THE PENS TRIAL

STATISTICAL ANALYSIS PLAN

Protocol Title (Number):	The Placebo-controlled Efficacy in iNPH Shunting (PENS) Trial. NCT05081128
Protocol Date:	January 31, 2022
SAP Authors:	Richard Holubkov, Ph.D. Hailey Jensen, M.Stat.
SAP Version:	1.1
SAP Date:	January 4, 2024

CONFIDENTIAL

Approvals:

Approved By: DocuSigned by: AEFA78C9944F4C6... 1/18/2024

Name: Mark Luciano, MD Date:
Principal Investigator for the PENS Trial

DocuSigned by: 677A811E45874DE... 1/16/2024

Name: Richard Hofubkov, PhD Date:
Principal Investigator, University of Utah DCC

Contents

1. PREFACE 6

2. PURPOSE OF SAP 6

3. STUDY OBJECTIVES AND ENDPOINTS 8

3.1 Study Objectives 8
3.1.1 Primary Objective 8
3.1.2 Secondary Objective 8
3.1.3 Exploratory Objectives 8
3.2 Study Endpoints 10
3.2.1 Primary Endpoint 10
3.2.2 Secondary Efficacy Endpoints 11
3.2.3 Tertiary Efficacy Endpoints 11
3.2.4 Safety Endpoints 12
3.2.5 CSF Endpoints 13
3.2.6 Imaging Endpoints 13

4. STUDY METHODS 13

4.1 Overall Study Design and Plan 13
4.2 Selection of Study Population 14
4.3 Method of Treatment Assignment and Randomization 15
4.3.1 Handling of Incorrect Randomizations in Study Analyses and Reports 16
4.3.2 Delivery of Randomization and Emergency Backup 16
4.4 Treatment Masking (Blinding) 17

5. SEQUENCE OF PLANNED ANALYSES 17

5.1 Interim Analyses 17
5.1.1 Frequency of and Timepoints for Interim Safety and Study Progress Analysis 17
5.1.2 Frequency of and Timepoints for Interim Efficacy Analysis 18
5.1.3 Blinding in the Interim Analysis 19
5.2 Final Analyses and Reporting 20

6. SAMPLE SIZE DETERMINATION 20

7. ANALYSIS POPULATIONS 21

8. GENERAL ISSUES FOR STATISTICAL ANALYSIS 21

8.1 Analysis Software 21

8.2 Methods for Withdrawals, Missing Data, and Outliers 21
8.3 Multicenter Studies 22
8.4 Multiple Comparisons and Multiplicity 22
8.5 Planned Subgroups, Interactions, and Covariates 22
8.6 Derived and Computed Variables 23

9. STUDY SUBJECTS 24

9.1 Disposition of Subjects and Withdrawals 24
9.2 Inclusion and Exclusion Criteria 24

***10.* EFFICACY ANALYSES 24**

10.1 Primary Efficacy Variable Analysis 24
10.2 Secondary Efficacy Variable Analysis 26
10.2.1 Change in Tinetti score 26
10.2.2 Change in MoCA total score 26
10.2.3 Change in OAB Symptom Score 27
10.2.4 Type I Error for Secondary Efficacy Analyses 27
10.3 Tertiary Study Efficacy Outcomes 28
10.3.1 Eligibility for tertiary analysis 28
10.3.2 Statistical approaches for tertiary analyses of efficacy 28
10.4 Handling of Missing Values 29
10.5 Additional Analysis Considerations 29

11. SAFETY ANALYSES 30

11.1 Adverse Events 30
11.2 Falls (All, and clinically significant) 30
11.3 Incidence of complications 31
11.4 Analysis of Binary Safety Outcomes 31
11.5 Analysis of Safety Outcomes as Rate Data 32
11.6 Handling of Patients Changing Treatment Arms 32

12. OTHER PLANNED ANALYSES 32

12.1 Outcomes after 12 months of active shunting in the entire PENS cohort 33
12.2 Neuropsychological Substudy Analyses 33
12.3 CSF Biomarker Study Analyses 34
12.4 Imaging Study Analyses 35

