STATISTICAL ANALYSIS PLAN

Protocol: HB-001

SAFETY AND EFFICACY OF HB-1 FOR PANIC DISORDER: A MULTICENTER, RANDOMIZED, DOUBLE BLINDED, PLACEBO CONTROLLED TRIAL

SPONSOR: Honeybrains Biotech
515 W 20th St, Ste 4e
New York, NY 10011

PRODUCT: HB-1

AUTHORS: Kavita Gohil, Ph.D., Jeffrey Finman, Ph.D.

DATE: 08 August 2022

STATUS: Version 1.2

Statistical Analysis Plan

Protocol: HB-001

Version: 1.2 Date: 08 August 2022

CRC REPRESENTATIVE DOCUMENT REVIEWED AND APPROVED BY:	
SIGNATURE:	DATE: 18-Aug-2022
NAME: Jeffrey Finman, PhD	
DEPARTMENT/TITLE: Chief of Biostatistics, Cognitive Research Corporation	

CRC REPRESENTATIVE DOCUMENT REVIEWED AND APPROVED BY:	
SIGNATURE:	DATE: 18 Aug 2022
NAME: Gary Kay, PhD	
DEPARTMENT/TITLE: Chief Scientific Officer, Cognitive Research Corporation	

STATISTICAL CRC REPRESENTATIVE DOCUMENT REVIEWED AND APPROVED BY:	
SIGNATURE:	DATE: 18-Aug-2022
NAME: Kavita Gohil, PhD	
DEPARTMENT/TITLE: Lead Statistician, Cognitive Research Corporation	

SPONSOR COMPANY REPRESENTATIVE DOCUMENT REVIEWED AND APPROVED BY:	
SIGNATURE: Dr Alon Seifan	DATE: 8/18/2022
NAME: Alon Seifan, MD, MS	
DEPARTMENT/TITLE: Chief Medical Officer, Honeybrains Biotech	

1 TABLE OF CONTENTS

2 AMENDMENT HISTORY

Not applicable.

3 LIST OF ABBREVIATIONS

Abbreviation	Definition
ADHD	Attention Deficit Hyperactivity Disorder
AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
BMI	Body Mass Index
CAARS	Conners' Adult ADHD Rating Scale
CBC	Complete Blood Count
CGI-S	Clinical Global Impression-Severity Scale
CI	Confidence Interval
CMP	Comprehensive Metabolic Panel
CS	Clinically Significant
C-SSRS	Columbia Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events
DBP	Diastolic Blood Pressure
DSM-5	Diagnostic and Statistical Manual of Mental Disorders 5th edition
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EOS	End of Study
FDA	Food and Drug Administration
GCP	Good Clinical Practice
GLP	Good Laboratory Practice
HAM-A	Hamilton Anxiety Scale
ICF	Informed Consent Form
ICH	International Conference on Harmonization
ITT	Intent-to-Treat
LS	Least Squares

Abbreviation	Definition
LSAS	Leibowitz Social Anxiety Scale
MADRS	Montgomery-Asberg Depression Rating Scale
MAR	Missing at Random
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Affairs
MIDAS	Migraine Disability Assessment Questionnaire
MINI	Mini International Neuropsychiatric Interview for DSM-5
MMRM	Mixed Model Repeated Measures
MNAR	Missing Not at Random
MPV	Mean Platelet Volume
NCS	Not Clinically Significant
PCRS	Placebo-Control Reminder Script
PDSS	Panic Disorder Symptom Severity Scale
PE	Physical Examination
PI	Principal Investigator
PSQI	Pittsburgh Sleep Quality Index
PT	Preferred Term
PTSD	Post-Traumatic Stress Disorder
QD	Once a Day
QHS	Nightly at Bedtime
Q-LES-Q-SF	Quality of Life Enjoyment and Satisfaction Scale Questionnaire – Short Form
RDW	Red Cell Distribution
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SDC	CogScreen Symbol Digit Coding

Abbreviation	Definition
SF-MPQ	Short Form McGill Pain Questionnaire
SOC	System Organ Class
STD	Standard Deviation
TAS-20	Toronto Alexithymia Scale
TEAE	Treatment-Emergent Adverse Event
TLFs	Tables, Figures, and Listings
ULN	Upper Lower Normal
WOCBP	Women of Childbearing Potential

4 INTRODUCTION

This statistical analysis plan (SAP) describes the planned statistical analysis to be conducted for the study entitled “Safety and Efficacy of HB-1 for Panic Disorder: A Multicenter, Randomized, Double-Blind, Placebo-Controlled Trial” version 4.0, dated 02 December, 2021. Mock shells will be produced as a separate working document to facilitate programming of Tables, Figures, and Listings (TFLs) according to SAP. The SAP is to be interpreted in conjunction with the protocol.

