Statistical Analysis Plan

[redacted]

Sponsor	VistaGen Therapeutics, Inc.
Protocol Title:	A Phase 3 Open-label Safety Trial of PH94B Nasal Spray in the Acute Treatment of Anxiety in Adult Subjects with Social Anxiety Disorder (SAD)
Protocol Number:	PH94B-CL030
[redacted]	[redacted]
Document Version:	Draft Version 1.0
Document Date:	17-NOV-2022
NCT:	NCT05030350

Approvals

Role	Signatures
Biostatistician	Print Name: [redacted] Sign Name: [redacted]
Peer Reviewer	Print Name: [redacted] Sign Name: [redacted]

Role	Signatures
<Sponsor> Representative	Print Name: [redacted]
	Sign Name: [redacted]

[redacted]

Document History

Not applicable

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
AE	adverse event
ANCOVA	analysis of covariance
ASA	American Statistical Association
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
C-SSRS	Columbia-Suicide Severity Rating Scale
ccITT	Complete Case Intent-To-Treat
CGI-I	Clinical Global Impression – Improvement scale
CI	confidence interval
COVID-2019	coronovirus 2019
CSR	clinical study report
ECG	Electrocardiogram
eCRF	electronic case report form
EMA	European Medicines Agency
FCS	fully conditional specification
FDA	Food and Drug Administration
[redacted]	[redacted]
HAM-D	Hamilton Depression Scale

Abbreviation	Definition
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
IP	investigational product
ITT	Intent-to-Treat
IWRS	interactive web response system
LS	least squares
LSAS	Liebowitz Social Anxiety Scale
MAR	missing at random
MedDRA	Medical Dictionary for Regulatory Activities
MINI	Mini-International Neuropsychiatric Interview
PGI-C	Patient Global Impression of Change
Q1 / Q3	first quartile / third quartile
QTcF	QT interval corrected for heart rate using Fridericia's formula
RSS	Royal Statistical Society
SAD	social anxiety disorder
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SE	standard error
SOC	system organ class

Abbreviation	Definition
SUDS	Subjective Units of Distress Scale
TEAE	treatment emergent adverse event
WHO-DD	World Health Organization Drug Dictionary

1. Overview

This statistical analysis plan (SAP) describes the planned analysis and reporting for VistaGen Therapeutics, Inc. protocol number PH94B-CL030 (A Phase 3 Open-label Safety Trial of PH94B Nasal Spray in the Acute Treatment of Anxiety in Adult Subjects with Social Anxiety Disorder (SAD)), dated 14-Mar-2022 Version 2.0. Reference materials for this statistical plan include the protocol and the accompanying sample data collection documents. Operational aspects related to collection and timing of planned clinical assessments are not repeated in this SAP unless relevant to the planned analysis.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the Food and Drug Administration (FDA), European Medicines Agency (EMA), and International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use: Guidance on Statistical Principles in Clinical Trials (ICH, 1998). All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association (ASA, 2018) and the Royal Statistical Society (RSS, 2014), for statistical practice.

The planned analyses identified in this SAP may be included in clinical study reports (CSRs), regulatory submissions, or future manuscripts. Also, post-hoc exploratory analyses not necessarily identified in this SAP may be performed to further examine study data. Any post-hoc or unplanned, exploratory analysis performed will be clearly identified as such in the final CSR.

The statistical plan described hereafter is an *a priori* plan. It will be approved before any analysis of data pertaining to VistaGen Therapeutic Inc.'s study PH94B-CL030.

2. Study Objectives and Endpoints

2.1. Study Objectives

2.1.1. Primary Objective

The primary objective is to evaluate the safety and tolerability of repeated dosing of PH94B over a period of up to 12 months as assessed by documenting adverse events (AEs) and standard clinical measurements (physical examination, vital signs, clinical chemistry, hematology, suicidality, level of depression, and 12-lead electrocardiogram [ECG]).

2.1.2. [redacted]

[redacted]

- [redacted]

2.2. Study Endpoints

2.2.1. Safety Endpoints

The safety endpoints of this study include the following:

- Incidence and severity of AEs, AEs leading to discontinuation, and SAEs
- Changes in vital signs results
- Changes in clinical laboratory evaluation (hematology, chemistry, and urinalysis) results
- Changes in 12-lead ECG results
- Changes in physical examination findings
- Changes in HAM-D17 scores
- Changes in C-SSRS scores
- PWC-20 scores after termination of PH94B

2.2.2. Exploratory Efficacy Endpoints

The exploratory efficacy endpoints of this study include the following:

- Change in total LSAS scores from baseline (Visit 2) to end of treatment (Visit 14)
- Change in proportion of subjects with CGI-S scores >4 from baseline (Visit 2) to end of treatment (Visit 14)
- Change in proportion of subjects with CGI-I scores of 1(Very much improved) or 2 (Much improved) from first month (Visit 3) to end of treatment (Visit 14)
- Change in proportion of subjects with PGI-C scores of 1 (Very much improved) or 2 (Much improved) from first month (Visit 3) to end of treatment (Visit 14)
- [redacted]

3. Overall Study Design and Plan

Subject participation in the study will last up to approximately 12 to 13 months, depending on

the duration of the screening period and intervals between Visits. Upon signing the informed consent form (ICF), subjects will complete Visit 1 and enter a screening period lasting between 7 and 35 days. Subjects who meet all eligibility criteria will return to the clinic to complete Visit 2 (Baseline, Day 0) and undergo baseline measurements of clinical laboratory assessments (clinical chemistry, hematology, urinalysis, pregnancy test [if female and of childbearing potential], and urine drug screening), 12-lead ECG, vital signs, physical examination, Hamilton Rating Scale for Depression (HAM-D17), [redacted] Liebowitz Social Anxiety Scale (LSAS), Columbia-Suicide Severity Rating Scale (C-SSRS), and Clinical Global Impression Scale of Severity (CGI-S). At Visit 2 (Baseline), subjects will be trained on how to use and store the PH94B nasal spray delivery device, and how to use an electronic diary (eDiary) application; subjects will be instructed to record each incidence of PH94B use in the eDiary application. Adverse events (AEs) and concomitant medications will be reported and recorded. The subject will receive a [redacted] of PH94B [redacted] for use (up to 4 times per day) at the subject's discretion, as needed for social anxiety.