13. REFERENCES 37

ABBREVIATIONS

ABBREVIATION	DEFINITION
AD	Alzheimer's Disease
ADL/IADL	Lawton Activities of Daily Living/Independence in Activities of Daily Living
AHCRN	Adult Hydrocephalus Clinical Research Network
BDI	Beck Depression Inventory, 2nd edition
CRF	Case Report Form
CSF	Cerebrospinal Fluid
DCC	Data Coordinating Center
DSMB	Data Safety Monitoring Board
FDA	United States Food and Drug Administration
ICH	International Conference on Harmonisation
iNPH	Idiopathic Normal Pressure Hydrocephalus
ITT	Intent-To-Treat
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
MoCA	Montreal Cognitive Assessment Test
MRI	Magnetic Resonance Imaging
mRS	Modified Rankin Scale
NIH	National Institutes of Health
OAB	Overactive Bladder Questionnaire, Short Form
PENS	Placebo-controlled Efficacy in iNPH Shunting Trial
PP	Per-Protocol Population
SAP	Statistical Analysis Plan
SDMT	Symbol Digit Modalities Test

1. PREFACE

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for the NIH-funded Placebo-controlled Efficacy 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 NIH Trial Protocol, Version 0.05 October 31 2021
- Updates to be made to Protocol Version 0.05. To facilitate DSMB review of this SAP, any facets that reflect updates to Version 0.05 are presented as **bolded and underlined**. It is noted that these updates do not affect analytic strategies for the primary and secondary efficacy outcomes of the randomized component of PENS; the most substantial update is the addition of a study visit 12 months after active shunting in both treatment arms.
- Draft case report forms (CRFs) for the PENS Protocol

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.

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 Trial. 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.

Specifically, Version 1.1 of this SAP is an update to Version 1.0, specifying that valid time windows for follow-up evaluations have been expanded to the time interval from 14 days before to 45 days after the scheduled follow-up date, from the previously specified interval from 14 days before to 30 days after scheduled follow-up. This modification has been made prior to any efficacy analysis (including interim efficacy analysis) solely because several participants were noted to have 3-month follow-up times lightly more than 30 days after their 3-month anniversary of surgery, due to scheduling conflicts and similar issues.

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 3 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 balance at 3 months after shunt surgery, as assessed by the Tinetti score

ii. Improved cognition at 3 months after shunt surgery, as assessed by the Montreal Cognitive Assessment Test (MoCA)

ii. Improved bladder control at 3 months after shunt surgery, as assessed by the Overactive Bladder Questionnaire, short form (OAB-q)

3.1.3 Exploratory Objectives

The exploratory objectives of the PENS Trial include the following:

Efficacy-Related Objectives:

i. Evaluate the effect of shunting between active and placebo-controlled groups at three months using gait, modified Rankin Scale (mRS), and Core neuropsychology measures of cognition and mood.

ii. Evaluate the various efficacy measures described above at three, six, and **twelve** months of active shunting in all participants.

iii. Evaluate the clinical improvement of all study participants at three, six and **twelve** months of active shunting.

CSF-Related Objectives

iv. Evaluate the role of neurodegenerative (abeta42/40, ptau181, total-tau, NFL) and neurovascular (CSF VEGF) biomarkers, as diagnostic or prognostic biomarkers in differentiating iNPH from other neurodegenerative disorders that present with an iNPH phenotype.

v. Identify cutoffs for established neurodegenerative CSF biomarkers that differentiate participants with iNPH from common neurodegenerative disorders of aging including Alzheimer's disease (AD), Lewy body dementia, Parkinson's disease, and vascular dementia, adjusting the typical cut-offs of common biomarkers for altered volume of distribution when CSF is available from clinical testing.

vi. Identify cutoffs for CSF biomarkers associated with improved 12-month response to active shunting, enabling improved selection of participants in future trials.

Imaging-Related Objectives

vii. Measure the morphological features in and around the brain that are associated with gait improvement after shunt surgery.

viii. Determine physiological features measured by magnetic resonance imaging (MRI) that are associated with gait changes after shunting.

