A Pilot Multicenter, Randomized, Double-Blind Study Comparing the Proportion of Responders to PSD502 and to Placebo Using the PEBEQ™ in Subjects with Premature Ejaculation

Plethora Solutions Ltd Study No: PSD502-PE-008
Syne qua non Ltd Study No: PNS18001

Statistical Analysis Plan

Version: 5.0

Date: 29 April 2021

For Syne qua non Ltd – Lead Statistician

Paula McSkimming

SignNow e-signature ID: 6a14951714...
04/29/2021 08:45:26 UTC
Paula McSkimming (Signer)

For Plethora Solutions Ltd.

John Breddy

SignNow e-signature ID: f48663206e...
04/29/2021 08:55:03 UTC
John Breddy (Signer)

5 STATISTICS

5.1 DETERMINATION OF SAMPLE SIZE

In the Plethora study, "Development of the Premature Ejaculation Bothersome Evaluation Questionnaire (PEBEQ): Quantitative Research Stage[24], 10/56 = 18% of subjects had improvements of 1 unit or more on Item 3 (event-specific bother). This data was collected in the development and testing of the PEBEQ itself. No treatments were involved, so this can be considered to be a control success rate. Treatment change in "distress" scores of the IPE measured throughout the Sponsor study PSD502-PE-002[28], a double-blind placebo study, appear to support the anticipation of a 30% treatment change in the "bother" score of the PEBEQ in the current pilot study. Although there is no established criterion for a clinically meaningful success rate, it is hoped that treatment will result in more than doubling the success rate in the present study.

Randomizing 100 subjects (50 per treatment arm) is feasible for this pilot study in order to assess the proportion of subjects who achieve a 1-point or greater improvement over 4 weeks in the mean PEBEQ Item 3 (event-specific bother). This sample size is sufficient to detect a difference of 30% in success rates when the placebo rate is 20%, with power 80% using a chi square test with a 5% Type I error rate. It allows for a 10 percent dropout rate.

In order to assess the psychometric properties of the PEBEQ Item 3 (event-specific bother), the 100 subjects will be used. This sample size is sufficient to provide estimates of meaningful change; in a validation study conducted on 56 males to test the psychometric performance of the full PEBEQ including the General and Event Specific subscales, the standard error associated with correlations of about 0.50 on a number of PROs was about 0.12[24]. With a sample size of 100, the standard error is expected to decrease to around 0.05 to 0.08.

5.2 STATISTICAL AND ANALYTICAL PLANS

5.2.1 Data management

All data will be collected on paper CRFs and diary cards.

All data from subjects who give written informed consent and for whom screening data are collected will be entered into a validated electronic database with a full audit trail of edits made. All data management will be performed in accordance with the current Standard Operating Procedures of the designated CRO. Adverse events and concomitant medications will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and the World Health Organization (WHO) drug dictionary respectively.

The database will be locked when there are no more data queries outstanding and all defined pre-database lock procedures are completed, including the identification of protocol deviations according to the pre-defined protocol deviation specifications.

The data will then be extracted into SAS® datasets for the purpose of statistical analysis and the generation of data presentations.

5.2.2 Response variables

Efficacy variables are defined in Section 4.1 and separated into primary, secondary and exploratory variables. It is of interest to estimate the effect of treatment on the success rate among study subjects and to test the null hypothesis that this effect is zero, versus that it is greater than zero. Also, subjects will be circumcised or uncircumcised, and it is of interest to evaluate the extent to which circumcision status might affect the overall success rate and whether the treatment effect is affected by circumcision status. For this pilot study subjects will be randomized to treatment based on site and circumcision status.

The psychometric variable is defined in Section 4.2.

Safety assessments will be based on study drug exposure, the nature, incidence, relationship and severity of AEs and SAEs, incidence of and reasons for withdrawals, laboratory parameters, use of concomitant medications, vital signs, findings on physical examination, 12-lead ECG and visual examination of the penis.