Subjects will be instructed to return to the clinic for Visit 3 after 30 days (±3 days). At Visit 3, subjects will undergo clinical laboratory assessments (clinical chemistry, hematology, urinalysis, pregnancy test, and urine drug screening), vital signs, 12-lead ECG, physical examination, HAM-D17, [redacted] LSAS, C-SSRS, CGI-S, CGI-I, and self-evaluation by the patient-reported outcome measure, Patient Global Impression of Change (PGI-C). At Visit 3, PH94B use recorded in the eDiary application will be reviewed and AEs and concomitant medications will be reported and recorded. Subjects will receive a 1-month supply of PH94B (6 vials).

Subjects will return to the clinic at monthly intervals (±7 days, Visits 4 to 13), and will repeat the same assessments as those performed at Visit 3. PH94B use recorded in the eDiary application will be reviewed, AEs and concomitant medications will be reported and recorded, and at each Visit subjects will receive a 1-month supply of PH94B (6 vials).

At Visit 14 (1 year ±2 weeks), treatment will conclude, and subjects will undergo clinical laboratory assessments (clinical chemistry, hematology, urinalysis, pregnancy test, and urine drug screening), vital signs, 12-lead ECG, physical examination, HAM-D17, [redacted] LSAS, C-SSRS CGI-S, CGI-I, and PGI-C. Also, PH94B use recorded in the eDiary application will be reviewed, and AEs and concomitant medications will be recorded and reported. Subjects will also be informed of their follow-up Visit (Visit 15).

Subjects will return to the clinic for a follow-up Visit (Visit 15, 54 weeks ±3 days), approximately 2 weeks after the last PH94B administration, to undergo C-SSRS and vital signs assessments, and record AEs.

Safety Considerations:

Safety and tolerability of PH94B (≤ 4 doses per day up to 12 months in up to 600 subjects) will be assessed and summarized through changes from baseline (Visit 2) to end of treatment (Visit 14) in AEs, laboratory values, 12-lead ECGs, physical examinations, and vital sign assessments following exposure to PH94B.

3.1. Overall Design

3.2. Sample Size and Power

Based on feedback from the FDA, the sample size for this study (up to 600 subjects) has been selected to attain a PH94B safety database per the ICH E1 Guideline.

3.3. Study Population

The study population consists of male or female subjects, 18 through 65 years of age, inclusive, with a current diagnosis of SAD wand with no prior or current conditions that would increase risk of treatment or confound the results of the study.

3.4. Treatments Administered

The treatment to be administered is PH94B nasal spray, one spray to each nostril for at total dose of 3.2µg, up to four times a day.

[redacted]

3.5. Schedule of Events

Table 1 Schedule of Events

	Study Periods					
	Visit 1	Visit 2	Visit 3	Visits 4 to 13	Visit 14 / ET	Visit 15
	Screening	Baseline	1st Month	Monthly	End of Treatment	Follow-up
Assessment/Evaluation	Day -35 to Day -7	Day 0	Day 30 ±3 days	±7 days	52 Weeks ±2 weeks	54 Weeks ± 3 days
Informed Consent Form[a]	X	X				
Subject Demographics	X					
Medical and Psychiatric History	X					
Record Concomitant Medication Use	X	X	X	X	X	
SAD Diagnostic Criteria Checklist	X					
Clinical Chemistry, Hematology, and Urinalysis	X	X	X	X	X	
Urine Drug Screen	X	X	X	X	X	
Pregnancy Test[e]	X	X	X	X	X	
12-Lead Electrocardiogram[d]	X	X	X	X	X	
Physical Examination[e]	X[f]	X	X	X	X	

[redacted]

Assessment/Evaluation	Study Periods					
	Visit 1	**Visit 2**	**Visit 3**	**Visits 4 to 13**	**Visit 14 / ET**	**Visit 15**
	Screening	**Baseline**	**1st Month**	**Monthly**	**End of Treatment**	**Follow-up**
	Day -35 to Day -7	**Day 0**	**Day 30 ±3 days**	**±7 days**	**52 Weeks ±2 weeks**	**54 Weeks ± 3 days**
Vital Signs	X	X	X	X	X	X
MINI 7.0.2	X					
LSAS	X	X	X	X	X	
HAM-D17[redacted]	X	X	X	X	X	
C-SSRS	X	X	X	X	X	X
CGI-S	X	X	X	X	X	
CGI-I			X	X	X	
PGI-C			X	X	X	
PWC-20					X	X
Training on Use of Spray Device		X				
Storage Instructions		X	X	X		
Training on eDiary App to Record PH94B Use		X				
Adverse Event Recording		X	X	X	X	X
Dispense PH94B to Subject		X	X	X		
eDiary Review			X	X	X	

