Statistical Analysis Plan

Sponsor Name	Sumitomo Pharma America Holdings, Inc.
Protocol ID	BVR100-102
Protocol Version	1.00 (date: 21-JUN-2023)
Study Name	BVR100-102
Study Title	BVR-100 and BES-100 Validation Trial: A Randomized, Double-Blind, Parallel-Group, Controlled Study of Two At-Home Self-Guided Virtual Reality Interventions for Adults with Social Anxiety Disorder.
CRF Version	2.0 (date: 20-SEP-2023)
Document ID	BVR100-102_Statistical Analysis Plan
Document Version	1.00

	Signature	**Date**
[redacted]		

Table of Contents

1 Definitions

Abbreviation	Definition
AE	Adverse Event
BES100	Active at-home VR Intervention Comparator
BSL	Baseline
CGI-I	Clinical Global Impression - Improvement
CGI-S	Clinical Global Impression - Severity
CBT	Cognitive Behavioral Therapy
C-SSRS	Columbia Suicide Severity Rating Scale
CEQ	Credibility/ Expectancy Questionnaire
CSP	Clinical Study Protocol
DDP	Data Display Plan
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
DTx	Digital Therapeutics
ET	Early Termination
eCRF	Electronic Case Report Form
EOS	End of Study
FAS	Full Analysis Set
ICF	Informed Consent Form
ITT	Intent-to-Treat
IWRS	Interactive Web Responses System
LSAS	Liebowitz Social Anxiety Scale
MDD	Major Depressive Disorder
MedDRA	Medical Dictionary for Regulatory Activities
MINI	Mini-International Neuropsychiatric Interview
m-ITT	Modified Intent-to-Treat
PGI-I	Patient Global Impression - Improvement
PHQ-9	Patient Health Questionnaire-9
PRO	Patient-Reported Outcome
PP	Per-Protocol
RND	Randomization
RCT	Randomized Controlled Trials
SCR	Screening
SAE	Serious Adverse Event
SAD	Social Anxiety Disorder
TEAE	Treatment Emergent Adverse Events

FDA	U.S. Food and Drug Administration
UADE	Unanticipated Adverse Device Effects
VR	Virtual Reality
VRET	Virtual Reality Exposure Therapy
V	Visit
W	Week

2 Introduction

This statistical analysis plan (SAP) provides the detailed methodologies and statistical analyses of the data collected in the study BVR100-102 associated with protocol version 1.00 (dated 21-JUN-2023) and Electronic Case Report Form (eCRF) version 2.0 (date: 20-SEP-2023). This document may modify the plans outlined in the protocol; however, any major modifications of the study endpoints definition or the analysis will also be reflected in a protocol amendment. If a discrepancy is found between the description in the statistical section of the protocol and this document, the description in this document supersedes the descriptions in the statistical section of the protocol. Any deviations from the final analysis plan or from what is outlined in the protocol will be discussed in the final study report.

The SAP includes prespecified analyses of all data from each subject through the End of Study (EOS) or Early Termination (ET). The EOS is defined as the time at which subjects completed all study visits, have failed screen, are lost to follow-up, have withdrawn of consent for further participation in the study, have adverse event, including Unanticipated Adverse Device Effects (UADE), have protocol deviation or death, or other reasons. ET is defined as the time subjects who prematurely discontinue from the study.

2.1 Background

Social anxieties and fears exist along a continuum of intensity, ranging from mild phenomena such as shyness to psychopathological levels of fear and anxiety that can be classified as social anxiety disorder (SAD) [1]. SAD is characterized by overwhelming anxiety and excessive self-consciousness in everyday social situations, which causes significant distress and often leads to impairment in important areas of functioning [2].

SAD can be a chronic and naturally unremitting disorder if not treated adequately. Despite the extent of suffering and impairment, only about half of adults with the disorder ever seek treatment. Guidelines from the National Institute for Health and Care Excellence recommend individual Cognitive Behavioral Therapy (CBT) as the first-line treatment for adults with social anxiety [4].

Virtual reality (VR) technology offers a unique opportunity to disseminate CBT and exposure therapy. Studies indicate that patients prefer to receive VR exposure therapy (VRET) to traditional exposure therapy [5]. Because of its potential to shield subjects' attention from distraction and induce strong feelings of presence, VR can enhance the practices of mindfulness and facilitate emotion regulation therefore reducing symptoms of anxiety. Furthermore, VR may be combined with engaging game-design mechanics to deliver highly automated unguided interventions, thus showing an enormous potential for improving public health due to its cost-effectiveness and scalability.

The current study will test BVR-100, a VR-based multi-session, self-led intervention for social anxiety symptoms including CBT techniques against BES-100, a VR-based intervention intended as a control-arm for BVR-100.

2.2 Rationale

Digital therapeutics (DTx) are evidence-based software products used to prevent, manage, or treat medical conditions. DTx need to adhere to the same standards of evidence and regulatory oversight as

prescription medications and traditional medical devices. However, DTx face additional challenges to establishing appropriate control conditions that support evidence generation in randomized controlled trials (RCT). In particular, DTx interventions are subject to "digital placebo" effects, involving technology-related non-specific effects (e.g., beliefs about technology, design of the hardware and software, information provided, medium of delivery, Time-on-Task, etc.), that are not present in drug trials. Furthermore, DTx delivering psychological interventions such as CBT are faced with the additional challenge of controlling for more complex placebo effects that typically occur in the context of psychotherapy, such as the effect of the frequency and intensity of therapist–patient interactions. While the FDA acknowledges that "it may be challenging to construct a placebo control that appears to function like the investigational device but delivers no therapy" (Center for Devices & Radiological Health, n.d.), it is still recommended that Sponsors evaluate placebos in terms of their face-validity (i.e., Time-on-Task), blinding, and subject retention. Pilot testing BVR-100 against a "sham" control such as BES-100 will provide the Sponsors with the opportunity to test the face-validity of the BES-100, preliminarily assess the effectiveness of BVR-100 and BES-100, and inform the subsequent large-scale RCT.

3 Study Objectives

The objective of this study is to compare two Virtual Reality (VR)-based interventions, BVR-100 and BES-100, for the treatment of Social Anxiety Disorder (SAD).

3.1 Primary Objective

The primary objective of this study is to evaluate the credibility and expectation of benefit of BVR-100 and BES-100 for subjects with SAD, as measured with the Credibility/Expectancy Questionnaire (CEQ).

3.2 Secondary Objective

The secondary objective of this study is to evaluate subject retention and Time-on-Task of BVR-100 and BES-100.