5 STUDY OBJECTIVES

5.1 PRIMARY OBJECTIVE

To determine the safety and efficacy of HB-1 versus placebo in male and female adult patients aged 18 to 60 years, inclusive, with Panic Disorder.

5.2 EXPLORATORY OBJECTIVE

To explore the effectiveness of HB-1 on Common Neuropsychiatric Co-Morbidities.

6 STUDY DESIGN

6.1 DURATION OF STUDY

The study will consist of 12-week treatment period and a one-week post- treatment period for safety follow-up.

6.2 NUMBER OF PARTICIPANTS

Approximately 80 patients with Panic Disorder will be enrolled and randomized (using central randomization) in a 1:1 ratio of active treatment to control at approximately 8-10 clinical research sites in the United States.

6.3 DESIGN

This is a multicenter, randomized, double-blind, placebo-controlled trial. All patients with Panic Disorder, with or without specified co-morbidities, who meet all of the inclusion and none of the exclusion criteria will be eligible. Patients and researchers will be blinded to their treatment group. Approximately 80 patients who meet DSM-5 criteria for the diagnosis of Panic Disorder will be recruited at approximately 8-10 clinical research sites.

HB is a single fixed dose formulation of Verapamil (288 mg) and Telmisartan (96 mg) (extended-release formula) developed for the treatment of Panic Disorders. Patients meeting all inclusion criteria and none of the exclusion criteria, will receive an oral dose of HB-1 or matching placebo once a day (QD) nightly at bedtime (QHS) with or without food. Once enrolled, each patient will be assigned a unique subject identification number. This number will be recorded on the patient’s electronic case report form (eCRF) pages and used to identify the patient throughout the study.

An Informed Consent Form (ICF) must be signed by prospective patients prior to initiating any study-specific procedures. Inclusion and exclusion criteria will be reviewed for each potential

patient. Eligibility assessments will include the Mini International Neuropsychiatric Interview for DSM-5 (MINI) and Panic Disorder Symptom Severity Scale (PDSS). Eligibility will be documented in the electronic case report form (eCRF). Complete medical history will be obtained, including demographics, neurological and behavioral health information. At Screening, concomitant medications will be recorded with height, weight, and BMI. Physical exams and orthostatic vital signs checks will be conducted at the time points specified in the Schedule of Assessments (Table 1). A decline of ≥20mm Hg in systolic or ≥10 mm Hg in diastolic blood pressure after 3 minutes of standing indicates orthostatic hypotension. Clinically significant orthostatic hypotension should be treated per standard of care and the Medical Monitor should be contacted.

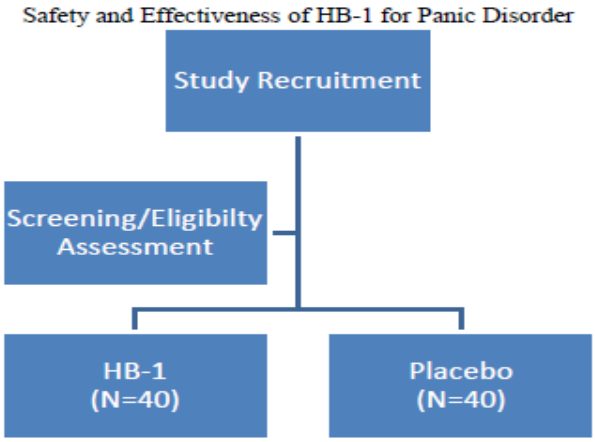
Any abnormal change from the initial screening physical examination (PE) must be assessed by the investigator as not clinically significant (NCS) or clinically significant (CS) and recorded in the source document and electronic case report form (eCRF).

Any clinically significant (CS) change after Visit 2, as determined by the investigator, will be recorded as an adverse event (AE) in source documentation and on the AE eCRF.

At Visit 1/Screening, Visit 4/Week 4, Visit 5/Week 8 and Visit 6/Week 12 a standard 12-lead Electrocardiogram (ECG) will be conducted. following an approximate 10-minute rest period in all patients and will be interpreted by the Investigator to assess clinical significance.

A Weber Test will be conducted at Baseline/Visit 2 and Visit 6/Week 12. The Weber Test is a test of lateralization of auditory function. Weber Test results will be documented on the eCRF.

