Title Page

Protocol Title: A PHASE 2, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL-GROUP, STUDY TO EVALUATE THE SAFETY AND EFFICACY OF EMB-001 IN SUBJECTS WITH MODERATE-TO-SEVERE COCAINE USE DISORDER

Protocol Number: ERL-003, Version 5.0, Amendment 4

Compound Number: EMB-001

Sponsor Name: Embera NeuroTherapeutics, Inc.

Legal Registered Address: [redacted]

Table of Contents

Title Page **1**

Table of Contents **2**

Version History **4**

1. Introduction **5**
1.1. Objectives and Endpoints 5
1.2. Study Design 6
1.3. Randomization 6

2. Statistical Hypotheses **7**

3. Sample Size Determination **7**

4. Analysis Populations **7**
4.1 All Enrolled Population 7
4.2 Safety (SAF) population 8
4.3 Modified Intent-To-Treat (mITT) population 8
4.4 Per Protocol (PP) population 8

5. Statistical Analyses **8**
5.1. General Considerations 8
5.2. Methodology for Missing Data 9
5.3. Participant Dispositions 9
5.4. Demographic and Baseline Characteristics 9
5.5. Surgical and Medical History 10
5.6. Prior, Concomitant, and Post Medications 10
5.7. Primary Efficacy Endpoint Analysis 10
5.7.1. Definition of Endpoint 10
5.7.2. Primary Analysis 10
5.7.3. Secondary Analysis 11
5.7.4. Sensitivity Analysis 11
5.8. Secondary Efficacy Endpoints Analysis 12
5.8.1. Secondary Endpoint: Weekly Cocaine Non-Use Days 12
5.8.2. Secondary Endpoint: Quantitative Measurements of Urine BE 13
5.8.3. Secondary Endpoint: Use of Other Drugs 13
5.8.4. Secondary Endpoint: Cocaine Craving Questionnaire-Brief (CCQ-Brief) 13
5.8.5. Secondary Endpoint: Pittsburgh Sleep Quality Index (PSQI) 14
5.8.6. Secondary Endpoint: Hospital Anxiety and Depression Scale (HADS) 14
5.8.7. Secondary Endpoint: Sheehan Disability Scale (SDS) 14
5.9. Subgroup analysis 14
5.10. Tertiary/Exploratory Endpoints Analysis 14
5.11. Safety Analyses 15
5.11.1. Extent of Exposure 15
5.11.2. Adverse Events 15

5.11.3. Physical Examination....16
5.11.4. Vital Sign....16
5.11.5. ECG....16
5.11.6. Laboratory investigations....17
5.11.7. ACTH stimulation testing....17
5.11.8. Glucocorticoid Insufficiency Review Checklist (GIRC)....17
5.11.9. Sheehan-Suicidality Tracking Scale (S-STS)....18
5.12. Interim Analyses....18
References....19

Version History

This Statistical Analysis Plan (SAP) for study ERL-001 is based on the protocol dated 22 Sep 2020 and an administrate letter dated 26 Aug 2021.

SAP Version	Approval Date	Change	Rationale
1			Original version



1. Introduction

This document describes the planned analysis that will be used for Study ERL-003, a phase 2 randomized, double-blind, placebo-controlled, parallel group study designed to assess the safety and efficacy of EMB-001 in subjects with moderate to severe cocaine use disorder (CUD).

The database lock will occur when all patients have completed the Week 18 visit, or exited earlier.

1.1. Objectives and Endpoints

Objectives	Endpoints to support the objectives
Primary	
• The primary objective of this study is to evaluate the efficacy of EMB-001 compared with placebo in treating moderate-to-severe CUD.	Primary endpoint: • Continuous abstinence from cocaine use over the final 3 weeks (Week 11 to Week 13) of the Treatment Period as assessed by self-reports using the timeline follow-back (TLFB) interview and confirmed by urine BE. Secondary endpoints: • Weekly cocaine non-use days by self-report • Quantitative measurements of urine BE • Use of other drugs assessed in drug screens • Cocaine Craving Questionnaire-Brief (CCQ-B) • Pittsburgh Sleep Quality Index (PSQI) • Hospital Anxiety and Depression Scale (HADS) • Sheehan Disability Scale (SDS)
Secondary	
• The secondary objectives of this study are to evaluate the safety and tolerability of EMB-001 in subjects with moderate-to-severe CUD.	• AEs, including SAEs • Physical examinations • Vital sign measurements

