month” and “since last visit”, he/she would have dozed in eight specific situations that are commonly met in daily life (0 = would never doze; 3 = high chance of dozing). Thus, the participant is asked to characterize, retrospectively, part of his usual behavior in a variety of situations that are more or less soporific. The ESS score is the sum of eight item-scores and can range from 0 to 24. This assessment takes approximately 5 minutes to perform.

ESS for Children and Adolescents (ESS-CHAD)

The Epworth Sleepiness Scale for Children and Adolescents (ESS-CHAD; Johns 2015) will be completed as part of the safety assessments. The ESS-CHAD is a brief, support person-administered eight-item questionnaire that measures daytime sleepiness in children and adolescents. Support persons of adolescents and participants with an IQ score <70 will be asked to rate on a scale of 0-3 the chances that, "Over the past month," and “since last visit”, “your child” would have dozed in eight specific situations that are commonly met in daily life (0 = would never doze; 3 = high chance of dozing). Thus, the support person is asked to characterize, retrospectively, part of his usual behavior in a variety of situations that are more or less soporific. The ESS score is the

sum of eight item-scores and can range from 0 to 24. This assessment takes approximately 5 minutes to perform.

Pediatric Quality of Life Inventory (PedsQL) Generic Core Scale

The Pediatric Quality of Life Inventory (PedsQL™) Version 4.0 Generic Core Scale assessment consists of a 23-item questionnaire encompassing four core scale domains: Physical Functioning (8 items), Emotional Functioning (5 items), Social Functioning (5 items) and School Functioning (5 items). Each item utilizes a 5-point response scale ranging from "never a problem" (0) to "almost always a problem" (4).

Each item will be reversed scored and linearly transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0. Scores will be calculated as the sum of all the items over the number of items answered within each domain. In addition, a Psychosocial Health Summary score, Physical health summary score and total score will be calculated. The scores range from 0 to 100, with higher scores indicating better quality of life.

Pediatric Quality of Life Inventory (PedsQL) Cognitive Functioning Scale

The Pediatric Quality of Life Inventory (PedsQL™) Version 3.0 Cognitive Functioning Scale consists of 6-items. Each item utilizes a 5-point response scale ranging from "never a problem" (0) to "almost always a problem" (4).

Each item will be reversed scored and linearly transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0. PedsQL Cognitive Functioning Score will be calculated as the sum of all the items over the number of items answered. The score ranges from 0 to 100, with a higher score indicating better cognitive functioning.

Sudden Onset of Sleep Questionnaire

The Sudden Onset of Sleep Questionnaire has been developed by the Sponsor for this study and will be completed as part of the safety assessments. Each participant (or support person for adolescents aged 15 to 17 years and for low-functioning participants) will be asked to answer a series of up to 8 questions to assess the occurrence of daytime somnolence, its suddenness (level of awareness), frequency and circumstances in which the episode(s) took place, impact on work/daily activities/ social life, if worried by episode and ease of waking up (see Appendix 2 for reporting of AEs). This questionnaire will be administered as an interview. This assessment takes approximately 5 minutes to perform.

Neurocognitive Battery

The cognitive tests will be applied as part of the safety assessments. The total battery takes approximately 45 minutes to perform, the selection of attention tasks (detection, identification, and one back tests) takes approximately 10 minutes to perform. For low-functioning participants, the selection of attention tasks (detection, identification, and one back tests) should be prioritized and performed first. All effort should be made to also collect the remainder of the tasks if the study participant is able to comply. If not all tasks can be completed, an explanation should be recorded on the “Additional Observations” eCRF. Two test sessions of shorter duration will be carried out at the pre-baseline visit for habituation to the test environment.

Detection Test (Psychomotor Function)

The Detection Test (DET) is a measure of psychomotor function and uses a well-validated simple reaction time paradigm with playing card stimuli. In this test, the playing cards all depict the same joker. The participant is asked to press the “Yes” key as soon as the card in the center of the screen turns face up. The software measures the speed and accuracy of each response. This assessment takes approximately 3 minutes to perform.

Identification Test (Attention)

The Identification Test (IDN) is a measure of visual attention and uses a well-validated choice reaction time paradigm with playing card stimuli. In this test, the playing cards are all either red or black jokers. The participant is asked whether the card displayed in the center of the screen is red. The participant responds by pressing the “Yes” key when the joker card is red and No when it is black. The software measures the speed and accuracy of each response. This assessment takes approximately 3 minutes to perform.

One Card Learning Test (Visual Learning)

The One Card Learning Test (OCL) is a measure of visual learning and uses a well-validated pattern separation paradigm with playing card stimuli. In this test, the playing cards are identical to those found in a standard deck of 52 playing cards (without the joker cards). The participant is asked whether the card displayed in the center of the screen was seen previously in this test. The participant responds by pressing the “Yes” or “No” key. The software measures the speed and accuracy of each response. This assessment takes approximately 6 minutes to perform.

One Back Test (Working Memory)

The One Back Test (ONB) is a measure of working memory and uses a well-validated n-back paradigm with playing card stimuli. In this test, the playing cards are identical to those found in a standard deck of 52 playing cards (without the joker cards). The participant is asked whether the card displayed in the center of the screen is the same as the card presented immediately previously. The participant responds by pressing the “Yes” or “No” key. Because no card has been presented yet on the first trial, a correct first response is always “No”. The software measures the speed and accuracy of each response. This assessment takes approximately 4 minutes to perform.

Two Back Test (Working Memory)

The Two Back Test (TWOB) is a measure of working memory and uses a well-validated n-back paradigm with playing card stimuli. In this test, the playing cards are identical to those found in a standard deck of 52 playing cards (without the joker cards). The participant is asked whether the card displayed in the center of the screen is the same as the one presented two cards previously. The participant responds by pressing the “Yes” or “No” key. Because no card has been presented two-back on the first and second trials, a correct response on these trials is always “No”. The software measures the speed and accuracy of each response. This assessment takes approximately 4 minutes to perform.

The International Shopping List Test (Verbal Learning)

The International Shopping Test (ISL) is a measure of verbal learning and uses a well-validated list-learning paradigm. The test is administered using a computer. High frequencies, high imagery, concrete nouns (items from a shopping list) are read to the participant by the test supervisor at the rate of one word every two seconds. Once all 12 words have been read, the participant is asked to recall as many of the words as he/she can as quickly as possible. The test supervisor uses a mouse or stylus to mark the words recalled by the participant on the computer screen. When the subject can recall no more words, the same list is read a second time. The test supervisor records the words recalled by the participant on this trial. This is then repeated a third time. 30 minutes later: without having the list read again, the participant is asked to recall the words from the list (delayed recall condition). The recall test is followed by a recognition condition test where the test supervisor reads a shopping list item that may or may not have been on the original list and the participant has to respond either affirmatively (if the item was on the original list) or negatively (if it was not). The recall and recognition test is repeated the next day about 24 hours after the list learning. If the clinic visit is split into two days (which is possible for Week 6, Week 12, and 6-week follow-up visits), the repeats are performed at the clinic. If the participant returned home after the first day the repeats are performed over the phone as close in time to the 24h timepoint as possible. The software registers the number of correct responses as recorded by the test supervisor. This assessment takes approximately 15 minutes to perform (5 minutes for each of the three parts).