Abbreviations: C-SSRS = Columbia-Suicide Severity Rating Scale; CGI-I = Clinical Global Impression Scale of Improvement; CGI-S = Clinical Global Impression Scale of Severity; eDiary = electronic diary; ET = Early Termination [redacted] HAM-D17 = Hamilton Rating Scale for Depression (17 items); LSAS = Liebowitz Social Anxiety Scale; MINI = Mini-International Neuropsychiatric Interview; PGI-C = Patient Global Impression of Change; PWC-20 = Penn Physician Withdrawal Checklist; SAD = social anxiety disorder

a Subjects entering the study having participated in a previous PH94B Phase 3 SAD study will sign an informed consent form for the present study at either Visit 1 or Visit 2, depending on when they enter the study (see **Error! Reference source not found.**).

b For subjects entering the study having participated in a previous PH94B Phase 3 SAD study, a review of inclusion and exclusion criteria will be conducted at either Visit 1 or Visit 2, depending on when they enter the study (see **Error! Reference source not found.**).

c Pregnancy tests will only be given to female subjects of childbearing potential and not post-menopausal females.

d 12-Lead electrocardiograms will be administered to subjects at the Screening and Baseline Visits and then every month during the study.

e Physical examination will include measurement of weight and an examination of the nasal passages. Height will be measured at Screening and at Visit 14.

f If total anosmia is suspected during review of the nasal passages at Visit 1, the Quick Olfactory Test (QOT) is to be administered.

Note: Long-term C-SSRS will be administered at screening. All other C-SSRS will be "Since Last Visit".

Note: If central laboratory urine drug screen testing is positive for a subject, then the participation of the subject in the study may have to be discontinued. The Medical Monitor should be consulted on these cases. An unscheduled repeat urine drug screen may be obtained to investigate whether drug use is continuing.

4. Statistical Analysis and Reporting

4.1. Introduction

Data processing, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will primarily use SAS (release 9.4 or higher). If the use of other software is warranted, the final statistical methodology report will detail what software was used for what purposes.

Continuous (quantitative) variable summaries will include the number of subjects (n) with non-missing values, mean, standard deviation (SD), minimum, and maximum.

Categorical (qualitative) variable summaries will include the frequency and percentage of subjects who are in the particular category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of subjects in the study population with non-missing values.

The minimum and maximum will be reported with the same degree of precision (ie, the same number of decimal places) as the observed data. Measures of location (mean and median) will be reported to 1 degree of precision more than the observed data and measures of spread (SD) will be reported to 2 degrees of precision more than the observed data.

Percentages <100% will be presented to 1 decimal place and percentages of 100% will be reported with no decimal place. Counts of zero will be presented without percentages.

Unless otherwise indicated, all statistical tests will be conducted at the 0.05 significance level using 2-tailed tests, and *P* values will be reported. Corresponding 95% confidence intervals (CIs) will be presented for statistical tests. The statistical tests will be descriptive with no adjustment for multiplicity.

This study is considered an exploratory study. Statistical tests will be interpreted in an exploratory sense only and will not be considered formal hypothesis tests.

4.2. Interim Analysis and Data Monitoring

No interim analyses are planned.

5. Analysis Populations

The following analysis populations are planned for this study:

- **Safety Population:** All subjects who are dispensed IP during Visit 2 (Baseline) will be included in the Safety population. All subjects will receive PH94B in an open label fashion. Adverse events (AEs) will be grouped by associated length of exposure and level of PH94B use.

Inclusion of subjects in the Safety Population will be determined prior to database lock.

6. General Issues for Statistical Analysis

6.1. Statistical Definitions and Algorithms

6.1.1. Baseline

The last observation recorded before the first dose of PH94B will be used as the baseline observation for all calculations of change from baseline.

Table 1: Baseline Definitions

Parameter	Baseline
Safety Variables (clinical laboratory assessments, vital signs, ECGs, physical examinations)	The last non-missing assessment before the first dose of study drug given at Visit 2 (Baseline).
C-SSRS Scores	Scores on the Baseline version of the C-SSRS given at Visit 1 (Screening).
LSAS total score, HAM-D total score[redacted]	Total score at Visit 2 (Baseline).

Abbreviations: CGI-I = Clinical Global Impression of Improvement; ECG = electrocardiogram; C-SSRS = Columbia-Suicide Severity Rating Scale; [redacted] HAM-D = Hamilton Depression Scale; LSAS = Liebowitz Social Anxiety Scale

6.1.2. Adjustments for Covariates

No adjustments will be made for covariates.

6.1.3. Multiple Comparisons

No adjustments will be made for multiple comparisons.

6.1.4. Handling of Dropouts or Missing Data

Only observed data will be included in the final report.

6.1.5. Analysis Visit Windows

Analysis Visits will be assigned based on the nominal Visit recorded in the study database.

6.1.6. Pooling of Sites

Sites will not be pooled for analysis.

6.1.7. Derived Variables

6.1.7.1. General

- **Age** = Integer difference in years between the subject's date of informed consent and date of birth.
- **Change from Baseline to Visit X in Parameter Y** = Value at Visit X minus the value at Baseline (as defined in Table 2).
- **Percent Change from Baseline to Visit X in Parameter Y** = Change from Baseline to Visit X in Parameter Y divided by the value at Baseline (as defined in Table 2) multiplied by 100.
- **Shift from Baseline to Visit X in Parameter Y Categories** = Cross-tabulation of below the lower limit of the normal range, within the limits of the normal range, and above the upper limit of the normal range at Baseline (as defined in Table 2) versus at Visit X.
- **Height (cm) and Weight (km)** = Height (cm) will be calculated as height (in) times 2.54. Weight (kg) will be calculated as weight (lb) times 0.4536.
- **Temperature (C°)** = Temperature (F°) -32 times 5 divided by 9.
- **Treatment duration** = Days in study (date of last Visit minus date of Visit 2 (Baseline).
- **Days of exposure** = Date of Visit 14 (End of Treatment) minus date of Visit 2. [In the case Visit 14 is missing then the date of study completion will be used.]
- **PH94B use** = average weight of all 6 vials when dispensed – weight of all 6 vials when returned at subsequent visit, at each visit.

