[redacted]

BIONOMICS LIMITED

STATISTICAL ANALYSIS PLAN

Protocol Title:	A Phase 3, Randomized, Double-blind, 2-arm, Parallel-group, Placebo-controlled Study to Evaluate the Efficacy and Safety of BNC210 for the Acute, As-needed Treatment of Anxiety in Adults with Social Anxiety Disorder
Protocol Number:	BNC210.014
Acronym:	AFFIRM-1
Phase:	3
Sponsor:	Bionomics Limited [redacted]
Author:	[redacted]
Document Version:	Final V3.0
Document Date:	25Sep2025

Document History

Final V1.0	12Feb2025 – Initial Version
Final V2.0	[redacted]
Final V3.0	[redacted]

Approval Signature Page

Bionomics Limited:

_____ _____

Sponsor: Date

[redacted]

[redacted]

_____ _____

Author: Date

[redacted]

Table of Contents

List of Abbreviations

Abbreviation	Definition
AE	adverse event
ANCOVA	analysis of covariance
AR	Auto-regressive
ASA	American Statistical Association
ATC	anatomical therapeutic chemical
BA	bioanalytical
BMI	body mass index
CB-PMM	control-based pattern mixture model
CI	confidence interval
CGI-S	Clinical Global Impressions - Severity
CRF	case report form
CSR	clinical study report
C-SSRS	Columbia-Suicide Severity Rating Scale
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ECG	electrocardiogram
EMA	European medicines agency
FCS	Fully Conditional Specification
FDA	food and drug administration
HIV	Human immunodeficiency virus
HR	heart rate
ICF	Informed Consent Form
ICH	International Conference on Harmonization
ITT	intent-to-treat
IWRS	interactive web response system
LLOQ	lower limit of quantification
LS	least squares
LSAS	Liebowitz Social Anxiety Scale
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MedDRA	medical dictionary for regulatory activities
MI	multiple imputation
MMRM	mixed model for repeated measurements
MNAR	missing not at random
PGI-I	Patient Global Impressions-Improvement
PMM	Predictive Mean Matching
PP	per-protocol
PT	preferred term
QTcF	Fridericia's correction
REML	restricted maximum likelihood

Abbreviation	Definition
RR	respiratory rate or relative rate
RSS	Royal Statistical Society
SAD	social anxiety disorder
SAE	serious adverse event
SAF	safety population
SAP	statistical analysis plan
SCID-5-CT	Structured Clinical Interview for DSM-5 Disorders – Clinical Trials version
SD	standard deviation
SE	Standard Error
SI	International System of Units
SIGH-D-17	Structured Interview Guide for the Hamilton Depression Rating Scale
SoA	Schedule of Activities
SOC	system organ class
STAI	State-Trait Anxiety Inventory
SUDS	Subjective Units of Distress Scale
TEAE	treatment-emergent adverse event
ULOQ	upper limit of quantitation
WHO-DD	world health organization drug dictionary

1. Overview

This statistical analysis plan (SAP) describes the planned analysis and reporting for Bionomics Limited protocol number BNC210.014 (*A Phase 3, Randomized, Double-blind, 2-arm, Parallel-group, Placebo-controlled Study to Evaluate the Efficacy and Safety of BNC210 for the Acute, As-needed Treatment of Anxiety in Adults with Social Anxiety Disorder*) dated 12-Nov-2024 Version 1.1. Reference materials for this statistical plan include the protocol and Version 2.0 of the case report forms (CRFs) dated 03-July-2024. 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.

This SAP describes the data that will be analyzed to reveal participant characteristics, efficacy, and study safety. Whereas the protocol presents a general discussion of statistical procedures, this SAP provides greater detail and further clarifies a particular application of a general procedure. In the event of any discrepant directives between the protocol and this SAP, the SAP will supersede the protocol.

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 unblinded inferential or descriptive analysis of data pertaining to Bionomics Limited's study BNC210.014. Any future amendments made to the protocol will not necessitate amendments to the SAP unless changes in the protocol result in changes to key analyses.

2. Study Objectives and Endpoints

2.1. Study Objectives

2.1.1. Primary Objective

The primary objective is to evaluate the efficacy of 225 mg BNC210 for the acute, as-needed treatment of anxiety in social anxiety disorder (SAD).

2.1.2. Secondary Objective

The secondary objective is to evaluate the safety and tolerability of a single 225 mg dose of BNC210.

[redacted]

2.2. Study Endpoints

2.2.1. Efficacy Endpoints

2.2.1.1. Primary Efficacy Endpoint

The primary efficacy endpoint of this study is the change in Subjective Units of Distress Scale (SUDS) scores from Baseline to the average of the performance phase of a public speaking challenge.

2.2.1.1.1 Primary Efficacy Estimand

The primary efficacy estimand is defined by the following components:

1. The population is restricted to participants with SAD who meet all the study eligibility criteria.
2. The variable is the change in SUDS score from Baseline (average of 2 Baseline SUDS scores) to the average of the performance phase of a public speaking challenge.
3. During the study, participants may be exposed to possible known or unknown intercurrent events that could perhaps impact the estimand. The treatment policy strategy will be adopted for handling participants who refuse to take study drug, receive incorrect study drug, or take prohibited medications during the study. In the unlikely event of a death of a participant, or if a participant withdraws consent, a hypothetical strategy will be utilized.
4. The population-level summary is the treatment difference in least squares (LS) mean change in SUDS score from Baseline to the average of the performance phase of a public speaking challenge, analyzed using a restricted maximum likelihood (REML)-based mixed model for repeated measures (MMRM) with multiple imputation (MI) [redacted]

[redacted]

The analysis model will include fixed effects for treatment, phase (resting, anticipation, or performance), interaction between treatment and phase, sex, and a covariate for Baseline SUDS score. The outcome value of the model will be the SUDS change from Baseline score during resting and the average of the SUDS change from Baseline scores measured for each participant within each phase (each participant will have an average score within the anticipation phase and an average score within the performance phase). A linear contrast of the MMRM parameters will be used to compare treatments on the change from Baseline to the average of the performance phase.

