THE PENS TRIAL

STATISTICAL ANALYSIS PLAN

NCT03350750

Protocol Title (Number):	The Placebo-controlled Effectiveness in INPH Shunting (PENS) Trial.
Protocol Date:	September 8, 2017
SAP Author:	Richard Holubkov, Ph.D.
SAP Version:	1.0
SAP Date:	May 14, 2018

CONFIDENTIAL

Approvals:

Approved By:

Name:	Mark Luciano, MD Principal Investigator for the PENS Trial	Date:
Name:	J. Michael Dean, MD, MBA Principal Investigator, University of Utah DCC	Date:

Table of Contents

ABBREVIATIONS

ABBREVIATION	DEFINITION
SAP	Statistical Analysis Plan
ITT	Intent-To-Treat Population
PP	Per-Protocol Population
FDA	United States Food and Drug Administration
DSMB	Data Safety Monitoring Board
SNP	Single-Nucleotide Polymorphism
PENS	Placebo-controlled Effectiveness in INPH Shunting Trial.
INPH	Idiopathic Normal Pressure Hydrocephalus
AHCRN	Adult Hydrocephalus Clinical Research Network
SDMT	Symbol Digit Modalities Test
BDI	Beck Depression Inventory, 2nd edition
ADL/IADL	Lawton Activities of Daily Living/Independence in Activities of Daily Living
MoCA	Montreal Cognitive Assessment Test
OAB	Overactive Bladder Questionnaire, Short Form
DCC	Data Coordinating Center
CSF	Cerebrospinal Fluid
CRF	Case Report Form

1. PREFACE

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for the Placebo-controlled Effectiveness in INPH Shunting (PENS) Trial.

The PENS Trial is a multi-center, randomized, placebo-controlled design investigation of Cerebrospinal Fluid (CSF) shunt surgery.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the FDA and International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH): Guidance on Statistical Principles in Clinical Trials.[1] All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association[2] and the Royal Statistical Society,[3] for statistical practice.

The following documents were reviewed in preparation of this SAP:

- PENS Protocol, Version Sep 08 2017
- Draft case report forms (CRFs) for the PENS Protocol
- ICH Guidance on Statistical Principles for Clinical Trials.
- Minutes of the PENS DSMB initial protocol review meeting of January 8, 2018

The reader of this SAP is encouraged to also read the clinical protocol for details on the conduct of this study, and the operational aspects of clinical assessments and timing for completing a patient in this study.

It is possible that, due to updates or identification of errors in specific statistical software discussed below, the exact technical specifications for carrying out a given analysis may be modified. This is considered acceptable as long as the original, prespecified statistical analysis approach is completely followed in the revised technical specifications.

2. PURPOSE OF SAP

The purpose of this SAP is to outline the planned analyses to be completed for the PENS trial, including analyses to be reported to the study DSMB. The planned analyses identified in this SAP will be included in future study abstracts and manuscripts. Also, exploratory analyses not necessarily identified in this SAP may be performed. Any *post hoc*, or unplanned, analyses not explicitly identified in this SAP will be clearly identified as such in any published reports from this study.

This SAP may be updated in response to additional developments, either within or outside of the PENS trials. In such circumstances, an updated version number will be assigned to the revised SAP. Previous SAP versions will be archived, and changes to each SAP revision tracked.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1 Study Objectives

The primary and secondary objectives of the PENS trial are described below.

3.1.1 Primary Objective

The primary objective of the PENS trial is to test the hypothesis that treatment of idiopathic normal pressure hydrocephalus (INPH) with an open shunt results in improved gait velocity, compared to placebo treatment with a closed shunt, at 4 months after shunt surgery.

3.1.2 Secondary Objective

The secondary objective of the PENS trial is to test the hypotheses that treatment of idiopathic normal pressure hydrocephalus (INPH) with an open shunt, compared to placebo treatment with a closed shunt, results in:

i. improved cognition at 4 months after shunt surgery, as assessed by the Montreal Cognitive Assessment Test (MoCA)

ii. improved bladder control at 4 months after shunt surgery, as assessed by the Overactive Bladder Questionnaire, short form (OAB)

3.1.3 Exploratory Objectives

The exploratory objectives of the PENS trial are as follows:

i. Compare changes in additional assessments of cognition, mood, and function [Symbol Digit Modalities Test (SDMT), Beck Depression Inventory, 2nd edition (BDI-II), Lawton Activities of Daily Living/Independence in Activities of Daily Living (ADL/IADL), and Modified Rankin Scale] after 4 months of treatment with shunt surgery versus placebo.

ii. Assess frequency of adverse events (all AEs, with specific examination of falls and surgical and non-surgical complications) associated with shunt surgery and compared between treatment arms

iii. In analyses combining treatment arms, evaluate clinical improvement after 8 months of active shunting with respect to the primary endpoint of gait velocity, as well as with respect to secondary and tertiary clinical outcome measures.

