Internal Approval Page

Protocol Number: MAA-1
Document Under review: Statistical Analysis Plan Version 1 August 31, 2017

This document was subject to review. The information it contains is consistent with:

- The current version of the Investigator's Brochure
- The moral, ethical and scientific principles governing clinical research as set out in the Good Clinical Practice guidelines

The investigator will be supplied with details of any significant or new findings, including adverse events.

MPBC
Author
Alli Feduccia

DocuSigned by:
Alli Feduccia
D7349712DF3349D...

8/31/2017

[Author Approval] [Date]

MPBC
Internal Reviewer
Berra Yazar-Klosinski

DocuSigned by:
Berra Yazar-klosinski
5A2B6B874725441...

8/31/2017

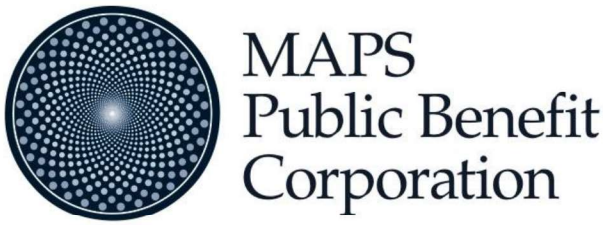
[Internal Reviewer Approval] [Date]

MPBC
Statistician
Scott Hamilton

DocuSigned by:
Scott Hamilton
26918922ABA7426...

8/31/2017

[Statistician Approval] [Date]



MAA-1 Statistical Analysis Plan

Version 1: 31 August, 2017

CORPORATION	MAPS Public Benefit Corporation (MPBC) 1115 Mission Street Santa Cruz, CA 95060
CORPORATION DESIGNEE	Amy Emerson Executive Director and Director of Clinical Research
USE	In conjunction with relevant FDA guidance
STUDY TITLE	A Placebo-Controlled, Randomized, Blinded, Dose Finding Phase 2 Pilot Safety Study of MDMA-Assisted Psychotherapy for Social Anxiety in Autistic Adults
LATEST PROTOCOL	Version 2, May 21, 2013
INVESTIGATOR	Charles S. Grob, M.D.
MEDICAL MONITOR	Michael C. Mithoefer, M.D.
STATISTICIAN	Scott Hamilton, Ph.D.
PLAN PREPARED BY	Scott Hamilton, Ph.D. Allison Feduccia, Ph.D. Berra Yazar-Klosinski, Ph.D.

Table of Contents

List of Abbreviations

AE(s)	Adverse Event(s)
ANOVA	Analysis of Variance
AVP	Arginine Vasopressin
BDI-II	Beck Depression Inventory-II
BP	Blood Pressure
BT	Body Temperature
CORT	Cortisol
C-SSRS	Columbia Suicide Severity Rating Scale
DBP	Diastolic Blood Pressure
DES-II	Dissociative Experiences Scale-II
ERQ	Emotion Regulation Questionnaire
ES	Effect Size
GWB	General Well Being
HR	Heart Rate
IR	Independent Rater
IRI	Interpersonal Reactivity Index
ITT	Intent To Treat
LSAS	Liebowitz Social Anxiety Scale
MAPS	Multidisciplinary Association for Psychedelic Studies
MDMA	3,4-methylenedioxymethamphetamine
MPBC	MAPS Public Benefit Corporation
OT	Oxytocin
PI	Principal Investigator
PP	Per Protocol
PSS	Perceived Stress Scale
PTSD	Posttraumatic Stress Disorder
QoL-Q	Quality of Life Questionnaire
RCT	Randomized Controlled Trial
RSES	Rosenberg Self Esteem Scale
SAE(s)	Serious Adverse Event(s)
SBP	Systolic Blood Pressure
SRR	Spontaneously Reported Reactions
STAI	State-Trait Anxiety Index
SUD	Subjective Units of Distress
TAS-20	Toronto Alexithymia Scale – 20
TASIT	(The) Awareness of Social Inference Test

1.0 Definitions of Terms

Categorical data: refers to discrete (indivisible) variables, such as gender or ethnicity; data will be presented as total numbers of each category as needed to describe the sample

Descriptive data: includes mean, median, standard deviation, minimum, and maximum of numerical data used as needed to describe the sample

Difference scores: consist of scores computed by subtracting one value from another, as subtracting Baseline from End of Stage 1 score, used to test for differences between and within groups to determine change as a function of experimental treatment over time