6.4 STUDY SCHEMA



6.5 SCHEDULE OF EVENTS

Table 1: Schedule of Assessments

	Screening	Baseline	Treatment Period				Safety Follow-up
Procedure or Assessment	**Visit 1**	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6/ EOS Visit[k]**	**Safety Follow-up[l]**
Study Week	*Week - 1*	*Week 0*	*Week 2*	*Week 4*	*Week 8*	*Week 12*	*Week 13*
Visit Windows	*V2 within 7 days of V1*		*+/- 3 days*				
Informed Consent	X						
Medical History and Demographics[a]	X						
Eligibility Assessment	X						
Number of Panic Attacks[b]	X	X	X	X	X	X	
Columbia Suicide Severity Rating Scale[c]	X	X	X	X	X	X	
Clinical Laboratory Assessments[d]	X	X	X			X	
Physical Exam	X			X	X	X	
Height/Weight/BMI	X						
Urine Pregnancy Test[e]	X	X	X	X	X	X	
Screening Questionnaires (3)[f]	X						
Orthostatic Vitals[g]	X	X	X	X	X	X	
12-Lead ECG	X			X	X	X	
CogScreen Symbol Digit Coding Test [o]	X						
Weber Test		X				X	
Placebo-Control Reminder Script (PCRS)[n]			X	X	X	X	
Efficacy Endpoint Questionnaires (5)[h]		X	X[m]	X	X	X	
Exploratory Endpoint Questionnaires (8)[i]		X				X	
Study Treatment Dispensation/Return/ Accountability [j]		X	X	X	X	X	
Concomitant Medications	X	X	X	X	X	X	
Adverse Event Monitoring		X	X	X	X	X	X
Safety Follow-up Phone Call							X[l]

[a] Demographics will include: Age, Sex, Race, Ethnicity, Education, Gender Identity, Lateral Preference Questions, and Parental Education

[b] Number of Panic Attacks. The clinician will ask the patient "How many panic attacks have you had in the last week?" and will confirm the reported panic attack(s) meet DSM-5 criteria (APA 2013).

[c] Columbia Suicide Severity Rating Scale (C-SSRS) (Oquendo MA 2003). At screening, use the Screening/Baseline C-SSRS; at subsequent visits, use the Since Last Visit C-SSRS.

[d] Clinical laboratory assessments will include a comprehensive metabolic panel, CBC including Vitamin B12, urinalysis.

[e] Urine pregnancy test to be administered for women of child-bearing potential (WOCBP) within 7 days prior to Week 0, at Baseline/Week 0, and at each subsequent study visit

[f] Screening Questionnaires will include: Mini International Neuropsychiatric Interview for DSM-5 (MINI) (Sheehan, DV. 1998; Sheehan, DV. 1992-2016), the Clinical Global Impression-Severity Scale (CGI-S) (Guy W. 1976), and the Panic Disorder Symptom Severity Scale (PDSS) (Furukawa TA. 2009).

[g] Orthostatic vital signs should be obtained per Section 6.1.7 in the protocol.

[h] Efficacy Endpoint Questionnaires will include:

- Panic Disorder Symptom Severity Scale (PDSS) (not repeated at Week 0) (Furukawa TA. 2009)
- Clinical Global Impression-Severity Scale (CGI-S) (not repeated at Week 0) (Guy W. 1976)
- Quality of Life Enjoyment and Satisfaction Questionnaire Short Form (Q-LES-Q-SF) (Endicott J. 1993)
- Self-Report Leibowitz Anxiety Scale (LSAS) (Heimberg RG. 1999, Baker SL. 2002)
- Assessment of Overall Sleep Quality -- Question #6 from the Pittsburgh Sleep Quality Index (Buysse DJ. 1989)

[i] Exploratory Endpoint Questionnaires will include:

- Toronto Alexithymia Scale (TAS-20) (Bagby RM. 2020)
- Hamilton Anxiety Rating Scale (HAM-A) (Bech P. 1990)
- CogScreen Symbol Digit Coding test (SDC) (Kay G. 1995)
- Montgomery-Asberg Depression Rating Scale (MADRS) (Fantino B., Moore N. 2009) only for patients with a medical history of depression
- Conners' Adult ADHD Rating Scale (CAARS) (Walls BD. 2017) only for patients with a medical history of ADHD
- The PTSD Checklist for DSM-5 (Blake DD. 1995, Blevins CA. 2015) only for patients with a medical history of PTSD
- Short Form McGill Pain Questionnaire (SF-MPQ) (Melzack R. 1987) only for patients with a medical history of chronic pain
- Migraine Disability Assessment Questionnaire (MIDAS) (Lipton RB. 2001) only for patients with a medical history of migraine

[j] Patients will be randomly assigned to either the HB-1 or placebo treatments groups and will be provided with study drug each month while on treatment. Patients will be instructed to take the study drug once a day (QD) nightly at bedtime (QHS) with or without food, at the same time each day with a glass of water.

[k] Patients who discontinue from the study prior to Week 12: End of study visit (EOS) and all EOS assessments should be performed within 7 days following the discontinuation from the study. Patients who discontinue study treatment prior to Week 12: Patients may discontinue study treatment and remain in the study. Complete the EOS study visit and all EOS assessments within 7 days of study treatment discontinuation. The patient should continue to follow the schedule of assessments through the Safety Follow-Up Phone Call with modified assessments at interim visits including Week 12/Visit 6 as follows:

- Number of Panic Attacks
- C-SSRS (Oquendo MA 2003)
- Efficacy Endpoint Questionnaires
- Exploratory Endpoint Questionnaires (when applicable)
- Concomitant Medications
- Adverse Event Monitoring

[l] Safety Follow-up Phone Call is performed approximately 7 days after Visit 6/EOS.