iv. Identify novel CSF biomarkers that differentiate patients with INPH from AD patients and other dementias, adjusting the typical cut-offs of traditional AD related biomarkers for altered volume of distribution.

v. Identify specific biomarkers associated with improved medium term (8 month outcomes) response to active shunting, enabling improved selection of participants in future trials.

vi. Identify the prevalence of common neurodegenerative disorders in those selected for the PENS trial vs those in the broader AHCRN registry and to discover if the lack of responsiveness in those receiving shunts at 12 months could be explained by the prevalence of any single nucleotide polymorphisms (SNPs) related to the above neurodegenerative disorders.

In this version of the SAP, the final three objectives have not been explicitly addressed. The analyses of specific biomarkers, SNPs, and degenerative disorders will be appropriate to include in an updated version of the SAP, after information is known on specific biomarkers and SNPs to be examined (specifics may evolve during the course of recruitment in the trial).

3.2 Study Endpoints

3.2.1 Primary Endpoint

The primary efficacy endpoint of the PENS Trial is the change in gait velocity from baseline (pre-shunt-surgery) assessment to an assessment performed at 4 months after shunt surgery.

At each evaluation timepoint, the gait velocity used for analysis is the best velocity occurring in up to three trials of a 10-meter walk test.

3.2.1.1 Eligibility for Primary Analysis

The primary endpoint is to be evaluated in all randomized patients who have both a valid pre-surgical assessment of gait velocity (required for randomization, as gait velocity ≤ 1 *m/s* is a study eligibility criterion) and a valid assessment of velocity 4 months after shunt surgery.

More specifically, pre-surgical gait velocity assessments will be considered valid if they occurred before the lumbar puncture (LP) assessing eligibility and no more than 6 weeks prior to shunt surgery. The gait velocity assessment that occurs closest to the lumbar puncture will be used in the primary analysis. If the shunt surgery is scheduled more than 6 weeks after the pre-LP gait velocity, further gait velocity assessments must be performed within 6 weeks of the shunt surgery, but not within 2 weeks post-LP. In this case, the gait velocity assessment closest to the shunt surgery will be used in the primary assessment. Post-surgical assessments of gait velocity will be considered valid if they occurred between 14 days prior and 30 days after each follow up time point. In the event that a patient has multiple follow-up assessments of gait velocity during a follow-up time window, the assessment closest to the exact time window (e.g., closest to the day exactly 4 months after shunt surgery) will be used in the primary intention to treat analysis.

3.2.2 Secondary Endpoints

The secondary efficacy endpoints of the PENS trial are:

1. Montreal Cognitive Assessment Test (MoCA) composite score.
2. Overactive Bladder Questionnaire Composite Symptom Score (OAB).

3.2.2.1 Eligibility for Secondary Analysis

All patients randomized in the PENS Trial with valid assessment for the secondary outcome are to be included in the analyses of the secondary endpoints.

3.2.3 Exploratory Endpoints

Tertiary endpoints of the PENS trial include:

i. Symbol Digit Modalities Test (SDMT)

ii. Beck Depression Inventory, 2nd edition (BDI-II)

iii. Lawton Activities of Daily Living/Independence in Activities of Daily Living (ADL/IADL)

iv. Modified Rankin Scale

v. Levels of CSF biomarkers **[clarify and specify]**

vi. Prevalence of common neurodegenerative disorders **[clarify and specify]**

3.2.3.1 Eligibility for Tertiary Analysis

All patients randomized in the PENS trial with valid assessments of appropriate outcomes will be eligible for analyses of tertiary efficacy outcomes (SDMT, BDI, OAB, Rankin Scale).

3.2.4 Safety Endpoints

The safety endpoints of the PENS trial are:

1. Frequencies of adverse events and serious adverse events
2. Rates of falls
3. Rates of complications, overall and classified as requiring versus not requiring neurosurgery.

3.2.4.1 Eligibility for Safety Analysis

All patients randomized in the PENS trial receiving a shunt will be eligible for safety analyses.

4. STUDY METHODS

4.1 Overall Study Design and Plan

The PENS trial will randomize eligible and consented patients to shunt surgery, using an FDA-approved programmable CSF shunt, the Certas Plus with Siphonguard (Codman, Johnson & Johnson, Raynham, MA). Patients will be randomized in a 1:1 ratio, to active (open shunt group) (setting 4) (110 mm H2O) or placebo (closed shunt group) (setting 8) (>400 mm H2O) shunt settings.