5. Treatment: There will be one active treatment arm (225 mg BNC210) and one control arm (placebo).

2.2.1.2. Key Secondary Efficacy Endpoints

The key secondary endpoints of this study are as follows:

1. Change in SUDS scores from Baseline to the average of the anticipation phase of a public speaking challenge.
2. Change in Clinical Global Impressions-Severity (CGI-S) score from Baseline to the end of the performance phase of a public speaking challenge.
3. Difference in Patient Global Impressions-Improvement (PGI-I) scores at the end of the performance phase of a public speaking challenge.
4. Change in State-Trait Anxiety Inventory-State component (STAI-State) scores from Baseline to the end of the performance phase of a public speaking challenge.

2.2.1.2.1 Key Secondary Efficacy Estimands

The <u>first</u> key secondary efficacy estimand is defined by the following components:

1. The population is restricted to participants with SAD who meet all the study eligibility criteria.
2. The variable is the change in SUDS score from Baseline (average of 2 Baseline SUDS scores) to the average of the anticipation phase of a public speaking challenge.
3. During the study, participants may be exposed to possible known or unknown intercurrent events that could perhaps impact the estimand. The treatment policy strategy will be adopted for handling participants who refuse to take study drug, receive incorrect study drug, or take prohibited medications during the study. In the unlikely event of a death of a participant, or if a participant withdraws consent, a hypothetical strategy will be utilized.
4. The population-level summary is the treatment difference in LS mean change in SUDS score from Baseline to the average of the anticipation phase of a public speaking challenge, analyzed using a REML-based MMRM with MI [redacted]

[redacted]

The analysis model will include fixed effects for treatment, phase (resting, anticipation, or performance), interaction between treatment and phase, sex, and a covariate for Baseline SUDS score. The outcome value of the model will be the SUDS change from Baseline score during resting and the average of the SUDS change from Baseline scores measured for each participant within each phase (each participant will have an average score within the anticipation phase and an average score within the performance phase). A linear contrast of the MMRM parameters will be used to compare treatments on the change from Baseline to the average of the anticipation phase.

5. Treatment: There will be one active treatment arm (225 mg BNC210) and one control arm (placebo).

The second key secondary efficacy estimand is defined by the following components:

1. The population is restricted to participants with SAD who meet all the study eligibility criteria.
2. The variable is the change in CGI-S score from Baseline to the end of the performance phase of a public speaking challenge.
3. During the study, participants may be exposed to possible known or unknown intercurrent events that could perhaps impact the estimand. The treatment policy strategy will be adopted for handling participants who refuse to take study drug, receive incorrect study drug, or take prohibited medications during the study. In the unlikely event of a death of a participant, or if a participant withdraws consent, a hypothetical strategy will be utilized.
4. The population-level summary is the treatment difference in LS mean change in CGI-S score from Baseline to the end of the performance phase of a public speaking challenge, analyzed using an analysis of covariance (ANCOVA) model with MI [redacted] The analysis model will include fixed effects for treatment, sex, and a covariate for Baseline CGI-S. The outcome value of the model will be the change in CGI-S score.
5. Treatment: There will be one active treatment arm (225 mg BNC210) and one control arm (placebo).

The third key secondary efficacy estimand is defined by the following components:

1. The population is restricted to participants with SAD who meet all the study eligibility criteria.
2. The variable is the difference in PGI-I score at the end of the performance phase of a public speaking challenge.
3. During the study, participants may be exposed to possible known or unknown intercurrent events that could perhaps impact the estimand. The treatment policy strategy will be adopted for handling participants who refuse to take study drug, receive incorrect study drug, or take prohibited medications during the study. In the unlikely event of a death of a participant, or if a participant withdraws consent, a hypothetical strategy will be utilized.

4. The population-level summary is the treatment difference in LS mean change in PGI-I score at the end of the performance phase of a public speaking challenge, analyzed using an ANCOVA model with MI [redacted] The analysis model will include fixed effects for treatment and sex. The outcome value of the model will be the PGI-I score.
5. Treatment: There will be one active treatment arm (225 mg BNC210) and one control arm (placebo).

The fourth key secondary efficacy estimand is defined by the following components:

1. The population is restricted to participants with SAD who meet all the study eligibility criteria.
2. The variable is the change in STAI-State score from Baseline to the end of the performance phase of a public speaking challenge.
3. During the study, participants may be exposed to possible known or unknown intercurrent events that could perhaps impact the estimand. The treatment policy strategy will be adopted for handling participants who refuse to take study drug, receive incorrect study drug, or take prohibited medications during the study. In the unlikely event of a death of a participant, or if a participant withdraws consent, a hypothetical strategy will be utilized.
4. The population-level summary is the treatment difference in LS mean change in STAI-State score from Baseline to the end of the performance phase of a public speaking challenge, analyzed using a REML-based MMRM with MI [redacted] The analysis model will include fixed effects for treatment, phase (resting or performance), interaction between treatment and phase, sex, and a covariate for Baseline STAI-State score. The outcome value of the model will be the STAI-State change from Baseline score during resting, and the change from Baseline at the end of the performance phase.
5. Treatment: There will be one active treatment arm (225 mg BNC210) and one control arm (placebo).

2.2.2. Secondary Endpoints

The secondary endpoints of this study are the incidence and severity of treatment-emergent adverse events (TEAEs), TEAEs leading to discontinuation and serious adverse events (SAEs), as well as vital signs, electrocardiograms (ECGs), suicidality (Columbia Suicide Severity Rating Scale, C-SSRS), and laboratory (hematology, biochemistry, and urinalysis) parameters.

[redacted]

[redacted]

3. Overall Study Design and Plan

3.1. Overall Design

This is a randomized, double-blind, placebo-controlled, 2-arm, parallel group, multicenter study to compare the effects of a single 225 mg dose of BNC210 and placebo in approximately 332 participants with SAD during a public speaking challenge.