6.1.7.1. Clinical Global Impression Scale of Improvement (CGI-I)

The CGI-I is a clinician-rated instrument that measures the clinician's evaluation of change in subjects' overall improvement with treatment on a scale where 1 = "Very much improved" and 7 = "Very much worse." In this study, the CGI-I will be completed by the study investigator at each monthly Visit to the clinic.

The derived variable from this assessment will be the following:

- **Proportion CGI-I Responders** = Proportion of subjects with a score of 1 (Very much improved) or 2 (Much improved) on the CGI-I.

6.1.7.2. Patient Global Impression of Change (PGI-C)

The PGI-C is a self-administered instrument that measures change in subjects' overall improvement with treatment on a scale where 1 = "Very much improved" and 7 = "Very much worse." Subjects will rate their impression of change at each monthly visit to the clinic.

The derived variable from this assessment will be the following:

- **Proportion PGI-C Responders** = Proportion of subjects with a score of 1 (Very much improved) or 2 (Much improved) on the PGI-C.

6.1.7.3. Liebowitz Social Anxiety Scale (LSAS)

The LSAS is a clinician-rated scale that has been shown to be sensitive to treatment-related change in social phobia symptoms. The time frame for rating symptomatology is the past week. Note, there (instead of these) are monthly Visits to the clinic. The scale consists of 24 items. Each item is given 2 ratings: fear or anxiety on a scale of 0 to 3 and avoidance on a scale of 0 to 3, with a total maximum overall score of 144. The total LSAS score for each subject is the sum of all ratings across both scales. The derived variable for this assessment is change from the LSAS baseline (Visit 2 [Baseline]) at each monthly Visit in total LSAS rating.

6.1.7.4. Hamilton Depression Scale (HAM-D)

At Visit 1 (Screening) and each monthly Visit, subjects will be asked to complete the clinician-rated HAM-D questionnaire to assess depression. The HAM-D questionnaire consists of 17 items, each scored on either a 3- or 5point scale.

The Hamilton Depression Scale (HAM-D) total score is calculated as the sum of the 17 individual item scores and ranges from 0 to 52.

6.1.7.5. [redacted]

[redacted] (HAM-A) total score is calculated as the sum of the 14 individual scores, ranging from 0 to 56.

6.1.8. Data Adjustments/Handling/Conventions

All collected data will be presented in listings. Data not subject to analysis according to this plan will not appear in any tables or graphs but will be included only in the data listings.

Medical history and AEs (all types) will be coded using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA), version 25.1. Prior and concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD) Anatomical Therapeutic Chemical (ATC) classes and preferred terms, version WHODrug B3 Global Sep 2021.

If partial dates occur, the convention for replacing missing dates for the purpose of statistical analysis is as follows: if just day is missing then the day assigned is the first day of the month or the date of first dose (if in the same month), whichever is later; if just month is missing then the month assigned is the month of the first dose, unless that results in a date before the first dose in

which case the month after the first dose is used; and if both month and day are missing then the month assigned is the month of the first dose and the day assigned is either the first day of the month or the first dose date, whichever is later.

If partial times occur, the convention is as follows: if the missing time occurs on the day of the first dose and both the hour and minute are missing then the time assigned is the time of the first dose, otherwise if both the hour and minute are missing and the date is not the date of first dose the time assigned is 12:00; if the date is the same as the date of the first dose and only hour is missing the hour assigned is 12 or the hour of first dose, whichever is later, and if the date is the same as the date of first dose and only the minute is missing the minute assigned is 30 or the minute of first dose, whichever is later. Otherwise the hour assigned is 12 if the hour is missing and the date is not the same as the date of first dose and the minute assigned is 30 is the the date is not the same as the date of first dose.

These conventions will be applied only to AE and medication onset dates and times with the following precaution: if the missing date and time reflect the date and time of onset of an AE, the modified date and time will be constructed to match the first documented date/time post drug administration while preserving the order in which the AE/medication was reported in the eCRF.

7. Study Subjects and Demographics

7.1. Disposition of Subjects and Withdrawals

The numbers of subjects enrolled, completing, and withdrawing, along with reasons for withdrawal, will be tabulated by duration of time in the study. Data will also be presented in a by-subject listing.

7.2. Protocol Violations and Deviations

Protocol deviations will be identified via the eCRFs and classified as important or non-important. Data will also be presented in a by-subject listing.

7.3. Demographics and Other Baseline Characteristics

Demographic variables will include race, ethnicity, age, age categories (18-35 years, 36-55 years, 56-65 years), sex, height, weight, and body mass index (BMI). Age will be calculated as an integer in years as the difference between the subject's date of informed consent and the date of birth. Demographic and baseline characteristic data will be summarized. Data will also be presented in a by-subject listing.

Prior medications will be summarized, by the number and percentage of subjects taking each medication and classified by ATC class and preferred term. This analysis will be conducted for the Safety population. Data will also be presented in a by-subject listing, which will include ATC Level 2 classification, preferred name, and verbatim name.