After 4 months, shunts for subjects in the closed shunt group will be adjusted to the active setting of 110 mm H2O. After this 4-month adjustment, all subjects in both groups will continue to be followed, for a total of 12 months after shunt surgery. During the remainder of follow-up, subjects will have shunt adjustments according to clinical standards at each center.

4.2 Selection of Study Population

Patients will be eligible for enrollment in PENS if they meet all of the following inclusion criteria:

A. age ≥ 60 years,

B. diagnosis of INPH based on clinical criteria and testing as described in the iNPH Guidelines (Relkin et al, Neurosurgery. 2005 Sep;57(3 Suppl):S4-16),

C. one positive supplementary test (infusion test, large volume LP or extended CSF drainage (Marmarou et al, Neurosurgery. 2005 Sep;57(3 Suppl):S17-28), and

D. duration of gait impairment ≥ 6 months.

Additionally, PENS patients must not meet any of the following exclusion criteria:

E. unable to walk 10 meters with or without an assistive device,

F. baseline gait velocity >1 m/s, with or without an assistive device,

G. unable to return to the study center for follow up evaluation and shunt programming,

H. not medically cleared for shunt surgery per local standards,

I. secondary NPH (prior encephalitis, meningitis, subarachnoid hemorrhage, traumatic brain injury (including concussion), brain abscess, brain tumor, obstructive hydrocephalus (including acquired aqueductal stenosis and carcinomatous meningitis)),

J. prior or existing shunts, endoscopic third ventriculostomy, or any other previous surgical intervention for hydrocephalus,

K. previous intracranial neurosurgical procedure,

L. current treatment with anticoagulation medications or expected to be on anti-coagulation medications in future based on clinician evaluation,

M. large cerebral or cerebellar infarction (asymptomatic lacunar infarctions are permitted),

N. hemiparesis, cerebellar signs or neurological deficits (e.g., cervical or lumbar myelopathy, previous stroke) precluding gait assessment,

O. diagnosis of Parkinson's disease,

P. diagnosed clinical depression,

Q. diagnosis of schizophrenia or any other psychiatric diagnosis which in the clinician's judgment will complicate the outcome evaluation,

R. sensory or functional deficit (e.g., uncorrectable severe visual or hearing impairment) that does not allow full clinical evaluation,

S. dementia, documented with a MOCA score of 21 or less, taken at standard initial evaluation, or

T. conditions impairing gait that are considered to be unrelated to hydrocephalus, such as hemiparesis, spasticity, cerebellar ataxia or musculoskeletal and joint disease.

4.3 Method of Treatment Assignment and Randomization

Randomization to open versus closed shunt will be in a 1:1 ratio. Randomization will be stratified by clinical center. Randomized blocks of varying lengths will be used for randomization. As there is the possibility of relatively small numbers of patients within some centers, smaller block lengths will be used preferentially at the beginning of each generated sequence. As this trial is unblinded to the

surgeon performing the initial shunt procedure, specific block length probabilities are not included in this SAP but will be kept on file at the DCC, in order to limit predictability of subsequent treatment assignments.

More specifically, for each site participating or potentially participating in PENS, sequences shall be prepared as follows:

Block length will be selected randomly (length 2 or length 4, with probability of selection changing as the sequence lengthens).

1. For the selected block length, the treatment sequence (for block length 2, one of each treatment assignment; for block length 4, two of each treatment assignment) will be randomly shuffled, and the resulting sequence of treatments added to the existing randomization sequence for the center stratum.

2. The process in Steps 1 and 2 above will be repeated until a minimum of 40 treatment assignments have been generated for each clinical center.

Randomization sequences will be prepared by the study biostatistician.

4.3.1 Handling of Incorrect Randomizations in Study Analyses and Reports

It is not expected, but possible, that a patient may be randomized into the incorrect treatment arm, due to a technical randomization error. Any patients who are assigned incorrect randomizations solely due to technical malfunctions of the primary Web-based randomization system will be given the assigned treatment and treated as such in all analyses (recognizing that extremely rare "good faith" errors of this type may occur, and that the source of any such malfunctions must be immediately rectified).As the above violations affect within-stratum treatment balance, they are to be reported to the DSMB as part of all reports.

4.3.2 Delivery of Randomization and Emergency Backup

Randomizations will be delivered directly to the clinical centers using a Web-based system administered by the DCC. This system will use each enrolled patient's clinical center to deliver the next assigned treatment in the particular center stratum. It is expected that for each consented and eligible subject, the randomization will be performed immediately prior to the shunt implantation procedure.

Given the importance of randomizing all available patients in the PENS trial, an "emergency backup" system will be implemented using sealed opaque envelopes. A single such envelope, containing a treatment assignment randomly generated

with a 50% chance of randomization to either closed or open shunt, will be kept at each participating PENS site at all times. This "emergency" backup is intended for use only in situations where the Web-based system is not functional or accessible.

