Extended Release Naltrexone for Opioid-Dependent Youth

NCT01843023

Statistical Analysis Plan for Aims 1 and 2

Study Protocol Version 14

IRB approved May 27, 2016

Submitted to ClinicalRrials.gov PRS March 27, 2020

Statistical Analysis Plan for Aims 1 and 2 (from the protocol, version 14)

1.1. Analyses for Aims 1 and 2

Aim 1: To determine the relative effectiveness of XR-NTX compared to TAU for opioid-dependent youth in terms of the following 3- and 6-month outcomes: a) days in treatment; b) opioid use; c) other drug (cocaine and marijuana) and alcohol use; d) criminal behavior and arrests; e) relapse to DSM-5 defined opioid use disorder.

Hypothesis 1A: Participants in XR-NTX Condition will be retained in treatment longer than those in TAU.

Hypothesis 1B: Participants in XR-NTX will have lower rates of heroin and prescription opioid and other drug use, alcohol use, criminal behavior and arrest, and opioid dependence criteria.

Aim 2: To examine the impact of XR-NTX on HIV drug- and sex-risk behaviors at 3- and 6-month follow-up, and HIV infection status at 6-months.

Hypothesis 2: Participants in the XR-NTX will be less likely to use other drugs and will therefore report lower HIV drug- and sex-risk behaviors.

Outcome Measures. Outcome variables will be of three distinct types: 1) discrete random variables (e.g., number of days in treatment), assumed to follow a Poisson distribution; 2) dichotomous variables (e.g., urine drug test results for opioids and other drugs), assumed to follow a binomial distribution; and, 3) continuous random variables (e.g., RAB scale scores), assumed to follow a normal distribution.

Explanatory Variables. The predictor variables in all statistical models can be categorized as either *Treatment Variables* or *Control Variables*.

Treatment Variables. There will be a single treatment variable: Treatment Condition (XR-NTX vs. TAU).

Control Variables. Two additional predictor variables – participant age and gender – will be included as "main effects" in all analyses in order to examine for potential differences in treatment outcome as a function of these two participant characteristics. The interaction between Treatment Condition and each of these participant characteristics will also be estimated and tested for significance. These interaction effects test the extent to which responsiveness to Treatment Condition varies as a function of age and/or gender.

Time. Finally, the "repeated factor" in the statistical analysis of all outcome variables measured repeatedly will be assessment Time point, which will allow for the evaluation of both differential course and impact of the interventions as a function of the "between-subjects" Treatment Group factor. (For a summary of assessment time points for each outcome, see Figure 1, above.).

Intent-to-Treat Approach. All analyses will be conducted on available study-related data from all participants, regardless of whether or when they drop out of treatment.

Statistical Method. A Generalized Linear Mixed Model (GLiMM) will be used to conduct all analyses.

Power

To acknowledge the possibility that there might be 2 equally-spaced interim analyses, α will be set equal to .0379, based on the O'Brien-Fleming spending function,[1,2] for the test of all effects. Power for the primary outcome of number of days in treatment, assumed to follow a Poisson distribution, does not involve a Time effect, as it is measured at the

conclusion of the 6-month treatment period. Assuming the base rate (β) for TAU and odds ratio for XR-NTX of 1.15, power ($1 - \beta$) = .80. This odds ratio implies a 15% difference in number of days retained in treatment between the two treatment conditions.

Stroup [3,4] has outlined a comprehensive four-step procedure to estimate power that can be applied to GLiMMs. This procedure was implemented in the current case, and involves estimation of power using a variety of covariance structures for the Time (T) effect. Heterogeneous covariance structures were defined in all three cases, in which the variance component at T1 (baseline) was arbitrarily set at 1, at T2 (3-month follow- up) was specified as 90% of the variance of Time 1, while at T3 (6-month follow-up) was specified as 95% of the variance of T2. This pattern reflects the frequent occurrence that groups become more homogeneous over time following an intervention. Correlation over Time was specified as .6 for the compound symmetric heterogeneous structure. The correlations of adjacent time points were set to .6 for the first-order autoregressive structure, thereby setting the T1–T3 correlation to .36. For the UN case, the T1-T2 correlation was specified as .8, T2-T3 as .7, while T1-T3 as .6. This pattern was chosen to acknowledge the change in variances over Time (which thereby reduces the corresponding correlations). Second, two datasets were created, assuming the dependent variable was (1) continuous and normally distributed or (2) binary and distributed binomially, respectively. In each dataset (1), the mean of the TAU condition was set at 1, .8, and .9 at T1-T3, respectively, indicating a small treatment effect associated with TAU at 3 months, then with an expected decrement in success at 6 months, while the means for the XR-NTX condition was set at 1 at T1, and then set to .5 at 3 months, indicating a medium treatment effect, with the mean set to .65 at T3 to reflect a small-to-medium effect, again taking into account an expected decrement. Similar procedures were followed for dataset (2). The only non-null effect in each dataset was the hypothesized effect of interest. Thus, simulations were conducted under what might be considered "worst-case scenarios." Finally, observations were dropped from the data consistent with the expected loss of approximately 10% of participants at T2 and an additional 5% at T3 (see **Follow-up and Attrition**, in Protocol version 14). Therefore, in each dataset, 34 (approximately 10%) of the observations were chosen at random and their observations at T2 and T3 were dropped. An additional 17 of the observations at T3 were chosen at random and dropped. Hence, power was estimated for a design in which the effects were unbalanced in a manner similar to what is expected in the proposed research. Resulting power to test all hypotheses exceeded .9 in most simulations, except, not surprisingly, for an unstructured covariance matrix, for which power fell into the low .80s.