3.3 Exploratory Objectives

This study will pursue the following exploratory objectives:

- *To explore subjects' experiences and acceptability of BVR-100 and BES-100 in a Qualitative Interview.*
- *To evaluate the therapeutic response of the two interventions for SAD on:*
    - *Fear and avoidance symptoms, as measured by the clinician administered Liebowitz Social Anxiety Scale (LSAS).*
    - *Overall symptomatology and functioning, using the Clinical Global Impression Severity scale (CGI-S), the Clinical Global Impression Improvement scale (CGI-I), and the Patient Global Impression Improvement scale (PGI-I).*
    - *Depression symptoms, measured with the Patient Health Questionnaire (PHQ-9).*
    - *To evaluate the number of required sessions completed during the study.*

3.4 Safety Objectives

This study will evaluate the safety of BVR-100 and BES-100, as measured by:

- *Incidence of adverse events (AEs),*
- *Incidence of serious adverse events (SAEs),*
- *Incidence of unanticipated adverse device effects (UADEs), and*
- *The Columbia Suicide Severity Rating Scale (C-SSRS).*

4 Study Design

4.1 Overview

This is a parallel-group, double-blind, randomized, controlled, decentralized, study. The aim of the study is to evaluate the credibility and acceptability, subject retention, and Time-on-Task of BVR-100 and BES-100 in a sample of subjects with Social Anxiety Disorder (SAD).

The study will include 3 periods as shown in table below.

SCR	RND	Intervention Period (Double-Blind)				
		BSL / W1	W2	W4	W6	W8/EOS/ET
V1	V2	V3	V4	V5	V6	V7
Up to 14 days	Up to 7 days	Days 1-7	Days 8-14	Days 22-28	Days 36-42	Days 50-56

Abbreviations: SCR = Screening; RND = Randomization; BSL = baseline; EOS = End of Study; ET = Early Termination; V = telehealth / call visits; W = week.

The total study duration will be approximately 11 weeks from Screening through EOS (Week 8).

4.2 Sample Size Justification

Due to the pilot nature of the study, no formal sample size calculations will be conducted. However, fifteen people per arm are recommended. In order to account for a possible attrition rate of 25%, twenty people per arm (total n = 40) will be recruited. In order to ensure adequate data for evaluation, early discontinued subjects may be replaced at Sponsor discretion.

4.3 Randomization and Blinding

Randomization will occur after confirmation of eligibility and continued subject interest are verified. Subjects will be randomly allocated to either the treatment group (BVR-100) or the control (BES-100) group. Randomization will be assigned via an interactive web responses system (IWRS) in a 1:1 ratio to either BVR-100 or to BES-100.

Once a randomization number has been assigned, it cannot be reused. The randomization schedule will be generated by an independent, non-study biostatistician.

Subjects, Investigators, investigator site staff, persons performing the assessments, clinical operations personnel, data analysts and Sponsor, will remain blinded to the identity of the VR intervention from the time of randomization until database lock and unblinding.

4.4 Inclusion/Exclusion Criteria

For details of Inclusion/Exclusion criteria, please refer to the protocol section 1 – Inclusion and Exclusion Criteria.

4.5 Treatment Allocation

Subjects will be assigned via an IWRS in a 1:1 ratio to either BVR-100 or to BES-100. A VR headset pre-loaded with either BVR-100 or BES-100 will be shipped to the subject's address.

BVR-100 is a non-invasive, self-guided, at-home, VR intervention which offers educational modules and immersive environments created using both computer-generated and 360º videos. The intervention includes immersive and automated VR experiences that guide the user through educational and experiential learning modules.

BES-100 is a non-invasive, self-guided, multisession, at-home, VR intervention designed to be used as a sham control condition for the evaluation of BVR-100 in the treatment of SAD. BVR-100 and BES-100 are comparable in all technological, design, and programmatic characteristics, except for the mechanisms of action. The computer-generated environment will be interactive to increase engagement and provide subjects with a similar interactive experience as BVR-100.

Both are intended as an 8-week program with a recommended minimum frequency of 3 weekly sessions. Each session will be approximately 5-20 minutes long (depending on type and amount of content experienced).

5 Assessment Schedule

			Investigational Intervention Use Period				
Study Period	**Screening**	**Randomization**	**Baseline [a]**				
Study Visit	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**	**V7**
Study Week	NA	NA	W1	W2	W4	W6	W8 (EOS/ET)[b]
Study Visit Days	Up to 14 days	Up to 7 days	1-7	8-14	22-28	36-42	50-56
Procedure							
Obtain Informed Consent	X						
Review Inclusion/Exclusion Criteria	X	X					
Record Demographics and Baseline Characteristics	X						
Record Medical/ Psychiatric History	X						
Administer MINI	X						
Randomization [c]		X					
Dispense Investigational Intervention via IWRS [c]		X					
Subject Self-guided Weekly VR Sessions [d]			X	X	X	X	X
Clinician Administers the LSAS	X		X	X	X	X	X
Clinician Administers the C-SRSS	X		X		X		X
Subject Completes the PHQ-9 [e]	X		X	X	X	X	X
Subject Completes the CEQ			X			X	X
Clinician Administers the CGI-I				X	X	X	X
Clinician Administers the CGI-S			X	X	X	X	X
Subject Completes the PGI-I				X	X	X	X
Conduct Qualitative Interview [f]							X
Prior/Concomitant Medications	X	X	X	X	X	X	X
Pretreatment Event and Adverse Event Monitoring	X	X	X	X	X	X	X
Subject Returns Investigational Intervention [f]							X
Telephone Contact [g]	X	X	X	X	X	X	X

Abbreviations: AE = Adverse Event; CEQ = Credibility/Expectancy Questionnaire; D = Day; EOS = End of Study; ET = Early Termination; C-SRSS = Columbia Suicide Severity Rating Scale; MINI = Mini-International Neuropsychiatric Interview; CGI-I = Clinical Global Impression – Improvement scale; CGI-S = Clinical Global Impression – Severity scale; LSAS = Liebowitz Social Anxiety Scale; PGI-I = Patient Global Impression – Improvement scale; PHQ-9 = Patient Health Questionnaire; PROs = Patient-Reported Outcomes VR = Virtual Reality; W = Week.

[a] The Baseline visit should be scheduled within 5-7 days of the Randomization visit. The Baseline visit will consist of two parts. The first part will be conducted by investigator site staff via a telehealth call and will include the administration of clinician-administered outcome measures, collection of AEs, concomitant medication, and delivery of the PHQ-9. The second part will be self-administered by the subject and will include the video-guided VR onboarding session, the first self-guided VR session, and the completion of the CEQ.