4.4 Treatment Masking (Blinding)

The PENS trial is necessarily performed in a partially unblinded fashion. The neurosurgeon performing the surgery will pre-set the adjustable valves to the designated open or closed settings, as specified according to the randomization system that the neurosurgeon will access directly. The setting will be performed and verified by the neurosurgeon alone, without assistance.

To maintain blinding among other study staff, an independent assessor will perform the gait and cognitive tests during the follow-up visits and will remain blinded to the treatment assignment. Other support staff will enter data for the trial.

While the PENS biostatistician will be unblinded to treatment assignments, knowledge of arm-specific treatment results will be limited to biostatisticians involved in the interim and final analyses. Moreover, for all interim analyses, such materials will be prepared with (for example) one arm randomly labeled as "Setting A" and one arm as "Setting B", with knowledge of arm identities limited to the biostatistician(s) presenting such materials to the Data Safety Monitoring Board (DSMB). The PENS DSMB may request to be unblinded to treatment assignment at any time.

5. SEQUENCE OF PLANNED ANALYSES

5.1 Interim Analyses

5.1.1 Frequency of and Timepoints for Interim Analysis

The PENS trial has DSMB meetings, involving review of study safety data and other facets of study progress and adherence, scheduled to occur after 10 patients, and then again after approximately two thirds of the target sample size of 40, have been randomized, treated, and have 4-month evaluations available.

The PENS DSMB is an independent body appointed by the Sponsor. The DSMB is scheduled to examine patient safety data at these meetings, without review of interim data on efficacy outcomes. The DSMB will, at their discretion, be able to request analyses additional to those described in this SAP. Full operations of the DSMB will be specified in a DSMB Charter.

At all meetings, the DSMB is to review interim data and make recommendations regarding continuation, modification, or termination of the PENS trial.

At this time, there are not formal subgroups of patients (based on clinical criteria)

scheduled for review by the DSMB.

5.1.2 Blinding in the Interim Analysis

As noted above, the number of DCC biostatisticians who are unblinded to results by treatment arm and identity of treatment arms will be limited as much as possible, while other DCC personnel shall be blinded to all safety and efficacy data, prior to final analysis or decision to unblind all PENS investigators to study results.

All interim analysis tables and analyses involving treatment arms will have treatment not explicitly identified, but referred to in a coded fashion, for example as "Setting A" and "Setting B", consistently throughout the report presented to the DSMB. The DSMB will have the option of being unblinded to treatment arm identity at any time, by opening a sealed envelope containing these identities.

5.2 Final Analyses and Reporting

All final, planned analyses identified in the protocol and in this SAP will be performed only after all randomized patients have completed the protocol, and results of all significant queries have been resolved. A blinded data review meeting will be held prior to final database lock and completion of the final analyses. In addition, no database may be locked, random code unblinded, or analyses completed until this SAP has been approved.

Any *post hoc*, exploratory analyses completed to support planned study analyses, which were not identified in this SAP, will be documented and reported as such in all study publications.

6. SAMPLE SIZE DETERMINATION

Pilot data on change in gait velocity, from 96 patients in a multisite NPH database (NPH centers at Cleveland Clinic, Sinai Hospital and Umea University) with baseline gait velocities of 1.0 m/s or less, had a standard deviation of 0.31 m/s for the post-treatment gait velocity, with a Pearson correlation of 0.44 between baseline and follow-up gait velocity. Using these estimates, the proposed analysis of covariance approach for the primary analysis requires the following number of patients completing the study for adequate power, using a two-sided testing approach with 0.05 Type I error:

Assumed 4-Month Mean Velocity Change: Closed Setting	Assumed 4-Month Mean Velocity Change: Open Setting	Total Number Required for 80% Power	Total Number Required for 85% Power	Total Number Required for 90% Power
0.05 m/s	0.3 m/s	40	46	54
0.05 m/s	0.325 m/s	34	38	46
0.05 m/s	0.35 m/s	30	32	38
0.05 m/s	0.375 m/s	26	28	32

Therefore, under these assumptions, a target sample size of 34 patients, 17 per study arm, will yield an estimated 80% power to detect a significant treatment effect if the true benefit of Open Setting versus Closed Setting on gait velocity at 4 months is at least 0.275 m/s, and over 85% power if the true benefit is at least 0.3 m/s. The study will plan on enrolling 40 patients to account for possible attrition to no less than 34.

7. ANALYSIS POPULATIONS

The following analysis populations are planned:

- **Screening Population (SCREEN):** The Screening Population includes all subjects who are screened for eligibility into PENS, regardless of randomization into the trial or treatment status. This population will be equivalent to all subjects who received a study identification number for this study and represents all patients who meet all of the inclusion criteria outlined in the study protocol. This population will be used for reporting of study flow per CONSORT guidelines.