Efficacy: type of analysis used to assess therapeutic effects or benefits

Exploratory analyses: inferential or descriptive analysis of the data to determine trends that might lead to hypotheses for further study

Frequency listing: tabular listing of numbers and/or percentages of events used as needed to describe the sample or data characteristics

Outcome measures: primary and secondary study measures that are used to test the study hypotheses

Process measures: study measures or qualitative observations collected during the study that may increase depth of understanding and that are not necessarily related to safety or efficacy

Protocol deviation: event that represents significant divergence from the intended study design as described in the protocol

Safety: assessment of the condition of study subjects that examines potential risks, adverse events, and reactions

Safety measures: study measures that assess safety, such as heart rate monitoring, that are used to assess safety of the study drug

Spontaneously reported reactions, reactions: specific expected reactions gathered from the literature on MDMA

Study design: all elements of a research project that define the study question, experimental methods, study procedures including randomization and blinding, measurement techniques, data workflow, and statistical analysis

Tabular listing: list of each variable or item for each individual subject either in total or by condition in a table format

2.0 Introduction

This document presents a Statistical Analysis Plan (SAP) for MAPS study protocol MAA-1, a Phase 2 pilot clinical trial designed to evaluate the safety, feasibility, and efficacy of MDMA-assisted psychotherapy in treating social anxiety in adults on the autism spectrum.

The Multidisciplinary Association for Psychedelic Studies (MAPS) is a non-profit research and educational organization working to obtain approval for the prescription use of 3,4-methylenedioxymethamphetamine (MDMA)-assisted psychotherapy in adults on the autism spectrum with social anxiety. This is an initial pilot study to assess the efficacy of MDMA for this indication, following promising results from MAPS-sponsored studies of MDMA-assisted psychotherapy for posttraumatic stress disorder (PTSD).

The Phase 2 pilot study will examine the safety, feasibility, and efficacy of MDMA-assisted psychotherapy for social anxiety in 12 MDMA-naïve adults on the autism spectrum. This study is designed as a dose escalation study to assist with the exploration of safety and finding the most effective dose in this population. Subjects assigned to the MDMA group will receive two of three different doses, either 75 mg, 100 mg, or 125 mg MDMA. Overall, eight subjects (66.7%) will be randomized to the MDMA group and four subjects (33.3%) will be randomized to the placebo group. Subjects who received the placebo during Stage 1 will have the opportunity to cross over and take part in a second study segment, referred to as Stage 2, with two open-label experimental sessions, after completing the six-month follow-up visit at the end of Stage 1.

This study is designed to obtain safety data, and secondarily estimates of effect size for efficacy. The data will be used to determine study design and dosing for future studies for this indication. MAA-1 is also intended to contribute to the development of a manualized psychotherapeutic approach to this potential treatment. This study is exploratory. As a result, the sponsor will perform tests on all statistical comparisons without correcting for multiple comparisons.

3.0 Analysis Plan Objectives

3.1 Primary Objective

The primary objective is to assess changes in social anxiety symptoms between MDMA and placebo groups, as measured by the clinician-administered administered Liebowitz Social Anxiety Scale (LSAS) administered by a blinded Independent Rater (IR), at baseline, one day, two weeks, and one month after each experimental session, and at 6- month follow-up.

3.2 Secondary Objectives

The following objectives will compare MDMA and placebo groups in the blinded Stage 1:

- To assess effects on levels of oxytocin (OT), arginine vasopressin (AVP), and cortisol (CORT) in peripheral blood at baseline, once during each experimental session between hours 2 – 3, one month after the second experimental session, and at 6-month follow-up.
- To assess effects on social perception using The Awareness of Social Inference Test (TASIT) at baseline, once during each experimental session between hours 2-3, one month after the second experimental session, and at 6-month follow-up.

- To assess effects on emotion labeling using the Toronto Alexithymia Scale-20 (TAS-20) at baseline, one month after each experimental session, and 6-month follow-up.
- To assess effects on emotion regulation using the Emotion Regulation Questionnaire (ERQ) at baseline, one month after each experimental session, and 6-month follow-up.
- To assess effects on depression symptoms as measured by the Beck Depression Inventory-II (BDI-II) at baseline, one day, two weeks, one month after each experimental session, monthly by mail during the 5 months leading up to the 6- month follow-up assessment.
- To assess effects on perceptions of stress as measured by the Perceived Stress Scale (PSS) at baseline, within 24 hours after each experimental session, one month after each experimental session and monthly by mail leading up to the 6- month follow-up assessment.
- To assess effects on four aspects of empathy, as measured by the Interpersonal Reactivity Index (IRI) at baseline, within 24 hours after each experimental session, one month after each experimental session, and at the 6-month follow-up assessment.
- To assess effects on self-esteem as measured by the Rosenberg Self-Esteem Scale (RSES) at baseline, within 24 hours after each experimental session, one month after each experimental session, and at the 6-month follow-up assessment.
- To assess effects on state and trait anxiety as measured by the State-Trait Anxiety Index (STAI Form Y) at baseline, one day, two weeks, and one month after each experimental session, and monthly by mail leading up to the 6-month followup assessment.