[m] Only the PDSS will be repeated at the Week 2 visit.

[n] Placebo-Control Reminder Script (PCRS) (Cohen, EA. 2021). The PCRS will be administered immediately prior to administering the PDSS at Visits 3-6.

[o] CogScreen Symbol Digit Coding Test (SDC) (Kay G. 1995); SDC to be done twice at Screening/Visit 1 and no passing score is required.

6.6 TREATMENT

Once enrolled, each patient will be assigned a unique subject identification number. This number will be recorded on the patient's eCRF pages and used to identify the patient throughout the study. Once a subject number is assigned, it cannot be reassigned to any other patient.

6.7 RANDOMIZATION

All patients will either be assigned to receive HB-1 or placebo using a central randomization system. Patients and Investigators will be blinded to the study treatment assignment.

Study treatment will be dispensed to patients in blinded bottles. Bottle 1 will be dispensed at Visit 2/Week 0; patient will bring it back for accountability at Visit 3/Week 2 and will take it back home; patient will return Bottle 1 at Visit 4/Week 4 and will be dispensed Bottle 2 (with return of Bottle 2 at Visit 5/Week 8); then Bottle 3 will be dispensed at Visit 5/Week 8 (and returned at Visit 6/Week 12).

6.8 ASSESSMENTS

Assessments and questionnaires included are:

- Mini International Neuropsychiatric Interview for DSM-5 (MINI) at screening
- Placebo-Control Reminder Script (PCRS) at Visits 3-6
- Panic Disorder Symptom Severity Scale (PDSS)
- Number of Panic Attacks at Visits 1-6
- Assessment of Overall Sleep Quality at Visit 2 and Visits 4-6
- Columbia Suicide Severity Rating Scale (C-SSRS) at screening and each post screening visit including Baseline
- Clinical Global Impression-Severity Scale (CGI-S)
- Quality of Life Enjoyment and Satisfaction Questionnaire Short Form (Q-LES-Q-SF)
- Leibowitz Social Anxiety Scale (LSAS)
- Toronto Alexithymia Scale (TAS-20)
- Hamilton Anxiety Rating Scale (HAM-A)
- Montgomery-Asberg Depression Rating Scale (MADRS)
- CogScreen Symbol Digit Coding test (SDC)
- Conners' Adult Attention Deficit Hyperactivity Disorder (ADHD) Rating Scale (CAARS)
- Post-Traumatic Stress Disorder (PTSD) Checklist for DSM-5
- Short Form McGill Pain Questionnaire (SF-MPQ)
- Migraine Disability Assessment Questionnaire (MIDAS)

6.9 EFFICACY

6.9.1 PRIMARY AND SECONDARY EFFICACY ENDPOINTS

- Change in Panic Disorder Symptom Severity Scale (PDSS)- primary efficacy endpoint.
- Number of panic attacks
- Change in Clinical Global Impression-Severity Scale (CGI-S)
- Change in Quality-of-Life Enjoyment and Satisfaction Questionnaire Short Form (Q-LES-Q-SF)
- Change in Leibowitz Social Anxiety Scale (LSAS)
- Change in Assessment of Overall Sleep Quality -- Question #6 from the Pittsburgh Sleep Quality Index

6.9.2 EXPLORATORY ENDPOINTS

- Change in Toronto Alexithymia Scale (TAS-20)
- Change in Hamilton Anxiety Rating Scale (HAM-A)
- Change in Montgomery-Asberg Depression Rating Scale (MADRS) only for patients with a medical history of depression
- Change in CogScreen Symbol Digit Coding test (SDC)
- Change in Conners' Adult ADHD Rating Scale (CAARS) only for patients with a medical history of ADHD
- Change in the PTSD Checklist for DSM-5 only for patients with a medical history of PTSD
- Change in Short Form McGill Pain Questionnaire (SF-MPQ) only for patients with a medical history of chronic pain
- Change in Migraine Disability Assessment Questionnaire (MIDAS) only for patients with a medical history of migraine

6.10 SAFETY

6.10.1 SAFETY ENDPOINTS

Incidence of Treatment Emergent Adverse Events (TEAEs), clinically significant changes in clinical laboratory values, changes in electrocardiogram (ECG), and Columbia Suicide Severity Rating Scale (C-SSRS), and orthostatic vital signs.