- **Intention to Treat Population (ITT):** The Intention-to-Treat Population includes all subjects who provide informed consent and who are randomized into the trial, regardless of whether treatment was initiated or adherence to the protocol. **The ITT population will be used for the primary efficacy analyses in the study.**
- **Safety Population (SAFETY):** The Safety Population includes all randomized patients undergoing a shunt implantation procedure, for either treatment strategy regardless of assigned treatment. This population will be used for analysis of adverse events and other safety outcomes.

8. GENERAL ISSUES FOR STATISTICAL ANALYSIS

8.1 Analysis Software

Analysis will be performed using SAS® Software version 9.4 or later whenever possible. Other software packages, including R and StatXAct, may be used in instances where a particular specialized procedure is not available in SAS®.

8.2 Methods for Withdrawals, Missing Data, and Outliers

Per the intention-to-treat principle, any subjects who withdraw from the study will have all available data used in the analysis. In the event that a substantial number of subjects are withdrawn, baseline characteristics and available information on hospital and post-hospital course will be reviewed and compared to subjects not withdrawn, to assess empirically if these subjects differ from those remaining in the study for the scheduled treatment and follow-up time.

Outliers will be reviewed for validity. Outliers that are valid, for example, values indicating very slow gait velocity that are confirmed for accuracy by the clinical center, will be included in all primary reports from this trial.

8.3 Multicenter Studies

The randomization sequences will be stratified by clinical center, to assure approximate balance of sites between the treatment arms at all times. It is not planned to adjust for the effect of center in the primary analysis. It is planned to examine variability in treatment effect by clinical center, although the very small number will prevent statistical detection of all but the most profound differences by center.

8.4 Multiple Comparisons and Multiplicity

As there is a single primary efficacy endpoint for this study, adjustment for multiple comparisons will not be required for the primary analysis.

For the two secondary efficacy endpoints, a Bonferroni-Holm stepdown test will be used for assessing significance, with a total alpha level of 0.05 for these two outcomes.

Tertiary efficacy outcomes will be reported using unadjusted *p*-values for each individual comparison. However, all reports of these outcomes, to the DSMB and in published reports, will explicitly note that multiple tertiary efficacy outcomes have been evaluated.

Safety outcomes for this study will be reported using unadjusted *p*-values for each individual comparison. However, all reports of these outcomes, to the DSMB and in published reports, will explicitly note that multiple safety outcomes have been evaluated. Formal multiplicity-adjusted significance assessments for these outcomes will be performed upon request of the DSMB or other reviewers. The Bonferroni-Holm procedure will be used for such assessments.

8.5 Planned Subgroups, Interactions, and Covariates

There are no prespecified formal efficacy analysis subgroups for the PENS trial at this time.

8.6 Derived and Computed Variables

All derived and computed variables will be outlined in the analysis dataset specifications for this study. These datasets are independently programmed by two statisticians. The SAS COMPARE procedure will be used to verify that the dual programmed analysis datasets are identical for each variable and each observation.

The primary, secondary, and tertiary outcome measures are based on standardized measurement regimens, surveys or assessments to be completed by the patient or by skilled neurobehavioral experts and neuropsychologists. Many of these surveys have multiple outcome measures associated with them. A separate, comprehensive summary of these measures with details on scoring will be developed to indicate which measures will be used for each study endpoint and which measures will only be used for exploratory analyses.

9. STUDY SUBJECTS

9.1 Disposition of Subjects and Withdrawals

All subjects who provide informed permission will be accounted for in this study. The frequency and percent of subjects in each population, study withdrawals, subgroups, and major protocol violations will also be presented. While the final definition of "major protocol violation" will be determined during the course of the PENS trial, in all instances randomizations of subjects who were later found not to meet eligibility criteria, and instances where subjects received the opposite of the assigned treatment strategy will be reported in the subject disposition reports.

9.2 Inclusion and Exclusion Criteria

For excluded subjects, percentage meeting each exclusion criterion will be presented. These data will be presented overall and by clinical center. In addition, data will be presented regarding subjects who are initially considered eligible, but are found to be ineligible after randomization (such subjects would generally be included in the ITT analysis).

10. EFFICACY ANALYSES

10.1 Primary Efficacy Variable Analysis

The analysis of the primary PENS study outcome, change in gait velocity at 4 months after initial shunting, will include the following test of hypothesis:

H_0: the average change in gait velocity rate at 4 months after initial shunting is equal in subjects assigned to open shunt and subjects assigned to closed shunt.

The alternative hypothesis is:

H_1: the average change in gait velocity rate at 4 months after initial shunting is different between in subjects assigned to open shunt and subjects assigned to closed shunt.