The following objectives will compare effects in specified subjects:

- To determine if subjects who receive placebo in Stage 1 show a reduction in LSAS scores after two open-label sessions in the Stage 2 crossover
- To explore whether TAS-20 scores and OT, AVP, and CORT levels at baseline predict LSAS scores at the primary endpoint

The following objectives will include exploratory analyses intended to inform protocol design:

- To assess the ability of the investigators and subjects to accurately guess condition assignment in Stage 1.
- To assess perceptions about dose and study procedures at baseline and 6-month follow-up using a qualitative interview and feedback questionnaire.
- To determine feasibility of recruiting, testing and treating subjects on the autism spectrum with MDMA-assisted therapy as measured by enrollment rates, dropout rates, and completion of assessments and other study-related procedures.

3.3 Safety Objectives

The safety objectives of the study are to monitor and assure safety of subjects during and after the experimental sessions by assessing physiological effects, psychological distress, adverse events, spontaneously reported reactions and suicidality.

- Vital signs (BP, HR, BT) will be measured at baseline and every hour during experimental sessions until they return to baseline.
- Participant distress will be measured by the Subjective Units of Distress (SUD) scale at baseline and every hour during experimental sessions until return to baseline.
- Risk of suicide will be measured by the Columbia Suicide Severity Rating Scale (C-SSRS) during all in person visits, as well as twice during each experimental session at two hours after drug administration and the end of the experimental session, and on Day 2 and Day 7 after each experimental session.
- SAEs and AEs will be assessed throughout the study and spontaneously reported reactions, defined as expected AEs, will be assessed during each experimental session and for seven days after each experimental session, or until resolution.
- In addition to spontaneously reported reactions, symptoms specific to the study population will be collected at baseline and tracked throughout the study

4.0 Study Design

This randomized, double-blind, placebo-controlled study will examine the safety, feasibility, efficacy of MDMA-assisted psychotherapy in twelve adults on the autism spectrum with social anxiety. Each of the 12 subjects will participate in two experimental drug-assisted (MDMA or placebo) therapy sessions of seven hours in duration within a brief course of non-drug therapy, including three preparatory sessions at the start of the study and three integrative sessions during the month after each experimental session.

Upon enrollment, the first six subjects (Group 1) will be randomized to receive one dose of either placebo (N=2) or 75 mg MDMA (N=4). In the second experimental session one month later, Group 1 subjects randomized to MDMA will escalate to 100 mg MDMA, unless contraindicated by safety and tolerability concerns. The second six subjects (Group 2) will be randomized to receive one dose of either placebo (N=2) or 100 mg MDMA (N=4). In the second experimental session one month later, Group 2 subjects randomized to MDMA will escalate to 125 mg MDMA, unless contraindicated by safety and tolerability concerns. Subjects randomized to placebo in Group 1 and Group 2 will receive placebo in both experimental sessions also scheduled one month apart. The blind will be maintained through the six-month follow-up for each individual subject, after which the blind will be broken. Placebo subjects will be offered an open-label extension (Stage 2) with two experimental sessions scheduled one month apart, in addition to three hour-long integrative sessions during the month after each experimental session. Stage 2 subjects will receive 75 mg MDMA in the first experimental session and will escalate to 125mg MDMA in the second experimental session, unless contraindicated by safety and tolerability concerns from other subjects in Stage 1. Subjects will attend a 1-hour follow-up integrative therapy session on the day after the experimental session. Two additional integrative sessions will be conducted within the month following each experimental session, prior to outcome assessments being conducted one month after each experimental session.