	• Laboratory investigations, including serum sodium at each visit • ACTH stimulation testing • Glucocorticoid Insufficiency Review Checklist (GIRC) • ECGs • Sheehan-Suicidality Tracking Scale (S-STS)

1.2. Study Design

This study is a Phase 2, randomized, double-blind, parallel-group study to compare the safety and efficacy of EMB-001 with a placebo control in subjects with moderate-to-severe CUD. Subjects will receive investigational medicinal product (IMP) during a 12-week, randomized, double-blind Treatment Period (Week 2 through Week 13) and a 1-week, double-blind Taper Period (Week 14).

After undergoing study procedures during the Screening and Baseline Periods, subjects who meet inclusion and exclusion criteria will then be randomized in a 1:1 ratio (approximately 40 patients/arm) to one of the following for the treatment arms (Treatment period Weeks 2 – 13):

- EMB-001 720 mg metyrapone/24 mg oxazepam BID, for a total daily dose of 1440 mg metyrapone and 48 mg oxazepam (Active Group)
- Placebo BID, (Placebo Group)

During the Taper Period (week 14), subjects in the Active Group will receive EMB-001 240/8 BID; and the Placebo Group will continue to receive placebo. Both groups will change from taking three capsules BID (twice daily) to one capsule BID (twice daily).

There will be a follow-up visit for safety assessments at Week 18.

Safety and efficacy assessments will be performed during the Screening Period (Day -30 to -1), the Baseline Period (Day 1 to 7) and on Study Day 8, and safety and efficacy will be assessed at least twice a week during Treatment and Taper Periods, and at follow-up/early termination. Cocaine use will be evaluated by self-report and by urinary cocaine metabolite benzoylecgonine (BE), assessed twice weekly, by study personnel using samples collected at the study site. Several methods will be employed to optimize and confirm adherence to IMP and protocol activities, including a medication adherence monitoring and optimization service/methodology. All subjects will receive a referral for cognitive behavior therapy (CBT) at the final study visit or at early termination.

The Screening and Baseline Periods are up to 37 days prior to initiation of treatment (Days -30 to 7). The total duration of the study will be up to 22 weeks.

1.3. Randomization

Patients will be equally randomized to one of the following treatment groups:

- Placebo BID
- EMB-001 720/24mg BID

Randomization will be stratified by level of cocaine use at baseline. The randomization schedule will be generated by TAB Clinical prior to the start of the study, using validated software.

To ensure adequate representation of patient population with different cocaine use levels, enrollment for less frequent use (<15 days in the prior 28 days) will be capped at 30%. Once 30% of less-frequent use patients have been randomized, future enrollment for this cohort will be closed and only subjects with relatively more frequent cocaine use will be randomized.

To maintain the integrity of the double-blinded study, access to randomization codes and unblinded information will be strictly limited to both patients and study team.

2. Statistical Hypotheses

The statistical hypothesis for testing superiority of EMB-001 to placebo with regard to abstinence rates over the final 3 weeks (Week 11 through Week 13) of the treatment period states as follows:

Null hypothesis: the cocaine abstinence rates of EMB-001 over the final 3 weeks of the treatment period is equal to placebo;

Alternative hypothesis: the cocaine abstinence rates of EMB-001 over the final 3 weeks of the treatment period is higher than placebo.

Hypothesis tests structured using one-sided alpha are consistent with a narrow interest to determine whether an active treatment has better efficacy than placebo, as such testing will use a one-sided test for two-sample, independent statistics with a significance level of 5%.

3. Sample Size Determination

Sample-size calculations were based on a comparison between EMB-001 and placebo. In a one-sided test for two-sample, independent proportions, at the significance level of 5%, 25 completers per group will provide at least 80% power to detect a difference in abstinence rates of 10% in the placebo arm and 40% in the active arm. Including a dropout rate of approximately of 37.5%, randomization of 80 subjects (approximately 40 subjects in each group) will provide ample power to meet study objectives.