[b] For subjects who prematurely discontinue from the study, every effort should be made to complete the final evaluation procedures at the Early Termination (ET) visit within 5 days of the last exposure to the Investigational Intervention, including the PROs if they have not been completed by the subject for the given week.

[c] Randomization cannot occur until the subject has been deemed eligible by the site and Inclusion and Exclusion criteria re-confirmed with the subject via a telehealth contact and after subject interest to participate in this study is confirmed. Upon Randomization using the IWRS, the Investigational Intervention shipment to the subject will be initiated.

[d] The subject self-guided weekly VR sessions begin after the first part of the Baseline visit is completed. During weeks 2, 4, 6 and 8, 3 required VR sessions should be completed prior to the PROs. In addition to the required practices, subjects will be encouraged to also engage with optional content and to review content completed in previous sessions.

[e] The PROs will be completed by the subject during the Telehealth contact, for Screening and Baseline Visits, as applicable, in an unassisted manner. During Weeks 2, 4, 6, and 8, these are to be completed after the self-guided at home VR sessions.

[f] The W8/EOS/ET assessments will consist of three parts. The first part will be the Subject's at home completion of the required 3 VR modules and electronic PRO assessments. The second part will include the telehealth visit with the investigator site staff to complete all applicable assessments. The third part will be conducted by the unblinded site support team where the Qualitative Interview will be completed as the final study assessment on the same day or within 48 hours. During this visit, the unblinded site support team will instruct the subject to return the device and provide any necessary guidance.

[g] The Telehealth contact during weeks 2, 4, 6, and 8, to collect concomitant mediation and/or AEs, clinician administered C-SRSS, LSAS, CGI-S, CGI-I, and other applicable assessments, should be completed on the last 3 days of each week. At any point in the study, an unscheduled Telehealth contact can be performed if requested by the investigator site staff/subject.

6 Interim Analysis

There is no interim analysis planned for this study.

7 Efficacy and Safety Endpoints

7.1 Primary Endpoint

The primary endpoint of this study is the measurement of intervention credibility and expectancy of benefit, assessed with the Credibility / Expectancy Questionnaire (CEQ) scores, at Baseline, Week 6, and Week 8/EOS.

7.2 Secondary Endpoints

Secondary endpoints for this study include:

- Subject retention in each group will be assessed by percentage of subjects who remain on the study at Week 6, and Week 8/EOS.
- Time-on-Task, averaged over all VR sessions, will be assessed at Week 6 and Week 8/EOS. Time-on-Task will be automatically recorded by the VR system after each session.

7.3 Safety Endpoints

- Incidence of AEs.
- Incidence of SAEs.
- Incidence of UADEs.
- Frequency of subjects with suicidal ideation or suicidal behavior based on the C-SSRS.

7.4 Exploratory Endpoints

- The overview of qualitative input obtained during Qualitative Interviews at Week 8.
- The mean change from baseline in LSAS scores at Week 2, Week 4, Week 6, and Week 8.
- The proportion of “responders”, that is, subject who are rated 1 (very much improvement) or 2 (much improvement) in the CGI-I at end-of-treatment (Week 8).
- The mean change from baseline in CGI-S score at Week 2, Week 4, Week 6, and Week 8.
- The mean PGI-I score at Week 2, Week 4, Week 6, and Week 8.
- The mean change from baseline in PHQ-9 score at Week 2, Week 4, Week 6, and Week 8.
- The number of VR sessions completed during the study.

8 Statistical Methods

8.1 General Conventions

Level of significance for statistical tests will be 0.05 two-sided. All analyses will be done using SAS (version 9.4, SAS Institute Inc., Cary, NC, USA).

Safety endpoints will be analyzed in a descriptive manner. Continuous data will be reported using the following descriptive statistics:

- Number of observations (n)
- Mean and Standard deviation (SD)
- Median
- Minimum (min) and Maximum (max)
- 25^{th} percentile and 75^{th} percentile

Categorical data will be presented using frequency (n = number of subjects; m = number of events) and percentage (%).

Listings will be provided for all data recorded in eCRF to study subject profiles. All listings will be sorted by treatment, subject ID and date (if applicable). Unscheduled visit data will only be listed and not included in summaries.

All p-values will be displayed to four decimal places, with p-values less than 0.0001 presented as '<0.0001' and p-values greater than 0.9999 presented as '>0.9999'. Minimum and maximum values will be reported in the units of collection with 3 decimals being maximum value; the mean, median, 25^{th} percentile and 75^{th} percentile will be presented with 1 decimal place more and the standard deviation 2 decimal places more than the units of collection. Percentages for categorical summaries will be represented to 1 decimal place.

8.2 Multiplicity Adjustment

Not applicable.

9 Data handling Procedures

9.1 Handling Missing/Incomplete Data

9.1.1 Missing Data in Efficacy Analysis

Missing data will not be imputed for efficacy endpoints analyses.

9.1.2 Missing Data in Safety Analysis

Missing data will not be imputed for safety endpoints analyses.

9.2 Partial Dates

Partial dates may be imputed for medications to determine whether they are concomitant medications using the algorithm given below.

End dates will be imputed as follows:

- If marked "ongoing" with no end date, then this will be considered as concomitant medication.
- If only year is present, then impute date and month as 31-DEC.
- If month and year are present, then impute date with last date of the month.
- If date and year are present, then impute month as DEC.
- If year is missing, then impute subject's last visit date in the study.

If the imputed end date is later than subject's last visit date, then the subject's last visit date will be used.

- In case the end date is partial, and we know that the medication has started after the last VR session, then the medication will not be considered as a concomitant medication.

In case of ambiguity, the worst case is considered (that is, concomitant medication) for safety analyses.

Example,

Partial End Date	Imputed End Date
--JUN2019	30JUN2019
-----2019	31DEC2019
12---2018	12DEC2018
14MAR----	Subject's last visit date

Imputed dates will be used for analyses and collected dates will be recorded in database will be displayed in listings.

9.3 Clean File / Blinded Review Meeting

Pre database lock blinded data review and protocol deviation review is out of scope for this SAP. Sumitomo statistician will be part of these review and meetings.

10 Analysis Populations

The membership of the analysis populations will be reviewed and finalized prior to database lock with blinded data.

10.1 Enrolled Population

Enrolled Population includes all subjects who sign the informed consent form and are successfully screened.

The Enrolled Population will be used for summaries of subject disposition.

10.2 Full Analysis Set (FAS)

The Full Analysis Set (FAS) consists of all subjects enrolled and randomized into the study.