Outcome measures will be assessed prior to treatment, during the treatment period, one month after each experimental session, and during the 6-month follow-up according to the Time and Events table. The sponsor will conduct an ongoing review of videos of therapy sessions, entry criteria, vital signs, and reaction data for completed sessions and any AEs. The sponsor will provide ongoing feedback to the co-therapist teams to ensure subject safety. Blood samples will be obtained for biomarker analysis at baseline and during the study according to the Time and Events Table. Biomarker analyses including plasma OT, AVP, and CORT will be explored as

surrogate endpoints at baseline, two hours after drug administration during the second experimental session, one month after the second experimental session, and at 6-month follow-up. Blood will be collected during the open-label extension Stage 2 two hours after drug administration during both experimental sessions and one month after both experimental sessions. Biomarkers will be evaluated for predictive and prognostic power. Subjects will complete a 6-month follow-up visit after completing Stage 1 or 2.

Table 1: Dose Regimen

Group # /Randomization /Sample	**Stage 1**		**Stage 2 Open-Label**	
	Session 1 Month 1	**Session 2 Month 2**	**Session 1 Month 6**	**Session 2 Month 7**
Group 1 Active (n=4)	75 mg	100 mg	N/A	N/A
Group 1 Placebo (n=2)	Placebo	Placebo	75 mg	125 mg
Group 2 Active (n=4)	100 mg	125 mg	N/A	N/A
Group 2 Placebo (n=2)	Placebo	Placebo	75 mg	125 mg

Table 2: Time & Events Stage 1

Study Phase	Baseline	Baseline	Preparatory Sessions			Experimental Session	Integrative Sessions			Experimental Session	Integrative Sessions			Follow-up	6-Month Follow-up
Visit #	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	(remote)	V14
Visit Timing			Week 1	Week 2	Week 3		1 day later	2 Wks later	2 Wks later	Appr.1 Month after V6	1 day later	2 Wks later	2 Wks later	Monthly By Mail	6 Months after V10
Informed Consent	✓														
Medical/Psychiatric History & Evaluation	✓														
SCID	✓														
General Physical Exam (includes neurological)		✓													
Screening Clinical Lab Tests, EKG		✓													
ADOS-2		✓ H													
LSAS		✓ G, H					✓ G	✓ G	✓ G		✓ G	✓ G	✓ G		✓ G
C-SSRS		✓			✓	✓ A, C, D	✓ E	✓	✓	✓ A, C, D	✓ E	✓	✓		✓
Feedback Questionnaire, Qualitative Interview		✓													✓
Study Enrollment if Eligible			✓												
Collect Concomitant Medication	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
Start Medication Taper (if applicable)			✓												✓ I
Collect Serology Samples (OT, AVP, CORT)		✓								✓			✓		✓
Record to Video			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
General Wellbeing			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
Pregnancy Screen (if applicable), Drug Screen		✓				✓				✓					
BDI-II, STAI			✓				✓	✓	✓		✓	✓	✓	✓	✓
PSS			✓				✓				✓		✓	✓	✓
IRI, RSES			✓				✓				✓		✓		✓
TASIT				✓		✓				✓			✓		✓
ERQ, TAS-20					✓								✓		✓
QoL-Q					✓										✓
Administer Drug + Therapy						✓				✓					
Monitoring of BP, Pulse, and Temperature, SUD						✓ A, B, C				✓ A, B, C					
Belief of Condition Assignment							✓				✓				
Integrative Therapy Session							✓	✓	✓		✓	✓	✓		
Start 7 Days Phone Contact							✓				✓				
Spont. Reported Reactions & All Adverse Events						✓	✓ F	✓		✓	✓ F	✓			
Symptoms at Baseline	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
Adverse Events Requiring Doctor's Visit			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓
Adverse Events Causing Withdrawal, SAEs			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Unblinding & Termination															✓ J
Approximate Hours for Study Visit	3	1.8	1.5	1.5	1.8	8.0	1.9	1.7	2.1	8.0	1.9	1.7	2.6	0	2.2

A = At the beginning of the session; B = Approximately every sixty minutes; C = As needed; D = Approximately six hours post drug; E =Day 2 and Day 7 phone calls only; F = Reactions

collected for seven days post experimental session; G = Independent Rater administered; H = May be recorded to video; I = Stage 2 subjects only; J = Only MDMA subjects will terminate.