5.2.3 Planned Statistical Methods

Statistical analyses will be conducted according to a Statistical Analysis Plan (SAP) for the efficacy analyses and a Psychometric Analysis Plan (PAP) for the assessment of psychometric properties of the PEBEQ Item 3 (event-specific bother). These analysis plans will be finalized before the database is "locked". Statistical analysis will be performed using the SAS system or an equivalent validated analysis system.

5.2.3.1 Missing Data

No individual PEBEQ Item 3 (event-specific bother) data will be imputed.

For each subject, the mean of PEBEQ Item 3 (event-specific bother) responses observed during the baseline period will serve as the baseline period score. In order to continue in the study, the subject must have at least three (3) observations of PEBEQ Item 3 (event-specific bother) in the baseline period. Similarly, the mean of PEBEQ Item 3 (event-specific bother) responses observed during the treatment period will serve as the treatment period score. At least one (1) treatment period PEBEQ Item 3 (event-specific bother) response is required. Subjects who do not have a treatment response for PEBEQ Item 3 (event-specific bother) will be dropped from all efficacy analyses and psychometric analyses concerning the responsiveness to change and meaningful change. The treatment period score must be at least 1 unit larger than the baseline period score for that subject to be considered a "success".

If a missing IELT for an individual encounter occurs due to anteportal ejaculation (ejaculation happening before penetration), the IELT for that encounter will be set to 0 seconds. If a missing IELT for an individual encounter occurs due to the non-ejaculation (penetration achieved but no ejaculation occurring), the IELT for that encounter will be set to the maximum achieved for that subject.

No safety data will be imputed.

5.2.3.2 Efficacy Population

All randomized subjects who receive at least one dose of a study drug and who have at least three observations of PEBEQ Item 3 (event-specific bother) in the baseline period and at least one in the treatment period will be included in the modified intent-to-treat (MITT) population. The MITT population will be used for the analysis of all baseline characteristics and the analysis of all efficacy data.

5.2.3.3 Psychometric Population

To adequately assess the reliability and validity of the PEBEQ Item 3 (event-specific bother), all eligible subjects entering the baseline period who indicate a 4 or greater on the Screening 11-point NRS Scale to Assess Bother will be included in the analyses (psychometric population). The PEBEQ Item 3 (event-specific bother) analysis set will consist of all subjects with valid sexual event data completed between screening and randomization.

The analysis of responsiveness to change and the derivation of meaningful change will be performed on the MITT population.

5.2.3.4 Safety Population

All subjects who are treated will be included in the safety population. The safety population will be used for the analysis of all safety data.

5.2.3.5 Interim Analysis

No interim analysis is planned for this study.

5.2.3.6 Demographics and Baseline Characteristics

Demographic and other baseline characteristics will be summarized for each treatment group and subgroup (site, circumcision status). The number of subjects who are randomized, receive study treatment, complete treatment and withdraw from the study will be presented, along with the reasons for withdrawal.

5.2.3.7 Treatment Compliance

The subjects' compliance for the return of the canisters at the end of treatment will be summarized by treatment group. Compliance is calculated as 100*(total number of canisters returned up to the end of treatment/total number of canisters dispensed at the start of treatment).

The percentage of encounters where the study medication is used will be summarized by treatment group. For encounters where the study medication is used, the percentage of encounters where the protocol-planned dose is used and the mean dose expressed as a percentage of the protocol-planned dose will also be summarized by treatment group.

5.2.3.8 Efficacy Analyses

The number of encounters, penetrations, ejaculations, anteportal ejaculations and penetrations without ejaculation will be summarized descriptively over the 4 weeks of treatment overall and for subgroup categories (treatment, site, circumcision status). Additionally the number of individual IELTs that are ≤30 seconds, >30 seconds-≤1 minute, >1 minute-≤2 minutes, and >2 minutes will also be summarized descriptively over the 4 weeks of treatment.

Descriptive statistics will include sample size, number of missing observations, counts or mean values and standard deviations for discrete and continuous variables, respectively.

