Statistical Analysis Plan

Phase II Trial of Tesamorelin for Cognition in Aging HIV-Infected Persons

NCT02572323

05/19/2020

STATISTICAL CONSIDERATIONS

General Design Issues

This is a phase 2, randomized, controlled, open-label two-parallel-groups-study. A total of 100 participants are randomized 3:2 to active (Period A: baseline to 6 months) versus. deferred treatment (Period B: 6 to 12 months) with tesamorelin (Egrifta) 2mg subcutaneous daily injection. The participants are HIV+ individuals with abdominal obesity, well controlled HIV infection, and a minimum level of neurocognitive impairment on cognitive tests. The primary analysis, as detailed below, compares changes in neurocognition at the end of Period A between the two groups. Secondary analyses include within-group changes in neurocognition during tesamorelin treatment in the immediate and delayed treatment arms, and for both arms combined.

Analysis of Baseline Data

Evaluation of Baseline Characteristics

Baseline demographic, medical, and laboratory characteristics of participants will be summarized overall and by treatment arm using N (%) for categorical variables and mean, standard deviation, and range for numeric variables. After checking for skewness and kurtosis and applying appropriate transformations, means, standard deviations and ranges will be used to summarize baseline systemic biomarkers of immune activation and neuroimaging outcomes. Fisher's exact test (FET) for categorical variables and Wilcoxon rank-sum test (WRST) for numeric variables will be used to compare these values between treatment arms, with the exception of neuroimaging measures. They will be compared between treatment arms using linear regressions that will include relevant measures of brain size or composition as covariates, and scanner, since these values are known to influence variability of MRI outcomes.

Laboratory abnormalities

The subjects with laboratory abnormalities will be tabulated and summarized overall and by treatment arm. The number of subjects with serious abnormalities will be compared between treatment arms using FET.

Randomization, Enrollment, Drop-out, and Treatment Discontinuation

Randomization, study enrollment, and drop-out

The number (%) of subjects randomized, N (%) of subjects completing Period A (week 24 visit) and Period B (week 48 visit), and the number (%) of patients lost to follow-up will be summarized for the entire study and by treatment arm, and compared between treatment arms using FET. For lost-to-follow-up participants we will also record dates and reasons (if known) for loss to follow-up. Baseline demographic and medical characteristics, treatment assignment distribution, and the prevalence and timing of adverse events, will be compared between dropouts and completers using FET and WRST for categorical and numeric variables, respectively. The comparison of the timing of adverse events will be done using the log-rank test.

Treatment adherence and discontinuation

The number (%) of subjects who discontinue study treatment permanently, dates and reasons for treatment discontinuation will be summarized overall and by treatment

arm. The iTab adherence measures will be summarized and by study period; these include: the proportion of successfully administered daily injections; the number (%) of subjects requiring motivational interviewing (MI) visits, and the number of MI per subject.

Adverse Events and Treatment Safety and Tolerability

The number (%) of subjects with adverse events will be tabulated and summarized overall and by treatment arm, with the grade of the event included. The study week and reasons for adverse events will be included. The number of subjects with severe adverse events will be compared between treatment arms, at week 24 and week 48, using FET.