6.10.2 ADVERSE EVENTS

All AEs should be captured on the eCRF from Visit 2/Week 0 through the Safety Follow-up visit. Adverse events considered at least possibly related to study treatment should be followed until resolution, return to baseline, or deemed chronic or stable.

All serious adverse events (SAEs) will be immediately reported to the Sponsor from the time of Visit 2/Week 0 (dosing) through the Safety Follow-up/End of Study Visit.

6.10.3 LABORATORY ASSESSMENTS

The following laboratory parameters will be measured at the time points specified in the Schedule of Assessments and will be analyzed locally:

- Pregnancy test should be administered for women of child-bearing potential (WOCBP) within 7 days prior to Week 0 and then at each subsequent study visit.
- Comprehensive Metabolic Panel (CMP): albumin, blood urea nitrogen, calcium, carbon dioxide, chloride, creatinine, creatinine phosphokinase, glucose, potassium, sodium, total bilirubin, total protein, alanine aminotransferase (ALT), alkaline phosphatase (ALP), aspartate aminotransferase (AST)
- Complete Blood Count (CBC): white blood cell count, red blood cell count, hematocrit, hemoglobin, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red cell distribution (RDW), platelet count, mean platelet volume (MPV), Vitamin B12
- Urinalysis: microscopic examination, specific gravity, pH, protein, glucose, ketones, blood and urobilinogen

6.11 STOPPING CRITERIA

Table 2: Subject Specific Stopping Criteria

Cardiac Disorders	
Hypotension	• SBP <90 mmHg and DBP <60 mmHg for three consecutive readings 10 minutes apart
Orthostatic Hypotension	• Decline of >20mm Hg in systolic or >10 mm Hg in diastolic blood pressure after 3 minutes of standing.
New-onset 2nd or 3rd degree AV block	• Grade 1 or higher per CTCAE V5.0*
Sick sinus syndrome	• Grade 1 or higher per CTCAE V5.0
Atrial flutter	• Grade 1 or higher per CTCAE V5.0
Atrial fibrillation	• Grade 1 or higher per CTCAE V5.0
New-onset bradycardia	• Resting heart rate below 50
New-onset tachycardia	• Resting heart rate above 120
Onset of angina pectoris	• Grade 1 or higher per CTCAE V5.0
Congestive heart failure	• Grade 1 or higher per CTCAE V5.0
Pulmonary edema	• Grade 1 or higher per CTCAE V5.0
Peripheral edema	• Grade 1 or higher per CTCAE V5.0
Pulmonary edema	• Grade 2 or higher per CTCAE V5.0

General Disorders	
Urticaria	• Grade 2 or higher per CTCAE V5.0
Hyperkalemia	• Potassium values that are 2X the ULN for the local lab normal refences ranges. • Values may be repeated and confirmed.
Gout exacerbation	• Hyperuricemia Grade 1 or higher per CTCAE V5.0 • Values may be repeated and confirmed.
Impaired hepatic function	• ALT, AST, ALP values that are twice the upper or lower limit of local lab normal reference range. • Values may be repeated and confirmed.
Impaired renal function	• Doubling in creatinine value. • Values may be repeated and confirmed.
Thrombocytopenia	• Platelet count values below 150,000/mL • Values may be repeated and confirmed.

*Common Terminology Criteria for Adverse Events (CTCAE) V5 dated 27Nov2017

7 STATISTICAL ANALYSES

7.1 GENERAL CONSIDERATIONS

All analysis dataset preparations and statistical analyses will be performed using SAS® version 9.4 or higher. Study data will be provided in listings and sorted by patient number, study period, and assessment time. The study results will be displayed using descriptive summary tables and inferential analyses completed as described below. Summaries will be displayed by treatment group for all tables. For measurements that have multiple intermediate visits, results for observed data will be summarized for each time point.

Standard numeric descriptive statistics include number of patients or records observed, mean, standard deviation, median, minimum, and maximum values. Standard categorical descriptive statistics include the count and percentages of patients.

7.2 STUDY DAY AND VISIT WINDOW DEFINITIONS

The baseline assessment from which to include baseline in the analysis model and for the calculation of change scores for all efficacy endpoints is defined as last assessment prior to the first dose of study treatment.

Data obtained during unscheduled and early discontinuation visits will be allocated to the scheduled visit corresponding to the visit window in which they fall in as specified in Table 3**.**

Data will be analyzed based on the nominal visits and nominal time points. If the data from the nominal visit or time point is missing, data from unscheduled visits or an early discontinuation visit for the same nominal visit or time point will be used.

If multiple visits among unscheduled or early termination assessments fall in the same visit window or time point, the non-missing assessment closest to target time point will be selected for analysis. If multiple values are the same number of days away from the target study day, then the latter value will be used. In the unlikely event an unscheduled or early discontinuation visit, associated with a particular visit window, falls either prior to the actual previous nominal visit date or after the subsequent nominal visit date, it will not be used.