Rejection of the null hypothesis, with a significantly higher gait velocity in the open shunt compared to closed shunt study arm, would be considered to be a successful demonstration of efficacy of immediate treatment of immediate shunting among patients meeting criteria for PENS.

To carry out the primary efficacy analysis, an analysis of covariance approach will be used to maximize statistical power. Specifically, a linear regression model will be fit with four-month gait velocity as the dependent outcome variable, and assigned treatment arm as a binary predictor along with continuous baseline gait velocity. Significance and magnitude of treatment effect will be assessed via the estimated coefficient of treatment in the model and its standard error. For purposes of assessing a significant treatment effect, the primary outcome of gait velocity will be compared between the assigned treatment arm using a two-sided test with a Type I error of 0.05.

If a subject is unable to complete a gait velocity test, or does not have gait velocity evaluation initiated, in a setting where there is unequivocal evidence of an excessively slow gait velocity, the subject will have a gait velocity of 0 used in the primary analysis. This "unequivocal evidence" criterion should be verified and applied conservatively; for example, an enrolled subject who is mobile, but who does not have follow-up gait velocity evaluation due to development of severe dementia, would be treated as having missing follow-up gait velocity.

While this analysis approach is robust to modest departures from normality in the model errors, the small sample size in the PENS trial and the possibility of outliers in change in gait velocity necessitate assessment of the effect of outliers. In the presence of gross departures from normality in the model errors, an empirical covariance matrix will be used to calculate standard errors of model coefficients in the analysis of covariance model. In addition, appropriate robustness assessments, which may include use of rank-based measures of treatment effect such as Hodges-Lehmann confidence intervals, will be implemented as appropriate in the case of excessive non-normality. In all published reports of the data from this pilot trial, however, the *a priori* method for the primary analysis above will be initially described as the intended approach.As the gait velocity used for the primary analysis is the best (fastest) observed in up to three trials of a 10-meter walk test performed at each study timepoint, a robustness analysis will repeat the primary analysis using average, rather than best, gait velocity at each evaluation timepoint.

10.2 Secondary Efficacy Variable Analysis

Secondary efficacy endpoints of this study are:

1. Montreal Cognitive Assessment Test (MoCA) composite score.
2. Overactive Bladder Questionnaire Composite (OAB) Symptom Score.

10.2.1 Change in MoCA composite score

The analysis of the secondary efficacy MoCA outcome will include the following test of hypothesis:

H_0: the average change in MoCA Composite Score at 4 months after initial shunting is equal in subjects assigned to open shunt and subjects assigned to closed shunt.

The alternative hypothesis is:

H_1: the average change in MoCA Composite Score at 4 months after initial shunting is different between in subjects assigned to open shunt and subjects assigned to closed shunt.

This secondary analysis will be carried out using an analysis of covariance approach, in a fashion completely analogous to the analysis of the primary gait velocity outcome above, with the same attention to robustness of results from any departures from assumptions of the regression model.

10.2.2 Change in OAB Symptom Score

The analysis of the secondary efficacy OAB outcome will include the following test of hypothesis:

H_0: the average change in OAB Symptom Score at 4 months after initial shunting is equal in subjects assigned to open shunt and subjects assigned to closed shunt.

The alternative hypothesis is:

H_1: the average change in OAB Symptom Score at 4 months after initial shunting is different between in subjects assigned to open shunt and subjects assigned to closed shunt.

This secondary analysis will be carried out using an analysis of covariance approach, in a fashion completely analogous to the analysis of the primary gait velocity outcome above, with the same attention to robustness of results from any departures from assumptions of the regression model.

10.2.3 Type I Error for Secondary Efficacy Analyses

It is desired to have a total Type I error of 0.05 for the two secondary efficacy analyses described in the above sections. To achieve this overall error a

Bonferroni-Holm step-down procedure[7] will be implemented. The two *p*-values from the analyses discussed in the above two sections will be considered, as detailed below.

If the smaller of the two *p*-values is less than 0.05/2=0.025, then the null hypothesis corresponding to that *p*-value will be rejected. In this instance, the remaining *p*-value will be examined, and if this remaining *p*-value is less than 0.05, then the null hypothesis corresponding to that *p*-value will be rejected as well. If the remaining *p*-value is 0.05 or greater, then the null hypothesis corresponding to that *p*-value will not be rejected. If the smaller of the two *p*-values is greater than or equal to 0.05/2=0.025, then neither of the two null hypotheses for the secondary analyses will be rejected.

10.3 Tertiary Study Efficacy Outcomes

The tertiary efficacy outcomes identified in this protocol (Symbol Digit Modalities Test (SDMT), Beck Depression Inventory, 2nd edition (BDI-II), Lawton Activities of Daily Living/Independence in Activities of Daily Living (ADL/IADL), and Modified Rankin Scale) will be treated as tertiary outcomes and will be clearly reported as such in all publications.