Continuous Paired Associate Learning Test (Paired Associate Learning)

The Continuous Paired Associate Learning Test (CPAL) is a measure of visual associate memory and uses a well-validated paired associate learning paradigm in which the subject must learn the locations of a number of amoeba-like shapes on the computer screen. This test consists of a single amoeboid shape displayed in the center of the screen surrounded by a number of blue- filled circles. Beneath all but two of the blue spheres are amoeboid shapes, one of which matches the central display; the two remaining circles are distractors. In the exposure phase of the test all of the to-be-remembered pattern-location associations are presented on the computer screen simultaneously. After a five second delay, a pattern is shown in the central location and this signals that the subject should touch the location in the periphery that contains the same pattern. This process continues until the participant has acknowledged all of the pattern-location associations. The learning phase begins with the same test display presented during the exposure phase except that now all of the peripheral locations are shown as blue spheres. One of the patterns presented in the exposure phase is presented in the center location. With the presentation of this pattern, the participant is required to select the peripheral location where an identical pattern is hidden beneath the blue sphere. This process continues until the correct location of each pattern is found. Finding the correct location for all patterns in the set is defined as a learning trial. There are six learning trials (a single trial delayed recall condition will be done after a 30-minute delay). The software records each move as an error or as a correct move. This assessment takes approximately 7 minutes to perform.

Groton Maze Learning Test (Executive Function)

The Groton Maze Learning Test (GMLT) measures executive function using a maze learning paradigm. A 10 x 10 grid of tiles is presented to the participant on the screen. A 28-step pathway is hidden among these tiles. A blue tile indicates the start and a tile with red circles indicates the finish. The participant must move one step at a time from the start toward the end by touching a tile next to their current location. If the correct move is made a green checkmark appears and if the move is incorrect a red cross is revealed. Once completed, they are returned to the start location to repeat the test and must try to remember the pathway they have just completed. The software records the total number of errors made in attempting to learn the same hidden pathway on 5 consecutive trials during a single session. This assessment takes approximately 7 minutes to perform.

Table 7 Efficacy Endpoints Details and Description

Efficacy Endpoint	Description
Vineland-3 ABC Composite GSV	A raw score will be calculated for each subdomain of the Socialization (interpersonal relationships, play and leisure time, coping skills), Communication (receptive, expressive, written) and Daily Living Skills (personal, domestic, community) domain as the sum of the scores for each item within the subdomain. The raw scores for each of the 9 subdomains will be mapped to GSV scores. A conversion table for mapping raw scores to GSV scores is found in Appendix 3. This was based on Table B.2 in the Vineland-3 manual. The Vineland-3 ABC Composite GSV score will be calculated as the mean GSV score (i.e. summing the 9 GSV subdomain scores and dividing by 9).
Vineland-3 Socialization GSV	A raw score will be calculated for each subdomain of the Socialization domain (interpersonal relationships, play and leisure time, coping skills) as the sum of the scores for each item within the subdomain. The raw scores for each of the 3 subdomains will be mapped to GSV scores. The Vineland-3 Socialization GSV score will be calculated as the mean GSV score (i.e. summing the 3 GSV subdomain scores and dividing by 3).

Vineland-3 Communication GSV	A raw score will be calculated for each subdomain of the Communication domain (receptive, expressive, written) as the sum of the scores for each item within the subdomain. The raw scores for each of the 3 subdomains will be mapped to GSV scores using the conversion table in Appendix 3. The Vineland-3 Communication GSV score will be calculated as the mean GSV score (i.e. summing the 3 GSV subdomain scores and dividing by 3).
Vineland-3 Daily Living Skills GSV	A raw score will be calculated for each subdomain of the Daily Living Skills domain (personal, domestic, community) as the sum of the scores for each item within the subdomain. The raw scores for each of the 3 subdomains will be mapped to GSV scores. The Vineland-3 Daily Living Skills GSV score will be calculated as the mean GSV score (i.e. summing the 3 GSV subdomain scores and dividing by 3).
Repetitive-Behaviors Scale-Revised (RBS-R) total score	The RBS-R total score will be calculated as the sum of the scores for each of the 43 items.
Repetitive-Behaviors Scale-Revised (RBS-R) subscale score	The RBS-R subscale score will be calculated as the sum of the raw scores for each item within each of the following subscales, a 6-factor conceptual model; Bodfish et al. 1999: - Stereotyped (items 1-6) - Self-injurious (items 7-14) - Compulsive (items 15-22) - Ritualistic (items 23-28) - Sameness (items 29-39) - Restricted behavior (items 40-43) And a 3-factor model (Mirenda et al. 2010) confirmed by internal analyses: ● Repetitive Sensory Motor (items 1-6, 40-43) ● Self-injurious (items 7-14)

	• Insistence on Sameness (items 15-39)
Caregiver-Reported Routines Inventory (CRRI) total score	The CRRI total score will be calculated as the sum of the scores for each of the 62 items.
Caregiver-Reported Routines Inventory (CRRI) subscale score	The CRRI subscale score will be calculated as the sum of the raw scores for each item within each of the following subscales. The subscale structure was derived from the oRBiting study (Crawley et al. manuscript in preparation): - 'Just right' behaviors (items 5,9,15,16,18,19,23,29,30,37,40,61) - Routines and Sameness behaviors (items 1,2,4,6,7,8,10,20,42,58,59) - Sensory & Motor Behaviors (items 11,13,14,17,24,25,27,28,31,32,33, 34,35,43,44,45,46,48,50,51) - Restricted Interests / Stereotypies (items 26,36,38,39,47,53,54,55,56,57,62) - Food Rigidities (items 3,12,21,22,41,49,52,60) Of note, items 52 and 60 will be reverse coded.
Clinical Global Impressions Scale - Severity (CGI-S) Domain score	An impairment rating score with 7 categories (no symptoms to very severe symptoms) for each of the following domains: - communication/speech - activities of daily living - social functioning - disruptive behavior - restricted and repetitive behavior Note: The customizable "other" symptom domain will not be summarized.
Clinical Global Impressions Scale - Severity (CGI-S) Overall score	A rating score with 7 categories (no symptoms to very severe symptoms) for the overall rating.

Clinical Global Impressions Scale - Improvement (CGI-I) Domain score	A rating score with 7 categories (very much improved to very much worse) for each of the following domains: - communication/speech - activities of daily living - social functioning - disruptive behavior - restricted and repetitive behavior Note: The customizable "other" symptom domain will not be summarized.
Clinical Global Impressions Scale - Improvement (CGI-I) Overall score	A rating score with 7 categories (very much improved to very much worse) for the overall rating.
Vineland-3 ABC standard score	A raw score will be calculated for each subdomain of the Socialization (interpersonal relationships, play and leisure time, coping skills), Communication (receptive, expressive, written) and Daily Living Skills (personal, domestic, community) domain as the sum of the scores for each item within the subdomain. The raw scores for each of the 9 subdomains will be mapped to V-scale scores relative to a normative age group as described in the Vineland-3 scoring manual (Sparrow et al. 2016). A standard score will be calculated for each of the three domains (Communication, Socialization and Daily Living Skills) as the sum of the V-scale scores within each domain. The Vineland-3 ABC standard score will be calculated as the sum of the standard scores for the three domains.
Vineland-3 Socialization standard score	A raw score will be calculated for each subdomain of the Socialization domain (interpersonal relationships, play and leisure time, coping skills) as the sum of the scores for each item within the subdomain. The raw scores for each of the 3 subdomains will be mapped to standard (V-scale) scores relative