Medical and psychiatric history will be summarized for the Safety population classified by system organ class (SOC) and preferred term. Data will also be presented in a by-subject listing, which will include SOC, preferred term, and the verbatim term.

Results from the Mini-International Neuropsychiatric Inteview (MINI) will be presented in a by-subject listing.

7.4. Exposure

PH94B will be supplied in [redacted] Each spray delivers 1.6 µg per 100 µL spray. At Visit 2 (Baseline) and each monthly Visit thereafter, site personnel will instruct the subject to self-administer the IP by 1 spray into each nostril (right and left nasal passages), for 2 total sprays per dose up to 4 times per day.

Number of days of exposure to PH94B (date of Visit 14 (or completion date) minus date of Visit 2 (Baseline)) will be summarized overall. Average difference in vial weights at each visit (for all subjects completing each visit) will be summarized overall.

7.5. Exploratory Efficacy Analysis

All exploratory endpoints will be summarized for descriptive purposes only.

The following variables will be summarized for all subjects completing each visit. All derived variables will be computed as described in Section 0.

Categorical data will report frequency of responses and proportion of responders by study visit:

1. Proportion of CGI-I responders will be assessed from Visit 3 to Visit 14 by study visit.
2. Change in proportion of PGI-C responders will be assessed from Visit 3 to Visit 14 by study visit.
3. Proportion of CGI-S scores > 4 by from Visit 2 to Visit 14 by study visit.

Continuous data will report descriptive statistics of observed values and change from baseline by study visit:

4. Change in LSAS total scores by visit.
5. [redacted]

[redacted]

8. Safety and Tolerability Analysis

Safety in this study will be evaluated based on data collected on AEs, clinical laboratory assessments, vital signs, 12-lead ECGs, PWC-20 (at Visit 14/15 End of Study/Follow-up), and physical examinations including examination of the nasal passages. As a precautionary measure, subjects will also be assessed using the C-SSRS and HAM-D at each study Visit.

All safety analysis reporting will be based on the Safety Population (as defined in Section **Error! Reference source not found.**). No formal statistical testing will be conducted for the safety analyses. Descriptive statistics will be used to evaluate safety data. Summaries will present data by overall subjects. Listings will list safety data by subject.

8.1. Adverse Events

Treatment emergent AEs (TEAEs) are defined as:

- AEs with onset at the time of or following the start of treatment with IP through final Visit or follow-up phone call.
- AEs starting prior to the start of treatment with IP but increasing in severity or relationship at the time of, or following, the start of treatment with IP through Visit 14 (End of Treatment/Follow-Up).
- AEs that occur between Visit 14 and Visit 15.

Any medical condition that is present at the time that the subject is screened will be considered as baseline and not reported as an AE. However, if the study subject's condition deteriorates at any time during the study, it will be recorded as an AE.

The Investigator will record all reportable events with start dates occurring any time after informed consent is obtained until 7 days after the last day of study participation. Events will be followed for outcome information until resolution or stabilization. Missing and partially missing start dates will be imputed for the purpose of analysis as described in Section **Error! Reference source not found.**.

All AEs, TEAEs, and serious AEs (SAEs) will be coded using the MedDRA dictionary, as discussed in Section **Error! Reference source not found.**.

All AE data will be categorized as Severe > Moderate > Mild. Relationship of all AEs to IP will be categorized as related or not related. In the summaries showing severity and relationship to study drug, the event with the maximum severity or strongest relationship to study drug will be reported. If a particular event is missing the severity and/or relationship, then the strongest possible severity or relationship will be assumed for analysis (severity = severe, relationship = related).

An overall summary of TEAEs will be provided; this will present number and percent of subjects who reported at least 1 of the following: TEAE (including any TEAE, TEAEs by maximum severity, and related TEAEs), TEAE leading to discontinuation of the study drug, TEAE leading to discontinuation from the study, serious TEAEs, and SAEs leading to death. TEAEs will be summarized by two time periods (TEAEs which occur prior to Visit 14 and those that occur during or after Visit 14).

In addition to the overall summary, summaries of the number and percentage of subjects with TEAEs will be displayed by overall by SOC and preferred term by time period. Summaries of TEAEs by severity (Severe > Moderate > Mild) and relationship to IP (Related > Not Related)

will also be provided. Serious TEAEs and TEAEs resulting in discontinuation of IP will be summarized separately in a similar manner.

Subject listings of AEs, SAEs, and AEs causing discontinuation of IP will be produced. Adverse events that are treatment emergent will be flagged.

The AE summary tables will include counts of subjects. Therefore, if a subject experiences more than 1 episode of a particular AE, the subject will be counted only once for that event. If a subject has more than 1 AE that is coded to the same preferred term, the subject will be counted only once for that preferred term. Similarly, if a subject has more than 1 AE within a SOC, the subject will be counted only once in that SOC. This will be handled separately for each time period (post-Visit 2 through Visit 15 [End of Treatment/Follow-up]).

8.1.1. Adverse Events Leading to Withdrawal

A summary of incidence rates (frequencies and percentages) of TEAEs leading to discontinuation of study drug will be summarized by period (Visit 2 [Baseline] and Visit 14 [End of Treatment/Follow-Up]), and by SOC and preferred term for the Safety population. No inferential statistical tests will be performed.

A data listing of AEs leading to discontinuation of study drug will also be provided, displaying details of the event(s) captured on the eCRF.

8.1.2. Deaths and Serious Adverse Events

Any deaths that occur during the study will be listed.