Participants will attend a Screening Visit to conduct the assessments required to evaluate eligibility. Eligible participants will return to the clinic within 21 days after the Screening Visit and will be randomized in a 1:1 allocation ratio to either 225 mg BNC210 or matched placebo. Randomized participants will receive a single oral dose of their allocated study drug and, approximately 1 hour later, will participate in a public speaking challenge. A follow-up assessment will be conducted by telephone or video conference approximately 7 days after administration of the study drug to assess suicidal ideation and behavior, and record any AEs since the Treatment Day visit.

[redacted]

[redacted]

[redacted]

3.3. Study Population

The study population is comprised of male and female participants who are 18 to 65 years of age, with a current diagnosis of SAD as defined in the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) and confirmed by the Structured Clinical Interview for DSM-5 Disorders – Clinical Trials version (SCID-5-CT). Furthermore, the participants must receive a total score of ≥60 on the Leibowitz Social Anxiety Scale (LSAS) at the Screening Visit and at Baseline with no >20% change in score from Screening to Baseline, and a sub-score of 2 or 3 on Question 6 of the LSAS (Fear and Anxiety subscale) at the Screening Visit and at Baseline.

3.4. Treatments Administered

Participants will be randomized to one of two treatment groups:

- 225 mg BNC210
- Placebo

[redacted]

3.5. Method of Assigning Participants to Treatment Groups

Participants who meet inclusion criteria and none of the exclusion criteria will be randomly assigned in a 1:1 ratio to receive 225 mg BNC210 or placebo. Randomization will not be stratified by any factors.

3.6. Blinding and Unblinding

This is a double-blind study in which all participants, investigators, raters, and study personnel involved in the conduct of the study, including data management, are blinded to treatment assignment, with the exception of a specified unblinded statistician and interactive web response system (IWRS) manager, who will have access to the randomization code. The unblinded study personnel will not participate in study procedures or data analysis prior to unblinding of the study data to all study related personnel.

For emergency unblinding, study personnel will use the IWRS. The IWRS will be programmed with blind-breaking instructions. If a participant's drug assignment is unblinded, the Sponsor must be notified within 24 hours of this occurrence. The date and reason for the unblinding must be recorded.

Overall unblinding will take place at the end of the study only after database lock.

3.7. Schedule of Events

Detailed schedules of events for the study are provided in Tables 1 and 2.

Table 1 Schedule of Activities – Overview

Assessment[1]	Screening (≤21 + 3 Days)[2]	Treatment (1 Day)[3]				Follow-up via Telephone or Video Conference (7 ± 3 Days Post-Treatment)[4]
		Baseline/ Predose	Dosing	Public Speaking Challenge	Post-Challenge	
Entrance and History						
Informed consent	X					
Inclusion/exclusion criteria	X	X[5]				
Demography	X					
Medical history, including prior/current medications	X					
Structured Clinical Interview for DSM-5 Disorders - Clinical Trials version (SCID-5-CT)	X					
Liebowitz Social Anxiety Scale (LSAS)	X	X				
Structured Interview Guide for the Hamilton Depression Rating scale (SIGH-D-17)	X					
Physical examination[6]	X	X				
Concomitant medications		X				
Randomization		X				

[redacted].

Assessment[1]	Screening (≤21 + 3 Days)[2]	Treatment (1 Day)[3]				Follow-up via Telephone or Video Conference (7 ± 3 Days Post-Treatment)[4]
		Baseline/ Predose	Dosing	Public Speaking Challenge	Post-Challenge	
Efficacy Assessments[7]						
Subjective Units of Distress Scale (SUDS)		X		X		
State-Trait Anxiety Inventory-State Component (STAI-State)		X		X		
Clinical Global Impressions – Severity (CGI-S) Scale		X		X		
Patient Global Impressions – Improvement (PGI-I) Scale				X		
Safety Assessments						
Vital signs[8]	X	X			X	
Triplicate 12-lead electrocardiograms (ECG)[9]	X				X	
Adverse event recording			X	X[10]	X	X
Columbia-Suicide Severity Rating Scale (C-SSRS)[11]	X	X				X
Assessment of participant's readiness for discharge[12]					X	
Laboratory Assessments[13]						
Clinical laboratory tests (hematology, biochemistry, urinalysis)	X	X			X	

[redacted]

Assessment[1]	Screening (≤21 + 3 Days)[2]	Treatment (1 Day)[3]				Follow-up via Telephone or Video Conference (7 ± 3 Days Post-Treatment)[4]
		Baseline/ Predose	Dosing	Public Speaking Challenge	Post-Challenge	
Serology (Human Immunodeficiency Virus [HIV])	X					
Serum pregnancy test (if applicable)	X					
Urine pregnancy test (if applicable)[14]		X				
Urine drug screen[15]	X	X				
Blood sample for BNC210 concentration[16]					X	
Study Intervention						
Dosing with study intervention[17]			X			

[redacted]

Table 2. Schedule of Activities (SoA) - Efficacy Assessments on Treatment Day

Assessment	Baseline/ Predose (~-5 min)		Dosing (0 min)	Resting/ Post-dose (55 min)	Instructions on Challenge (60 min)	Public Speaking Challenge								
						Anticipation Phase			Performance Phase					
	Minutes					Minutes			Minutes					
	0	1				0	1	2	0	1	2	3	4	5
Subjective Units of Distress Scale (SUDS)	X	X		X		X	X	X	X	X	X	X	X	X
Clinical Global Impressions – Severity (CGI-S) Scale		X												X
Patient Global Impressions – Improvement (PGI-I) Scale														X
Anxiety Inventory – State component (STAI-State)		X		X										X

4. Details of Efficacy Endpoints

[redacted]

4.2. Subjective Units of Distress Scale (SUDS)

The SUDS measures the subjective intensity of disturbance or distress currently experienced by an individual and is a standard instrument for rating social and performance anxiety in patients with SAD during role-playing situations. The individual self-assesses where they are on the scale, between 0-100, with higher numbers indicating more anxiety/greater discomfort (Wolpe 1969).