Outcome Measures

Primary Outcome Comparisons

Study Aim #1

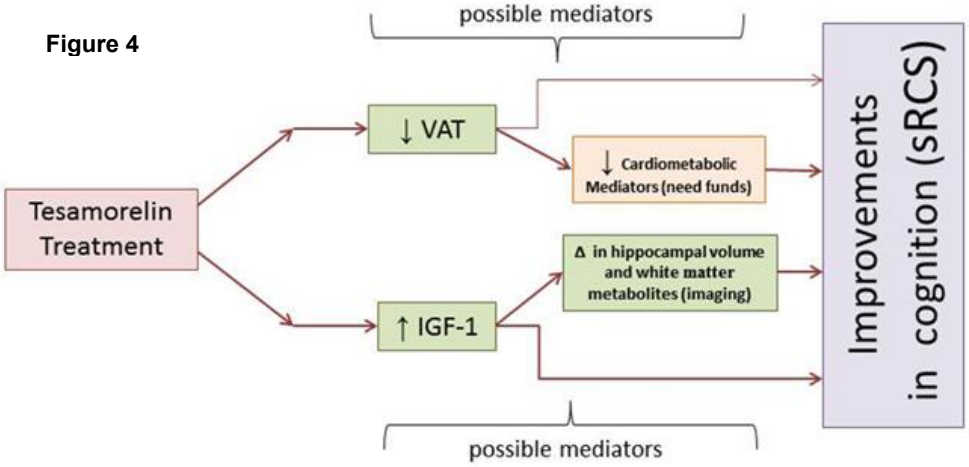
The primary study measure is the change from baseline in neurocognitive performance. This will be evaluated in each participant via the summary regression change score (sRCS) from baseline to week 24. The sRCS measures the change in neurocognitive performance relative to that of a large neurocognitively normal (control) group comparable on demographic characteristics. Higher sRCS values are associated with better neurocognitive performance. Improvement (yes or no) in cognitive testing from baseline to week 24 will be calculated for each participant based on the top 10% cut-off of sRCS values in the neurocognitively normal control group, from previously developed norms [1].

The mean sRCS at 6 months of treatment will be compared between the tesamorelin and no treatment during Period A using the linear model (ANCOVA), adjusting for site (UCSD VS. USC). The analysis will use all randomized subjects, under intent-to-teat (ITT) paradigm. For subjects with missing 6-month endpoint, these will be multiply imputed from the lower (worse) half of the distribution from their arm.

Study Aim #2

A conceptual Path model has been developed for testing (Figure 4 below):

Figure 4



A path analysis will be performed to analyze the mediation effect of systemic immune activation and neurodegeneration on the mechanistic pathway of tesamorelin to

neurocognitive outcomes (see **Fig. 4**). Changes in immune activation markers at week 24 will be compared between Tes and no treatment using t-tests for independent groups. IGF-1 markers will be regressed on treatment and immune activation markers. Finally, sRCS outcomes will be regressed on IGF-1 values and treatment arm. Specifically, the mediation effects will be evaluated using bootstrap methods [2]. Additionally, we will explore interactions used to establish mediation effects. Multivariable linear models will regress changes in measures of CNS immune activation on treatment arm and on measures of systemic immune activation. Additionally, we will explore interactions between treatment and systemic immune activation, which will measure how responses to treatment are affected by varying levels of these biomarker covariates. Baseline and change values of systemic immune activation values will be analyzed separately. Prior to all parametric analyses, appropriate assumptions will be tested and non-parametric alternatives or power transformations will be used. All comparisons are two-sided, at level α=0.05.

Secondary Outcome Measures

Neurocognitive comparisons between arms

The following testing will be performed. 1. Mean sRCS values at week 24 will be compared between the tesamorelin and no treatment groups of all randomized participants, using the independent samples T-test, under the following scenarios: (a) as treated (AT), with all subjects for whom sRCS is available included in their allocated group; (b) on-study treatment (OST), including only subjects for whom the week 24 sRCS is available and who remained on tesamorelin study treatment for the entire primary follow-up period. 2. The proportion of sRCS improvers will be compared between the two groups using logistic regression under the ITT, AT, and OST scenarios, as described above. Additional multivariable secondary analyses will control for demographic and baseline clinical characteristics that are significant at the 0.15 level in backward model selection.

9.4.2.2. Site effects on treatment

We will test for differential treatment effects between the two sites, UCSD and USC, using a 2x2 ANOVA model with effects for treatment arm, site, and their interaction on the sRCS at week 24. (Site effects may be due to demographic differences between sites, or to the fact that the primary outcome assessors are blinded at UCSD, but not at USC.)