The first date on which subject receives study treatment is defined as Study Day 1. Study days for other visits will be calculated as follows:

Before Study Day 1 visit: Study Day = date of assessment – date of Study Day 1.

On or after Study Day 1 visit: Study Day = date of assessment – date of Study Day 1 + 1.

The target study days of Visits are summarized below.

Table 3: Time Windows for Assessments (unscheduled and early discontinuation visits)

Scheduled Visit Number	Nominal Visit (label)	Time Window (day)	Target Time Point (day)
1	Screening	≤ -1	≤ -1
2	Visit 2 (Baseline/Day 1)	1	1
3	Visit 3 (Week 2)	2 to 21	14
4	Visit 4 (Week 4)	22 to 42	28
5	Visit 5 (Week 8)	43 to 70	56
6	Visit 6 (Week 12)	71 to < Safety Follow-up Visit	84

7.3 ANALYSIS POPULATIONS

Analyses will be performed comparing treatment groups for each of the following populations:

- All Patients – all patients who sign informed consent and are assigned a unique patient identification number.
- Intent-to-Treat (ITT) – all patients who are successfully randomized to either the HB-1 or control treatment groups. This will be the primary population used for efficacy analyses. Patients will be analyzed according to randomized treatment group.
- Safety Population – patients who are randomized and receive at least one dose of study drug. This will be the primary population used to summarize all safety data. Safety summaries will be based on the treatment that patients receive.

7.4 SAMPLE SIZE DETERMINATION

Assuming a standard deviation of 4.5 for the PDSS, 80 patients will provide more than 95% power to detect a difference in the mean change (from baseline) in the PDSS score of 3.75 at an

alpha-level of 0.05 at 12 weeks. This corresponds to a 50% reduction in mean PDSS scores on HB-1 and a 25% reduction in mean PDSS scores on placebo, assuming baseline mean PDSS scores of 15 for both groups.

In addition, with respect to adverse events, the study has approximately 95% power to detect at least one patient with an event that occurs in 7.2% of patients in either treatment group.

7.5 TREATMENT GROUPS

The analyses will be conducted by treatment group:

- HB-1
- Placebo

7.6 STATISTICAL HYPOTHESIS TESTS

The hypotheses for testing the primary analysis are:

$$\text{Ho: } \mu_a = \mu_p \text{ vs. Ha: } \mu_a \neq \mu_p$$

where μ_a is the mean change from baseline to Week 12 in the PDSS for the HB-1 arm and μ_p is the mean change from baseline to week 12 in the placebo arm.

Testing will be done at the 0.05 alpha-level. Any confidence intervals (CI) provided will be at 95% confidence.

7.7 MISSING DATA

Data may be missing either because an assessment is not completed or if patients have withdrawn consent from further participation in the study. To minimize missing data, participants who may discontinue treatment, but do not withdraw consent from further participation are expected to continue in the study, with primary and key secondary efficacy assessments conducted.

The approach to handling missing data for the primary endpoint and key secondary endpoints is dependent upon the extent of missing data and potential reasons for missingness. Therefore, prior to database lock, an assessment of missing data will be conducted to determine potential reasons for missingness, where missing data will be considered missing not at random (MNAR), unless sufficient evidence is available to reasonably conclude the data is missing at random (MAR).

The mixed model analysis is considered valid under a missing at random assumption, however alternative approaches such as multiple imputation using within-treatment group imputation, control-based imputation and a tipping point analysis with a within-treatment imputation will be considered if a sufficient amount of missing data are present and considered potentially MNAR.

7.8 MULTIPLICITY ADJUSTMENTS

To control for multiplicity of select analyses, Hochberg's procedure will be used to test the protocol designated primary endpoint (PDSS) as well as the key secondary endpoint CGI-S,

which are not expected to be negatively correlated. Accordingly, if the largest p-value among these two endpoints (at the primary time point) does not exceed 0.05, then both endpoints will be considered statistically significant. If the largest p-value exceeds 0.05 and the smallest p-value does not exceed 0.025, then the endpoint with the smaller p-values will be considered statistically significant. This provides control of the familywise error rate at $\alpha = 0.05$ such that either of the 2 endpoints statistically significant according to the above procedure will constitute a successful efficacy outcome for the trial.

Irrespective of the significance of these specific measures, planning for future studies will be based on the totality of safety and efficacy data and nominal p-values will be reported for all inferential tests.

7.9 INTERIM ANALYSIS

Interim analyses are not planned for this study.

7.10 ENDPOINTS AND ESTIMANDS

Endpoints are based on measurement scales collected over the visits of the study. Each of the scales may have domains or subdomains. The treatment policy strategy is planned for all intercurrent events.