Table 3: Time and Events Stage 2

Study Phase	Experimental Session	Integrative Sessions			Experimental Session	Integrative Sessions		
Visit #	V15	V16	V17	V18	V19	V20	V21	V22
Visit Timing		1 day later	2 Wks later	2 Wks later	Appr.1 Month after V15	1 day later	2 Wks later	2 Wks later
LSAS		✓ G	✓ G	✓ G		✓ G	✓ G	✓ G
C-SSRS	✓ A, C, D	✓ E	✓	✓	✓ A, C, D	✓ E	✓	✓
Feedback Questionnaire, Qualitative Interview								✓
Collect Concomitant Medication	✓	✓	✓	✓	✓	✓	✓	✓
End Medication Taper (if applicable)	✓ I							
Collect Serology Samples (OT, AVP, CORT)	✓			✓	✓			✓
Record to Video	✓	✓	✓	✓	✓	✓	✓	✓
General Wellbeing	✓	✓	✓	✓	✓	✓	✓	✓
Drug Screen	✓				✓			
Pregnancy Screen (if applicable)	✓				✓			
BDI-II, STAI		✓	✓	✓		✓	✓	✓
PSS		✓				✓		✓
IRI, RSES		✓				✓		✓
TASIT	✓				✓			✓
ERQ, TAS-20								✓
Administer Drug + Therapy	✓				✓			
Monitoring of BP, Pulse, and Temperature, SUD	✓ A, B, C				✓ A, B, C			
Integrative Therapy Session		✓	✓	✓		✓	✓	✓
Start 7 Days Integrative Phone Contact		✓				✓		
Spont. Reported Reactions & All Adverse Events	✓	✓ F	✓		✓	✓ F	✓	
Symptoms at Baseline	✓	✓	✓	✓	✓	✓	✓	✓
Adverse Events Requiring Doctor's Visit	✓	✓	✓	✓	✓	✓	✓	✓
Adverse Events Causing Withdrawal	✓	✓	✓	✓	✓	✓	✓	✓
SAEs	✓	✓	✓	✓	✓	✓	✓	✓
Termination								✓
Approximate Hours for Study Visit	8.0	1.9	1.7	2.1	8.0	1.9	1.7	2.7

A = At the beginning of the session; B = Approximately every sixty minutes; C = As needed; D = Approximately six hours post MDMA; E =Day 2 and Day 7 phone calls only; F = Reactions collected for seven days post experimental session; G = Independent Rater administered; H = May be recorded to video; I = At least 5 half-lives and 1 week before the experimental sessio

5.0 Randomization and Blinding

This is a randomized, double-blind, placebo-controlled study with an open-label cross-over segment. For Stage 1, a randomization list will be prepared at the beginning of the study. If needed, subsequent randomization lists will be created to replace subjects who withdraw from the study to ensure that replacement subjects are not added to the end of the initial list. Each subject will be assigned to one of the two dose conditions: 0 mg (placebo) or 125 mg (full dose). Eight evaluable subjects will be assigned to the full dose condition, and four to the placebo condition. The study will employ a blinded randomization procedure that will maintain the 8:4 ratio while maintaining the blind and ensuring that each subsequent condition assignment is not predicated on the previous assignment.

6.0 Sample Size and Power Considerations

This study is a pilot investigation intended to gather preliminary data on the safety and efficacy of MDMA in 12 subjects on the autism spectrum with social anxiety. Because of their exploratory nature, pilot studies are often not powered for detecting the desired effect. Because it is a pilot study in a small sample, results will used to collect effect size estimates for statistical power calculations for adequately powered subsequent studies.

7.0 Measures

7.1 Outcome Measures

- Liebowitz Social Anxiety Scale (LSAS), total score
- Perceived Stress Scale (PSS), total score
- Beck Depression Inventory-II (BDI-II), total score
- Interpersonal Reactivity Index (IRI), composite scores
- Rosenberg Self Esteem Scale (RSES), total score
- State Trait Anxiety Index (STAI), composite scores
- The Awareness of Social Inference Test (TASIT), composite and total scores
- Toronto Alexithymia Scale 20 (TAS-20), total score
- Emotional Regulation Questionnaire (ERQ), composite scores

- Quality of Life Questionnaire (QoL-Q), composite and total scores

7.2 Safety Measures

- Columbia Suicide Severity Rating Scale (C-SSRS)
- Subjective Units of Distress (SUD)
- General Well-being (GWB)
- Symptoms Specific to Population
- Vital signs (systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate (HR), body temperature (BT))
- Adverse Events (AE), including Spontaneously Reported Reactions (SRR)

7.3 Process Measures

- Belief of Treatment Group Assignment
- Subject perceptions of Experimental Sessions
- Feedback Questionnaire

8.0 Analyses

In general, nominal variables will be described in terms of frequencies and percentages and analyzed using chi square analysis. Ordinal and non-normal continuous variables will be described using sample median and range, and analyzed by non-parametric statistical tests, and approximately normal variables will be described using sample mean and standard deviations and analyzed by parametric statistical tests. All statistical tests will be two-sided. All p-values will be rounded to and displayed in four decimals. If a p-value less than 0.0001 occurs it will be shown in tables as <0.0001. For any analysis that is called out as parametric, e.g. t-test, the parametric assumptions will be examined. If necessary, nonparametric analyses will be utilized, e.g. Wilcoxon rank-sum test.