9.4.2.3. Changes in neurocognition over time

The change in neurocognition during tesamorelin treatment will be evaluated within each treatment arm and for the two arms combined, as follows: i) for the immediate treatment arm the mean sRCS from baseline to end of Period A will be compared to a mean of 0, using the one-sample T-test. Ii) Similarly, the mean sRCS will be evaluated in the delayed arm during Period B (tesamorelin treatment) and compared to a mean of 0. Iii) Finally, the mean sRCS will be evaluated and compared to 0 for the two arms combined (baseline to week 24 for immediate treatment arm and week 24 to week 48 for the delayed treatment arm). 95% confidence intervals will be reported in each case.

Additional Secondary Analyses

An additional secondary analysis will be done using the Principal Stratification framework of Frangakis and Rubin [3]. This analysis compares the treatment arms separately for four subsets determined based on whether the subjects would complete the study on any of the two study treatments or not. Accordingly, the four sets are:

completers on both Tes, no treatment group (those randomized to deferred therapy for Period A; non-completers on both Tes and no treatment, completers on Tes but not in the no treatment group; completers on no treatment but not on Tes. In practice data are available on only one of the two statuses, and the second status is effectively imputed statistically. Groups with less than five members may be ignored. This innovative analysis allows separate evaluation of effects of Tes for those who can and those who cannot tolerate Tes and the daily injections, and will give insights into differential effectiveness of Tes.

ample Size and Accrual

Power analysis: We will randomize N=100 subjects total in a 3:2 ratio, that is, 60 subjects to Arm A and 40 to Arm B. With a 20% dropout rate we anticipate 48 and 32 completers at week 24 in the immediate and delayed tesamorelin treatment arms, respectively. For the primary analysis (Aim 1) we have at least 80% power to detect an effect size (Cohen's D) = 0.647. The secondary analyses have 80% power to detect an effect size (Cohen's D) = 0.647 (AT analysis) and 0.699 (OST analysis, assuming a further treatment discontinuation in 15% of study subjects). .For binary outcomes the detectable difference between groups is of 33 percentage points (e.g., 20% vs. 53%). Multiple regression will have 80% power to detect a partial correlation coefficient R= 0.304 or larger between the outcome and the predictor of interest, e.g. between sRCS and plasma IGF-1 levels. For the within-group evaluation we have 80% power to detect a mean sRCS corresponding to Cohen's d = 0.413 for immediate arm, d = 0.539 for the delayed treatment arm, and d = 0.323 for the two arms combined.

Monitoring

Interim Safety and Efficacy Review. Interim Safety and Efficacy Review. A Data Monitoring Committee (DMC) will review the study for safety yearly. The study will only be stopped early if there are serious safety concerns. After consultation with the DMC, it was mutually decided not to stop the study early for differences in efficacy based on the following considerations:

- Stopping has no treatment implications, since such limited data cannot/should not be used for ad hoc treatment by practitioners and could not be basis for requesting FDA approval for an expanded labeling indication for treatment. It would only prompt a bigger confirmatory study (e.g. phase III).
- The study is a reasonably small phase II study and not like larger phase III anti-retroviral studies where early benefit would have major treatment implications. Therefore, continuing the study is not an ethical issue.
- By stopping earlier, this limits the opportunity to learn if positive effects were due to 1) changes in hippocampal volume or CNS metabolites or 2) reductions in VAT or some combination of these effects.

Similarly, because validation of the proposed pathophysiological mechanisms through imaging and biomarkers is scientifically meritorious independent of statistically significant treatment benefits for neurocognitive function, there will be no futility analysis.

The following items will be included in the yearly DMC meetings:

- Enrollment and study status including,

  - Number of premature discontinuations of study treatment and reasons
  - Data on premature study discontinuations
- Compliance with study treatment
- Baseline characteristics
- Baseline laboratory values
- Safety and treatment toxicity including,
  - Summary of grade 3 and 4 SAEs
  - Summary of unexpected grade 1 and 2 SAEs