Serious AEs will be listed. Serious TEAEs and serious TEAEs related to IP will be tabulated by SOC and preferred term and presented by period (post-consent and prior to IP administration at Visit 2 [Baseline] and post-Visit 2 through Visit 15 [End of Treatment/Follow-up).

8.2. Clinical Laboratory Evaluations

Descriptive summaries of observed values and change values will be presented for continuous hematology, chemistry, and urinalysis results for overall subjects at each study Visit. Ninety-five percent CIs for the change and percent change will be included. Categorical urinalysis results will be summarized using frequencies by overall subjects and by study Visit.

The number of subjects with clinical laboratory values categorized as below, within, or above normal ranges will be tabulated showing change from baseline (shift tables) for each clinical laboratory analyte by overall subjects and by study Visit.

Laboratory values that are outside the normal range will also be flagged in the data listings and presented with the corresponding normal ranges. Any out-of-range values that are identified by the investigator as being clinically significant will also be shown in a data listing.

For women only, pregnancy status will be determined by evaluation of urine pregnancy test.

Data on pregnancy status will be listed by subjects. Subjects who are pregnant are excluded from the study. Subjects who become pregnant during the study will be discontinued.

8.3. Vital Signs

Descriptive summaries of observed values and changes from baseline will be calculated for systolic blood pressure (mmHg), diastolic blood pressure (mmHg), heart rate (beats/min), respiratory rate (breaths/min), and temperature (°C) by overall subjects and by study Visit.

8.4. Electrocardiograms

The number and percentage of subjects with normal and abnormal 12-lead ECG findings will be summarized for overall subjects at each study Visit. Abnormal results will be grouped as clinically significant and not clinically significant.

Descriptive summaries for the observed results and change from baseline will be presented for 12-lead ECG measures of PR interval (msec), QRS interval (msec), and heart rate for overall subjects at each study Visit. If triplicate reads are taken, the mean of the reads will serve as the observed value at the specific Visit.

A comparison of QT and QTcF results will be presented. Summary statistics for baseline values at post-consent and prior to IP administration at Visit 2 [Baseline] and post-Visit 2 through final Visit/follow-up phone call will be displayed by overall subjects for QT interval (msec) and QT interval corrected for heart rate using Fridericia's formula (QTcF) (msec).

Fridericia's Correction (QTcf)

$$QTc_f = \frac{QT_{msec}}{\sqrt[3]{RR}}$$

8.5. Concomitant Medication

Concomitant medications will be summarized descriptively using counts and percentages for the Safety population. Medications will be summarized overall, by the number and percentage of subjects taking each medication and classified by ATC class and preferred term. Data will also be presented in a by-subject listing, which will include ATC Level 2 classification, preferred term and verbatim name.

Prior medications will be presented separately from concomitant medications. All medications taken in the month prior to Visit 1 (Screening) or in the time interval between Visit 1 (Screening) and Visit 2 (Baseline) will be considered prior medications, whether or not they were stopped before Visit 2 (Baseline). Any medications continuing or starting after Visit 2 (Baseline) will be considered concomitant. If a medication starts before Visit 2 (Baseline) and continues after Visit 2 (Baseline) it will be considered both prior and concomitant.

Medications will be coded using WHO-DD, as described in Section **Error! Reference source not found.**.

8.6. Columbia-Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is a semi-structured interview that captures the occurrence, severity, and frequency of suicide-related thoughts and behaviors during the assessment period. The interview includes definitions and suggested questions to solicit the type of information needed to determine if a suicide-related thought or behavior has occurred. The C-SSRS has a "baseline" version, which will be completed at Visit 1 (Screening), and a "since last Visit" version that will be completed at all subsequent Visits.

There are a maximum of 19 items to be completed: 7 that are required, 10 potential additional items if there is a positive response to a required item, and 2 items for suicide/suicide behavior present during the interview. The C-SSRS uses dichotomous scales (i.e., yes or no), Likert scales, and text or narrative to further describe the thoughts or behaviors.

The number and percentage of subjects reporting "Yes" for any of the 5 suicidal ideation questions (Categories 1 to 5) or any of the 5 suicidal behavior questions (Categories 6 to 10) will be displayed for overall subjects, along with the number and percentage reporting suicidal ideation OR behavior (Categories 1 to 10). Treatment-emergent suicidal ideation and behavior (an increase in maximum score during treatment) will also be presented.

8.7. Penn Physician Withdrawal Checklist

The PWC-20 consists of 20 questions related to withdrawal from benzodiazepines or other anti-anxiety compounds. The symptoms measured are based on those that are potentially related to anxiolytic withdrawal: gastrointestinal, mood, sleep, motor, somatic, perception, and cognition. The questions will cover any symptoms that might occur during the week following discontinuation of PH94B.

Summaries of PWC-20 total scores and frequency and percentage of subjects with withdrawal symptoms will be presented overall for only those subjects who completed both Visit 14 and Visit 15. Descriptive statistics for total scores and frequencies of withdrawal symptoms will be presented by Visit 1 and Visit 15.

8.8. Hamilton Depression Scale (HAM-D)

Total scores and change from baseline will be summarized for the Safety population.

8.9. Coronavirus 2019 (COVID-19) Impact

The SARS-CoV-2 virus and variants may still be a threat and safety precautions may still be in place during study enrollment. To assess the impact of the COVID-19 pandemic on study

procedures, study staff will complete the COVID-19 Impact questionnaire. Should a study Visit be affected by COVID-19, study staff will record how the Visit was affected (missed, abbreviated, delayed, performed remotely, or otherwise deviating from the planned format in the protocol) and which procedures were impacted. All COVID-19 protocol deviations will be provided in a separate listing. This listing will include the subjects, affected Visits, and affected assessments.