From a more rudimentary perspective, power calculations based on the set correlation method [5,6] for a multivariate multiple-groups profile analysis can be used to calculate reasonable estimates of power for the primary outcome measures under the assumption they are normally distributed, with the estimates likely to be conservative in the case of Poisson or binomial outcome variables. Under the extremely limiting assumption that no other effect in the model was significant, and assuming N = 289 [= 340 – 51 (15% of the sample) to conservatively account for attrition] and α = .0379, an effect size in the population of $\mathbf{f}^2$=.0365 associated with Treatment Condition X Time interaction effect would yield a power of .8 for that effect. This effect size falls in the "small" range, with $\mathbf{f}^2$=.02 considered a "small" effect and $\mathbf{f}^2$=.15 a "medium" effect.[6] In other words, and imprecisely, under the assumption that the effect in the population was ≥ .0345 (or, alternately, that the population semi-partial r^2 associated with the Treatment Condition X Time effect was ≥ .0351), there is 80% chance of concluding that effect is significant if α is set to .0379 and 289 participants complete the trial.

1.2. Supplementary Analyses

Various follow-up and ancillary analyses will be considered, based on the results of the analyses of the primary outcome variables. In addition to the supplemental analyses mentioned above regarding treatment preference, treatment expectations, and the number of counseling sessions received, there are three sets of additional analyses that we intend to conduct. First, as noted in Study Design Issues (above), we intend to re-fit our statistical models, including Primary Drug of Abuse (heroin v. prescription opioids) and its interaction with Treatment Group to examine, on a post hoc basis, whether Primary Drug of Abuse moderates treatment effectiveness. Second, examination of the use of illicit drugs other than opioids, particularly cocaine and marijuana, at treatment entry may help to explicate the extent to which the use of such other drugs plays a role in treatment outcome. Third, it may be important to examine number of days in treatment from the perspective of a survival model, because such an analysis may lead to a better understanding of time-dependent barriers that lead to early dropout from treatment (e.g., if there is decided attrition in the XR-NTX group after three months, it would suggest an examination of treatment impediments that might exist following 90 days of treatment).

Addendum (11/12/2019)

This addendum updates the preliminary approach for data analysis articulated in the study protocol.

Primary Outcome. Consistent with the pre-registered trial information in clinicaltrials.gov on April 3, 2013, the primary outcome is defined as the number of days of opioid use at the 6 month follow-up. The change in primary outcome from the "number of days in treatment" to the "number of days of opioid use" was necessitated by the complexity in defining "days in treatment" in a consistent manner across the two study conditions. This difficulty stemmed from the differences between mode of administration of XR-NTX (monthly injection) and buprenorphine (daily sublingual administration) and the fact that all participants were planned to receive behavioral treatments. The other outcomes will be examined secondarily.

Final Sample Size. The final sample size for the study is 288.

Adherence to randomized condition and implications for interpretation. There was poor adherence to receiving the randomly assigned condition in the residential treatment program as planned. This occurred for a variety of reasons (e.g., patient and caregiver preferences, barriers) and persisted despite the best efforts of the research and clinical teams. There were participants who were randomized to XR-NTX but did not receive it prior to discharge (101/144), and there were participants who were randomized to TAU but changed their mind prior to discharge from the residential unit and decided to begin XR-NTX prior to discharge. (39/144). It was determined early in the study that participants would continue to be followed as planned on an intention-to-treat basis, and that outcomes would be examined two ways: First, based on which arm participants were assigned to (the original intention of the study, and the information reported within clinicaltrials.gov), and second, based on what treatment they were actually receiving at time of discharge from the treatment facility (which represents the main analysis of interest in disseminating findings in the scientific literature because it corresponds to actual exposure to the specific medication). For this reason, great caution is urged in interpreting outcomes based on study condition as assigned, as study condition did not accurately reflect what medication a given participant actually received.

Analytical Approach. Analyses "as-randomized" will be conducted on available study-related data from all participants, regardless of whether or when they drop out of treatment. Analyses "as-received" will use all available data but will analyze outcomes based on what treatment a participant actually received prior to discharge from the residential unit rather than on the condition to which s/he was randomized.

Statistical Method. A Generalized Linear Mixed Modeling (GLiMM) approach will be used to conduct the main analyses. Choices between different candidate models (e.g., family and link function; covariance structure) will be based on the distributional properties of the outcome variable, conceptual considerations, and metrics of model fit. In the case of secondary outcomes measured as endpoints, traditional regression approaches may be used (e.g., logistic regression to analyze discrete urine test results at 3-and 6-month follow-up; Poisson or negative binomial regression to examine count data).

References

1. Lan KKG, Demets DL. Discrete sequential boundaries for clinical trials. *Biometrika.* 1983;70(3):659-663.

2. Reboussin DM, DeMets DL, Kim KM, Lan KK. Computations for group sequential boundaries using the Lan-DeMets spending function method. *Control Clin Trials.* 2000;21(3):190-207.

3. Stroup WW. Mixed model procedures to assess power, precision, and sample size in the design of experiments. Lincoln, NE: University of Nebraska, American Statistical Association; 1999:19-24.

4. Littell RC, Millike GA, Stroup WW, Wolfinger RD, Schabenberger O. *SAS for mixed models.* Second ed. Cary, NC: SAS Institute, Inc.; 2006.

5. Borenstein M, Cohen J. *Statistical Power Analysis.* Hillsdale, NY: Erlbaum; 1988.

6. Cohen J. *Statistical power analysis for the behavioral sciences (Second edition)* Hillsdale, NJ: LEA; 1988.