Primary objective

The primary response variable is success on the PEBEQ Item 3 (event-specific bother), where success is defined *a priori* as having a 1-point or greater improvement between the mean response over the treatment period and the mean response during the baseline period. Success will be modelled using a generalized linear mixed effects model with treatment, circumcision status and baseline mean PEBEQ Item 3 (event-specific bother) score as fixed effects and site as a random effect. The model assumes a binomial family with logit link. The null hypothesis of no treatment effect will be tested after adjustment for these other covariates in the usual manner. From this model, estimated odds ratio of PSD502 to placebo and the corresponding 95% confidence interval will be calculated. The distributional assumptions of the model will be evaluated to ensure a satisfactory fit prior to statistical estimation and testing.

Additionally, the treatment effect for each circumcision status will be evaluated by adding the treatment-by-circumcision status interaction to the above model. From this model, estimated odds ratio of PSD502 to placebo and the corresponding 95% confidence interval for each circumcision status will be calculated.

Secondary objectives

Secondary response variables are listed in Section 4.1.2. Secondary objectives include evaluation of the treatment effect on the secondary responses of the study, regardless of the

primary response, and also evaluation of the relationship between the primary response to each of the secondary responses. Each of these will include evaluation of circumcision and site effects. In some cases subject effects will be included, as described below.

PGI-C and PGI-C-Bother are measures of change. The median treatment effect on the PGI-C and on the PGI-C-Bother will each be computed using a non-parametric test adjusted for covariates[38] of site and circumcision status. A linear model will be fitted with site and circumcision status as covariates.

A linear mixed effects model will be used to evaluate treatment effect on IELT during the study. The fixed effects are period (baseline or treatment period), treatment, period-by-treatment interaction and circumcision status and the random effects are site and subject within site. From this model, the mean difference between PSD502 and placebo in the change from baseline will be estimated and the corresponding 95% confidence interval calculated.

The same analysis will be conducted on PGI-S, Independent NRS Bother Item, and IEQI.

Scoring for IPE is shown in Appendix 8.7, Section 8.7.2, which results in three Domain Scores per subject. A linear mixed effects model will be used to evaluate treatment effect on the change from baseline in each of the IPE domains. The fixed effects are treatment, circumcision status and baseline, while site is the random effect. From this model, the mean difference between PSD502 and placebo in the change from baseline will be estimated and the corresponding 95% confidence interval calculated.

In each case with the secondary objectives, the distributional assumptions of the model will be evaluated to ensure a satisfactory fit prior to statistical estimation and testing.

Exploratory objectives

Exploratory efficacy variables include the:

- PEBEQ Multi-Item Event Specific Subscale Total
- PEBEQ Multi-Item General Subscale Total

A linear mixed effects model will be used to evaluate treatment effect during the study. The fixed effects are period (baseline or treatment period), treatment, period-by-treatment interaction and circumcision status and the random effects are site and subject within site. From this model, the mean difference between PSD502 and placebo in the change from baseline will be estimated and the corresponding 95% confidence interval calculated.

Further analyses as suggested by the data may also be undertaken.

5.2.3.9 Psychometric Analyses

Primary objective

The evaluation of test-retest reliability will be made using the intra-class correlation coefficient (ICC) using a two-way mixed effect model with absolute agreement for single

measures. These analyses will be conducted using one pair of successive valid diary entries per subject where subject indicated no change from the PGI-S on the second sexual encounter. The primary analysis will use diary entries for sexual encounter 1 and sexual encounter 2. If the subject indicates having a change between sexual encounter 1 and sexual encounter 2, then that subject's diary entries for sexual encounter 2 and sexual encounter 3 will be used for the analysis; or if there is a change between 2 and 3, then diary entries for sexual encounter 3 and sexual encounter 4 will be used.

Convergent validity (demonstrating that different measures of the same concept substantially correlate when assessed concurrently) will be evaluated by examining magnitude of correlations between the PEBEQ Item 3 (event-specific bother) and the Independent NRS bother item. Convergent validity would be supported when the PEBEQ Item 3 (event-specific bother) are substantially correlated ($\geq$0.40).