Clinical data will be presented in tabular format. Data not subject to analysis according to this plan will not appear in any tables or graphs, but will be included in the data listings. Analyses will be carried out with SAS Version 9.3 or higher. Selected results may be presented graphically using standard graphical software.

Post-hoc exploratory analyses not identified in this SAP may be performed to further examine the study data. These analyses will be clearly identified as such in the final clinical study report.

8.1 Analysis Populations

Intent-to-treat (ITT): all subjects who were randomized

Per protocol (PP): all subjects who completed Stage 1, underwent assessment of anxiety symptoms, and did not experience a major protocol deviation

Crossover: all subjects who completed Stage 2 in addition to completing Stage 1

Safety: all subjects who receive any study treatment

8.2 Handling of Dropouts, Missing Data

Early termination visit data for ITT and Safety variables will be analyzed at the closest scheduled visit. If the closest visit has valid data, the early termination data will be assigned to the next available visit.

8.3 Partial or Missing Dates

The following conventions will be used to impute missing portions of dates for adverse events and concomitant medications. Note that the imputed values outlined here may not always provide the most conservative date. In those circumstances, the imputed value may be replaced by a date that will lead to a more conservative analysis.

A. Start Dates

1) If the year is unknown, then the date will not be imputed and will be assigned a missing value.

2) If the month is unknown, then:

   i) If the year matches the first dose date year, then impute the month and day of the first dose date.

   ii) Otherwise, assign ‘January.’

3) If the day is unknown, then:

   i) If the month and year match the first dose date month and year, then impute the day of the first dose date.

   ii) Otherwise, assign the first day of the month.

B. Stop Dates

1) If the year is unknown, then the date will not be imputed and will be assigned a missing value.
2) If the month is unknown, then assign 'December.'
3) If the day is unknown, then assign the last day of the month.

8.4 Protocol Deviations

All protocol deviations will be included as a categorized listing. Safety and Intent to Treat analyses will include all enrolled subjects with all available data. Subjects with major deviations will be excluded from the per protocol analyses. Major deviations will be defined as anyone who completes through the primary endpoint and anyone who was enrolled but found to not meet inclusion/exclusion criteria during the course of the study. The number of subjects in each protocol deviation category listed below will be summarized by MDMA group, and individual subjects will be listed in the appendix.

Possible protocol deviations include the following seven categories:

- Subject entered study but did not meet criteria
- Subject developed withdrawal criteria but was not withdrawn
- Subject received excluded concomitant treatment
- Protocol procedure not performed per protocol
- Subject received incorrect treatment or incorrect dose
- Protocol procedure performed out of range
- Miscellaneous

8.5 Pooling of Investigator Centers

All subjects in this study come from one investigational center.

8.6 Baseline Values

Baseline values are from screening/baseline visit for all measures, except C-SSRS. For C-SSRS, baseline (visit 2) scores will be used as a measure of 'lifetime' suicidal ideation and behavior, and preparatory session 3 (visit 5) will used as 'baseline.'

8.7 Subject Disposition and Dosing Summary

All subjects enrolled in the study (i.e., who sign informed consent and complete inclusion/exclusion criteria) will be included in the summary of subject disposition and accountability. No inferential statistical tests will be performed. The tabulation of number of subjects in each treatment group and overall will be displayed for all subjects in the Safety Population, in the ITT Population, and in the PP Population. The number and percent of subjects

who completed or discontinued the study will be displayed for each treatment group and overall together with reasons for early termination, where the percent is with respect to the total number of randomized subjects in that treatment group. The timepoint of doses and total MDMA (mg) administered will be summarized by treatment group for the Safety, ITT and PP Populations.

8.8 Demographics and Baseline Characteristics

Subject demographic data and baseline characteristics will be summarized descriptively by treatment group and overall. The demographic data and baseline characteristics will be summarized for the ITT and Crossover Populations.

8.9 Prior and Concomitant Medications

The number and percent of subjects who took medications prior to and after signing informed consent will be summarized descriptively for each treatment group. Concomitant medications will be summarized similarly. Prior and concomitant medications will be summarized for the Safety Population. Psychiatric medications will be coded to common drug classes and terms.