	to a normative age group as described in the Vineland-3 scoring manual (Sparrow et al. 2016) The Vineland-3 Socialization standard score will be calculated as the sum of the V-scale scores within each subdomain.
Vineland-3 Communication standard score	A raw score will be calculated for each subdomain of the Communication domain (receptive, expressive, written) as the sum of the scores for each item within the subdomain. The raw scores for each of the 3 subdomains will be mapped to standard (V-scale) scores relative to a normative age group as described in the Vineland-3 scoring manual (Sparrow et al. 2016) The Vineland-3 Communication standard score will be calculated as the sum of the V-scale scores within each subdomain.
Vineland-3 Daily Living Skills standard score	A raw score will be calculated for each subdomain of the Daily Living Skills domain (personal, domestic, community) as the sum of the scores for each item within the subdomain. The raw scores for each of the 3 subdomains will be mapped to standard (V-scale) scores relative to a normative age group as described in the Vineland-3 scoring manual (Sparrow et al. 2016). The Vineland-3 Daily Living Skills standard score will be calculated as the sum of the V-scale scores within each subdomain.
Beck Anxiety Inventory (BAI)	The BAI total score will be calculated as the sum of the scores for each of the 21 items.
Parent Rated Anxiety Scale- ASD (PRAS-ASD)	The PRAS-ASD total score will be calculated as the sum of the scores for each of the 25 items.
Adolescent/Adult Sensory Profile (AASP)	Quadrant raw scores will be calculated as the sum of the scores for items within each of the following quadrants:

	- low registration (items 3, 6, 12, 15, 21, 23, 36, 37, 39, 41, 44, 45, 52, 55 and 59) - seeking (items 2, 4, 8, 10, 14, 17, 19, 28, 30, 32, 40, 47, 50 and 58) - sensitivity (items 7, 9, 13, 16, 20, 22, 25, 27, 31, 33, 34, 48, 51, 54 and 60) - avoiding (items 1, 5, 11, 18, 24, 26, 29, 35, 38, 43, 45, 49, 53, 56 and 57) Reference: Adolescent/Adult Sensory Profile
Short Sensory Profile (SSP-2)	The SSP-2 sensory raw score will be calculated as the sum of the scores for items 1-14. The SSP-2 behavioral raw score will be calculated as the sum of the scores for items 15-34. Quadrant raw scores will be calculated as the sum of the scores for items within each of the following quadrants: - seeking (items 6-8, 11, 14, 31 and 32) - avoiding (items 16-20 and 22-26) - sensitivity (items 1-5, 21, 25, 28, 29 and 33) - registration (items 9, 10, 12, 13, 15, 27, 30 and 34) Reference: Short-Sensory Profile-2 Caregiver Questionnaire
Pediatric Quality of Life Inventory (PedsQL) Generic Core Scale	The first step is to transform the scores. Each item will be reversed scored and linearly transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0. The second step is to calculate scores for each of the four dimensions: - physical functioning - emotional functioning - social functioning - school functioning

	If more than 50% of the items in the scale are missing, the scale scores should not be computed. Mean scores for each dimension will be calculated as the sum of the items over the number of items answered. Psychosocial Health Summary Score will be calculated as the sum of the items over the number of items answered in the Emotional, Social, and School Functioning Scales. Physical Health Summary Score is the Physical Functioning Scale Score PedsQL Total Score will be calculated as the sum of all the items over the number of items answered on all the scales. Reference: PedsQL scoring manual
Pediatric Quality of Life Inventory (PedsQL) Cognitive Functioning Scale	The first step is to transform the scores. Each item will be reversed scored and linearly transformed to a 0-100 scale as follows: 0=100, 1=75, 2=50, 3=25, 4=0. PedsQL Cognitive Functioning Score will be calculated as the sum of all the items over the number of items answered. Reference: PedsQL scoring manual
Patient-Reported Outcomes Measurements Information System (PROMIS) - Sleep-Disturbance Short Form Scale	A PROMIS Sleep-Disturbance raw score will be calculated as the sum of the scores for each of the 8 items. The raw score will be mapped to a T-score using the PROMIS Sleep Disturbance scoring manual. Reference: PROMIS Sleep Disturbance Scoring Manual
Patient-Reported Outcomes Measurements Information System (PROMIS) - Sleep-	A PROMIS Sleep-Related Impairment raw score will be calculated as the sum of the scores for each of the 8 items.

Related Impairment Short Form Scale	The raw score will be mapped to a T-score using the PROMIS Sleep-Related Impairment scoring manual. Reference: PROMIS Sleep-Related Impairment Scoring Manual

4.2 PRIMARY, SECONDARY AND EXPLORATORY EFFICACY ANALYSIS

The primary efficacy analysis evaluates the mean change from baseline to 12 weeks of treatment as measured by the Vineland-3 ABC three-domain GSV composite. The difference in the mean change from baseline between each of the two active doses and placebo will be extracted as contrasts from the ANCOVA model.

The efficacy analyses of the secondary endpoints and of the CRRI total score evaluate the mean change from baseline to 12 weeks of treatment. The difference in the mean change from baseline of the two active doses and placebo will be extracted as contrasts from the ANCOVA model.

The efficacy analysis of CGI-I evaluates the overall score at 12 weeks of treatment. The difference of the two active doses and placebo will be extracted as contrasts from the ANCOVA model.

4.2.1 Definition of Primary Endpoint

The primary endpoint is Change from baseline to Week 12 in the Adaptive Behavior Composite score of the Vineland™ Adaptive Behavior Scales, Third Edition (Vineland™-3), as defined in Section 1.1 of SAP (see Table 1).

4.2.2 Main Analytical Approach for Primary, Secondary, CRRI and CGI-I Endpoints

The analyses of primary, secondary, CRRI and CGI-I efficacy endpoints will include all participants in the Efficacy Analysis Population.

An analysis of covariance (ANCOVA) model will be used to compare treatment effects for the primary, secondary, CRRI and CGI-I efficacy endpoints across the overall Efficacy Analysis Population of adolescents and adults. The model will include treatment as main effect, individual age at screening, IQ, and baseline score as covariates. Treatment differences will be estimated together with their 80% confidence intervals as derived from the model and their associated p-values.

The ANCOVA model will be applied to change from baseline data y_{ij} of participants i, and treatment j ∈ {placebo (reference), low dose, high dose} as follows:

$$y_{ij}=\mu+t_j+a_i+IQ_i+BL_i+\varepsilon_{ij}$$

where μ gives the overall endpoint mean change, followed by the fixed effect *t* which models the effect of treatment, age *a, IQ,* the endpoint baseline value *BL*, and ε, which gives the overall error term with $N(0,\sigma_\varepsilon^2)$.

Table 8 Efficacy Statistical Analysis Methods

Endpoint	Statistical Analysis Methods
Primary Secondary CRRI, CGI-I	**ANCOVA**

4.2.2.1 Subgroup Analyses for Primary and Secondary Efficacy Endpoints

The generalizability of Vineland-3 ABC three-domain GSV composite and RBS-R total score results when comparing any active treatment arm to placebo will be investigated by estimating the treatment effect in subgroups based on geographic region (North America vs. Europe), rater change (yes vs. no), site type (academic vs. ASD research sites), age categories (adolescents vs. adults), and IQ categories (high functioning vs. low functioning).

The dependence of BAI total score and RBS-R total score results on a history of clinical diagnosis of anxiety at baseline (yes vs. no) will be investigated.

For RRB endpoints RBS-R total score, CRRI total score and CGII (RRB domain) subgroups of male vs female will be investigated.

Summaries of these subgroup analyses will be provided in forest plots.

