Limbic pallidum DBS for the treatment of severe alcohol use disorder

Protocol Number:	STUDY22030036
Version Date:	1/31/2024
Version Number:	8
IDE Sponsor:	Khaled Moussawi, MD, PhD
Principal Investigator:	Brian Coffman, PhD
IDE Number:	G220077
NCT Number:	05522751
Source of Funding:	National Institutes of Health (NIH)

12 Statistical Analysis

12.1 General Approach

The general analysis plan begins with data description, including basic summary statistics such as means, standard deviations, correlations, cross-tabulations, and checking for outliers and missingness patterns. Data will be evaluated for normality through goodness-of-fit tests and q-q plots. If systematic deviations from normality are detected, the corresponding variables will be appropriately transformed using logarithm or other Box-Cox transformations.

We will examine spaghetti plots for visualization of the individual trajectories over time for the longitudinal hypothesis. The analyses for the longitudinal data will consist of computation of summaries across time and at each time point. For any fitted models, we will evaluate uncertainty in the model parameter estimates and assess goodness of fit for each model using standard techniques such as residual analysis and regression diagnostics.

For testing hypotheses we will use standard methods such as the t-test; we will use the Wilcoxon and chi-square tests when appropriate. With small sample size, we will use exact methods to compute the significance level. For testing proportions, we will use the standard arcsin transformation of proportions to stabilize variance and improve the normal approximation.

For each analysis discussed, we will also consider models that control for socio-demographic descriptors for the individuals in the study when available.

Missing data:

We will enroll a replacement participant if withdrawal occurs. To account for missing data, we will use inverse weighted probability analyses and imputation-based methods, and conduct sensitivity analysis to assess the stability of conclusions across analyses.

12.2 Sample Size Determination

Since this study is primarily designed for the assessment of safety and feasibility, a cohort of 3 patients was decided upon per clinical judgement and experience.

12.3 Analysis of Primary Endpoints

The primary outcomes are:

1) *Safety and tolerability* measured as incidence of any serious adverse effect (AE), and incidence rate and severity of AEs that are definitely, probably, or possibly related to LP DBS throughout the study duration.
2) *Feasibility* measured as recruitment rate and completion rates of follow-up visits and assessments throughout the study duration.

The primary outcomes will be assessed continuously throughout the study.

H1. LP DBS for severe AUD will be deemed safe and well-tolerated by DSMB.

H1.a. LP DBS for severe AUD will not result in any serious adverse effects.

H1.b. The incidence rate of AE will be similar to standard of care DBS per the DSMB judgement.

H2. LP DBS for severe AUD will be deemed feasible.

H2.a. We will successfully recruit 3 participants for the study.

H2.b. Participants will complete >80% of all study follow-up visits and assessments.

Safety, tolerability, and feasibility will be assessed continuously throughout the study. We will screen for AEs using 4 approaches: 1) patient self-report, 2) clinical evaluations (medical, neurologic, and psychiatric), 3) questionnaires (e.g., screening for depression, suicidality, anxiety, apathy, mania) and 4) neuropsychological evaluation to screen for any LP DBS effects on various cognitive domains. Every AE will be evaluated for severity, expectedness and relatedness, and reported to the DSMB.

Safety and tolerability: The study will be deemed safe and tolerable if no serious *AE* that are definitely, probably, or possibly related to LP DBS occur during the study and if the AE severity and rate are not higher than standard of care DBS per DSMB's assessment.

Feasibility: the study will be deemed feasible if we successfully recruit the proposed number of participants and if participants complete >80% of all study follow-up visits and assessments.

12.4 Analysis of Secondary Endpoints

The secondary outcomes are:

1) *Alcohol use* measured as PDA, DDD, PEth, GGT at baseline and monthly after DBS implantation.
2) *Overall functioning* and disability measured with the WHODAS 2.0 scale at baseline and monthly after DBS implantation.
3) *Target engagement* measured with FDG-PET scans at 3 time points: pre-DBS baseline, post-DBS-OFF, and post-DBS-ON at 6 months.

The secondary outcome *PDA* as a measure of *alcohol use* is a proportion, collected at baseline and during follow-up and assessment visits (1 baseline and 12 monthly post-DBS measurements).

The other measures of alcohol use (*DDD, PEth, GGT*) and the overall functioning score *WHODAS 2.0* are continuous variables collected at baseline and during follow-up and assessment visits (1 baseline and 12 monthly post-DBS measurements).

Alcohol use and overall functioning

H3: LP DBS reduces alcohol use and improves overall functioning in subjects with severe AUD compared to baseline. Specifically, LP DBS increases PDA, and decreases DDD, PEth, GGT and WHODAS 2.0 score compared to baseline.