The FAS population will be used for summaries of subject disposition, demographic and baseline characteristics. Subjects will be analyzed according to the treatment group they were assigned at randomization.

10.3 Safety Population

The safety population consists of all randomized subjects who have completed at least one VR session.

The safety population will be used for the analysis of safety. Subjects will be analyzed as treated, regardless of the randomized treatment assigned, if this differs from that to which the subject was randomized.

10.4 Intent-to-Treat (ITT) Population

The Intent-to-Treat (ITT) population consists of all randomized subjects who have at least one VR session.

The ITT population is the primary population to analyze efficacy endpoints. Demographic and baseline characteristics of the ITT population will also be presented. Subjects will be analyzed according to the treatment they received.

10.5 Per-Protocol Population

The Per-Protocol population consists of all ITT subjects who have completed the 8-week intervention period without any important protocol deviations.

The PP population is the population used for supportive analysis of the primary endpoint. Demographic and baseline characteristics of the PP population will be presented. Subjects will be analyzed according to the treatment they were received.

11 Analysis Variables

11.1 Population Flags

Population flags will be finalized and authorized by the Study Statistician and Sponsor prior to database lock as per definitions provided in Section 10. These flags will be included in the analysis datasets.

11.2 Treatment groups

All analyses will be summarized by treatment groups. The following treatment groups will be defined in the analysis datasets.

- BVR-100
- BES-100

11.3 Analysis Visits

All analysis datasets with measurements taken at more than one visit/time point will have analysis visit and timepoint as defined below.

Collected Visit	Analysis Visit
Screening	Visit 1 Screening
Randomization	Visit 2 Randomization
Week 1	Visit 3 Week 1
Week 2	Visit 4 Week 2
Week 4	Visit 5 Week 4
Week 6	Visit 6 Week 6
Week 8/ End of Study/ Early Termination	Visit 7 Week 8

A subject is considered to have attended any visit only if at least one assessment has been taken at the visit. As per section 11.10, derived baseline results will also be presented in summary tables.

11.3.1 Repeat Assessments

For safety and efficacy endpoints, if an unscheduled visit occurred on the same day as the scheduled visit, then it will be considered as a repeat for that assessment. Repeat and Unscheduled visits will not be considered for summary purpose but will be included in listing and baseline calculation.

11.4 Completion Flags

Subjects who completed the study will be flagged in analysis datasets.

11.5 Study Duration and Treatment Duration

Study duration of VR intervention participation to the study is defined as:

- Study Duration of VR intervention participation (days) = (Date/time of Last Visit – Date/time of Informed Consent) + 1.
- Study Duration of VR intervention participation (weeks) = ((Date/time of Last Visit – Date/time of Informed Consent) + 1) / 7.

Treatment duration of VR intervention participation to the study is defined as:

- Treatment duration of VR intervention participation (days) = (Date/time of last required VR session – Date/time of first required VR session) + 1.
- Treatment duration of VR intervention participation (weeks) = ((Date/time of last required VR session – Date/time of first required VR session) + 1) / 7.

11.6 Discontinued Subjects

Discontinued subjects will have assessments at the ET visit within 5 days of the last exposure to the investigational intervention and will be summarized along with Visit 7 Week 8.

11.7 Replacement Subjects

Additional subjects may be enrolled as replacement to early discontinuation subjects into the study to ensure adequate data for evaluation. For such subjects, summaries will include data like any other subjects.

11.8 Rescreened Subjects

Subjects who fail screening may be rescreened into the study if they meet the eligibility criteria. For such subjects, summaries will include the latest data reported.

11.9 Common Derivations

The following variables will be derived in analyses datasets and used in summaries as per the formula below:

- Age (years) = Year of Informed consent – Year of Birth

These variables will be derived and verified against EDC calculated values. In cases where they do not match, queries will be raised. If not reconciled until database lock, derived values will be considered correct and will be used in TLFs.

11.10 Baseline

Baseline is defined as the latest non-missing measurement prior to first video-guided VR onboarding session on week 1 day 1 including repeat assessments and unscheduled visits.

For CEQ scores, baseline is defined as first non-missing measurement taken after the first self-guided VR session and before week 1 day 7.

11.10.1 Change from Baseline

Change between baseline and post-baseline result will be calculated as below.

- Change = Result at Visit X – Baseline Result

11.11 Study Day

Study Day 1 will be the day the first BVR-100 session. Study Day will be calculated as:

For events that occurred on the day of or after first BVR-100 session:

- Study Day = visit date – date of first BVR-100 session +1

For events that occurred on the day before first BVR-100 session:

- Study Day = visit date – date of first BVR-100 session

11.12 Protocol Deviations

Protocol deviations will be classified as Important and Non-Important. Important protocol deviations will be used for summarization.

11.13 Medical / Psychiatric History

If a condition has end date prior to first exposure to the investigative intervention or if the Ongoing status = No, then these will be flagged as Medical / Psychiatric History.

11.14 Prior and Concomitant Medications

If a medication ended prior to first use of BVR-100/BES-100, then it is considered as prior medication.

If the medication ended on or after the first use of BVR-100/BES-100 or is ongoing, it is considered as concomitant medication.

11.15 Adverse Events

Any Adverse Event (AE) is defined as any effect that occurs or worsens after the first VR session until EOS will be summarized. Any events that occur after the time of signing the Informed Consent Form (ICF) and prior to the first VR session are termed as pre-treatment events.

Related Adverse Events include Possibly, Probably and Definitely Related Adverse Events.

UADEs are serious adverse effects caused by, or associated with a device which were not previously anticipated to occur. UADEs collected in CRF data will be summarized.

11.16 Columbia-Suicide Severity Rating Scale (C-SSRS)

The following outcomes are C-SSRS categories. The categories have been re-ordered from the actual scale to facilitate the definition of composite endpoints:

Category 1	Wish to be Dead
Category 2	Non-Specific Active Suicidal Thoughts
Category 3	Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act
Category 4	Active Suicidal Ideation with Some Intent to Act, without Specific Plan
Category 5	Active Suicidal Ideation with Specific Plan and Intent
Category 6	Actual Attempt
Category 7	Has subject engaged in Non-Suicidal Self-Injurious Behavior?
Category 8	Interrupted Attempt
Category 9	Aborted Attempt
Category 10	Preparatory Acts or Behavior

Suicidal Ideation – A “yes” answer at a visit to any one of the 5 suicidal ideation questions (categories 1-5) on the C-SSRS.