10. Changes from Planned Analysis

Endpoints removed: analysis not performed – data not collected.

Study staff will complete the Coronavirus 2019 (COVID-19) Impact questionnaire to assess the impact of the COVID-19 pandemic on study procedures. The COVID-19 Impact questionnaire is described in Section 8.9. All COVID-19 protocol deviations will be provided in a separate listing.

Exploratory objectives (changes in pattern of PH94B use daily frequency and dosing interval) requiring the use of eDiary data were removed due to unreliability of eDiary data.

The original planned interim safety analysis was removed from the study.

A qualitative review of categorical data was done instead of an assessment of a change in proportion over time due to limited availability of data.

11. References

ASA. (2018) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2018. http://www.amstat.org/about/ethicalguidelines.cfm

Carpenter JR and Kenward MG. *Multiple Imputation and its Application (Statistics in Practice)*. Chichester, West Sussex, UK: John Wiley & Sons; 2013.

Graham JW, and Olchowski AE, Gilreath TD. (2007) "How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory." *Prev Sci.* 2007;8(3): 206-213.

ICH (1998). ICH Harmonised Tripartite Guideline. Statistical Principles for Clinical Trials E9; 1998. https://database.ich.org/sites/default/files/E9_Guideline.pdf

Little R and Rubin D. (1987). Statistical Analysis with Missing Data, New York: Wiley.

RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014. https://rss.org.uk/about/policy-and-guidelines/code-of-conduct/.

12. Tables, Listings, and Figures

All listings, tables, and graphs will have a header showing the sponsor company name and protocol and a footer showing the version of SAS, the file name and path, and the source of the

data (e.g., listing number).

The following are planned summary tables for protocol number PH94B-CL032. The table numbers and page numbers are place holders only and will be determined when the tables are produced.

The following are planned summary tables for protocol number <Protocol>. The table numbers and page numbers are place holders only and will be determined when the tables are produced.

12.1. Demographic Data Summary Tables and Figures

Table 3: Demographic Data Summary Tables and Figures

Table Number	Population	Table Title/Summary
14.1 Demographic Data Summary Tables and Figures		
Table 14.1.1.1	All Subjects	Subject Enrollment and Disposition
Table 14.1.1.2	All Subjects	Protocol Deviations
Table 14.1.2.1	Safety	Demographics and Baseline Characteristics
Table 14.1.3	Safety	Medical and Psychiatric History by System Organ Class and Preferred Term
Table 14.1.4	Safety	Prior Medications by ATC Class and Preferred Term
Table 14.1.5.1	Safety	Exposure to Investigational Product
Table 14.1.5.2	Safety	PH94B Use by Study Visit

12.2. Efficacy Data

12.3. Safety Data

Table 2: Safety Data

Table Number	Population	Table Title / Summary
14.3.1 Displays of Adverse Events		
Table 14.3.1.1	Safety	Overall Summary of Treatment Emergent Adverse Events
Table 14.3.1.2.1	Safety	Treatment Emergent Adverse Events by System Organ Class and Preferred Term
Table 14.3.1.2.2	Safety	Treatment Emergent Adverse Events by Preferred Term
Table 14.3.1.3	Safety	Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Severity

Table Number	Population	Table Title / Summary
Table 14.3.1.4	Safety	Treatment Emergent Adverse Events by System Organ Class, Preferred Term, and Maximum Relationship to Investigational Product
14.3.2 Summary of Deaths, Other Serious and Significant Adverse Events		
Table 14.3.2.1	Safety	Treatment Emergent Adverse Events Leading to Study Discontinuation by System Organ Class and Preferred Term
Table 14.3.2.2	Safety	Serious Treatment Emergent Adverse Events by System Organ Class and Preferred Term
Table 14.3.2.3	Safety	Serious Treatment Emergent Adverse Events Related to Investigational Product by System Organ Class and Preferred Term
14.3.3 Narratives of Deaths, Other Serious and Certain Other Significant Adverse Events		
Table 14.3.3.1	Safety	Subject Listing of Serious Adverse Events
Table 14.3.3.2	Safety	Subject Listing of Adverse Events Leading to Discontinuation of Investigational Product
Table 14.3.3.3	Safety	Subject Listing of Adverse Events Leading to Death
14.3.4 Abnormal Laboratory Values		
Table 14.3.4.1	Safety	Subject Listing of Abnormal Laboratory Values: Clinical Chemistry and Thyroid
Table 14.3.4.2	Safety	Subject Listing of Abnormal Laboratory Values: Hematology
Table 14.3.4.3	Safety	Subject Listing of Abnormal Laboratory Values: Urinalysis
14.3.5 Laboratory Data Summary Tables		
Table 14.3.5.1.1	Safety	Hematology: Observed Results and Change and Percent Change from Baseline by Study Visit
Table 14.3.5.1.2	Safety	Hematology: Shifts from Baseline Relative to the Normal Range
Table 14.3.5.2.1	Safety	Clinical Chemistry: Observed Results and Change from Baseline by Study Visit
Table 14.3.5.2.2	Safety	Clinical Chemistry: Shifts from Baseline Relative to the Normal Range by Study Visit
Table 14.3.5.3.1	Safety	Quantitative Urinalysis: Observed Results and Change from Baseline by Study Visit
Table 14.3.5.3.2	Safety	Quantitative and Qualitative Urinalysis: Shifts from Baseline Relative to the Normal Range
Table 14.3.5.3.3	Safety	Qualitative Urinalysis: Summary of Results by Study Visit
14.3.6 Other Safety Data Summary Tables		