4. Analysis Populations

The analysis populations are defined as follows:

4.1 All Enrolled Population

For the analysis purposes, the all enrolled subjects population (ENR) will contain all subjects who provided informed consent for this study and were deemed as meeting the

inclusion/exclusion criteria of this study by the Investigator before taking the first dose of study drug. This analysis population will be used to summarize subject disposition and inclusion in the analysis populations.

4.2 Safety (SAF) population

The Safety population will include all randomized subjects who receive study treatment.

4.3 Modified Intent-To-Treat (mITT) population

The mITT population will include all randomized subjects who receive study treatment and return for at least one postbaseline efficacy visit

4.4 Per Protocol (PP) population

The PP population will include all randomized subjects who complete the study (as defined by completing through Day 91 study visit) with no major protocol deviations.

All safety analyses will be performed on the safety population and all efficacy analyses will be performed on the mITT and PP populations.

5. Statistical Analyses

5.1. General Considerations

Statistical analyses will be reported with tables, figures, and listings and using recommended International Council on Harmonization (ICH) numbering. Output specifications for all tables, figures, and listings will be in conformance with guidelines specified by the ICH in Appendix 7 of the Electronic Common Technical Document Specification (Apr 2003).

Data from all participating investigational sites will be pooled prior to analysis.

Descriptive statistics will be used to summarize the data. Continuous variables will be presented showing number of observations available, mean, median, minimum, maximum, and standard deviations (or standard errors, depending on the variable) by visit. Categorical variables will be presented showing frequencies and percentages by visit.

Efficacy analyses will be done on the mITT and, potentially, also on the PP. If the PP differs from the mITT by more than 10%, the analyses will be replicated on the PP. If the PP and the mITT do not differ by more than 10%, analysis may not be done on the PP. Final judgments will be made at the blind data review meeting in accordance with ICH E9.

MedDRA v25 will be used. SAS version 9.4 will be used.

For efficacy and safety assessments (e.g., laboratory and vital signs), assessments performed on Day 8 (or the last available assessment prior to Day 8) will be used as baseline.

If subjects have no value as defined above, then the baseline results will be considered missing.

All participant data up to the end of study (EOS)/completion or to early termination/withdrawal from study will be included in the analyses, regardless of treatment status and respective of applicable analyses set.

Deviations from the original, planned analyses will be documented in the final study report.

See Appendix 4 for a Software and Analysis Formatting details.

See Appendix 5 for a list of analysis tables and data listing outputs to be generated.

5.2. Methodology for Missing Data

Missing efficacy data will be handled as described in Section 5.7 of this analysis plan. Missing questionnaire data will be handled as described in Section 5.8 of this analysis plan. Missing safety data will not be imputed.

5.3. Participant Dispositions

A tabulation of subject disposition will be presented by treatment group and overall, including the number screened, the number randomized in each population group, the number dosed in each population group, the number who withdrew prior to completing the study, and reasons for withdrawal.

Inclusion and exclusion from each analysis population, including reasons for exclusion, and major protocol violations will be presented based on the All Enrolled Population.

5.4. Demographic and Baseline Characteristics

Demographic (age, sex, and race) and baseline characteristics will be summarized descriptively using the SAF by treatment arm. Baseline characteristics include, but are not limited to:

- Age (years)
- Birth gender
- Race
- Ethnicity
- Weight
- Height
- BMI
- Cocaine use at screening (including number of days of cocaine use, level of cocaine use, form of cocaine, route of administration)
- ACE score (Day 8)
- CTQ score (Day 8)
- CCQ-B score (Day 8)
- Severity of cocaine use disorder

More details are provided in Appendix 2.

5.5. Surgical and Medical History

Medical/surgical history collected at screening and baseline will be coded using the MedDRA v25, and are defined as those medical conditions/ diseases which stopped prior to or on the date of informed consent form (ICF) signing. Medical history data will be summarized based on the SAF, by System Organ Class (SOC) and Preferred Term (PT). If a subject reports a past medical condition/disease, surgery or history of treatment more than once within a SOC or PT, the medical condition/disease, surgery or history of treatment will be reported only once for that SOC or PT.

5.6. Prior, Concomitant, and Post Medications

Medications will be coded using WHO DRUG dictionary, Version 23.