Suicidal Behavior – A “yes” answer at a visit to any one of the 5 suicidal behavior questions (categories 6-10) on the C-SSRS.

Suicidal Ideation or Suicidal Behavior – A “yes” answer at a visit to any one of the questions on the C-SSRS.

It is possible for subjects to have neither suicidal ideation nor suicidal behavior, or even to have both suicidal ideation and suicidal behavior.

11.17 Endpoints

11.17.1 CEQ Scores

The CEQ asks about the improvements that subjects believe will be achieved as a result of treatment, and how believable, convincing, and logical the treatment seems. It contains 6 items (each item is rated on a 1-9 scale). Total CEQ score is calculated for each subject at each visit by taking the sum of score of all 6 items of the scale ranging from 6 to 54. Higher score corresponds to better credibility and expectancy.

CEQ scores will be assessed at Baseline, Week 6 and Week 8. Based on the 6 items, two factors will be derived: the credibility factor and the expectancy factor.

Credibility Score

Credibility score is calculated for each subject at each visit by taking the sum of score of the first 3 items of the scale.

The items which are considered for calculating credibility score are:

- At this point, how logical does the treatment offered to you seem?
- At this point, how useful do you think the treatment will be in reducing your symptoms of social anxiety?
- How confident would you be in recommending this treatment to a friend who experiences similar problems?
- Expectancy Score

  Expectancy score is calculated for each subject at each visit by taking the sum of score of the last 3 items of the scale.

  The items which are considered for calculating expectancy score are:

- By the end of the therapy period, how much improvement in your symptoms of social anxiety do you think will occur?
- At this point, how much do you really feel that therapy will help you to reduce your symptoms of social anxiety?
- By the end of the therapy period, how much improvement in your symptoms of social anxiety do you really feel will occur?

11.17.2 Subject Retention Rate

Subject retention rate will be derived at Week 6 and Week 8/EOS for each of treatment groups.

$$Subject\ Retention\ Rate\ at\ Week\ 6\ or\ Week\ 8/EOS =$$

$$\frac{Number\ of\ subjects\ on\ the\ study\ at\ Week\ 6\ or\ Week\ 8/EOS}{Total\ number\ of\ subjects\ in\ that\ treatment\ group} * 100$$

Note: Subject who do not perform Week 6 but perform Week 8 will still be counted as subjects remaining on the study at Week 6.

11.17.3 Session-wise Time-on-Task

Task time is defined as the required video task time for BES-100 and required elements task time for BVR-100. Time-on-task calculation in each session for BVR-100 and BES-100 is given below:

Week	Session	BVR-100	BES-100
Week 1	Session 1	Required Journey Map Tutorial/SC Practice/Mindfulness Tutorial Task Time	Required Video Task Time + Required Campfire Task Time
	Session 2	Required Intro Cog Reframing Task Time + Required Cog Reframing Tutorial Task Time	Required Video Task Time
	Session 3	Required SC Practice Task Time	The same as Session 2
Week 2 - 8	Session 4 - 24	The same as Session 3	The same as Session 2

11.17.4 Average Time-on-Task

Average time on task at Week 6 and Week 8/EOS (and at other weeks as applicable) is calculated as the sum of the task time (calculated as mentioned in section 11.17.3) from each required session completed, divided by the number of required completed sessions at Week 6 and Week 8/EOS.

$$Average\ Time-on-Task\ (minutes) = \frac{Sum\ of\ the\ task\ time\ from\ each\ required\ VR\ session\ completed\ at\ Week\ X}{Total\ number\ of\ completed\ required\ VR\ sessions}$$

X is 1, 2, …, 8.

Overall average Time-on-Task per session across the entire study will be calculated as

$$Overall\ Average\ Time-on-Task\ (minutes) = \frac{Sum\ of\ the\ task\ time\ from\ each\ required\ VR\ session\ completed}{Total\ number\ of\ completed\ required\ VR\ sessions}$$

11.17.5 Total Number of Active Days

Total number of active days across the study is the total number of days that subjects have at least VR session. If a subject has 3 sessions each week and 1 session per day, total number of active days is 24. Repeated session in the same day does not count.

11.17.6 LSAS Score and Subscores

The LSAS is a clinician administered instrument that assesses both anxiety and avoidance across a number of social situations. The scale consists of 24 items each depicting different social situations. The anxiety scale ratings range from 0 (no anxiety) to 3 (severe anxiety). The avoidance ratings also range from 0 (never avoid) to 3 (usually avoid). In addition to the anxiety and avoidance subscales, the

LSAS is further divided into two subscales for scoring, including social interaction (11 items) and performance situations (13 items).

Items dealing with socializing are:

- Talking to people in authority
- Going to a party
- Calling someone you don't know very well.
- Talking with people you don't know very well.
- Meeting strangers
- Being the center of attention
- Expressing a disagreement or disapproval to people you don't know very well.
- Looking at people you don't know very well in the eyes.
- Returning good to a store
- Giving a party
- Resisting a high pressure salesperson

Items dealing with performance are:

- Telephoning in public
- Participating in small groups
- Eating in public places
- Drinking with others in public places
- Acting, performing or giving a talk in front of an audience
- Working while being observed
- Writing while being observed
- Urinating in a public bathroom
- Entering a room when others are already seated.
- Speaking up at a meeting
- Taking a test
- Giving a report to a group

- Trying to pick up someone.

An LSAS total score along with six additional scores based on anxiety and avoidance is derived at Baseline, Week 2, Week 4, Week 6 and Week 8: total fear, fear of social interaction, fear of performance situations, total avoidance, avoidance of social interaction, and avoidance of performance situations.

11.17.6.1 LSAS Total Score

Total LSAS score is calculated for each subject at each visit by taking the sum of both anxiety and avoidance score of all 24 items of the scale. Scores may range from 0 to 144.

11.17.6.2 Total Fear

Total fear is evaluated using total fear subscore which is calculated for each subject at each visit by taking the sum of fear scores of all 24 items. Scores range from 0 to 72.

11.17.6.3 Fear of Social Interaction

Fear of social interaction is evaluated using fear of social interaction subscore which is calculated for each subject at each visit by taking the sum of fear scores of the items dealing with socializing. Scores range from 0 to 33.

11.17.6.4 Fear of Performance Situations

Fear of performance is evaluated using fear of performance subscore which is calculated for each subject at each visit by taking the sum of fear scores of the items dealing with performance. Scores range from 0 to 39.

11.17.6.5 Total Avoidance

Total avoidance is evaluated using total avoidance subscore which is calculated for each subject at each visit by taking the sum of avoidance scores of all 24 items. Scores range from 0 to 72.