Neuropsychological Substudy Objectives

For participants in the neuropsychology substudy:

ix. Evaluate the effect of shunting on sustaining improvements in cognition, mood, behavior, and quality of life between active and placebo-controlled groups at three months post initial shunt surgery, **and after 12 months of active shunting** in the entire randomized cohort using core and extended measures of cognition and mood.

x. Identify broad cognitive composites (using baseline information for both Core and extended batteries).

xi. Assess treatment effect (and overall outcomes as appropriate) **after 12 months of active shunting** for these composite measures.

This SAP addresses in detail the analysis of efficacy and safety outcomes. Separate full SAPs will be developed for the CSF biomarker and Imaging studies, and the Neuropsychological substudy. Summaries of analytic approaches for these

substudies are, however, included in this document.

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 3 months after shunt surgery.

At each evaluation timepoint, the gait velocity used for analysis is the best (fastest) velocity occurring in up to three valid 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 verifying eligibility criteria prior to randomization) and a valid assessment of velocity 3 months after shunt surgery.

More specifically, pre-surgical "baseline" gait velocity assessments will be considered valid if they occur at least 4 weeks after the drainage trial assessing eligibility. It is expected that the baseline assessments will occur no more than 6 weeks prior to shunt surgery (valid assessments can be performed up to and including the day of surgery, if prior to surgery). In the case of multiple pre-surgical assessments, the gait velocity assessment that occurs closest to the lumbar puncture (which is still at least 4 weeks post drainage trial) will be used in the primary analysis.

Post-surgical assessments of gait velocity will be considered valid if they occurred between 14 days prior to and 45 days after each scheduled 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 3 months after shunt surgery) will be used in the primary intention to treat analysis.

3.2.2 Secondary Efficacy Endpoints

The three formal secondary efficacy endpoints of the PENS Trial are:

1. Tinetti score
2. Montreal Cognitive Assessment Test (MoCA) total score.
3. Overactive Bladder Questionnaire Composite Symptom Score (OAB-q).

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. Validity "windows" for these endpoints are the same as for the primary outcome above.

3.2.3 Tertiary Efficacy Endpoints

Tertiary endpoints of the PENS Trial include:

i. Boon Scale (assessment of gait)

ii. Modified Rankin Scale (assessment of function)

Several core measures of cognition and mood, including:

iii. Symbol Digit Modalities Test (SDMT)

iv. Flanker Inhibitory Control and Attention Test (from NIH Toolbox)

v. Dimensional Change Card Sort Test (from NIH Toolbox)

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

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

The following additional measures of cognition, mood, behavior, and quality of life, to be assessed among participants in the neuropsychology substudy:

viii. Wide Range Achievement Test - Fifth Edition (WRAT-5) Reading

ix. Controlled Oral Word Association Test (COWAT)

x. Hopkins Verbal Learning Test -Revised (HVLT-R)

xi. Boston Naming Test (BNT)- short form

xii. WAIS-III- Digit Span subtest

xiii. Trail making Test A and B

xiv. Animal Fluency

xv. Stroop Color Word Test

xvi. Judgment of Line Orientation (JLO)-short form

xvii. Grooved Pegboard

xviii. Finger Tapping Test

xix. Frontal Systems Behavior Scale (FrSBe)

xx. Neuropsychiatric Inventory Questionnaire

xxi. Medical Outcomes Survey Short Form (SF-12)

3.2.3.1 Eligibility for Tertiary Analysis

Patients with valid assessments of appropriate outcomes will be eligible for tertiary analyses; assessment "windows" for validity will be identical to that of the primary efficacy outcome.

3.2.4 Safety Endpoints

The safety endpoints of the PENS Trial are:

1. Adverse events and serious adverse events
2. Falls, overall and clinically significant

3. Rates of complications, overall and classified as requiring versus not requiring neurosurgery. Specific complications examined will include, but not be limited to, the following: subdural hematomas, hygromas, stroke and death.

3.2.4.1 Eligibility for Safety Analysis

All patients randomized in the PENS Trial and having shunt surgery initiated will be eligible for safety analyses.