8.10 Efficacy Analyses

For all primary, secondary and exploratory endpoints descriptive statistics (n, mean, standard deviation, median, range, or counts and percentages where appropriate) will be provided by treatment group.

8.10.1 Primary Efficacy Analyses

<u>Liebowitz Social Anxiety Scale (LSAS)</u>

The primary efficacy evaluation is the change from baseline to the primary outcome timepoint (visit 13, one month post experimental session 2) in the LSAS score of Social Anxiety (difference score). The primary efficacy comparison on difference scores will be made with independent sample t-tests at an alpha level of 0.05. If the primary hypothesis rejects, pairwise comparisons among the treatment groups will be made with t-tests. If the parametric assumptions for the t-test analyses are not met, the analogous nonparametric methods will be used (Wilcoxon Rank Sum Test). If data has normal distribution, Cohen's d placebo-subtracted effect size will be estimated on mean difference scores from baseline to Visit 13 with confidence intervals.

8.10.2 Secondary Efficacy Analyses

The secondary efficacy analyses will be made with independent sample t-test comparing change from baseline (Baseline) to the primary outcome timepoint (visit 13, one month post experimental session 2) of all the secondary measures, following the same methodology used for the primary endpoint. If the parametric assumptions for the t-test analyses are not met, the analogous nonparametric methods will be used (Wilcoxon Rank Sum Test).

Liebowitz Social Anxiety Scale (LSAS) composite scores will be analyzed in the same manner as the LSAS total scale primary analysis, and LSAS total score will also be analyzed for effect size with Cohen's *d* analysis.

Perceived Stress Scale (PSS) total score will be analyzed in the same manner as the LSAS total scale primary analysis.

Beck Depression Inventory-II (BDI-II) total score will be analyzed in the same manner as the LSAS total score primary analysis.

Interpersonal Reactivity Index (IRI) composite scores will be analyzed in the same manner as the LSAS total score primary analysis.

Rosenberg Self Esteem Scale (RSES) total score will be analyzed in the same manner as the LSAS total score primary analysis.

State Trait Anxiety Index (STAI) subscale scores will be analyzed in the same manner as the LSAS total score primary analysis.

The Awareness of Social Inference Test (TASIT) composite and total scores will be analyzed in the same manner as the LSAS total score primary analysis.

Toronto Alexithymia Scale 20 (TAS-20) total score will be analyzed in the same manner as the LSAS total score primary analysis.

Emotional Regulation Questionnaire (ERQ) composite scores will be analyzed in the same manner as the LSAS total score primary analysis.

Quality of Life Questionnaire (QoL-Q) composite and total scores will be analyzed in the same manner as the LSAS total score primary analysis.

Biomarker (oxytocin, vasopressin, and cortisol) levels will be analyzed in the same manner as the LSAS total score primary analysis.

8.10.2.1 Secondary Efficacy Analyses at Secondary Endpoints

Crossover Subject Analyses
LSAS data from subjects assigned to the placebo group, i.e. 'crossover population,' will be analyzed by t-tests on difference scores from Stage 1 endpoint (Visit 13) to Stage 2 endpoint (Visit 22) only if there is sufficient data to examine.

6-month Follow-up Analyses
To determine durability of effects, LSAS data will be analyzed by t-tests on difference scores from Stage 1 endpoint (Visit 13) to 6-month follow-up. Additionally, the following measures' scores will be analyzed in this same fashion: PSS, BDI-II, IRI, RSES, STAI, TASIT, TAS-20, ERQ, QoL-Q

8.10.3 Exploratory Analyses

Exploratory analyses of biomarkers will examine their ability to predict clinical outcome. Analyses will be performed for OT, AVP, CORT, and BDNF. Changes in biomarker concentrations from pre-treatment levels will be used to analyze effects of acute dosing and to determine the time-course of elevated biomarker levels post-drug. Pre-treatment biomarker concentrations will be evaluated in regression analysis with condition assignment in Stage 1 and biomarker concentrations at baseline as predictors and treatment outcomes as dependent variables. Additionally, the effect of biomarker levels taken at the second experimental session, and two weeks and six months after the second experimental session, will be examined for correlation with LSAS outcomes.

Distributional characteristics will be examined for outliers and extreme values. If outliers are found in primary or secondary outcome measures, the sponsor will perform analyses with and without the outlying data.