4.3. Clinical Global Impressions – Severity (CGI-S) Scale

The CGI-S is a widely accepted tool, frequently employed in clinical trials for neuropsychiatric disorders that measures overall disease severity of the participant's symptoms as scored by a clinician (Guy 1976). Severity is rated on the CGI-S from 1 (normal, not at all ill) to 7 (among the most extremely ill of patients). In this study, the CGI-S will be used for the clinician to rate the severity of the anxiety or distress during the public speaking challenge rather than the overall severity of the SAD disease state.

4.4. Patient Global Impressions – Improvement (PGI-I) Scale

The PGI-I is a patient self-reported counterpart of the CGI designed to assess the patient's impression of their perceived change in overall symptoms. Improvement is rated on the PGI-I from 1 (very much improved) to 7 (very much worse). In this study, the PGI-I will be used for the participants to rate their perceived change in symptoms compared to similar past experiences.

4.5. State-Trait Anxiety Inventory (STAI)

The STAI (Spielberger 1983) is a self-report measure of subjective anxiety. It comprises 2 scales of 20 items asking participants to report their feelings of anxiousness: the STAI-State scale requires reporting of feelings at the time of questionnaire completion and the STAI-Trait scale requires reporting of how the subject generally feels. Responses are marked on a 4-point scale from "not at all" to "very much so". The STAI-State component only will be used in this study.

The STAI-State total score is calculated by summing the 20 STAI-State items after reverse scoring items 1, 2, 5, 8, 10, 11, 15, 16, 19, and 20. For reverse scoring items, so that lower scores indicate a better outcome for all items, a value of 1 is scored as a 4, a 2 is scored as a 3, a 3 is scored as a 2, and a 4 is scored as a 1.

5. Statistical Analysis and Reporting

5.1. 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 participants (n) with non-missing values, mean, standard deviation (SD), median, minimum, and maximum.

Categorical (qualitative) variable summaries will include the frequency and percentage of participants who are in the category or each possible value. In general, the denominator for the percentage calculation will be based upon the total number of participants in the study population for the treatment groups, unless otherwise specified. The denominator for by-visit displays will be the number of participants in the relevant study population with non-missing data at each visit.

The minimum and maximum will be reported with the same degree of precision (i.e., 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 will be presented to 1 decimal place, unless otherwise specified.

Unless otherwise indicated, all testing of statistical significance will be 2-sided and a difference resulting in a p-value ≤ 0.05 will be considered statistically significant.

5.2. Interim Analysis and Data Monitoring

No interim analysis will be conducted for this study.

6. Analysis Populations

The following analysis populations are planned for this study:

- **Safety Population (SAF):** The Safety population includes all participants who receive study drug. The Safety population will be used for the analysis of the safety endpoints. Participants will be analyzed by actual treatment received.

- **Intent-to-Treat Population (ITT):** The ITT population includes all randomized participants. The ITT population will be the primary population used for all efficacy analyses. Participants will be analyzed by randomized treatment.

- **Per Protocol Population (PP)**: The PP population includes all participants in the ITT population who have no major protocol deviations [redacted]

- [redacted]

7. General Issues for Statistical Analysis

7.1. Statistical Definitions and Algorithms

7.1.1. Baseline

The SUDS is assessed twice at the Baseline/Predose timepoint of the Treatment visit (once at 0 minutes, once at 1 minute). Baseline for the SUDS score is defined as the average of the two scores recorded at this visit. If only one score is available, then the single available score will be considered as the Baseline value.

For all other assessments, the last non-missing observation recorded before receiving the dose of study drug will be used as the Baseline observation and for all calculations of change from Baseline.

7.1.2. Adjustments for Covariates

The Baseline score collected for the various efficacy endpoints will be utilized as a covariate in the efficacy analyses. Sex will also be included in efficacy models. Other covariates may be used for exploratory analyses.

7.1.3. Subgroups

The following subgroup will be used for the primary and key secondary efficacy endpoints. Additional subgroups may also be utilized for exploratory analyses.

- Sex
- Race

7.1.4. Multiple Comparisons

To control for multiplicity, a sequential testing procedure will be used on the key secondary efficacy endpoints. Thus, if the primary efficacy endpoint is statistically significant, then the first key secondary efficacy endpoint can be tested. Then if the first key secondary efficacy endpoint is statistically significant, the next key secondary efficacy endpoint can be tested. This will continue until all key secondary efficacy endpoints have been tested, or until non-significance is observed for an endpoint. The ordering of the tests will be as follows:

1. SUDS (performance phase – primary endpoint)
2. SUDS (anticipation phase)
3. CGI-S (performance phase)
4. PGI-I (performance phase)
5. STAI-State (performance phase)

Following this procedure will not result in any need for α-adjustments to conclude statistical significance.

7.1.5. Handling of Dropouts or Missing Data

All possible efforts will be made to ensure that participants stay in the study and all data are collected as scheduled. However, the occurrence of missing data cannot be completely eliminated.

[redacted]

7.1.5.1. Multiple Imputation Methods

Multiple imputation (MI) is a simulation-based approach where missing values are replaced using an appropriate stochastic model given the observed data and covariates, creating multiple completed data sets. These completed datasets are then analyzed using standard analysis methods, and the different parameter estimates across the datasets are then combined to produce unique point estimates, standard errors, and confidence intervals (CIs) taking into account the uncertainty of the imputation process.

Data are considered MAR if, conditional upon the independent variables in the analytic model, the missingness depends on the observed outcomes of the variable being analyzed but does not depend on the unobserved outcomes of the variable being analyzed. This assumption implies that the behavior of the post dropout observations can be predicted from the observed variables, and therefore that treatment effect can be estimated without bias using the observed data (European Medicines Agency, 2010). For studies of missing data in a controlled clinical trial setting, MAR is usually considered as a plausible underlying missing mechanism (Molenberghs and Kenward, 2007; Siddiqui et al., 2009; Mallinckrodt et al., 2008; Mallinckrodt et al., 2013). The assumption of MAR is often reasonable because, particularly in longitudinal studies wherein the evolution of treatment effects is assessed by design over time, the observed data and the models used to

analyze them can explain much of the missingness (Little and Rubin, 1987; Verbeke and Molenberghs, 2000). This point may be especially relevant in well-controlled studies, in which extensive efforts are made to observe all outcomes and factors that influence them while participants are following protocol-defined procedures. Thus, longitudinal clinical trials by their very design aim to reduce the amount of MNAR data (missingness explained by unobserved responses), thereby increasing the plausibility of MAR (Mallinckrodt et al., 2008).