3.2.5 CSF Endpoints

CSF outcomes will include levels of the following biomarkers: VEGF, totaltau, abeta42, abeta40, NFL, ptau181, as well as ratios of abeta42/abeta40 and ptau181/abeta42.

3.2.6 Imaging Endpoints

Imaging outcomes examined as biomarkers will include the following: ventricular volume, Evans Index, morphology, CSF flow, blood flow, and connectivity.

4. STUDY METHODS

4.1 Overall Study Design and Plan

Clinics participating in the PENS Trial will prescreen patients presenting for NPH assessment and evaluate them for inclusion and exclusion criteria (listed immediately below). Patients potentially eligible for the trial based on initial evaluation of inclusion and exclusion criteria will be approached to provide pre-consent for CSF collection, and will have 10 *mL* of CSF collected during their standard lumbar puncture or extended lumbar drainage. This drainage trial will assess potential clinical benefit of CSF shunting as part of the screening process. Final PENS eligibility will be determined by the investigator based on a final evaluation including results of the drainage trial.

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 H_2O) or placebo (closed shunt group) (setting 8) (>400 mm H_2O) shunt settings. Randomization will be stratified by age of the patient at time of surgery, categorized as 60-69 years, 70-79 years, or 80 years and older.

After 3 months, shunts for subjects in the closed shunt group will be adjusted to the active setting of 110 mm H_2O. After this 3-month adjustment, all subjects in

both groups will continue to be followed, **each patient being followed for a total of 12 months after initiation of active shunting**. 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 and recommendation for shunt surgery based on the Investigator's clinical judgement using criteria and testing as described in the iNPH Guidelines[4]

C. Evans Ratio ≥ 0.30

D. One positive supplementary test to include either large volume Lumbar Puncture or extended CSF drainage per institutional standards[5]

E. History or evidence of gait impairment (such as decreased step height or length, decreased speed, retropulsion as described in the iNPH Guidelines[4]) of duration ≥ 6 months

F. Participant has the sensory motor skills, communication skills and understanding to comply with the testing and reporting required in the PENS Trial

G. Participant is able to give written informed consent.

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

H. Unable to walk 10 meters with or without an assistive device

I. Screening fastest gait velocity > 1 m/sec prior to drainage trial, unless the patient's gait velocity improves > 30%

J. Unable to return to the study center for follow up evaluation and shunt programming

K. Participant is not medically cleared for shunt surgery per local standards

L. Secondary NPH (Prior encephalitis, meningitis, subarachnoid hemorrhage,

traumatic brain injury [including concussion] within two years or with brain injury or skull fracture on baseline imaging, brain abscess, brain tumor, obstructive hydrocephalus [including acquired aqueductal stenosis and carcinomatous meningitis])

M. Prior or existing shunts, endoscopic third ventriculostomy, or any previous surgical intervention for hydrocephalus including lumbar CSF shunting

N. Previous intracranial neurosurgical procedure

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

P. Symptomatic cerebral or cerebellar infarction occurring within 6 months from screening (asymptomatic lacunar infarctions are permitted)

Q. Diagnosis of Parkinsonian syndrome that, in the investigator's judgment, will impede the outcome evaluation

R. Diagnosis of schizophrenia or any psychiatric diagnosis (including depression) that, in the investigator's judgment, will impede the outcome evaluation (such as neuroleptic treatment for schizophrenia)

S. Diagnosis of dementia disorder where the investigator considers cognition deficit limits participation in the study

T. Conditions impairing gait that are considered to be unrelated to hydrocephalus, such as hemiparesis, spasticity, cerebellar ataxia or musculoskeletal and joint disease, which will interfere with gait assessment or the potential for gait improvement.

4.3 Method of Treatment Assignment and Randomization

Randomization to active versus closed shunt will be in a 1:1 ratio. Randomization will be stratified by age of the patient, categorized as 60-69 years, 70-79 years, and 80 years or older. Randomized blocks of varying lengths will be used for randomization. To facilitate added balance in allocations between arms early in the trial, 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 age stratum 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 100 treatment assignments have been generated for age stratum.