Selected categories used in this section, and their characteristic values are provided in the following table:

Table 9 Selected Categories and Characteristic Values

Category	Characteristic Value
Anxiety at Baseline	Preferred Terms: Anxiety, Anxiety Disorder, Generalized Anxiety Disorder, Social Anxiety Disorder
ASD Research Site	Southwest Autism Research And Resource Center, UNITED STATES Nathan Kline Institute, UNITED STATES APG Center for Autism & ADHD,

	UNITED STATES Okanagan Clinical Trials, CANADA

4.3 SUPPORTIVE EXPLORATORY EFFICACY ENDPOINTS ANALYSES

Additional exploratory analyses will be performed in a descriptive manner for continuous endpoints Vineland-3, RBS-R, CRRI, CGI, AASP, PRAS-ASD, BAI, SSP-2, PedsQL™, and PROMIS per treatment arm comprising baseline values, (up to) week 12 values and change from baseline to (up to) week 12 values.

For endpoint CGI, analyses will be performed to get the proportion of subjects in each of the rating categories per treatment arm and per time point.

For the assessment of differences between each Alogabat group and placebo regarding the CGI-I overall score at week 12, a mixed-effects MMRM analysis incorporating data up to 12 weeks of treatment will be used to utilize all the data collected over time with consideration of the variance-covariance matrix of the repeated measures. The model will include the CGI-I overall score as the dependent variable. The effects in the model will include independent variables of the fixed, categorical effects of treatment, assessment weeks relative to the first dose of study medication (i.e., time), and treatment-by-time interaction, along with covariates of the baseline value, age at screening and IQ at baseline. Time will be treated as the repeated variable within a subject. An unstructured variance-covariance structure will be applied to model the within-subject errors.

For the primary, secondary, CRRI and CGI-I endpoints, the ANCOVA will be repeated to model change from baseline data y_{ij} of participants i, and treatment j ∈ {placebo (reference), low dose, high dose} incorporating interactions to the model as follows:

$$y_{ij}=\mu+t_j+a_i+IQ_i+t_j*IQ_i+BL_i+\varepsilon_{ij}$$

and

$$y_{ij}=\mu+t_j+a_i+IQ_i+t_j*a_i+BL_i+\varepsilon_{ij}$$

where μ gives the overall endpoint mean change, followed by the fixed effect *t* which models the effect of treatment, age *a, IQ*, the interactions *t*IQ* and *t*a* respectively, the endpoint baseline value *BL*, and ε, which gives the overall error term with $N(0,\sigma_\varepsilon^2)$.

4.4 SAFETY ANALYSES

All safety analyses will be based on the safety analysis population.

Table 10 Safety Statistical Analysis Methods

Endpoint	Statistical Analysis Methods
Adverse events	The original terms recorded on the eCRF by the Investigator for AEs will be coded by the Sponsor. Adverse events will be summarized by mapped term and appropriate thesaurus level.
Clinical laboratory tests	All clinical laboratory data will be stored on the database in the units in which they were reported. Laboratory test values will be presented by tabular summaries, plots, and individual listings with flagging of abnormal results. Hereby, by default the International System of Units (SI units; Système International d'Unités) will be applied, and values will be transformed to common Roche standard reference range (following Roche Safety Lab Standardisation Guidance) where applicable.
Vital signs	Vital signs data will be presented by individual listings with flagging of values outside the normal ranges and flagging of abnormalities. In addition, tabular summaries of raw values and change from baseline will be used, as appropriate.
ECG	ECG data will be presented by individual listings. In addition, tabular summaries of raw values and change from baseline will be used, as appropriate.
Concomitant medications	The original terms recorded on the participants' eCRF by the Investigator for concomitant medications will be standardized by the Sponsor by utilizing a mapped term and appropriate drug dictionary level. Concomitant medications will be presented in summary tables and listings.
C-SSRS, Neurocognitive Battery, KSS, ESS or ESS-CHAD, and Sudden Onset of Sleep Questionnaire	Data will be presented in summary tables and listings

4.4.1 Extent of Exposure

A summary will be produced showing the treatment duration in days, the total cumulative number of doses and the total cumulative dose in mg per treatment arm. A histogram binned by hours will be produced showing the distribution of dosing times per treatment arm.

4.5 OTHER ANALYSES

4.5.1 Summaries of Demographics and Baseline Characteristics

Demographics and baseline characteristics data such as baseline values of relevant scales, medical history including medications and treatments, support person status, educational status of participant and parents, etc. will be summarized by treatment, age cohort, and IQ cohort.

Additional listings, summary tables of descriptive statistics, frequency tables, and corresponding graphs will be provided as appropriate.

4.5.2 Pharmacokinetic Analyses

For all participants, plasma concentrations of Alogabat will be presented descriptively by dose and age group.

Nonlinear mixed effects modeling will be used to analyze the sparse sampling dose versus concentration versus time data collected for Alogabat. The PK data collected in this study will be pooled with data from other clinical studies conducted with Alogabat. Population and individual PK parameters such as clearance and volume of distribution will be estimated, and the influence of various covariates such as age, gender, and body weight, on these parameters will be investigated, as appropriate. Secondary PK parameters will be derived from the model as appropriate. The results of the PopPK analysis will be reported separately from the CSR.

Subjects excluded from the PK Analysis Population will be documented in a listing.

Depending on results it will be decided post hoc after database lock whether analyses will be performed that investigate the participant's exposure to Alogabat, and whether correlations of exposure to Alogabat exist with selected efficacy, PD, biomarkers and safety endpoints.

4.5.3 Biomarker Analyses

Resting state EEG is recorded up to five times per patient: Pre-baseline, day 1 pre-dose, day 1 post-dose, week 2 post-dose (optional recording) and week 12 post-dose. EEG beta power is defined as the total power in the 13-30 Hz range averaged over the fronto-central channels Fz and Cz.

To estimate pharmacodynamic effects, beta power for post-dose recordings are normalized with baseline beta power to estimate relative change. Baseline beta power is defined as the average of the pre-baseline and day 1 pre-dose values. For each post-dose recording (day 1, week 2, week 12) and dose level, the mean change and SD of change will be reported. Additionally, two-tailed two sample t-tests comparing placebo to

each treatment arm will be performed for each post-dose recording. Hereby the null-hypothesis of no difference in beta power with Alogabat treatment compared to placebo will be rejected if the p-value resulting from a two-sided t-test is < 0.1 .

It will also be examined whether baseline beta power predicts clinical response. The hypothesis is that clinical efficacy increases as beta power decreases (patients with lower baseline beta power benefit more from alogabat than patients with higher baseline beta power). To this end beta power is normalized with power outside the beta range. The Baseline value is defined as the average of the pre-baseline and day 1 pre-dose values. The Pearson correlation coefficient between baseline beta power and change from baseline clinical efficacy scores will be computed for each active treatment arm and placebo. For active treatment arms the p-value will be derived from a two-sided test with null-hypothesis: r = 0. The null hypothesis will be rejected if the two-sided p-value is < 0.2. The test will be performed for the clinical efficacy endpoints Vineland-3 ABC GSV composite, Vineland-3 Socialization GSV composite, RBS-R total, and CRRI total.

4.6 INTERIM ANALYSES

Two interim safety analyses are planned for this study. The first interim analysis of accumulated safety, tolerability and PK will be performed once approximately 18 high-functioning adults have completed 6 weeks of treatment. The aim of this analysis is to evaluate the safety, tolerability, and PK in high-functioning adults before allowing the recruitment of the adolescents and low-functioning adult participants.

A second interim analysis of PK and accumulated safety and tolerability will be conducted once approximately 12 adolescents have completed 2 weeks of treatment. The aim of this analysis is to enable internal decisions to inform the design and dose selection for the adolescent participants and additional studies in the program.