Known-groups validity (the extent to which scores from a measure can discriminate between groups of participants that differ on a known relevant dimension, such as a measure/assessment of disease severity will be examined using an analysis of variance (ANOVA) model to ensure adequate variability in the bother grouping variables based on (1) IEQI, (2) PGI-S and (3) groups defined by the IELT $\leq$30 seconds, >30 seconds-$\leq$1 minute, >1 minute-$\leq$2 minutes, and >2 minutes).

Responsiveness to change of the PEBEQ Item 3 (event-specific bother) will be examined between the mean response over the treatment period and the mean response during the baseline period. Effect size estimate of change ([Mean $_{\text{treatment period}}$ - Mean $_{\text{baseline}}$] ÷ standard deviation (SD) $_{\text{baseline}}$), standardized response mean ([Mean $_{\text{treatment period}}$ - Mean $_{\text{baseline}}$] ÷ SD $_{\text{change}}$), standardized mean change difference ([Mean change $_{\text{PSD502}}$ - Mean change $_{\text{placebo}}$] ÷ SD $_{\text{baseline pooled}}$), and Guyatt's statistic ([Mean $_{\text{improved/worsened group in treatment period}}$ - Mean $_{\text{improved/worsened group at baseline}}$] ÷ SD $_{\text{change in stable group}}$) will be used in the evaluation of responsiveness. For each type of effect size, a larger absolute value always indicates a stronger effect. Cohen (1988) defines effect sizes as: small (0.20), medium (0.50) and large (0.80)[39].

Anchor-based assessments of change will also be evaluated. Differences in the PGI-S and IELT will be used to define subjects as responders (decrease in severity at end of study versus those that did not change or worsened at end of study) and descriptively compared to changes in PEBEQ Item 3 (event-specific bother). Similar analyses will be performed with PGI-S replaced with the other anchor scales.

To evaluate the consistency of effects across the entire distribution, a cumulative distribution plot (based on the cumulative distribution function [CDF]) will be generated showing the change from baseline by cumulative percent of subjects. This distribution curve will reveal the extent to which overall results are driven by outliers who improve or worsen more than others.

5.2.3.10 Safety Analysis

Study drug exposure

The length of exposure and number of doses of study drug used during the treatment period will be summarized descriptively.

Adverse events

The incidence of AEs occurring during the treatment period will be summarized by preferred term and body system for each treatment group. Similar summaries will be performed for treatment-related AEs, severe AEs, SAEs and AEs that lead to withdrawal.

The incidence of treatment-related AEs starting on the same day or the day after use of the study spray will also be summarized by body system and preferred term for each treatment group.

Adverse events reported by the subjects' partners will be listed and summarized descriptively as for the subject AEs.

Laboratory variables

Absolute values and changes from baseline at the Final Evaluation Visit in laboratory measurements will be summarized for each treatment group.

Concomitant medications

The incidence of concomitant medications taken during the treatment phase will be summarized for each treatment group.

Vital signs

Absolute values and changes from baseline at the Final Evaluation Visit in vital signs measurements will be summarized for each treatment group.

Physical examination

The number and percentage of subjects experiencing physical examination abnormalities will be summarized by treatment group.

12-lead electrocardiogram

The number and percentage of subjects with abnormal ECG findings at the Final Evaluation Visit will be summarized for each treatment group.

Visual examination of the glans penis

The number of subjects with abnormal findings on visual examination will be summarized for each treatment group.

5.2.3.11 Telephone Exit Interview

Qualitative analysis on the telephone exit interview findings will be conducted by HRA.

5.2.3.12 Amendments to the Statistical Analysis

Any changes to the planned statistical and/or psychometric analyses made prior to the database lock will be described in the SAP and/or PAP, along with the rationale for the changes, and noted in the final clinical study report.

Any changes made after database lock will be described in the final clinical study report, along with the rationale for the changes and a discussion of the potential biases in the analysis.

Statistical and psychometric analyses will be conducted according to separate plans (a SAP and PAP respectively) that will be finalized before the database is “locked”.