11.17.6.6 Avoidance of Social Interaction

Avoidance of social interaction is evaluated using avoidance of social interaction subscore which is calculated for each subject at each visit by taking the sum of avoidance scores of the items dealing with socializing. Scores range from 0 to 33.

11.17.6.7 Avoidance of Performance Situations

Avoidance of performance situation is evaluated using avoidance of performance situation subscore which is calculated for each subject at each visit by taking the sum of avoidance scores of the items dealing with performance. Scores range from 0 to 39.

11.17.7 Total PHQ-9 Score

The PHQ-9 is the depression module, which scores each of the 9 DSM criteria as “0” (not at all) to “3” (nearly every day). Total PHQ-9 score will be assessed at Baseline, Week 2, Week 4, Week 6 and Week 8.

Total PHQ-9 score is calculated for each subject at each visit by taking the sum of score of all 9 items of the scale. Scores may range from 0 to 27.

11.17.8 Responder Rate

Responder rate based on the CGI-I score is the proportion of responders at Week 8. Responders are subjects who rated 1 (very much improvement) or 2 (much improvement) in CGI-I at Week 8.

$$Responder\ Rate = \frac{Number\ of\ responders\ at\ Week\ 8\ ina\ treatment\ group}{Total\ number\ of\ subjects\ in\ that\ treatmnet\ group\ in\ mITT}$$

11.18 Self-Evident Corrections

In some instances, spelling errors may be corrected in programs for 'Other, Specify' fields (free text). These will be reconciled and confirmed with sponsor before database lock.

12 Statistical Analyses

Planned or Actual treatment groups will use the following conventions:

- Enrolled Population - Planned treatment group
- FAS Population – Planned treatment group
- Safety Population – Actual treatment group
- ITT Population – Actual treatment group
- Per-protocol Population – Actual treatment group

12.1 Subject Disposition

The number and percentage of subjects in screened, Enrolled Population, FAS population, Safety Population, Per-Protocol population and ITT population will be presented by treatment groups and total. Percentages for FAS population is based on Enrolled Population and percentages for all the other populations will be based on FAS population.

Number and percentage of subjects who completed the study, number and percentage of subjects who failed screening along with their reasons, number and percentage of subjects who terminated the study early along with their reasons will be summarized separately by treatment group and total. Reasons for screen failure will be displayed as in CRF. Percentages will be based on all screened subjects who signed ICF. Reason for discontinuation will be displayed in the descending order of "Total" column and percentages will be based on FAS population.

A flow diagram of subject disposition (CONSORT flow diagram) illustrating the progress of subjects throughout the study from screening to study completion/discontinuation will be provided.

Note that Enrolled Population and FAS population will be displayed by planned treatment and all other populations will be displayed by actual treatment.

Subject disposition information, status of inclusion exclusion criteria and subjects excluded from ITT analysis will be listed by subject who are screened.

Above analysis details are specified in DDP Table 14.1.1.1, Table 14.1.1.2, Figure 14.1.1.1, Listing 16.2.1.1, Listing 16.2.1.2 and Listing 16.2.3.1

12.2 Important Protocol Deviations

Number and percentage of subjects with any important protocol deviations and number of important protocol deviation events will be summarized by treatment groups and total based on FAS population. Important protocol deviation categories will be displayed in descending order of the "Total" column.

All the protocol deviations will be listed by subject who enrolled in the study.

Above analysis details are specified in DDP Table 14.1.2.1 and Listing 16.2.2.1.

12.3 Demographic and Baseline Clinical Characteristics

The following summaries will be presented by treatment groups and total based on Intent-to-Treat Population and Per-Protocol Population (as per ICH E3 Guidelines), unless otherwise specified.

Demographic characteristics and clinical characteristics at baseline will be summarized as follows:

- Demographic characteristics:
    - Age (years): summary statistics
    - Age categories
    - <18 years, ≥18 to ≤40 years, >40 to ≤65 years, >65 years
    - Sex assigned at birth: Male, Female
    - Ethnicity: Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown
    - Race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Other, Not Reported
- Clinical characteristics

The results collected at the baseline, as per section 11.10, will be considered as the baseline clinical characteristics. Following variables will be presented:

- CEQ score
- Credibility score
- Expectancy sore
- Total LSAS score
- CGI-S score
- PHQ-9 Score

Subject demographics will be listed by all subject who are screened. Baseline Clinical Characteristics will be listed by subject who are enrolled.

Above analysis details are specified in DDP Table 14.1.3.1, Table 14.1.3.2, Listing 16.2.4.1, Listing 16.2.6.1, Listing 16.2.6.3.1, Listing 16.2.6.3.2, Listing 16.2.6.4, Listing 16.2.6.5.

No statistical testing will be carried out for demographic or baseline clinical characteristics.

12.4 Psychiatric and Medical History

Number and percentage of subjects with any psychiatric or medical history and number of events will be summarized by treatment groups and total based on safety population. The same will be classified by System Organ Class (SOC) and Preferred Term (PT). The SOC and PT will be displayed in the descending order of frequency in "Total". Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 26.0 or Higher.

The summary will be presented by following categories:

- Comorbid psychiatric conditions
- General Medical History separately

Above analysis details are specified in DDP Table 14.1.5.1.

Psychiatric details collected for subject who are screened will be listed in DDP Listing 16.2.4.2.1

Psychiatric and Medical History will be listed by subject and body system based on Safety Population. List details are specified in DDP: Listing 16.2.4.2.2

12.5 Prior and Concomitant Medications

Number and percentage of subjects who have taken any prior / concomitant medication and number of medications will be summarized by treatment groups and total based on safety population. The same will be presented by Therapeutic Main group (ATC Level 3) and Preferred Term. The therapeutic main group and preferred term will be displayed in the descending order of frequency in "Total". Medications will be coded using the World Health Organization's Drug-Dictionary (WHODrug) version Mar-2023 or later.

The summary will be presented for prior and concomitant medications separately.

Above analysis details are specified in DDP: Table 14.1.6.1 and Table 14.1.6.2

All medications will be listed by subject and therapeutic subgroup.

Above analysis details are specified in DDP: Listing 16.2.4.3

12.6 Psychotropic Medications

Number and percentage of subjects who have taken any psychotropic medication 12 hours prior to a visit and number of medications will be summarized by treatment groups and total based on safety population. Only psychotropic medication taken 12 hours prior to a visit will be summarized. The same will be presented by Therapeutic Main group (ATC Level 3) and Preferred Term. The therapeutic main group and preferred term will be displayed in the descending order of frequency in "Total".