The CSR will document all interim analyses that occurred. All interim analyses will be performed and interpreted by members of the IMC, external experts (SOC), and appropriate senior management personnel. IMC and SOC members will be unblinded at the individual participant level, sponsor senior management will be unblinded at the treatment group level. Access to treatment assignment information will follow the Sponsor's standard procedures.

The "IMC Agreement" will describe the planned interim analyses in detail.

5. SUPPORTING DOCUMENTATION

The layout of TFLs will be according to Roche standards. Templates that deviate from these standards or which are tailored to meet special requirements are presented below in this section:

Change from Baseline to week 12	**Placebo**	**RO7017773** [redacted]	**RO7017773** [redacted]
n			
Adjusted Mean			
80% CI for Adjusted Mean			
Difference in Adjusted Means			
80% CI for Difference in Adjusted Means			
P-value			

6. REFERENCES

Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci. 1990;52:29-37.

Baulk SD, Reyner LA, Horne JA. Driver sleepiness–evaluation of reaction time measurement as a secondary task. Sleep. 2001;24:695–8.

Beck AT, Epstein N, Brown G et al. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol. 1988;56:893–97.

Bodfish, JW, Symons, FJ, Parker, DE et al. Varieties of repetitive behavior in autism: Comparisons to mental retardation. J Autism Dev Disord. 2000;30:237–43.

Brown C, Dunn W. Adolescent-adult sensory profile: User's manual. 2002. San Antonio: Therapy Skill Builders.

Busner J, Targum S. The clinical global impressions scale: applying a research tool in clinical practice. Psychiatry. 2007;4:28-37.

(as adapted from Kay, SR. Positive and Negative Syndromes in Schizophrenia:Assessment and Research. Clinical and Experimental Psychiatry, Monograph No. 5. Philadelphia, PA, US: Brunner/Mazel, 1991
and
Spearing, MK, Post RM, Leverich GS. Modification of the clinical global impressions (CGI) scale for use in bipolar illnesss (BP): the CGI-BP. Psychiatry

Res. 1997;73:159-71.)and
Spearing, MK, Post RM, Leverich GS. Modification of the clinical global impressions (CGI) scale for use in bipolar illnesss (BP): the CGI-BP. Psychiatry Res. 1997;73:159-71.)

Buysse D, Yu L, Moul DE et al. Development and Validation of Patient-Reported Outcome Measures for Sleep Disturbance and Sleep-Related Impairments. Sleep 2010 Jun1;33(6):781-792.

Dunn, W. Sensory profile. San Antonio, TX: The Psychological Corporation, 1999.

Dunn, W. Sensory profile 2: User's Manual. Bloomington, MN: Pearson, 2014.

Esbensen AJ, Seltzer MM, Lam KS, et al. Age-related differences in restricted repetitive behaviors in autism spectrum disorders. J Autism Dev Disord. 2009;39:57-66.

Evans DW, Uljarevic M, Lusk LG, Loth E, Frazier T. Development of two dimensional measures of restricted and repetitive behavior in parents and children. J Amer Acad Child Adolesc Psychiatry. 2017;56:51-8.

Guy W, ed. ECDEU Assessment Manual for Psychopharmacology. Rev. ed. Washington DC: US Department of Health, Education and Welfare. 1976.

Johns MW. A new method for measuring daytime sleepiness: the Epworth Sleepiness Scale. Sleep. 1991;14:540-5.

Johns MW. Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep. 1992;15:376-81.

Johns M. The assessment of sleepiness in children and adolescents. Sleep Biol Rhythm. 2015;13(Suppl1):97. Available at: epworthsleepinessscale.com. Accessed on 21 April 2020.

Kim SH, Lord C. Restricted and repetitive behaviors in toddlers and preschoolers with autism spectrum disorders based on the Autism Diagnostic Observation Schedule (ADOS). Autism Res. 2010;3:162-73.

Lam KS, Aman MG. The Repetitive Behavior Scale-Revised: independent validation in individuals with autism spectrum disorders. J Autism Dev Disord. 2007;37:855-66.

Mirenda P, Smith IM, Vaillancourt T, Georgiades S, Duku E, Szatmari P, Bryson S, Fombonne E, Roberts W, Volden J, Waddell C, Zwaigenbaum L; Pathways in ASD Study Team. Validating the Repetitive Behavior Scale-revised in young children with autism spectrum disorder. J Autism Dev Disord. 2010 Dec;40(12):1521-30. doi: 10.1007/s10803-010-1012-0. PMID: 20405194.

Pender, R., Fearon, P., Heron, J., & Mandy, W. (2020). The longitudinal heterogeneity of autistic traits: A systematic review. Research in Autism Spectrum Disorders, 79, 101671.

Richler J, Huerta M, Bishop SL, Lord C. Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. Dev Psychopathol. 2010 Winter;22(1):55-69.

Simonoff E, Kent R, Stringer D, Lord C, Briskman J, Lukito S, Pickles A, Charman T, Baird G. Trajectories in Symptoms of Autism and Cognitive Ability in Autism From Childhood to Adult Life: Findings From a Longitudinal Epidemiological Cohort. J Am Acad Child Adolesc Psychiatry. 2020 Dec;59(12):1342-1352.
Sparrow, SS, Cicchetti DV, Balla DA. Vineland Adaptive Behavior Scales-2nd edition – teacher rating form manual. Minneapolis, MN: NCS Pearson, Inc. 2005.

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). Vineland Adaptive Behavior Scales, Third Edition (Vineland-3). Bloomington: NCS Pearson.

Uljarević M, Frazier TW, Jo B, Billingham WD, Cooper MN, Youngstrom EA, Scahill L, Hardan AY. Big Data Approach to Characterize Restricted and Repetitive Behaviors in Autism. J Am Acad Child Adolesc Psychiatry. 2022 Mar;61(3):446-457. doi: 10.1016/j.jaac.2021.08.006. Epub 2021 Aug 12. PMID: 34391858.

Yu, L, Buysse DJ, Germain A et al. Development of Short Forms from the PROMIS Sleep Disturbance and Sleep-Related Impairment Item Banks. Behav Sleep Med. 2011 Dec 28; 10(1): 6–24. doi:10.1080/15402002.2012.636266.

APPENDIX 1

LIST OF PLANNED OUTPUTS FOR EFFICACY AND SAFETY

Provisional Output Name	Population	Title
l_ae_gloss	Safety	Glossary of Adverse Event Preferred Terms and Investigator-Specified Adverse Event Terms
l_ae	Safety	Listing of Adverse Events
l_ae_postdose	Safety	Listing of new onset AEs after last dose of study medication
l_ae_disc	Safety	Listing of Adverse Events Leading to Study Treatment Discontinuation
l_cm_gloss	Safety	Glossary of Concomitant Medication Class, Preferred Name, and Investigator-Specified Terms
l_cm	Safety	Listing of Concomitant Medication