Prior medications will be defined as medications with a start date and an end date prior to the first dose of study drug. Prior medications will be summarized by anatomical therapeutic chemical (ATC) level 3 term and preferred term based on the safety population.

Concomitant medications will be defined as medication with end date between the first dose of study drug and the the last dose date + 30 days. The start date of the medication may be before or after the first dose of study drug. Concomitant medications will be summarized by ATC level 3 term, preferred term based on the safety population. Procedures will be summarized by MedDRA SOC and PT. The summaries will be based on the safety population.

Post medications are defined as any medications with start date more than 30 days after the last study dose.

5.7. Primary Efficacy Endpoint Analysis

The primary efficacy parameter will be the continuous abstinence from cocaine use over the final 3 weeks of the Treatment Period (Week 11 to Week 13) as assessed by self-reports using the TLFB interview and confirmed by urine BE.

5.7.1. Definition of Endpoint

For this composite endpoint, data from the TLFB interview will be aligned with urine BE results across study timepoints Day 74, 77, 81, 84, 88, and 91.

A subject's primary endpoint will be classified as:

- **Abstinent = 1/Yes,** if at all timepoints the TLFB interview indicates non-use in the prescribed interval per the questionnaire and corresponding urine test indicates negative for cocaine use (urine BE result is < 150 ng/mL).
- **Abstinent = 0/No**, if at any timepoint the TLFB interview result or the urine BE test result indicate positive for cocaine use

5.7.2. Primary Analysis

Drug abstinent status over the last 3 weeks of treatment period will be summarized descriptively by treatment arm, and by baseline severity of cocaine use. The proportion of abstinent and non-abstinent subjects will be summarized. The treatment effect will be

assessed with logistic regression with baseline measures for severity of cocaine use and treatment as covariate.

Subjects who have completed through the end of Week 13, but who have an isolated (single) missing urine BE data (e.g. missing urine BE result at the Day 81 visit, but urine BE results for Day 77 and Day 84 are available) will have the missing result imputed. If the urine BE results before and after the missing datum are negative, then the missing urine BE will be imputed as negative. Otherwise, the missing urine BE evaluation will be imputed as positive.

Subjects with consecutive missing BE measurement will be considered as non-abstinent. No missing imputation will be conducted for TLFB data. The drug abstinent status for a visit with missing TLFB will be obtained from the non-missing TLFB assessment at future visit. Subjects who have discontinued prior to the end of Week 13 will be considered as non-abstinent.

The primary analysis will be based on the mITT population

5.7.3. Secondary Analysis

A general linear mixed effect model for repeated measurement will be used to model the weekly abstinence. The model will include treatment, time, and treatment by-time interaction and randomization stratification factor as fixed effects, and subject as a random effect. We will use AIC to select the variance covariance structure for mixed effect model. The difference in abstinence rate between treatment arms and the corresponding 2-sided 95% confidence interval will be calculated based on the general linear mixed effect model. The model will handle missing data and does not employ formal imputation.

The primary and secondary analysis will be repeated based on the Per Protocol population if that the analysis population is different from mITT. The primary and secondary analyses will also be repeated using a urine BE result of <300 ng/mL as negative for cocaine use.

5.7.4. Sensitivity Analysis

This section describes analyses to explore the potential impact of missing data. Importantly, the method of handling missing data described in section 5.7.2 are anticipated to be the most conservative approach as it would tend toward smaller treatment effect. The additional methods described in this section will be used to qualitatively evaluate the robustness of the primary analysis method.

The sensitivity analysis will be based on the mITT population.

5.7.4.1. Sensitivity Analysis with Last Observation Carried Forward (LOCF)

LOCF method is a simple and common method for handling missing data. LOCF is used extensively in spite of the fact that it requires the following two strong assumptions: (1) missing data due to dropouts completely at random (MCAR) following early discontinuation, and (2) response following a subject dropping out remains constant following the last observed value prior to drop out. For each subject, missing TLFB interview result and urine BE result will be replaced with the last non-missing value.

5.7.4.2. Sensitivity Analysis with Multiple Imputation (MI) Approach

Instead of filling in a single value for each missing value, Rubin's (1987) multiple imputation procedure replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute.