Medications will be coded using the World Health Organization's Drug-Dictionary (WHODrug) version Mar-2023 or latest.

All medications will be listed by subject and therapeutic subgroup on Safety Population. Above analysis details are specified in DDP Table 14.1.6.3 and Listing 16.2.4.4

12.7 Study and Treatment Duration

The subject will be encouraged to engage in minimum three VR sessions per week, for a period of 8 weeks. The following analysis will be summarized based on the Safety Population.

Study duration and treatment duration of VR intervention participation to the study will be summarized as a continuous variable as well as in categories by treatment groups. The categories for durations are >0 – 2, >2 – 4, >4 – 6, and >6 – 8 weeks.

Above analysis details are specified in DDP Table 14.1.7.1.

12.8 Efficacy Analyses

12.8.1 Primary Efficacy Analysis

12.8.1.1 Descriptive Summaries

The following summaries will be presented by treatment groups based on ITT and PP population. For Credibility, Expectancy, Total CEQ Score separately, descriptive summary of observed and change from baseline results for expectancy and credibility score will be summarized by visit (Baseline, Week 6 and Week 8/EOS). Also, descriptive summary of the total CEQ score and score for each item will be summarized based on treatment groups and visit. Mean Credibility and Expectancy scores will be plotted using line plots by visit. All listings are based on enrolled population.

12.8.1.2 Inferential Summaries

The independent t-test or Mann-Whitney U test will be used to check the significant difference between treatment groups. Exact p-value obtained will be presented in summary tables.

Normality test

Shapiro-Wilk test will be performed to test the normality of the two treatment groups and p-value will be interpreted. The test also assumes homogeneity of variances between the two treatment groups. Either one of the following tests will be performed:

- If Shapiro Wilk Test p-value > 0.05 then parametric T-Test will be used
- If Shapiro Wilk Test p-value < 0.05 then Mann-Whitney test will be used

Independent t-tests

Independent t-test will be used to compare mean scores (Observed and change from baseline scores) between treatment groups by visit.

Assumptions

- The samples are independent of one another.
- The sample is normally distributed within each of the two treatment groups.
- The variance should be homogeneous in the two treatment groups.

Reference SAS Program:

T-Test

```
proc ttest data = <INDATA>;
  class <TREATMENT>;
  var AVAL;
run;
```

Mann-Whitney test

```
proc npar1way data=<INDATA> wilcoxon;
    class <TREATMENT>;
    var AVAL;
run;
```

Hypothesis:

H_0: Mean of parameter is the same in two treatment groups.

vs.

H_1: Mean of parameter is not the same in two treatment groups.

Supportive analysis of the same will be presented for ITT and Per-Protocol population in a separate table.

Above analysis details are specified in DDP: Table 14.2.1.1, Table 14.2.1.2, Table 14.2.1.3, Table 14.2.1.4, Table 14.2.2.1, Table 14.2.2.2, Table 14.2.2.3, Table 14.2.2.4, Figure 14.2.2.5.

All CEQ scores will be listed in Listing 16.2.6.1

12.8.2 Secondary Efficacy Analyses

Secondary efficacy analyses will be presented for ITT population.

12.8.2.1 Subject Retention Rate

Descriptive summary of subject retention rate will be summarized for Week 6 and Week 8 by treatment groups. The analysis details are specified in DDP Table 14.2.3.1.

12.8.2.2 Time-on-Task

Total number of active days, total number of sessions, overall average Time-on-Task (minutes) will be summarized by treatment groups. Weekly average Time-on-Task (minutes) will be presented by treatment group and week. Above analysis details are specified in DDP: Table 14.2.3.2, Listing 16.2.6.2.

12.9 Exploratory Analyses

12.9.1 Descriptive summaries

Descriptive summary of change from baseline results for total LSAS score and subscores, CGI-S score and PHQ-9 score, and the observed scores will be summarized by treatment groups and week based on ITT population. Observed scores for CGI-I score and PGI-I score will be summarized by treatment groups and week based on ITT population. Descriptive summary of the number of required sessions completed during the study will be summarized by treatment groups based on ITT population. All scores will be listed based on all Enrolled Population

For LSAS summary parameters to be summarized are: LSAS Total Score, Fear of performance situations, Avoidance of performance situations, Fear of social interaction, Avoidance of social interaction, Total fear and Total avoidance. Average of LSAS score will be plotted using line plot by treatment group and week. All listings are based on the Enrolled Population.

12.9.2 Inferential Summaries

The independent t-test or Mann-Whitney U test will be used to check the significant difference between treatment groups. Exact p-value obtained will be presented in summary tables. Normality test will be performed based on the methods as described in section 12.8.1.2.

Above analysis details are specified in DDP: Table 14.2.4.1, Table 14.2.4.2, Table 14.2.4.3, Table 14.2.5.1, Table 14.2.5.2, Figure 14.2.4.1.

Fisher's Exact Test for Independence of Responders

Since sample size of the responders is expected to be small, the association of responders between the treatment groups will be investigated using Fisher's exact test. Fisher's exact test will be used to compare the proportion of respondents between the treatment groups.

Descriptive statistics and p-values from the test will be provided.

Assumptions

- Variables are categorical.
- All observations are independent.

Reference SAS Program:

```
proc freq data = <INDATA>;
  tables <TESTVAR1>*<TESTVAR2> / fisher;
run;
```

Hypothesis:

H_0: There is no association between treatment group and responders rate.

vs.

H_1: There is association between treatment group and responders rate.

Above analysis details are specified in DDP Table 14.2.6.1, Listing 16.2.6.3.1, Listing 16.2.6.3.2, Listing 16.2.6.4, Listing 16.2.6.5, Listing 16.2.6.6, Listing 16.2.6.7

12.10 Safety Analyses

12.10.1 Adverse Events

All Adverse events will be coded using the MedDRA version 26.0 or Higher. All summaries will be summarized by treatment groups and total based on safety population. Total will not be displayed in the table.

12.10.2 Overall Adverse Events

Number and percentage of subjects and number of events will be summarized by treatment groups and total for the following categories:

- Adverse Events
- Serious adverse events
- Related adverse events
- Related serious adverse events
- Unanticipated adverse device effects
- Adverse events led to death.
- Adverse events led to study discontinuation.
- Adverse events by severity
- Adverse events by relationship

All adverse events will be listed by subject and system organ class. Adverse events will be grouped by subjects. AEs leading to death and serious AEs will be listed separately.