l_cssrs	Safety	Columbia-Suicide Severity Rating Scale (C-SSRS)
l_ds	Safety	Listing of Patients who Discontinued from Study
l_dv	Safety	Listing of Major Protocol Deviations
l_eg	Safety	Listing of ECG Data
l_ex_ti	Safety	Listing of Drug Administration and Dosing Time
l_mh_gloss	Safety	Glossary of Medical History Coded Terms
l_mh	Safety	Listing of Medical History
l_vs	Safety	Listing of Vital Signs
t_ae_int	Safety	Adverse Events by Greatest Intensity
t_ae_oview	Safety	Safety Summary
t_ae	Safety	Adverse Events
t_cssrs	Safety	Columbia-Suicide Severity Rating Scale (C-SSRS)
t_dm	Safety	Demographic and Baseline Characteristics
t_ds	Safety	Patient Disposition
t_eg_cfb	Safety	ECG Parameters Change from Baseline by Visit
t_ess_cat	Safety	Epworth Sleepiness Scale (ESS) Total Score Categories
t_ess_cfb	Safety	Epworth Sleepiness Scale (ESS) Change from Baseline by Visit
t_ess_chad_cat	Safety	ESS for Adolescents Sleepiness Scale (ESS-CHAD) Total Score Categories
t_ess_chad_cfb	Safety	ESS for Adolescents Sleepiness Scale (ESS-CHAD) Change from Baseline by Visit
t_ex	Safety	Study Treatment Exposure
t_kss_cfb	Safety	Karolinska Sleepiness Scale (KSS) Change from Baseline by Visit
t_lb_abn_ls	Safety	Laboratory Abnormalities (Standardised parameters)
t_lb_abn_nls	Safety	Laboratory Abnormalities (Non-Standardised parameters)
t_lb_cfb_ls	Safety	Laboratory Tests and Change from Baseline by Visit (Standardised parameters)

t_lb_cfb_nls	Safety	Laboratory Tests and Change from Baseline by Visit (Non-Standardised parameters)
t_neuro_cfb	Safety	Neurocognitive Battery Change from Baseline by Visit
t_sudden	Safety	Sudden Onset of Sleep Questionnaire
t_vs_cfb	Safety	Vital Signs Change from Baseline by Visit
l_ae_ser	Safety	Listing of Serious Adverse Events
l_dm	Safety	Listing of Demographic and Baseline Characteristics
l_lb_abn	Safety	Listing of Laboratory Abnormalities
l_lb	Safety	Listing of Laboratory Data and Change from Baseline by Visit
t_ae_pt	Safety	Adverse Events by Preferred Term
t_ae	Safety	Adverse Events Related to Study Treatment
t_cm	Safety	Medications Started after Baseline by Medication Class and Preferred Name
t_cm	Safety	Medications Ongoing at Baseline by Medication Class and Preferred Name
t_cm	Safety	Medications Completed Prior to Baseline by Medication Class and Preferred Name
t_eg_abn	Safety	ECG Parameters outside Normal Limits Regardless of Abnormality at Baseline
t_eg_abnbl	Safety	ECG Parameters outside Normal Limits Among Patients without Abnormality at Baseline
t_eg_qtcf	Safety	QTcF Actual Values and Change from Baseline by Visit
t_mh	Safety	Previous and Concurrent Medical History
l_ae_nsaesi	Safety	Listing of Non-Serious Adverse Events of Special Interest
t_ae_nsaesi	Safety	Non-Serious Adverse Events of Special Interest
g_ex_hist	Safety	Histogram of Dosing Time
g_vs_mean	Safety	Mean Plot from of Vital Sign Parameters Over Time with 95% CI
g_vs_mean_cfb	Safety	Mean Change from Baseline Plot of Vital Sign Parameters Over Time with 95% CI

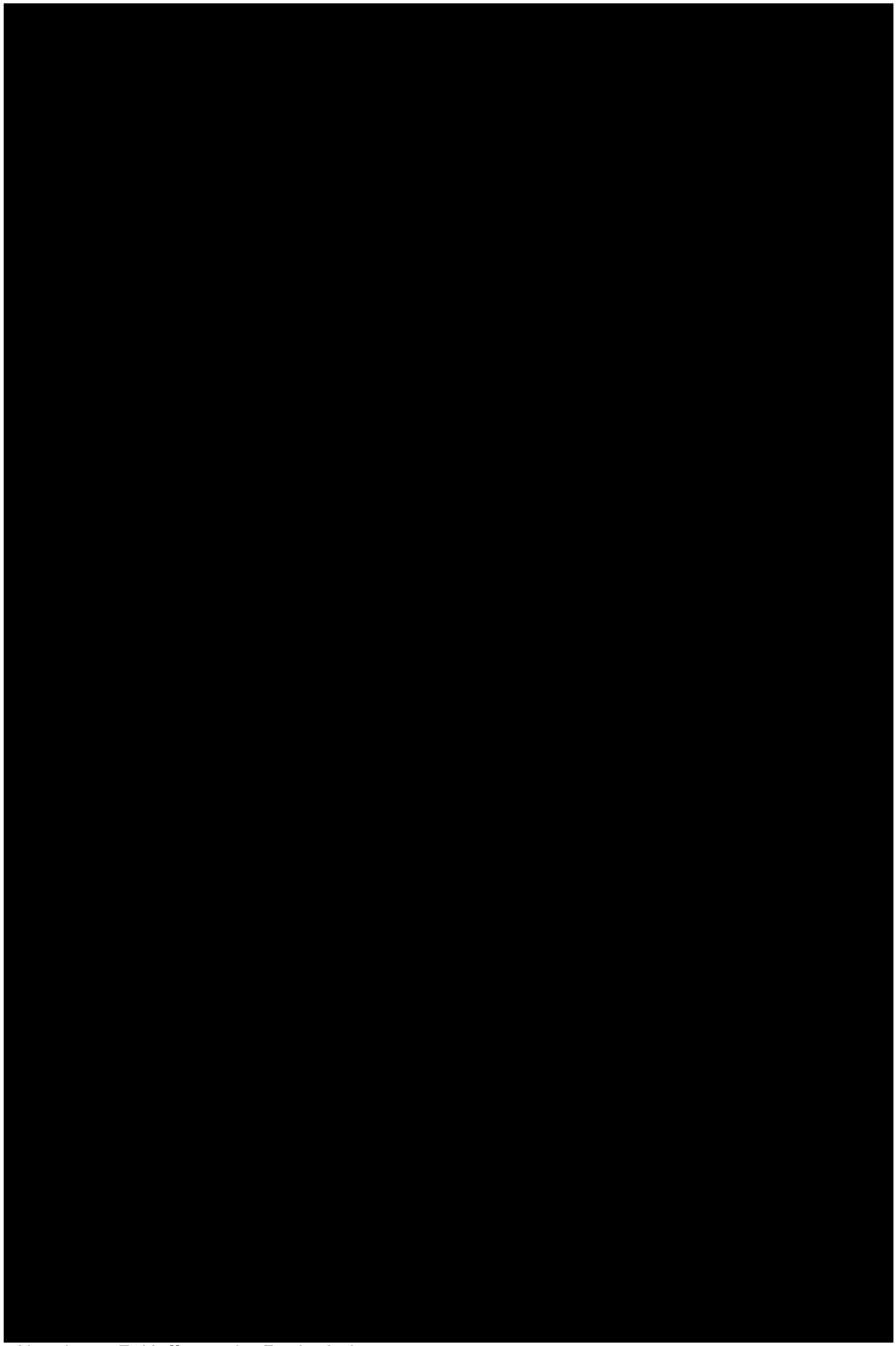
g_lb_mean_ls	Safety	Roche-Standardized Laboratory Data
g_lb_mean_nls	Safety	Non-Standardized Laboratory Data
g_eg_mean	Safety	Mean Plot of ECG Parameters Over Time with 95% CI
g_eg_mean_cfb	Safety	Mean Change from Baseline Plot of ECG Parameters Over Time with 95% CI
g_ess	Safety	ESS over Time
g_ess_cfb	Safety	ESS Change from Baseline over Time
g_kss	Safety	KSS over Time
g_kss_cfb	Safety	KSS Change from Baseline over Time
g_neuro	Safety	Neurocognitive Battery