Collected data are often incomplete, with a mixed monotone and non-monotone structure. For example, some subjects may drop out from the study, resulting in a monotone pattern of missing data. Alternatively, data may be missing intermittently due to invalid measurements; this non-monotone pattern is considered missing at random (MAR).

Two MI methods are planned:

Markov chain Monte Carlo (MCMC) method (Schafer 1997): For non-monotone missing data patterns, the MCMC assumes multivariate normality and generates multiple imputations using simulations from a Bayesian prediction distribution for normal data. Without prior information about the mean and covariance estimates, a noninformative prior will be used under the MCMC method.

Regression method: For monotone missing data patterns, a regression model is fitted for each variable with missing values. Based on the resulting model, a new regression model is simulated to impute subsequent missing values for the variable (Rubin 1987). The process is repeated sequentially for variables with missing values.

Based on these two MI methods, the missing data will be imputed using the following approaches.

Step 1: non-monotone missing data will be imputed using the MCMC method within each treatment group

Step 2: monotone missing data in each treatment group will be imputed using the regression method on the basis of the predicted future pattern for the same treatment group

Step 3: one hundred imputed datasets will be generated. Each imputed dataset is analyzed separately using the same primary analyses methods described in Section 5.7.2 and the final estimate of treatment effect will be the average of the estimates based on the 100 individual imputed datasets. The pooling of the individual estimates and inferences based on the combined estimate will be handled by SAS procedure MIANALYZE (Smith and Kosten 2017).

5.8. Secondary Efficacy Endpoints Analysis

5.8.1. Secondary Endpoint: Weekly Cocaine Non-Use Days

This secondary efficacy parameter will be the average number of non-use days during the final 3 weeks of the Treatment Period (Week 11 to Week 13) as assessed by self-reports using the TLFB interview. Data from the TLFB interview will be summarized for each study week, capturing the total number of days indicating drug use. The number of days of non-use will be calculated by total number of days looking back – number of days with drug use. These measures will be determined for each subject at study Week 11, Week 12 and Week 13.

The weekly cocaine no-use days will be summarized descriptively by week by treatment arm. Poisson regression will be used to compare the no-use days between treatment arm at the end of each study week.

5.8.2. Secondary Endpoint: Quantitative Measurements of Urine BE

This secondary efficacy parameter will be the level of urinary cocaine metabolite benzoylecgonine (BE) measured during the final 3 weeks of the Treatment Period (Week 11, Week 12, and Week 13). The endpoint will be the natural log-transformed quantitative urine BE measurements of Benzoylecgonine for each subject timepoint during study Week 11, Week 12 and Week 13.

The weekly BE measurement will be summarized descriptively by visit by treatment arm. The comparative analysis assessing treatment effect between treatment groups will use a mixed model repeated measure (MMRM) analysis with treatment, time, and treatment by-time interaction and randomization stratification factor as fixed effects, and subject as a random effect. We will use AIC to select the variance covariance structure for mixed effect model. The difference in no-use days between treatment arms and the corresponding 2-sided 95% confidence interval will be calculated based on the MMRM method. The model will handle missing data and does not employ formal imputation.

5.8.3. Secondary Endpoint: Use of Other Drugs

This secondary efficacy parameter will be the use of other drugs will be determined from the drug screens for drugs of abuse, which include tests for alcohol and opiates, oxazepam, non oxazepam benzodiazepines, barbiturates, amphetamine, methamphetamine, and phencyclidine during the treatment period of the study.

The results of the drug screens test will indicate use or non-use for each of the specific drugs at each timepoint. Missing screen test result will not be imputed and overall drug use status will be defined based on available screen test result. For each drug screened the endpoint will be defined as follows:

Drug Use = 1/Yes: if result is positive at any timepoint during the treatment period.

Drug Use = 0/No: if result is negative at any timepoint during treatment period.

The other drug use status will be summarized descriptively by treatment arm. The comparative analysis assessing treatment effect between treatment groups will use an exact test of proportions across the treatment groups for each drug screened.

Same analysis will be repeated for drug use status over the last 3 weeks of treatment period (Week 11,12,13).