Primary analysis of H3: We will compare the baseline and 6-month time point using a paired t-test for each of the measures (PDA, DDD, PEth, GGT and WHODAS 2.0).

For PDA, the unit of analysis is the proportion of days X_{ij} in a month that participant i is abstinent during month j. To do the paired t-test, we use the two paired proportions X_{i0} (0 refers to the baseline measure) and X_{i6} (6 refers to the 6-month measure). We can use a t-test on the difference $D_i = (X_{i6} - X_{i0})$. This should be a reasonable approximation because the averages X over roughly n = 30

days should be approximately Gaussian. The Gaussian approximation can be improved by using the variance stabilizing transformation

$$Y_{ij} = 2arcsin\sqrt{X_{ij}}$$

Thus, we will do a t-test using the differences $2[arcsin\sqrt{X_{i6}} - arcsin\sqrt{X_{i0}}]$. With n = 3 participants, we have little power unless the effect is large (e.g., if PDA moves from 0.05 to 0.50).

Secondary analysis of H3: To assess the course of PDA, DDD, PhEt, GGT, and WHODAS 2.0 over the duration of the study while taking into account the repeated measures (13 measures for each participant – one at baseline and 12 post-DBS), we will use a **mixed-effects model**:

$$Z_{ij} = \mu + \alpha_i + \tau_j + \epsilon_{ij}$$

with random effect α_i for subject, a fixed time effect τ_j, and the error term (unexplained variation) ϵ_{ij}. We will have 39 measures (3 subjects times 13 measures per subject at months 0 to 12). For PDA, we will use the arcsin-transformed proportions described above to improve Gaussian approximation by stabilizing the variance.

Target engagement

H4: LP DBS will result in circuit specific target engagement measured with FDG-PET.

H4a. CGM (cerebral glucose metabolism) as measured with [^{18}F]FDG-PET in pre-defined cortico-limbic ROIs will be increased in post-DBS-ON compared to pre-DBS and to post-DBS-OFF conditions.

H4b. DBS-induced increases in CGM (%) at the 6-month time point will be associated with decreased alcohol use (PDA, DDD).

Analysis of H4(a): For each ROI in each of the scan types we will run a one-way Repeated Measures (RM) ANOVA comparing CGM between the 3 time points (baseline pre-DBS, DBS-OFF, and DBS-ON at 6 months), with post-hoc comparisons of DBS-ON vs. post-DBS-OFF, and DBS-ON vs. baseline.

Analysis of H4(b): We will use Pearson's correlation to assess the relationship between changes in CGM between post-DBS-OFF and 6 months and corresponding changes in measures of alcohol (PDA, DDD).

12.5 Planned Interim Analysis

Not applicable.

12.6 Exploratory Analysis

Cognitive and behavioral tasks

The outcomes of the cognitive and behavioral tasks (Cue Reactivity (CR), Card Guessing Task (CGT), Delay Discounting task (DD), Balloon Analog Risk Task (BART), and Stop Signal Reaction Time task (SSRT)) are collected at baseline and at the 6-, and 12-month time points. These outcome measures include:

- craving, positive affect, negative affect, calm, and excitement scores for CR
- β- and γ-spectral power for CGT
- discounting rate (k) for DD
- adjusted pumps for BART
- SSRT

H5: LP DBS will normalize reward processing (reduce baseline and increase event-related spectral power in the β- and γ-range) and reduce alcohol CR, temporal discounting, risk taking, and motor impulsivity post-DBS compared to baseline.

Analysis of H5: The outcome measures will be analyzed like the PET measure of target engagement measures, using one-way RM ANOVA between baseline values, DBS-ON at 6 months, and DBS-ON at 12 months, with post-hoc comparisons to the baseline.

Intra-operative neurophysiology

H6. Different neuronal populations in the LP encode rewarding vs. aversive stimuli and these neurons respond differently for alcohol vs. neutral cues.

H6.1: a population of LP neurons will show significant increase in neuronal firing during reward expectation.

H6.2: a different population of LP neurons will show significant increase in responding during loss expectation.

Analysis of H6.1 and H6.2: A 2-way repeated-measure ANOVA will be used to compare the firing frequency of the different epochs (first epoch represents baseline firing) in different trial types (win vs. neutral vs. loss) (dependent variable = firing rate, factor 1 = epoch, factor 2 = trial type), with post-hoc comparisons. In task responsive neurons, we expect to see significant epoch X trial type interaction, as well as significant effect of epochs and trial types.