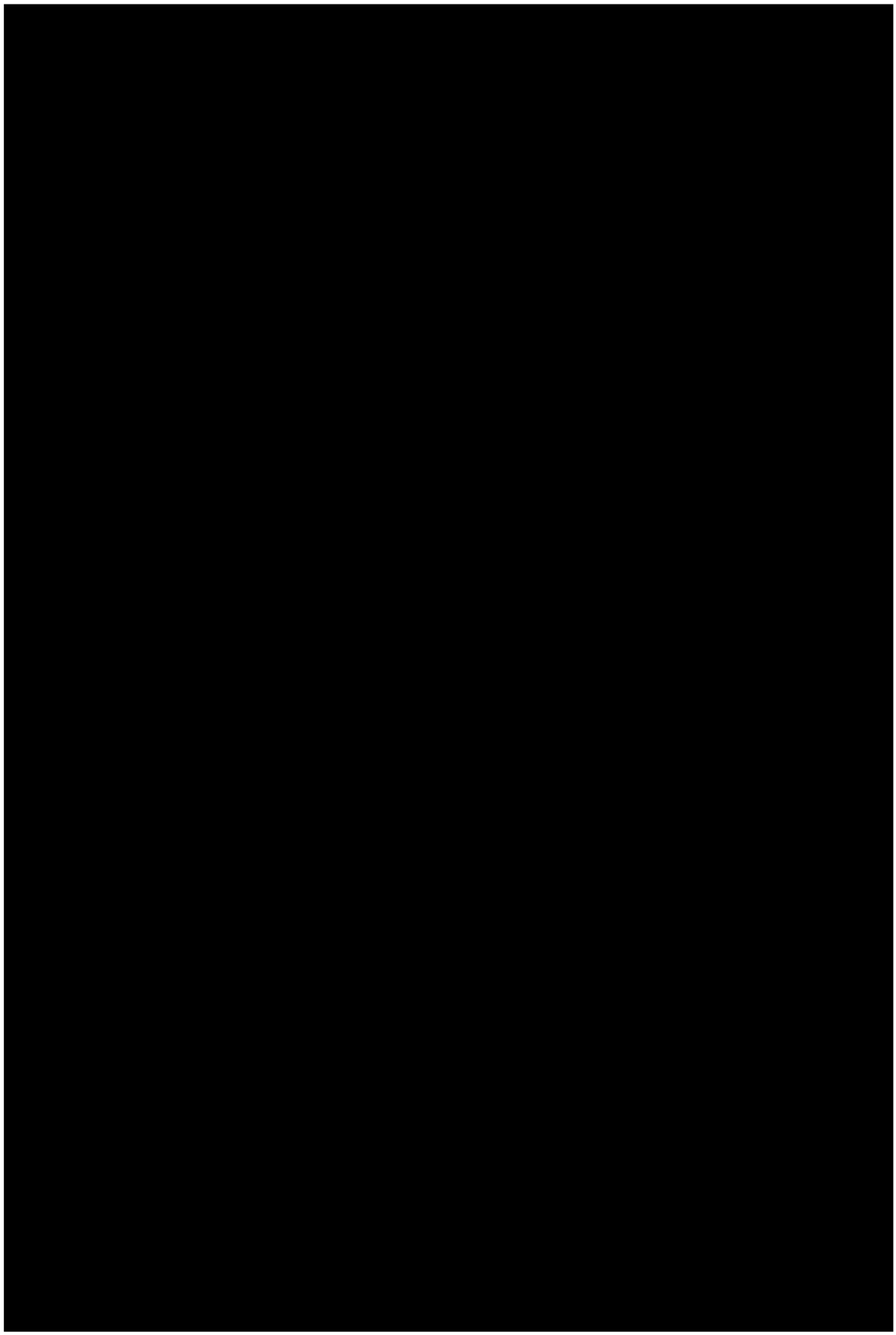
t_vabs3_ancova	Efficacy	Vineland-3 ABC GSV Score ANCOVA no interactions
t_vabs3_ancova	Efficacy	Vineland-3 ABC GSV Score ANCOVA age interaction
t_vabs3_ancova	Efficacy	Vineland-3 ABC GSV Score ANCOVA IQ interaction
t_vabs3_ancova	Efficacy	Vineland-3 SOC GSV Score ANCOVA no interactions
t_vabs3_ancova	Efficacy	Vineland-3 SOC GSV Score ANCOVA age interaction
t_vabs3_ancova	Efficacy	Vineland-3 SOC GSV Score ANCOVA IQ interaction
t_vabs3_ancova	Efficacy	Vineland-3 COM GSV Score ANCOVA no interactions
t_vabs3_ancova	Efficacy	Vineland-3 COM GSV Score ANCOVA age interaction
t_vabs3_ancova	Efficacy	Vineland-3 COM GSV Score ANCOVA IQ interaction
t_vabs_cfb	Efficacy	Vineland-3 ABC GSV Score Change from Baseline
t_vabs_cfb	Efficacy	Vineland-3 SOC GSV Score Change from Baseline
t_vabs_cfb	Efficacy	Vineland-3 COM GSV Score Change from Baseline
t_vabs_cfb	Efficacy	Vineland-3 DLS GSV Score Change from Baseline
t_vabs_cfb	Efficacy	Vineland-3 ABC Standard Score Change from Baseline
t_vabs_cfb	Efficacy	Vineland-3 SOC Standard Score Change from Baseline
t_vabs_cfb	Efficacy	Vineland-3 COM Standard Score Change from Baseline
t_vabs_cfb	Efficacy	Vineland-3 DLS Standard Score Change from Baseline
t_rbsr_ancova	Efficacy	RBS-R Total Score ANCOVA no interactions
t_rbsr_ancova	Efficacy	RBS-R Total Score ANCOVA age interaction
t_rbsr_ancova	Efficacy	RBS-R Total Score ANCOVA IQ interaction
t_rbsr_cfb	Efficacy	RBS-R Total Score Change from Baseline
t_rbsr_cfb	Efficacy	RBS-R Subscale Score (3FM) Change from Baseline

t_crri_ancova	Efficacy	CRRI Total Score ANCOVA no interactions
t_crri_ancova	Efficacy	CRRI Total Score ANCOVA age interaction
t_crri_ancova	Efficacy	CRRI Total Score ANCOVA IQ interaction
t_crri_cfb	Efficacy	CRRI Total Score Change from Baseline
t_crri_cfb	Efficacy	CRRI Subscale Score (5FM) Change from Baseline
t_cgii_ancova	Efficacy	CGI-I Overall Score ANCOVA no interactions
t_cgii_ancova	Efficacy	CGI-I Overall Score ANCOVA age interaction
t_cgii_ancova	Efficacy	CGI-I Overall Score ANCOVA IQ interaction
t_cgii_mmrm	Efficacy	CGI-I Overall Score MMRM
t_cgii	Efficacy	CGI-I Overall Score
t_cgii_cat	Efficacy	CGI-I Proportions per Rating Category
t_cgis	Efficacy	CGI-S Overall Score
t_cgis_cat	Efficacy	CGI-S Proportions per Rating Category
t_bai_cfb	Efficacy	BAI Total Score Change from Baseline
t_pras_cfb	Efficacy	PRAS-ASD Total Score Change from Baseline
t_ssp_cfb	Efficacy	SSP-2 Summary Score Change from Baseline
t_aasp_cfb	Efficacy	AASP Quadrant Summary Score Change from Baseline
t_promis_cfb	Efficacy	PROMIS Sleep-Disturbance T-Score Change from Baseline
t_promis_cfb	Efficacy	PROMIS Sleep Related Impairment T-Score Change from Baseline
t_<>	Efficacy	Change from Baseline in EEG beta power
t_<>	Efficacy	Correlation between Baseline EEG beta power and clinical efficacy
t_pql_tot_cfb	Efficacy	PedsQL™ Total Score Change from Baseline
t_pql_cfs_cfb	Efficacy	PedsQL™ Cognitive Functioning Score Change from Baseline
g_cgii	Efficacy	CGI-I Overall Score over Time
g_cgis	Efficacy	CGI-S Overall Score over Time
g_forest	Efficacy	Vineland-3 ABC GSV composite