10.3.1 Eligibility for tertiary analysis

All randomized patients who have a particular tertiary outcome assessment performed both at baseline and at the 4-month study timepoint (within 14 days prior and 30 days after), which is valid in the judgment of the examiner, will be eligible for analysis of that outcome.

10.3.2 General statistical approaches for tertiary analyses

The tertiary outcomes listed above have continuous distributions, and therefore the analysis of covariance approach specified in this SAP for the formal efficacy outcomes is appropriate.

In the event that changes from baseline to four months for an outcome have a distribution that is grossly disparate from the normality assumptions required for the validity of the parametric approach (for example, an outcome being effectively bimodal or trimodal for the entire PENS patient population), a fully nonparametric approach would be implemented. A two-sample Wilcoxon rank-based test would be used to assess differences between treatment arms with respect to the distributions of change, and an exact two-sided 95% Hodges-Lehmann estimate of the magnitude of differences between treatment arms presented along with the estimate of significance.

All reported tertiary analyses will explicitly list the total number of neuropsychological outcomes assessed, and either explicitly adjust the alpha level for multiplicity or note that results must be viewed with caution due to the substantial numbers of comparisons performed.

10.4 Handling of Missing Values

Patients with missing data will be excluded from the respective secondary and tertiary analyses.

10.5 Additional Analysis Considerations

Additional types of analyses will be performed in addition to the ITT analysis, in part to gauge the robustness of the ITT analysis to treatment nonadhererers. One such analysis, termed "per protocol", will exclude patients with off-protocol valve adjustments. Specifically, if a patient randomized to a closed shunt has an adjustment to the open position prior to the 4-month study timepoint (more specifically, prior to the time interval during which 4-month assessments are acceptable), he or she will be excluded from any analyses following the off-protocol shunt adjustment.

A very similar analysis will consider patients in an "as-treated" fashion. If a patient has a valve adjustment from one randomized position to the other prior to the 4-month study timepoint, the patient will be analyzed in the group according to the treatment that the patient was receiving at the beginning of the 4-month evaluation window. For example, any patients whose shunt was in the open position 14 days prior to the 4-month evaluation window would be treated as "open shunt" for the "as-treated" evaluation of treatment effect at 4 months.

11. SAFETY ANALYSES

In all reported analyses of safety in the PENS trial, rates and other measures will be reported for the set all patients who received an initial shunt in the trial, as well as by "treatment received", compared between patients whose initial shunt setting was in the open position versus the closed position.

The safety endpoints of the PENS trial are:

1. Frequencies of adverse events and serious adverse events
2. Rates of falls

3. Occurrence of other prespecified complications (subdural hematoma, subdural effusion, shunt infection, shunt obstruction, and wound infection., overall and classified as surgical versus non-surgical)

11.1.1 Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) central coding dictionary. Adverse events will be described in summary tables and in individual listings for the DSMB reports. Tables and listings will be repeated using the subset of adverse events that are classified as serious by the study investigator.

11.1.2 Falls

This safety outcome will be treated both as binary (occurred/did not occur) and as rates for purposes of analysis. For rate calculations, the "time at risk" will be the number of days between the time of the initial shunt procedure, and the date of the last visit that a physical examination was performed that reported fall(s) or confirmed that falls did not occur.

Any patients who expire during the follow-up period, or withdraw during the follow-up period, without having had a fall will be treated as having a "No" value for the binary outcome.

For the rate outcome, number of separate falls will be calculated from available physical examination forms.

11.1.3 Incidence of complications

Complications of interest that will be treated as binary include subdural hematoma, subdural effusion, shunt infection, shunt obstruction, and wound infection.

11.1.4 Analysis of Binary Safety Outcomes

Statistical comparisons between the safety "as treated" study arms will be made using chi-squared approaches, including the standard chi-squared test when expected counts are at least five for all cells in a table. When this is not the case, exact approaches will be implemented that maximize the statistical power to detect a treatment effect. Specifically, when exact approaches are appropriate, the mid-*p*-value correction[8] will be implemented, whereby the probability value of obtaining a result at least as extreme as observed is reduced by one half of the probability of obtaining the specific result observed. This is directly calculable in SAS, for example, as the Fisher's exact test output gives these two probabilities directly.

11.1.5 Analysis of Safety Outcomes as Rate Data

Rates of falls (per month of follow-up) will be reported along with 95% confidence intervals. Poisson regression will be used to analyze models, using exact inference in the case of small numbers of counts. Overdispersion will be evaluated, and if applicable, binomial regression, or zero-inflated Poisson, or negative binomial regression, will be implemented. Information-based criteria including Akaike's information criterion (AIC) will be used to gauge the fit of non-nested models considered for these data.