7.11 EFFICACY AND EXPLORATORY ANALYSES

7.11.1 PRIMARY AND SECONDARY EFFICACY ENDPOINTS

- Panic Disorder Symptom Severity Scale (PDSS)- primary efficacy endpoint
- Number of panic attacks – key secondary endpoint
- Clinical Global Impression-Severity Scale (CGI-S) – key secondary endpoint
- Quality-of-Life Enjoyment and Satisfaction Questionnaire Short Form (Q-LES-Q-SF)
- Leibowitz Social Anxiety Scale (LSAS)
- Assessment of Overall Sleep Quality -- Question #6 from the Pittsburgh Sleep Quality Index

The primary analysis of efficacy will be performed using the ITT population. The change from baseline to in PDSS will be compared between HB-1 and placebo using Mixed Model Repeated Measures (MMRM) model using SAS mixed procedure with repeated measures. The model will include the fixed effect of treatment, visit, the treatment-by-visit interaction, baseline, baseline-by-visit. The outcome will be the change from baseline to each on-treatment time point. All protocol-scheduled time points will be included in the model. The model will use the unstructured or Toeplitz (whichever has a better fit by AICC value) within-subject variance-covariance matrix. The denominator degrees of freedom will be according to the method of Kenward-Roger. The estimated least square (LS) means and their standard errors as well as the estimated treatment effect (differences between treatments at each time point) will be summarized by treatment. In addition, the 95% CIs of the LS means and the difference between treatments, p-values for treatment comparisons will be provided. The primary analysis

timepoint will be at 12 weeks. Secondary endpoints will be analyzed similarly. For endpoints without a baseline measurement, baseline and the visit-by-baseline interaction will be omitted from the model.

7.11.2 EXPLORATORY ENDPOINTS

Analyses will be conducted in the same manner as for the primary efficacy endpoint. In some cases, the analysis population will be restricted to only those patients with the underlying condition, as noted (e.g., MADRS, only for patients with a medical history of depression), The questionnaires which will be used for exploratory endpoints are referenced in Section 6.9.2.

Safety endpoints include TEAEs, clinical laboratory abnormalities, changes in ECG, and orthostatic vital signs between treatment groups. C-SSRS results will be provided in a listing. All summaries of safety and tolerability will be performed on data from the safety population using descriptive statistics only unless otherwise specified.

7.11.3 SUBGROUP ANALYSES

All efficacy analyses will be repeated for the subgroup of participants not concurrently being treated with benzodiazepines prior to enrollment.

7.12 SAFETY ANALYSES

7.12.1 ADVERSE EVENTS

All AEs will be coded to System Organ Class (SOC) and Preferred Term (PT) using the most up-to-date version of the Medical Dictionary for Regulatory Affairs (MedDRA) at the time of data reporting.

Pre-treatment adverse events are defined as AEs that started after signing the informed consent form up to the start of study medication. Pre-treatment AEs will be listed.

A TEAE is defined as any untoward medical event not present prior to the initiation of the study treatment, or any event already present which worsens either in intensity or frequency following the exposure to the study treatment.

Summaries of the number and percentage of patients in the safety population with at least one adverse event, classified according to preferred term and/or body system using the current version of the MedDRA will be provided for the following types of AEs:

- TEAEs
- Drug Related TEAEs
- SAEs
- Drug Related SAEs
- AEs by severity (mild, moderate, severe, life-threatening, death)
- AEs by relationship (not related, remote, possible, probable, definite)

7.12.2 CLINICAL LABORATORY EVALUATIONS

Clinical laboratory results will be collected at screening and at each visit specified in the Schedule of Assessments table. Clinical laboratory test results for pregnancy, comprehensive metabolic panel, complete blood count, and urinalysis, will be summarized by treatment group and presented descriptively. Descriptive summaries of the change from baseline will be provided along with 95% CIs within each group.

7.13 VITAL SIGNS

Vital signs check will be performed at the time points specified in the Schedule of Assessments. Vital signs parameters (including blood pressure and pulse rate after 3 minutes of lying flat in a supine position; followed by blood pressure and pulse rate after standing for 3 minutes) results will be summarized by treatment group and presented descriptively. Descriptive summaries of the change from baseline calculations will be provided along with 95% CIs for the mean change from baseline to detect potentially significant changes in vital sign values.

7.14 ELECTROCARDIOGRAMS

ECGs will be performed at the time points specified in the Schedule of Assessments. ECG results will be summarized by treatment group and presented in tabular and graphic formats where appropriate. ECG abnormalities will be summarized by timepoint and for post-baseline assessments, the number and percentage of patients with new abnormalities will be summarized.

7.15 PHYSICAL EXAMINATIONS

Physical examination findings will be listed for each patient at each assessment time point as specified in the Schedule of Assessments.