Table Number	Population	Table Title / Summary
Table 14.3.6.1	Safety	Vital Signs: Observed Results and Change from Baseline by Study Visit
Table 14.3.6.2.1	Safety	12-Lead Electrocardiogram: Summary of 12-Lead Electrocardiogram Interpretation by Study Visit
Table 14.3.6.2.2	Safety	12-Lead Electrocardiogram: Summary of 12-Lead Electrocardiogram Interpretation by Study Visit
Table 14.3.6.2.3	Safety	12-Lead Electrocardiogram: Summary of QT and QTcF Results by Study Visit
Table 14.3.6.4	Safety	Summary of Concomitant Medications by ATC Class and Preferred Term
Table 14.3.6.5	Safety	Summary of Columbia Suicide Severity Rating Scale (C-SSRS) by Study Visit

12.4. Other Data Summary Tables

Table 5: Other Data Summary Tables

Table Number	Population	Table Title / Summary
14.5 Other Data Summary Tables		
Table 14.5.1	Safety	Hamilton Depression Scale (HAM-D): Observed Results and Change from Baseline by Study Visit
[redacted]	[redacted]	[redacted]
Table 14.5.3	Safety	Summary of Physician Withdrawal Checklist (PWC-20)

12.5. Planned Listing Descriptions

The following are planned data and patient/subject data listings for protocol number <Protocol>.

In general, one listing will be produced per CRF domain. All listings will be sorted by treatment, site, and subject number. All calculated variables will be included in the listings.

In all listings a blank line will be placed between each subject. Within a data listing, if an item appears line after line (eg, repetition of subject number), then only the first occurrence will be displayed.

In data listings, the information for one subject will be kept on one page if at all possible, rather than splitting a subject's information across pages.

Table 6: Planned Listings

Listing Number	Population	Listing Title / Summary
16.2 Subject Data Listings		
16.2.1 Subject Discontinuations/Completions		
Listing 16.2.1.1	All Enrolled Subjects	Subject Disposition
16.2.2 Protocol Deviations		
Listing 16.2.2.1	All Subjects	Inclusion and Exclusion Criteria
Listing 16.2.2.2	All Enrolled Subjects	Protocol Deviations
Listing 16.2.2.3	Screen Failures	Reasons for Screen Failures
16.2.3 Subjects Excluded from the Efficacy Analyses		
Table 16.2.3.1	All Enrolled Subjects	Analysis Populations
16.2.4 Demographic Data and Other Baseline Characteristics		
Listing 16.2.4.1	All Enrolled Subjects	Subject Consent and Demographics
Listing 16.2.4.2	All Enrolled Subjects	Medical and Psychiatric History
Listing 16.2.4.3	All Enrolled Subjects	Mini-International Neuropsychiatric Inteview (MINI)
Listing 16.2.4.4	All Enrolled Subjects	Social Anxiety Disorder (SAD) Diagnostic Checklist
16.2.5 Compliance and/or Drug Concentration Data		
Listing 16.2.5.1	All Enrolled Subjects	Study Drug Dispensation
16.2.6 Individual Efficacy Response Data		
Listing 16.2.6.2	All Enrolled Subjects	Clinical Global Impression – Improvement scale (CGI-I)
Listing 16.2.6.3	All Enrolled Subjects	Patient Global Impression of Change (PGI-C)
Listing 16.2.6.4	All Enrolled Subjects	Liebowitz Social Anxiety Scale (LSAS)
16.2.7 Adverse Event Listings (by Subject)		
Listing 16.2.7.1	All Enrolled Subjects	Adverse Events
16.2.8 Laboratory Values (by Subject)		
Listing 16.2.8.1	All Enrolled Subjects	Clincal Chemistry and Thyroid Laboratory Evaluations
Listing 16.2.8.2	All Enrolled Subjects	Hematology Laboratroy Evaluations
Listing 16.2.8.3	All Enrolled Subjects	Urinalysis Laboratory Evaluations
Listing 16.2.8.4	All Enrolled Female Subjects	Pregnancy Test Results
Listing 16.2.8.5	All Enrolled Subjects	Laboratory Results for Drug Screening
16.2.9 Other Clinical Observations and Measurements (by Subject)		
Listing 16.2.9.1	All Enrolled Subjects	Vital Signs
Listing 16.2.9.2	All Enrolled Subjects	12-Lead Electrocardiogram Measurements

Listing Number	Population	Listing Title / Summary
16.2 Subject Data Listings		
Listing 16.2.9.3	All Enrolled Subjects	Physical Examination Measurements
Listing 16.2.9.4	All Enrolled Subjects	Quick Olfactory Test (QOT)
Listing 16.2.9.5	All Enrolled Subjects	Prior and Concomitant Medications
Listing 16.2.9.6	All Enrolled Subjects	Columbia Suicide Severity Rating Scale (C-SSRS) Screening
Listing 16.2.9.7	All Enrolled Subjects	Hamilton Depression Rating Scale (HAM-D)
[redacted]	[redacted]	[redacted]
Listing 16.2.9.9	Subecjts with Visit 14 and Visit 15	Physician Withdrawal Checklist (PWC-20)
[redacted]	[redacted]	[redacted]
16.2.10 Other Study Measurements or Assessments (by Subject)		
Listing 16.2.10.1	All Enrolled Subjects	COVID-19 Impact Assessment