When data are MAR, the missingness of the data does not depend on the missing value after conditioning on the observed data (i.e., prior assessments and Baseline covariates). Note that when the missingness of the data depends on the values of the missing variables after conditioning on the observed data, the data are called MNAR.

Note that for SUDS, the 2 Baseline (pre-dose) measurements will be used for the imputation steps along with the post-dose resting value, the 3 anticipation phase values (0, 1, and 2 minutes) and the 6 performance phase values (0, 1, 2, 3, 4, and 5 minutes). All other efficacy endpoints will include all scheduled timepoints in the imputation steps. Planned treatment group and sex will also be included in all the models for imputation.

[redacted]

[redacted]

Note: The seed for the primary imputation [utilizing Fully Conditional Specification (FCS) predictive mean matching (PMM) method] will be 12345. If the primary imputation fails to converge then the seed will be changed to 43439. If this does not fix the convergence issue the FCS PMM method will be changed to Markov Chain Monte Carlo (MCMC) utilizing seed #1 (12345). If MCMC with seed #1 fails to converge then seed #2 (43439) will be utilized. This procedure will be utilized for all analyses that utilize multiple imputation. The order of operations is summarized below.

1. FCS PMM with seed #1
2. FCS PMM with seed #2
3. MCMC with seed #1
4. MCMC with seed #2

If MCMC methods are utilized, any imputed values outside the possible range of values will be reassigned the minimum or maximum value. The minimum and maximum values for each endpoint where MI will be performed are as follows:

- SUDS: 0 and 100
- CGI-S: 1 and 7
- PGI-I: 1 and 7
- STAI: 20 and 80

[redacted]

[redacted]

- [redacted]
- [redacted]

[redacted]

group.

[redacted]

- [redacted]

[redacted]

[redacted]

[redacted]

7.1.6. Pooling of Sites

This is a multicenter study and there will be no pooling of a site factor.

7.1.7. Derived Variables

- **Change from Baseline** = value at current timepoint – value at Baseline.
- **Average of the SUDS change from Baseline scores within the anticipation phase** = the sum of the change from Baseline scores recorded during minutes 0, 1, and 2 of the anticipation phase, divided by the total number of SUDS scores recorded during the phase.
- **Average of the SUDS change from Baseline scores within the performance phase** = the sum of the change from Baseline scores recorded during minutes 0, 1, 2, 3, 4, and 5 of the performance phase, divided by the total number of SUDS scores recorded during the phase.
- **Treatment Emergent Adverse Event (TEAE)** = any AE that starts or deteriorates on or after the first dose of study medication and through Follow-up Day 7 visit.
- **Peak =** the highest (maximum) SUDS measurement of a phase. Note that higher SUDS scores indicate more anxiety/greater discomfort, so this endpoint is the worst anxiety/discomfort in each phase.

[redacted]

[redacted]

7.1.9. 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.

All p values will be displayed in four decimals and rounded using standard scientific notation (e.g., 0.XXXX). If a p value less than 0.0001 occurs it will be shown in tables as < 0.0001.

Adverse Events and Concomitant Medication Coding

An AE is any untoward medical occurrence in a clinical study participant, temporally associated with the use of study drug, whether or not considered related to study drug. All SAEs and AEs will be collected from the start of study drug until the Follow-up Visit. Any medical occurrence that begins before the start of study drug, but after obtaining informed consent will be recorded as medical history/current medical conditions and not as an AE.

AEs and medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 26.1 or higher. Concomitant medications will be coded using the World Health Organization Drug Dictionary (WHO-DD Global B3 version Sep 2023).

Partial Date Imputation

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 dosing (if in the same month), whichever is later; if just month is missing then the month assigned is the month of the dosing, unless that results in a date prior to dosing in which case the date of dosing is assigned; and if both month and day are missing then the month

assigned is the month dosing and the day assigned is either the first day of the month or the dosing date, whichever is later.

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

These conventions will be applied only to AE onset and concomitant medication 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 was reported in the CRF.

Lower and Upper Limit of Quantitation

In general, for quantitative laboratory values reported as "<" or "≤" the lower limit of quantitation (LLOQ), one-half of the reported value (i.e., LLOQ/2) will be used for analysis. The exception to this data treatment is for plasma concentrations of BNC210 that are reported as <LLOQ, where a value of zero will be used in calculating summary statistics.

For quantitative laboratory values reported as ">" or "≥" the upper limit of quantitation (ULOQ), the reported value (i.e., ULOQ) will be used for analysis.

Laboratory Test Results

For analysis purposes, repeat laboratory test results will not be used unless the original laboratory value is missing or indicated as invalid, in which case the first non-missing repeat laboratory value will be used for data analysis. Note that the only exception is a repeat laboratory test result that is collected prior to dosing. In this case, the repeat value can be used for the Baseline value for that parameter if the result is valid and it is the last recorded value prior to dosing.

The International System of Units (SI) will be used in reporting all safety laboratory values.

Treatment Duration and Exposure

Treatment duration and exposure will not be calculated given this is a single dose study with study drug given at the site. Participants who are not dosed will not be included in the safety analyses.

8. Study Patients/Participants and Demographics

8.1. Disposition of Participants and Withdrawals

The number of participants screened, screen failed, randomized, completing all visits through Follow-up, and withdrawing from the study early, along with reasons for withdrawal, will be tabulated overall and by treatment group. The number of participants in each analysis population will be reported.

A by-participant listing of disposition data will be provided.

8.2. Protocol Deviations

Protocol deviations will be classified as important or non-important for statistical analysis purposes prior to unblinding. Important protocol deviations that might affect efficacy data to a significant degree will be designated as major protocol deviations and used to exclude participants from the per-protocol analysis.