8 CHANGES FROM PROTOCOL

Hochberg's procedure is now used to control the familywise error rate for the primary endpoint and the key secondary endpoints of the number of panic attacks and the CGI-S. In addition, subgroup analyses of efficacy endpoints are added for participants who are not receiving concomitant benzodiazepines at baseline.

9 REFERENCES

1. Furukawa TA, Katherine Shear M, Barlow DH, Gorman JM, Woods SW, Money R, Etschel E, Engel RR, Leucht S. Evidence-Based Guidelines for Interpretation of the Panic Disorder Severity Scale. Depress Anxiety. 2009;26(10):922-9. doi: 10.1002/da.20532. PMID: 19006198

2. Guy, W. Clinical Global Impression, in ECDEU Assessment Manual for Psychopharmacology, W. Guy, Editor. 1976, National Institute of Mental Health: Rockville, MD. p. 218-21.

3. Bagby RM, Parker JDA, Taylor GJ. Twenty-five Years with the 20-item Toronto Alexithymia Scale. J Psychosom Res. 2020 Jan 23;131:109940. doi: 10.1016/j.jpsychores.2020.109940.

4. Heimberg RG, Horner KJ, Juster HR, Safren SA, Brown EJ, Schneier FR, Liebowitz MR. Psychometric properties of the Liebowitz Social Anxiety Scale. Psychol Med 1999;29(1):199-212.

5. Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, Keane TM. The development of a Clinician-Administered PTSD Scale. J Trauma Stress 1995;8(1):75-90.

6. Melzack R. The Short-Form McGill Pain Questionnaire. Pain 1987;30(2):191-197.

7. Lipton RB, Stewart WF, Sawyer J, Edmeads JG. Clinical Utility of an Instrument Assessing Migraine Disability: the Migraine Disability Assessment (MIDAS) questionnaire. Headache 2001;41(9):854-861.

8. Walls BD, Wallace ER, Brothers SL, Berry DT. Utility of the Conners' Adult ADHD Rating Scale Validity Scales in Identifying Simulated Attention-Deficit Hyperactivity Disorder and Random Responding. Psychol Assess 2017;29(12)1437-1446.

9. Baker SL, Heinrichs N, Kim HJ, Hofmann SG. The Liebowitz Social Anxiety Scale as a Self-report Instrument: A Preliminary Psychometric Analysis. Behav Res Ther. 2002 Jun;40(6):701-15. doi: 10.1016/s0005-7967(01)00060-2.

10. Blevins CA, Weathers FW, Davis MT, et al. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. Trauma Stress. 2015 Dec;28(6):489-98. doi: 10.1002/jts.22059. Epub 2015 Nov 25.

11. Endicott J., Nee J., Harrison W., Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: A New Measure. Psychopharmacol Bull. 1993;29(2):321-6.PMID: 8290681

12. Bech P. Psychometric developments of the Hamilton scales: the spectrum of depression, dysthymia, and anxiety. Psychopharmacol Ser. 1990;9:72-9. doi: 10.1007/978-3-642-75373-2_9.PMID: 2204923

13. Kay G. *Cogscreen Aeromedical Edition Professional Manual*. Psychological Assessment Resources, Inc.; Odessa, FL: 1995.

14. Fantino B., Moore N. The self-reported Montgomery-Asberg Depression Rating Scale is a useful evaluative tool in Major Depressive Disorder. BMC Psychiatry. 2009 May 27;9:26. doi: 10.1186/1471-244X-9-26.PMID: 19473506

15. American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5). American Psychiatric Publishing: Washington, DC.

16. Oquendo M. A., Halberstam B. & Mann J. J., Risk factors for suicidal behavior: utility and limitations of research instruments. In M.B. First [Ed.] Standardized Evaluation in Clinical Practice, pp. 103 -130, 2003.

17. Buysse D.J., Reynolds C.F., Monk T.H., Berman S.R., & Kupfer D.J. (1989). The Pittsburgh Sleep Quality Index (PSQI): A new instrument for psychiatric research and practice. Psychiatry Research, 28(2), 193-213.

18. Cohen E.A., Hassman H.H., Ereshefsky L. *et al.* Placebo response mitigation with a participant-focused psychoeducational procedure: a randomized, single-blind, all placebo study in major depressive and psychotic disorders. *Neuropsychopharmacol.* 46, 844–850

19. . Sheehan DV, Lecrubier Y, Harnett-Sheehan K, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar G: The Mini International Neuropsychiatric Interview (M.I.N.I.): The Development and Validation of a Structured Diagnostic Psychiatric Interview. *J. Clin Psychiatry*, 1998;59(suppl 20): 22-33. http://www.psychiatrist.com/JCP/article/Pages/1998/v59s20/v59s2005.aspx

20. Sheehan, D.V. (1992 - 2016). Mini International Psychiatric Interview, English Version 7.0.2 for DSM-5. Harm Research Institute.

21. Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 74, 800-2.