5.8.4. Secondary Endpoint: Cocaine Craving Questionnaire-Brief (CCQ-Brief)

The CCQ-Brief consists of 10 statements on respondent's feelings and thoughts about using cocaine as he or she is completing the questionnaire (i.e., right now). This instrument allows investigators to obtain a reflection of the respondent's general cocaine craving. Eight items of

the scale are scored ranging from 1 for "strongly disagree" to 7 for "strongly agree" and two items are scored from 7 for "strongly disagree" to 1 for "strongly agree."

CCQ-Brief score is the average of non-missing items. CCQ-Brief score will be summarized descriptively by visit by treatment arm. The comparative analysis assessing treatment effect between treatment groups will use similar MMRM method described in 5.8.1.

5.8.5. Secondary Endpoint: Pittsburgh Sleep Quality Index (PSQI)

The PSQI is a self-report questionnaire that assesses sleep quality over a 1-month time interval. The measure consists of 19 individual items, each weighted on a 0 to 3 interval scale. The total PSQI score is then calculated by totaling the 7 component scores, providing an overall score ranging from 0 to 21, where lower scores denote a healthier sleep quality.

The total PSQI will be summarized descriptively by visit by treatment arm. The comparative analysis assessing treatment effect between treatment groups will use similar MMRM method described in 5.8.1.

5.8.6. Secondary Endpoint: Hospital Anxiety and Depression Scale (HADS)

Anxiety and depression will be measured by the HADS. This is a 14-item questionnaire (7 items each that assess depression and anxiety) in which the response to each item is ranked from 0 to 3.

Total score depression domain and anxiety domain will be calculated and summarized descriptively by treatment arm at each visit. Frequency and percentage in each of the 3 levels (0-7 = Normal 8-10 = Borderline abnormal 11-21 = Abnormal) will be presented. The chi-square test will be used to test the treatment differences between arms.

5.8.7. Secondary Endpoint: Sheehan Disability Scale (SDS)

The SDS is a brief, 5-item self-report tool that assesses functional impairment in work/school, social life, and family life. Total score ranges from 0 to 30 (0=unimpaired, 30=highly impaired).

The SDS total score will be summarized descriptively by visit by treatment arm. The comparative analysis assessing treatment effect between treatment groups will use similar MMRM method described in 5.8.1.

5.9. Subgroup analysis

Subgroup analysis for the primary endpoint and Quantitative Measurements of Urine BE (Key secondary endpoint) will be conducted for the following subgroups:

- Patients with low and high level of cocaine use at baseline
- Patients with positive and negative BE result at Day 8
- Patients with low and high level of cocaine craving at baseline (split by median of CCQ-B score at Day 8)

5.10. Tertiary/Exploratory Endpoints Analysis

No tertiary or exploratory Endpoints.

5.11. Safety Analyses

The following safety endpoints include:

- AEs, including SAEs
- Physical examinations
- Vital sign measurements
- ECGs
- Laboratory investigations
- ACTH stimulation testing
- Glucocorticoid Insufficiency Review Checklist (GIRC)
- Sheehan-Suicidality Tracking Scale (S-STS)

Safety analyses is based on the safety population. No inferential statistics will be performed; only summary statistics will be provided unless otherwise noted. Missing safety data will not be imputed.

5.11.1. Extent of Exposure

Exposure to study drug (in days) is calculated as last dose date of study drug – first dose date of study drug + 1 by treatment group. In additional, overall exposure to EMB-001 (in days) is calculated as last dose date of study drug – first dose date of study drug + 1 regardless of the dose interruption or modification.

Study drug compliance (%) captured with AICure will be summarized over the treatment period. In addition, compliance rate will be calculated as

100 x (total number of capsules dispensed – total number of capsules returned)/total number of capsules expected to be taken during the study (once daily).

Compliance will be summarized by treatment group. In addition, the number and percentage of subjects in pre-specified compliance categories (< 75%, 75-100%,>100%) will be summarized by treatment group.

5.11.2. Adverse Events

Adverse events will be coded using MedDRA v25 to classify events under primary SOC and PT.

A TEAE is defined as an AE that either begins after the first dose of study drug or worsens after the first dose of study drug. TEAEs will be summarized by SOC and PT and by treatment groups with sorting based on alphabetical order for the SOC and frequency count (descending order) for the PT. In addition, the incidence of all TEAEs and serious TEAEs will be summarized by PT with sorting based on the frequency count (descending order).