[redacted]

Protocol deviations will be summarized by type (any deviation, major deviations, and important deviations) and by treatment group for the ITT population. All protocol deviations will be reported in a by-participant listing.

8.3. Demographics and Other Baseline Characteristics

Demographic variables will include age, sex, race, and ethnicity. Baseline participant characteristics will include medical history (by system organ class (SOC) and preferred term (PT)), smoking status, confirmation of SAD diagnosis, LSAS total score (Screening and Baseline), Structured Interview Guide for the Hamilton Depression Rating Scale (SIGH-D-17) score at Screening, Baseline scores for SUDS (average of 2 Baseline scores), CGI-S, and STAI-State. These data will be summarized for the SAF and ITT populations, with the exception of medical history data which will only be summarized for the ITT population.

By-participant listings of demographic data will be provided.

9. Efficacy Analysis

The ITT population will be considered the primary population for efficacy analyses.

For all efficacy analyses, participants will be analyzed by the randomized treatment group assignment (225 mg BNC210 or placebo). All efficacy data will be presented in participant listings.

9.1. Primary Efficacy Endpoint

9.1.1. Primary Analyses

The primary efficacy endpoint for this study is the change in SUDS scores from Baseline to the average of the performance phase of a public speaking challenge for participants receiving a single dose of 225 mg BNC210 compared to participants receiving placebo, where the performance phase is defined as up to six timepoints (0, 1, 2, 3, 4, 5 minutes) measured during a public speaking challenge and the average is defined as the mean of all measured timepoints during that phase.

The null hypothesis for the primary efficacy endpoint of the equality of a single dose of 225 mg BNC210 and placebo is:

- H_{01}: Mean changes in SUDS scores between Baseline and the average of the performance phase in a public speaking challenge between a single dose of 225 mg BNC210 and placebo are equal.

The null hypothesis of equal treatment effect will be rejected if the statistical analysis results in a 2-sided p-value for treatment through the average of the SUDS scores during the performance phase in a public speaking challenge of the study of less than or equal to 0.05. The trial will be claimed successful if the hypothesis of no treatment effect on the primary efficacy endpoint of the ITT population is rejected at the 0.05 (2-sided) significance level and 225 mg BNC210 is superior to placebo.

The observed and change from Baseline SUDS scores will be summarized using descriptive statistics for the ITT population. A REML-based MMRM model with MI [redacted]

The post-Baseline timepoints during the public speaking challenge (i.e., resting timepoint; 0, 1, and 2 minutes in the anticipation phase; and 0, 1, 2, 3, 4, 5 minutes in the performance phase) will be used to derive the dependent variable in the analysis. For each of these timepoints, a change from Baseline score will be derived as the difference from Baseline to that timepoint. After calculating the change from Baseline for each timepoint, the averages of these change from Baseline scores within each phase (i.e., average in the anticipation phase and average in the performance phase) will be used as the dependent variable in the analyses. Note that the resting change is calculated from a single measurement.

The MMRM model will include fixed effects for treatment, phase (resting, anticipation, or performance), interaction between treatment and phase, sex, and a covariate for Baseline score, using an unstructured covariance structure. If the unstructured covariance model will not converge using the Newton-Raphson algorithm, the following alternatives will be tried in the order presented, stopping if convergence occurs:

1) Fisher Scoring Algorithm

2) Factor Analytic Structure

3) Successive univariate analytic regression method as proposed in Lu and Mehrotra (2010)

If still the mentioned methods fail, a more specific covariance structure, first the Toeplitz covariance structure (alternative 4) then the auto-regressive order 1 (AR[1]) (alternative 5) structures will be tested. A sandwich estimator will be used with structured covariance models to control Type I error. If still the model does not converge, additional covariance structures will be explored and the final decision will be documented in the CSR.

A linear contrast of the MMRM parameters will be used to compare treatments on the change from Baseline to the average of the performance phase. The LS mean difference of averages along with the corresponding 95% CI and p-value will be reported.

[redacted]

9.2. Key Secondary Efficacy Endpoint Analysis

9.2.1. Primary Analyses

9.2.1.1. Change in SUDS Scores to the Average of the Anticipation Phase

The first key secondary efficacy endpoint is the change in SUDS scores from Baseline to the average of the anticipation phase of a public speaking challenge for participants receiving a single dose of 225 mg BNC210 compared to participants receiving placebo, where the anticipation phase is defined as up to three timepoints (0, 1, 2 minutes) measured during a public speaking challenge and the average is defined as the mean of all measured timepoints during that phase.

The null hypothesis for this key secondary efficacy endpoint of the equality of a single dose of 225 mg BNC210 and placebo is:

- H_{02}: Mean changes in SUDS scores between Baseline and the average of the anticipation phase in a public speaking challenge between a single dose of 225 mg BNC210 and placebo are equal.

The first key secondary efficacy endpoint will be summarized and analyzed for the ITT population using the same model that is used for the primary efficacy endpoint [redacted] A linear contrast of the MMRM parameters will be used to compare treatments on the change from Baseline to the average of the anticipation phase.

9.2.1.2. Change in CGI-S Score to the End of the Performance Phase

The second key secondary efficacy endpoint is the change in CGI-S score from Baseline to the end of the performance phase of a public speaking challenge for participants receiving a single dose of 225 mg BNC210 compared to participants receiving placebo.

The null hypothesis for this key secondary efficacy endpoint of the equality of a single dose of 225 mg BNC210 and placebo is:

- H_{03}: Mean changes in CGI-S scores between Baseline and the end of the performance phase in a public speaking challenge between a single dose of 225 mg BNC210 and placebo are equal.

The second key secondary efficacy endpoint will be summarized and analyzed for the ITT population using an ANCOVA model with MI [redacted]. The model will include fixed effects for treatment, sex, and a covariate for Baseline CGI-S.

9.2.1.3. Difference in PGI-I Score at the End of the Performance Phase

The third key secondary endpoint is the difference in PGI-I scores at the end of the performance phase of a public speaking challenge for BNC210 compared to placebo.