Post-hoc exploratory analyses not identified in this SAP may be performed to further examine the study data. These analyses will be clearly identified as such in the final clinical study report.

8.10.3.1 Process Measures

Belief of Treatment Group Assignment
In order to compare the therapists', independent raters', and subjects' belief of treatment group to actual dose received in each blinded session, the number and frequency of correct guesses will be calculated and depicted by treatment group and study role (subject, therapist, or independent rater).

Feedback Questionnaire
The Feedback Questionnaire nominal variables will be described in terms of frequencies and percentages, while ordinal and non-normal continuous variables will be described using sample mean, standard deviations, and range.

8.10.4 Safety Analyses

The primary measure of safety will be the reporting of adverse events. The Adverse events considered are Treatment Emergent Adverse Events (TEAE) defined as those AE's that occurred after dosing and those existing AEs that worsened during the study. Verbatim terms on case report forms will be mapped to preferred terms and system organ classes using the MedDRA dictionary. For incidence reporting if a subject has more than one adverse event mapped to the same preferred term, that adverse event will be reported only once using the highest severity and closest relationship to study drug. Subject incidence of adverse events will be displayed by Stage I treatment group and by system organ class. Adverse events will also be summarized by severity and relationship to study drug. Subject incidence of serious adverse events by Stage I treatment group will also be displayed. In addition to the listing of all adverse events, a listing of serious adverse events and a listing of adverse events leading to discontinuation of study drug will be included.

Summary tables of frequency listings of expected adverse events (Spontaneously Reported Reactions) mapped to preferred terms will be displayed during and after each experimental session by Stage I treatment group.

Columbia Suicide Severity Rating Scale (C-SSRS)
Suicidal ideation and behavior will be summarized according to suggestions made in the Columbia-Suicide Severity Rating Scale Scoring and Data Analysis Guide [3]. A positive response for suicidal ideation is counted when a subject answers "yes" to any one of the five suicidal ideation questions (Categories 1-5) on the C-SSRS, i.e. a score > 0 for suicidal ideation score. Serious suicidal ideation is a suicidal ideation score of 4 or 5. A positive response for suicidal behavior occurs when a subject answers "yes" to any one of the five suicidal behavior questions (Categories 6-10) on the C-SSRS, i.e. a score > 0 for suicidal behavior score. The number and percent of positive responses of Positive Ideation, Serious Ideation, and Positive Behavior will be tabulated by Stage I treatment group and time period. Compare lifetime serious suicidal ideation and positive behavior frequencies to cumulative frequencies anytime during the study until end of stage 1 and stage2.

Subjective Units of Distress (SUD)
Descriptive statistics for SUD scores will be calculated by Stage I treatment group and time period with counts and percentages.

Vital signs
Vital signs (heart rate, body temperature, systolic and diastolic blood pressure) will be summarized using descriptive statistics at baseline and at each post-baseline time point. Occurrences of systolic and diastolic blood pressure, heart rate, and body temperature readings above the pre-determined cutoff will be displayed with numbers and percentages by timepoint.

8.11 Timing of Analyses

The primary efficacy analysis will be conducted after all subjects complete Stage 2. Subsequent analyses on this data set will not be conducted after initial analyses are performed, unless for further exploratory post-hoc analyses. Changes to protocol will not occur after primary analysis.

9.0 Statistical Software

Data manipulation, tabulation of descriptive statistics, calculation of inferential statistics, and graphical representations will be performed primarily using SAS (release 9.3 or higher) for Windows. If the use of other software is warranted, the final clinical study report will detail what software was used and for what purposes.

10.0 References

1. Mithoefer, M.C., et al., *The safety and efficacy of {+/-}3,4-methylenedioxymethamphetamine-assisted psychotherapy in subjects with chronic, treatment-resistant posttraumatic stress disorder: the first randomized controlled pilot study.* J Psychopharmacol, 2011. **25**(4): p. 439-52.
2. Mithoefer, M.C., et al., *Durability of improvement in post-traumatic stress disorder symptoms and absence of harmful effects or drug dependency after 3,4-methylenedioxymethamphetamine-assisted psychotherapy: a prospective long-term follow-up study.* J Psychopharmacol, 2013. **27**(1): p. 28-39.
3. Nilsson, M.E., et al., *Columbia Suicide Severity Rating Scale Scoring and Data Analysis Guide*, in *CSSRS Scoring Version 2.0*. 2013: http://www.cssrs.columbia.edu/documents/ScoringandDataAnalysisGuide_Feb2013.pdf. p. 1-13.