An overview of adverse events will include the number and percent of subjects who had at least one TEAE, Serious TEAE, TEAE related/suspected to be related to study drug, study drug withdrawn due to TEAE, maximum severity of TEAE, or TEAE with fatal outcome.

Multiple occurrences of an AE are counted only once per subject per SOC and PT for summary tables.

The following TEAEs will be summarized:

- Incidence of all TEAEs
- Incidence of all TEAEs by maximum severity (severe, moderate and mild) specified by investigators
- Incidence of TEAEs related/suspected to be related to study drug specified by investigators
- Incidence of serious TEAEs
- Incidence of serious TEAEs related/suspected to be related to study drug specified by investigators
- Incidence of TEAEs leading to study drug dose modification (including dose reduction, interruption or discontinuation)
- Incidence of TEAEs leading to study drug dose reduction
- Incidence of TEAEs leading to study drug interruption
- Incidence of TEAEs leading to study drug withdrawn and withdrawn from study
- Incidence of TEAEs in which the outcome is fatal

Incidence of TEAEs occurring in ≥ 5 and ≥ 10% of subjects in any treatment group will be tabulated.

All data collected in the AE case report form (CRF) will be listed in by-subject listings.

5.11.3. Physical Examination

Physical examination results will be presented in listings for each patient.

5.11.4. Vital Sign

The observed values and change from baseline observations in each vital sign parameters will be summarized descriptively by visit by treatment group. Vital sign values will be presented in a by-patient listing. Descriptive statistics displays summarizing the distribution in vital signs test over time will be presented.

5.11.5. ECG

The observed values and change from baseline observations in ECG parameters [HR, PR, QRS, QT corrected with Fridericia's formula (QTcF)] will be summarized by visit by treatment group.

The frequency and percentage of the following categories will be summarized by visit by treatment group:

- Male patients with QTcF >= 450 msec or female patients with QTcF >= 460 msec
- Patients with QTcF >= 500 msec
- Change from baseline QTcF >=30 and < 60 msec;
- Change from baseline QTcF >= 60 msec;

Electrocardiogram results will be presented in listings for each patient.

5.11.6. Laboratory investigations

The observed data and change from baseline observations in hematology, serum chemistry and quantitative urinalysis test results will be summarized by visit by treatment group. In addition, shifts in categories from Baseline to each post-baseline time point for each parameter will be summarized by treatment group. Laboratory investigations results will be presented in listings for each patient.

5.11.7. ACTH stimulation testing

ACTH stimulation testing result will be summarized descriptively by treatment arm Frequency and percentage in each of the 3 levels (Normal, Abnormal and Intermediate) will be presented. Definition of risk level and actions to be taken based on ACTH stimulation test results are provided in the table below

ACTH stimulation testing results will be presented in listings for each patient.

Cortisol Result	**Action(s) to be Taken**		
Normal (≥18 µg/dL)	The subject can continue on study with no modifications.		
Abnormal (<12 µg/dL)	The subject must be withdrawn from the study.		
Intermediate (≥12 µg/dL but <18 µg/dL)	The subject must return to the clinic 2 times per week and be evaluated once by telemedicine per week for the following 4 weeks. At the additional visit each week the subject should be assessed for vital signs, AEs, concomitant medications, GIRC, clinically significant symptoms, and suicidality.		
	ACTH stimulation test 1 week after initial ACTH stimulation test	Normal cortisol (≥18 µg/dL)	The subject can return to the regular visit schedule (1 in-person visit and 1 telemedicine interaction per week) and discontinue further ACTH stimulation tests.
		Abnormal cortisol (<12 µg/dL)	The subject must be withdrawn from the study.
		Intermediate cortisol (≥12 µg/dL but <18 µg/dL)	The subject must continue the 3 interactions (2 in-person visits, 1 telemedicine) per week schedule and return in 1 more week for another ACTH stimulation test.
	If intermediate results were obtained in the 1-week ACTH stimulation test, 1 week later another ACTH stimulation test is performed.	Normal cortisol (≥18 µg/dL)	The subject can return to the regular visit schedule (1 in-person visit and 1 telemedicine interaction per week) and discontinue further ACTH stimulation tests.
		Abnormal cortisol (<12 µg/dL)	The subject must be withdrawn from the study.
		Intermediate cortisol (≥12 µg/dL but <18 µg/dL)	The subject must continue the 3 interactions (2 in-person, 1 telemedicine interaction) per week schedule and return in 2 more weeks for another ACTH stimulation test.
	If intermediate results were obtained in the 2-week ACTH stimulation test, 2 weeks later another ACTH stimulation test is performed	Normal cortisol (≥18 µg/dL) or intermediate (≥12 µg/dL but <18 µg/dL)	The subject can return to the regular visit schedule (1 in-person visit and 1 telemedicine interaction per week) and discontinue further ACTH stimulation tests.
		Abnormal cortisol (<12 µg/dL)	The subject must be withdrawn from the study.