The null hypothesis for this key secondary efficacy endpoint of the equality of a single dose of 225 mg BNC210 and placebo is:

- H_{04}: Mean differences in PGI-I scores at the end of the performance phase in a public speaking challenge between a single dose of 225 mg BNC210 and placebo are equal.

The third key secondary efficacy endpoint will be summarized and analyzed for the ITT population using an ANCOVA model with MI [redacted] The model will include fixed effects for treatment and sex.

9.2.1.4. Change in STAI-State Scores at the End of the Performance Phase

The fourth key secondary efficacy endpoint is the change in STAI-State from Baseline to the end of the performance phase.

The null hypothesis for this key secondary efficacy endpoint of the equality of a single dose of 225 mg BNC210 and placebo is:

- H_{05}: Mean changes in STAI-State scores between Baseline and the end of the

performance phase in a public speaking challenge between a single dose of 225 mg BNC210 and placebo are equal.

The fourth key secondary endpoint will be summarized and analyzed for the ITT population using a MMRM with MI [redacted]. The model will include fixed effects for treatment, phase (resting or performance), interaction between treatment and phase, sex, and a covariate for Baseline STAI-State score. A linear contrast of the MMRM parameters will be used to compare treatments on the change from Baseline to the end of the performance phase.

[redacted]

[redacted]

10. Safety and Tolerability Analysis

All safety analyses will be performed on the SAF population, unless otherwise specified.

For all safety and tolerability analyses, participants will be analyzed by the treatment received (225 mg BNC210 or placebo) and overall for all participants receiving the study medication.

Safety measures including AEs, C-SSRS, physical examination findings, ECGs, vital signs, and clinical laboratory test results (hematology, biochemistry, urinalysis) will be summarized descriptively. No formal inferential analyses will be conducted for safety variables.

All safety and tolerability data will be presented in by-participant listings.

10.1. Adverse Events

In general, TEAEs are defined as AEs that are newly occurring or worsening after the dose of

study drug. In this study they are defined as AEs that start or deteriorate on or after the administration of study drug and no later than 7 days following the dose of study drug or reported through the Follow-up Day 7. For any participants who die during the study and the date of death is between the date of study drug administration and the date of study discontinuation (as entered by the site), inclusive, all AEs (including those resulting in death) that occur during the study will be considered as TEAEs and will be included in the TEAE summaries. All summaries of AEs will be based on TEAEs unless specified otherwise.

An overall summary of TEAEs will be produced and will summarize the incidence of the following event types:

- Any TEAE
- Any TEAE leading to discontinuation of study
- Any severe TEAE
- Any treatment-related TEAE
- Any treatment-related TEAE leading to discontinuation of study
- Any serious TEAE
- Any serious TEAE leading to discontinuation of study
- Any treatment-related serious TEAE
- Any treatment-related serious TEAE leading to discontinuation of study
- Any TEAE leading to death

The following summary tables will be presented for TEAE data:

- TEAEs by descending incidence of PT
- TEAEs by SOC and PT
- TEAEs by maximum severity, SOC, and PT
- TEAEs by maximum relationship to study drug, SOC, and PT
- Treatment-related TEAEs by SOC and PT
- Treatment-related TEAEs by maximum severity, SOC, and PT

TEAEs with a relationship of definitely related, probably related, or possibly related will be categorized as “related”; TEAEs with a relationship of “unlikely related” or “not related” will be categorized as “not related”.

In the case of multiple occurrences of the same TEAE within the same participant, each participant will only be counted once for each PT. In the summaries showing severity and relationship to study drug, the event with the maximum severity (severe) or strongest relationship (definitely related) will be reported. If a particular event is missing severity and/or relationship, then the strongest possible severity or relationship will be assumed for analysis (severity = severe, relationship = definitely related).

In the AE data listings, all AEs will be displayed.

10.1.1. Adverse Events Leading to Discontinuation from Study

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

10.1.2. Deaths and Serious Adverse Events

Any deaths that occur during the study will be listed.

Serious adverse events (SAEs) will be listed.

An SAE is an AE defined as any untoward medical occurrence that, at any dose, meets one or more of the criteria listed:

1) Results in death.
2) Is life-threatening.
3) It requires inpatient hospitalization or prolongation of existing hospitalization.
4) Results in a congenital abnormality or birth defect.
5) It is an important medical event that may jeopardize the participant or may require medical intervention to prevent one of the outcomes listed above.
6) Results in persistent or significant disability/incapacity

10.2. Clinical Laboratory Evaluations

Central laboratory tests (hematology, biochemistry, urinalysis) will be performed at Screening, Baseline/Pre-Dose, and Post-Challenge.

Human Immunodeficiency Virus (HIV) testing will be performed at Screening. A serum pregnancy test (if applicable) will be conducted at Screening and a urine pregnancy test (if applicable) will be conducted at Baseline/Pre-Dose. Urine drug screening tests will be conducted at Screening and Baseline/Pre-Dose.

Laboratory test results will be summarized descriptively by treatment and timepoint as both observed values and change from Baseline values.

The number of participants with clinical laboratory values below, within, or above the normal range by timepoint and in relation to Baseline will be tabulated in shift tables for each clinical laboratory analyte by treatment group.

Listings of all laboratory analytes by category (hematology, biochemistry, urinalysis) will be provided. Serology, urine drug screen, and pregnancy test results will also be included in listings. The listings of urine drug screen and pregnancy test results will be based on the ITT population.

10.3. Vital Signs

Vital signs will be collected at Screening, Baseline/Pre-Dose, and Post-Challenge. Descriptive summaries of actual values and changes from Baseline will be calculated for body temperature, pulse rate, respiratory rate, and seated systolic and diastolic blood pressure (at 5 minutes rest). Vital sign measurements, including body temperature, will also be listed.

10.4. Electrocardiograms

Triplicate 12-Lead ECGs will be collected at Screening and Post-Challenge. Descriptive summaries and changes from Screening will be presented for ECG measures of PR interval, QRS interval, QT interval, QT_CF interval, and heart rate (HR). These summaries will be presented by visit and treatment group using the average of the triplicate measurements.