Above analysis details are specified in DDP Table 14.3.1.1, Listing 16.2.7.1, Table 14.3.2.1 and Table 14.3.2.2

12.10.3 Summary of Adverse Events by SOC and PT

Number and percentage of subjects and number of events will be summarized by treatment groups and total, system organ class and preferred terms. Percentages will be based on safety population. Adverse events are sorted by the MedDRA Internationally Agreed Order of SOC and descending frequency of preferred term in Total. Total will not be displayed in the table. Preferred terms with the same frequency are sorted alphabetically. The same will be repeated for serious AEs, related AEs, related serious AEs, AEs leading to death and AEs leading to study discontinuation.

Above analysis details are specified in DDP: Table 14.3.1.2, Table 14.3.1.3, Table 14.3.1.4, Table 14.3.1.5, Table 14.3.1.6 and Table 14.3.1.7

12.10.4 Summary of Adverse Events by Relationship and Severity

Number and percentage of subjects in each system organ class, preferred term will be summarized by relationship and severity for adverse events and treatment group. System organ class and preferred term will be displayed by descending order of frequency in "Total". Total will not be displayed in the table.

For summaries of relationship,

- AEs will be grouped as "related" or "not related." AEs assessed as "possible," "probable," or "definite," will be grouped as "related."
- If a subject had multiple occurrences of the same preferred term within a system organ class, then it is counted once with closest relationship under the preferred term within the system organ class.
- If a subject had different preferred terms within the system organ class with different relationship, it is counted once within closest relationship within the system organ class and "Any AE" categories.

For summaries of severity,

- If a subject has multiple occurrences of the same preferred term within a system organ class, then it is counted once with maximum severity under the preferred term within the system organ class.
- If a subject had different preferred terms within the system organ class with different grades of severity, it is counted once within maximum severity grade within the system organ class and "Any AE" categories.

Above analysis details are specified in DDP: Table 14.3.1.8, Table 14.3.1.9

12.10.5 Columbia- Suicide Severity Rating Scale (C-SSRS)

Number and percentage of subjects without suicidal ideation or suicidal behavior, with suicidal ideation, with suicidal behavior, and with suicidal behavior or ideation will be presented by treatment groups based on ITT Population. All C-SSRS data will be listed based on all Enrolled population.

Above analysis details are specified in DDP Table 14.3.3.1 and Listing 16.2.8.1.

13 Changes to Planned Analyses

1 Protocol (section 12.3) specifies 'Baseline demographic and clinical characteristics will be compared between groups with the use of the two-sample t-test and the Mann–Whitney test for continuous measurements with and without normal distribution, respectively'. But the statistical testing for baseline characteristics is being waived as per the recent discussion with the sponsor.

2 Protocol (section 12.9) mentions an analysis about differences in the mean change from CGI-S baseline score in the CGI-I and PGI-I at Weeks 2, 4, 6, and 8. This analysis is waived as per the recent discussion with sponsor.

3 An additional exploratory endpoint number 7 regarding the number of required BVR-100 and BES-100 sessions completed, is added in the section 7.4 as per the decision by sponsor clinical team.

4 Sentence "to ensure a reliable estimate of the standard deviations to power a future trial with 90% and an expected effect size between 0.3 and 0.7" in section 4.2 Sample Size Justification is removed because it is for future study not for this current study.

5 Per-protocol population definition in section 10.5 is updated to consider only ITT population as per the recent discussion with sponsor.

6 Baseline clinical characteristics is updated to follow the baseline definition in section 11.10 as per the recent discussion with sponsor.

7 In section 10.4, mITT (Modified Intent-to-Treat) population is removed from analysis and ITT population (efficacy population) has been added as per sponsor request. This population will be used for primary and secondary endpoint analysis.
Reason to change [9]: Per dataset with data cutoff date January 19, 2024 (Data used for Blinded Data Review Meeting), only 24 out of 52 subjects in the Safety Population met the mITT population definition in the CSP (Clinical Study Protocol) due to missing CEQ data. The reasons for missing CEQ data are a 48-hours visit window restriction for CEQ assessment and only 2 CEQ assessments post baseline. To rectify this issue and to ensure all efficacy subjects are included in the final analysis, the team decided to modify the mITT definition in the CSP to ITT which includes all randomized subjects who have at least one VR session. With this change, the ITT population will include all 52 subjects.

14 Index of Tables, Listings and Graphs

Refer BVR100-102_Data Display Plan (Ver:1.00) for the list of Tables, Listings and Graphs.

15 References

1 McNeil DW, Randall CL. Conceptualizing and describing social anxiety and its disorders. In S. G. Hofmann & P. M. DiBartolo (Eds.), Social Anxiety (pp. 3–26). San Diego, CA: Elsevier; 2014.

2 American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5 (R)) (5th ed.). Arlington, TX: American Psychiatric Association Publishing; 2013

3 Mayo-Wilson E, Dias S, Mavranezouli I, Kew K, Clark DM, Ades AE, Pilling S. Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. The Lancet – Psychiatry. 2014;1(5):368–376.

4 Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society. 2007;10(5):722–724.

5 Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. The Lancet – Psychiatry. 2016;3(2):100–102.

6 ICH. ICH Harmonized Tripartite Guideline: Statistical Principles for Clinical Trials E9. 1998.

7 ASA. Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics. 1999.

8 ICH. ICH Harmonized Tripartite Guideline: Structure and Content of Clinical Study Reports E3. 1995.

9 Changes_to_Planned_Analyses_Form BVR100-102_final_30Jan2024.

16 Appendices

Summary of Statistical Analyses

Endpoints	Analysis Population	Analyses
Primary Endpoint		
Credibility / Expectancy Questionnaire (CEQ) scores	ITT Population	Primary: Descriptive summary, Independent t-test
	Per-Protocol Population	Supportive: Descriptive summary, Independent t-test
Secondary Endpoints		
Subject Retention Rate	ITT Population	Descriptive summary
Time-on-task		Descriptive summary
Exploratory Endpoints		
LSAS Overall & Subscores	ITT Population	Descriptive summary, Independent t-test
Responder Rate		Descriptive summary, Fisher's exact test
CGI-S Score		Descriptive summary
CGI-I Score		Descriptive summary
PGI-I Score		Descriptive summary
PHQ-9 Score		Descriptive summary
Safety Endpoints		
Adverse Events	Safety Population	Descriptive summaries on overall AE, relationship, severity etc.
C-SSRS		Descriptive summary

17 Change Log

Version	Author	Change Date	Description	Reviewer	Review Date
1.00	[redacted]	31-Jan-2024	First Approved Version	[redacted]	31-Jan-2024