5.11.8. Glucocorticoid Insufficiency Review Checklist (GIRC)

The GIRC is an extensively edited (by the current Embera Endocrine Advisory Board) version of the previously employed Adrenal Insufficiency Review Checklist used in Studies ERL-001 and ERL-002. If focuses on surveillance for the key symptoms, signs, and laboratory changes seen during the onset of glucocorticoid insufficiency, which metyrapone could potentially induce (e.g., orthostatic hypotension, hyponatremia, fever, and prostration), as well as for other signs of acute intercurrent illness (nausea, vomiting, diarrhea, etc.) that might reflect glucocorticoid insufficiency or be a risk factor for it.

The frequency and percentage of positive subjects in each of the following category will be summarized by visit by treatment arm. The frequency and percentage of patients with positive outcome in any of the following category will also be summarized.

GIRC results will be presented in listings for each patient at each visit.

If any answer on any item is "yes," hold EMB-001 or placebo and perform a STAT local serum morning cortisol. If cortisol result is >18 mcg/dL, subject may continue in study. If cortisol result is <18 mcg/dL or unavailable, administer an ACTH stimulation test. If cortisol result is unavailable at the appropriate time for the ACTH stimulation test, it will be administered without knowledge of cortisol level. **Sign, Symptom, or Laboratory Test**	**Yes***	**No**
Orthostatic hypotension per protocol (20 mmHg systolic or 10 mmHg diastolic from sitting to standing for 2 minutes) (Replicate finding once to confirm.)		
Significant loss of appetite or gagging (severe food aversion, not choking or other forms of gagging)**		
Severe nausea/vomiting/diarrhea or other acute illness**		
Low serum sodium (>1 mmol/L below the minimum in the normal range)		
Unexplained abdominal pain (unexplained = not in the setting of GERD, constipation, IBS, or other known source)**		
Unexplained persistent fever >100°F (unexplained = not in the setting of URI or other known source)**		

GERD: gastrointestinal reflux disease; IBS: irritable bowel syndrome; URI: upper respiratory infection

5.11.9. Sheehan-Suicidality Tracking Scale (S-STS)

The S-STS is a prospective, subject self-report or clinician-administered rating scale that tracks both treatment-emergent suicidal ideation and behaviors. Number of patients with flags for suicidal ideation or suicidal behavior or self-injurious behavior/no suicidal intent will be summarized by treatment arm by visit.

Data for patients with a flag for any of the events will be provided by listing.

5.12. Interim Analyses

No interim analyses are planned for this study.

References

- ICH Harmonized Tripartite Guideline – Statistical Principles for Clinical Trials (E9). Step 4, 5 February 1998.
- ICH Harmonized Tripartite Guideline – Structure and Content of Clinical Study Reports (E3). Step 4, 30 November 1995.
- ICH Harmonized Tripartite Guideline – M2 eCTD Electronic Common Technical Document Specification. April 2023.
- Rubin, D. B. (2004). Multiple imputation for nonresponse in surveys (Vol. 81). John Wiley & Sons.
- Schafer, J. L. (1997). Analysis of incomplete multivariate data. CRC press.
- Smith, C., & Kosten, S. (2017). Multiple imputation: a statistical programming story. 2017) https://support.sas.com/rnd/app/stat/papers/multipleimputation. pdf.