11.1.6 Handling of Patients Changing Treatment Arms

While primary assessments of safety in the PENS trial will occur during the first 4 months of follow-up, comparing patients by initial treatment received, it is recognized that patients may cross over between treatment arms during this initial treatment period. A patient crossing over from closed to open shunt will technically be in the "at risk" population for AEs in both groups.

The optimal approach to presentation of AE risk with any crossovers will depend on specific numbers and circumstances (e.g., whether an event/complication occurred in a patient who crossed over). In combination with AE/complication rates by initial treatment received, corresponding rates will also be presented by "Treatment Received Including Crossovers" where numerators may include the same patient. As inference is challenging in this setting with a combination of extremely sparse and correlated binary data, presentation will focus on whether modification of denominators "at risk", or counting of a post-crossover event, affects key results and how safety data would be interpreted. Generalized estimating equations (GEE) can be considered for supportive analyses if inference is desired by the DSMB or a reviewer, although care must be taken in the interpretation of GEE results relative to chi-squared approaches.

For rate analyses, such as comparing rates of falls across the time of the trial including after patients have crossed over from open to closed shunt, GEE approaches incorporating "clustering" at the patient level will be implemented, carefully noting number of days each patient is at risk in each treatment arm.

12. OTHER PLANNED EXPLORATORY ANALYSES

The brief discussion of exploratory analyses presented in this Section is by no means exhaustive, as PENS will generate a database useful for exploratory analyses of clinical interest. We intend to use contemporary analytic approaches, including modification of existing approaches and derivation of novel techniques when appropriate, for such analyses. All exploratory, non-prespecified analyses will be clearly described as such in published reports. Whenever possible, the

exploratory analyses outlined in this Section should be explicitly prespecified in advance, in a separate Analysis Plan.

12.1 Outcomes after eight months of shunting in the entire PENS cohort

After all subjects have completed 12 months of follow-up, it will be possible to examine the entire cohort of patients after 8 months of open shunting, using the appropriate timepoint evaluation by treatment arm (for example, patients assigned to closed shunt would have their "8 months post shunting" evaluation at month 12, while patients assigned to open shunt would have their "8 months post shunting" evaluation at month 8 of their participation in PENS.

Analyses of "8 months post shunting" outcome may be done simplistically by reporting values at each timepoint, and changes between timepoints. It is preferable, however, to make optimal use of the multiple evaluations in the PENS dataset, using approaches to analyses of repeated measures such as the linear mixed models, to more accurately report trajectories across time, and to perform exploratory comparisons between treatment groups (which may be defined in various fashions in such analyses).

Specific approaches to these exploratory analyses, including technical specifications for carrying out the analyses and selecting models, will be delineated in Manuscript Analysis Plans for each analysis.

12.2 Assessment of "Delay Hypothesis"

It will also be of interest to assess whether outcome at month 8 and month 12 is different for patients by their initially assigned treatment arm. Differences in gait velocity and other measures at these later follow-up timepoints would indicate that immediate shunting is preferred to delaying shunting with respect to longer-term outcome. The linear mixed model, with appropriate contrasts to assess differences at 8 and 12 months making use of all available study data, will be implemented for this analysis.

13. REFERENCES

1. US Federal Register. (1998) International Conference on Harmonization; Guidance on Statistical Principles for Clinical Trials. Department of Health and Human Services: Food and Drug Administration [Docket No. 97D-0174]. Federal Register Volume 63, Number 179, pages 49583-49598. September 16, 1998.

2. ASA. (1999) Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics, August 7, 1999. http://www.amstat.org/profession/ethicalstatistics.html

3. RSS. (1993) The Royal Statistical Society: Code of Conduct, August 1993. http://www.rss.org.uk/about/conduct.html

4. Mehrotra DV, Li X, Gilbert PB. A comparison of eight methods for the dual-endpoint evaluation in a proof-of-concept HIV vaccine trial. Biometrica 2006;62:893-900.

5. van Elteren PH. On the combination of independent two sample tests of Wilcoxon. Bulletin of the Institute of International Statistics 1960;37:351–361.

6. Koch GG, Carr GJ, Amara IA, Stokes ME, Uryniak T J (1990). Categorical data analysis. In D. A. Berry (ed), Statistical Methodology in the Pharmaceutical Sciences, 389- 473. New York: Marcel Dekker.

7. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979;6:65–70.

8. Lancaster HO. The combination of probabilities arising from data in discrete distributions. Biometrika 1949; 36, 370-382.

9. Borm GF, Fransen J, Lemmens WAJG. A simple sample size formula for analysis of covariance in randomized clinical trials. Journal of clinical epidemiology. 2007 Dec;60:1234-1238.