		by Geographic Region with 80% CI
g_forest	**Efficacy**	**Vineland-3 ABC GSV composite by Rater Change with 80% CI**
g_forest	**Efficacy**	**Vineland-3 ABC GSV composite by Site Type with 80% CI**
g_forest	**Efficacy**	**Vineland-3 ABC GSV composite by Age Category with 80% CI**
g_forest	**Efficacy**	**Vineland-3 ABC GSV composite by IQ Category with 80% CI**
g_forest	**Efficacy**	**RBS-R total score by Geographic Region with 80% CI**
g_forest	**Efficacy**	**RBS-R total score by Rater Change with 80% CI**
g_forest	**Efficacy**	**RBS-R total score by Site Type with 80% CI**
g_forest	**Efficacy**	**RBS-R total score by Age Category with 80% CI**
g_forest	**Efficacy**	**RBS-R total score by IQ Category with 80% CI**
g_forest	**Efficacy**	**RBS-R total score by Diagnosis of Anxiety at Baseline with 80% CI**
g_forest	**Efficacy**	**RBS-R total score by Sex with 80% CI**
g_forest	**Efficacy**	**CRRI total score by Sex with 80% CI**
g_forest	**Efficacy**	**CGII (RRB domain) by Sex with 80% CI**
g_forest	**Efficacy**	**BAI total score by Diagnosis of Anxiety at Baseline with 80% CI**

[redacted]





DocuSign

Certificate Of Completion

Envelope Id: 00A25604873540C595AE83AF56584F0E
Status: Completed
Subject: Complete with Docusign: BP41316_SAP_V02.pdf
Source Envelope:
Document Pages: 51
Signatures: 1
Envelope Originator:
Certificate Pages: 4
Initials: 0
[redacted]
AutoNav: Enabled
Grenzachstrasse 124
EnvelopeId Stamping: Disabled
Basel, Basel-Stadt 4070
Time Zone: (UTC+01:00) Amsterdam, Berlin, Bern, Rome, Stockholm, Vienna
[redacted]
IP Address: [redacted]

Record Tracking

Status: Original
6/28/2024 1:54:12 PM
Holder: [redacted]
Location: DocuSign

Signer Events	Signature	Timestamp
[redacted] Statistician F. Hoffmann-La Roche Security Level: Email, Account Authentication (Required)	[redacted] Signature Adoption: Pre-selected Style Signature ID: 10EC3CF5-F664-481F-BAE0-AD933DC0A009 Using IP Address: [redacted] With Signing Authentication via DocuSign password With Signing Reasons (on each tab): I approve this document	Sent: 6/28/2024 1:54:56 PM Viewed: 6/28/2024 1:55:15 PM Signed: 6/28/2024 1:55:35 PM

Electronic Record and Signature Disclosure:
Accepted: 11/2/2020 2:09:16 PM
ID: 102e05f3-7156-49ba-9ce4-508e079f28ec
Company Name: F.Hoffmann-La Roche Ltd. - CFR Part 11 - Sub 3

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	6/28/2024 1:54:56 PM
Certified Delivered	Security Checked	6/28/2024 1:55:15 PM
Signing Complete	Security Checked	6/28/2024 1:55:35 PM
Completed	Security Checked	6/28/2024 1:55:35 PM

Payment Events	Status	Timestamps

Electronic Record and Signature Disclosure

Guidelines for the use of the DocuSign electronic signing system within F. Hoffmann-La Roche and companies of the Roche Group

LEGAL DISCLOSURE

By placing my electronic signature on this document, I expressly consent to use and rely on Electronic and Digital Signatures and I understand my signature will have the same binding effect as if I was providing a handwritten signature.

I also confirm the email address that I am using as a valid one to be notified and identified electronically, for example through the DocuSign System, and/or to identify me as the signer of the document. I should inform Roche in the case that the email address changes.

CONSUMER DISCLOSURE

From time to time, F. Hoffmann-La Roche and companies of the Roche Group (hereinafter referred to as we, us or Company) may be required by law to provide to you certain written notices or disclosures. Described below are the terms and conditions for providing to you such notices and disclosures electronically through the DocuSign, Inc. (DocuSign) electronic signing system. Please read the information below carefully and thoroughly, and if you can access this information electronically to your satisfaction and agree to these terms and conditions, please confirm your consent by clicking the 'I agree' button at the bottom of this document.

Consequences of changing your mind

To indicate to us that you are changing your mind, you must withdraw your consent using the DocuSign 'Withdraw Consent' form on the signing page of a DocuSign envelope instead of signing it. This will indicate to us that you have withdrawn your consent to receive required notices and disclosures electronically from us and you will no longer be able to use the DocuSign system to receive required notices and consents electronically from us or to sign electronically documents from us.

All notices and disclosures will be sent to you electronically

Unless you tell us otherwise in accordance with the procedures described herein, we will provide electronically to you through the DocuSign system all required notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you during the course of our relationship with you. To reduce the chance of you inadvertently not receiving any notice or disclosure, we prefer to provide all of the required notices and disclosures to you by the same method and to the same address that you have given us. If you do not agree with this process, please let us know as described below. Please also see the paragraph immediately above that describes the consequences of your electing not to receive delivery of the notices and disclosures electronically from us.

How to contact us

You may contact us to let us know of your changes as to how we may contact you electronically, to request certain information from us and to withdraw your prior consent to receive notices and disclosures electronically. To do so contact the sender of the envelope.

To advise us of your new email address

To let us know of a change in your email address where we should send notices and disclosures electronically to you, you must send an email message to the sender of the envelope and in the body of such request you must state: your previous email address, your new email address. We do not require any other information from you to change your email address. In addition, you must notify DocuSign, Inc. to arrange for your new email address to be reflected in your DocuSign account by following the process for changing email in the DocuSign system.

To withdraw your consent

To inform us that you no longer want to receive future notices and disclosures in electronic format you may decline to sign a document from within your DocuSign session, and on the subsequent page, select the check-box indicating you wish to withdraw your consent.

Acknowledging your access and consent to receive materials electronically

To confirm to us that you can access this information electronically, which will be similar to other electronic notices and disclosures that we will provide to you, please verify that you were able to read this electronic disclosure and that you also were able to print on paper or electronically save this page for your future reference and access or that you were able to email this disclosure and consent to an address where you will be able to print on paper or save it for your future reference and access. Further, if you consent to receiving notices and disclosures exclusively in electronic format on the terms and conditions described above, please let us know by clicking the 'I agree' button below.

By checking the 'I agree' box, I confirm that:

- I can access and read this Electronic CONSENT TO ELECTRONIC RECEIPT OF ELECTRONIC CONSUMER DISCLOSURES document; and
- I can print on paper the disclosure or save or send the disclosure to a place where I can print it, for future reference and access; and
- Until or unless I notify the account owner as described above, I consent to receive from exclusively through electronic means all notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to me by the account owner during the course of my relationship with you.
- I acknowledge the information included in the DocuSign Privacy Policy