The number and percentage of participants with normal and abnormal ECG results will be summarized by treatment group and visit. Summary statistics for all collected visits will be displayed for QT and both QT_C correction methods by treatment group.

Fridericia's correction (QT_CF) will be derived as follows:

Fridericia's Correction (QT_CF) $$QTc_f = \frac{QT_{msec}}{\sqrt[3]{RR}}$$

where: Relative Rate: RR = 60 / HR. HR = Heart Rate obtained from the ECG.

10.5. Physical Examinations

A full physical examination will be done at Screening and an abbreviated physical exam will be completed at Baseline/Pre-Dose. A listing of physical examination findings will be provided for complete and brief examinations.

Complete physical examinations will include, at a minimum, assessments of the general appearance, skin and lymphatics, eyes, ears, nose, throat, cardiovascular system, respiratory system, abdomen/gastrointestinal system, musculoskeletal and neurological systems. Other body systems may also be examined as required. Height and weight will also be measured at Screening only.

Abbreviated physical examinations will include, at a minimum, assessments of the eyes, ears, nose, throat, cardiovascular system, respiratory system, and abdomen/gastrointestinal system. Other body systems may also be examined as required.

10.6. Columbia Suicide Severity Rating Scale

The C-SSRS is an instrument to assess both suicidal behavior and ideation. The Screening version will be administered at Screening and the Since Last Visit version of the scale will be administered at Baseline/Pre-Dose and at the Follow-Up Visit. C-SSRS data will be listed.

10.7. Prior and Concomitant Medications

Treatments that started and stopped before administration of study drug will be considered prior treatments. Any treatments continuing or starting after the dose of study drug was administered will be considered to be concomitant.

A by-participant listing of prior and concomitant medications will be provided for the ITT population.

[redacted]

[redacted]

[redacted].

12. Changes from Planned Analysis

None.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

```
%mend;
%analyse(indata=cb_pmm);
```

```
run;
```

[redacted]

14. References

1. ASA. (2018) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, April 2018. http://www.amstat.org/about/ethicalguidelines.cfm
2. Cnaan, A., Laird, N.M., & Slasor, P. (1997). Using the General Linear Mixed Model to Analyse Unbalanced Repeated Measures and Longitudinal Data. *Statistics in Medicine*. 16, 2349-2380.
3. European Medicines Agency. 2010. Guideline on Missing Data in Confirmatory Clinical Trials.
4. Guy W, editor. ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: US Department of Heath, Education, and Welfare Public Health Service Alcohol, Drug Abuse, and Mental Health Administration; 1976.
5. ICH (1998). ICH Harmonized Tripartite Guideline. Statistical Principles for Clinical Trials E9; 1998. https://database.ich.org/sites/default/files/E9_Guideline.pdf
6. Li Lingling. SAS® v9.4 MNAR statement for multiple imputations for missing not at random in longitudinal clinical trials. PharmaSUG. 2019. Paper ST-103.
7. Liebowitz, M.R., Salman, E., Nicolini, H., Rosenthal, N., Hanover, R., & Monti, L. (2014). Effect of an Acute Intranasal Aerosol Dose of PH94B on Social and Performance Anxiety in Women with Social Anxiety Disorder. *American Journal of Psychiatry,* 171, 675-682.
8. Little R., Rubin D. 1987. *Statistical Analysis with Missing Data*. New York: John Wiley.
9. Lu, K., Mehrotra, D.V. 2010. Specification of covariance structure in longitudinal data analysis for randomized clinical trials. *Statistics in Medicine,* 29, 474-488.
10. Mallinckrodt C. H., Clark W. S., David S. R. (2001). Accounting for dropout bias using mixed-effects models. *Journal of Biopharmaceuticals Statistics*. 11, 9-21.
11. Mallinckrodt C.H., Lane P.W., Schnell D., Peng Y., Mancuso J.P. 2008. Recommendations for the Primary Analysis of Continuous Endpoints in Longitudinal Clinical Trials. *Drug Information Journal*, 42, 303-319.
12. Mallinckrodt C.H., Roger J., Chuang-Stein C., Molenberghs G., O'Kelly M., Ratitch B., Jannsens M., Bunouf P. 2013. Recent Developments in the Prevention and Treatment of

Missing Data. *Drug Information Journal,* 48, 68-80.

13. Moelnberghs, G., Thijs, H., Jansen, I., Beunckens, C., Kenward, M.G., Mallinckrodt, C., & Carrol, R.J. (2004). Analyzing incomplete longitudinal clinical trial data. *Biostatistics.* 5, 445-464.
14. Molenberghs G., Kenward M.G. 2007. *Missing Data in Clinical Studies*. New York: Wiley.
15. Ratitch, B., O'Kelly, M., Tosiello, R. 2013. Missing data in clinical trials: from clinical assumptions to statistical analysis using pattern mixture models. *Pharmaceutical Statistics.* 21(6), 337-347.
16. RSS. (2014) The Royal Statistical Society: Code of Conduct, 2014. https://rss.org.uk/about/policy-and-guidelines/code-of-conduct/.
17. Siddiqui O., Hung H.M., O'Neil R. 2009. MMRM vs. LOCF a comprehensive comparison based on simulation study and 25 NDA datasets. *Journal of Biopharmaceuticals Statistics*, 19(2), 227-246.
18. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory: STAI (form Y). Palo Alto (CA): Consulting Psychological Press; 1983.
19. US Federal Register. (1988) International Conference on Harmonization; Guidance on Statistical Principles for Clinical Trials. Department of Health and Human Services: Food and Drug Administration [Docket No. 97D-0174]. Federal Register Volume 63, Number 179, pages 79583-49598. September 16, 1998.
20. Verbeke G., Molenberghs, G. 2000. *Linear Mixed Models for Longitudinal Data*. New York: Springer.
21. Wolpe J. The practice of behavior therapy. New York: Pergamon Press; 1969.