Randomization sequences will be prepared by the study biostatistician.

4.3.1 Handling of Incorrect Randomizations in Study Analyses and Reports

It is possible that a patient may be randomized using the incorrect age stratum due to incorrect age calculation. If this occurs, such a patient will be analyzed as being in their correct age stratum in all study analyses. The same modification will be made in the unexpected event that a patient's age stratum changes between date of randomization and date of surgery.

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 age 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 into the PENS Trial at a timpoint immediately prior to initiation of surgery, 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 (and thus not incorporating the patient's age into the randomization procedure), 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. All other study staff will remain blinded to initial treatment assignment until 9 months of follow-up. Following the evaluation at 9 months post surgery (which we note does not affect the integrity of the initial randomized 3-month comparisons), the site coordinator will be unblinded to initial treatment assignment for purposes of appropriately scheduling the participant's remaining study visit **to occur 12 months after initiation of active shunting**.

To maintain blinding among other study staff, an independent assessor will perform the gait and cognitive tests during all 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 "Arm A" and the other as "Arm 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 Safety and Study Progress

Analysis

The PENS study will have a DSMB appointed by NIH. The DSMB will have a charter, will review and approve the protocol prior to implementation, and will review interim analyses as applicable. The purpose of the DSMB is to advise NIH and the Principal Investigators regarding the continuing safety of study participants and the continuing validity and scientific merit of the study. The DSMB is responsible for monitoring accrual of study participants, adherence to the study protocol, assessments of data quality, performance of individual clinical center and review of serious adverse events and other safety issues.

The PENS Trial will have DSMB meetings, involving review of study safety data and other facets of study progress and adherence, scheduled to occur at approximate six-month intervals after the initial "kick-off" DSMB meeting prior to beginning of patient enrollment.

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.

5.1.2 Frequency of and Timepoints for Interim Efficacy Analysis

There will be two formal efficacy monitoring looks at the primary efficacy outcome, after approximately one-third and again after approximately two-thirds of the target population of 100 participants have been enrolled and have three-month efficacy data available. As early stopping for efficacy should be based on very strong evidence in our judgment, we will implement prespecified conservative O'Brien-Fleming-type efficacy boundaries, implemented using the approach of Lan and DeMets to allow for flexibility in the actual number of outcomes available for review at each meeting. This conservative monitoring strategy will preserve Type I error at 5%, while having nearly-negligible effect on study sample size requirements.

Table 1 below shows the probability of stopping for a finding of superiority by each of the three looks (two interim looks and the final look) under various assumed magnitudes of treatment effect (including no effect) for the primary efficacy outcome of change in gait velocity from baseline to three months, assuming a standard deviation of 0.29 for three-month change velocity in each arm.

Table 1: Probability of Stopping for Superiority according to Treatment Effect Magnitude

Assumed True Treatment Effect Magnitude	Probability of Stopping with Superiority Finding at First Look	Probability of Stopping with Superiority Finding by Second Look	Probability of Stopping with Superiority Finding by Third Look (Study Type I error or Power)
0.0 m/s (null hypothesis)	<0.1%	1.4%	5.0 %
1.0 m/s	<0.1%	14.8%	40.0%
1.5 m/s	2.4%	36.7%	72.7%
2.0 m/s (assumed alternative)	6.9%	64.2%	92.8%
2.1 m/s (observed effect magnitude in PENS pilot study)	8.4%	69.3%	94.9%
2.5 m/s	16.3%	85.7%	99.0%
3.0 m/s	31.4%	96.2%	99.9%

Futility will also be assessed during interim efficacy looks; while boundaries for futility monitoring can be formally prespecified along with the efficacy boundaries, we would recommend the conditional power approach, where the trial is terminated for futility if estimated conditional power to detect a significant effect calculated by appropriate criteria (assuming the true treatment effect is as initially estimated, assuming effect is at its current point estimate, and calculated using predictive power) is below a prespecified threshold (20%) under all realistic scenarios.the study design assumptions of a treatment effect of 0.2 m/s and within-arm standard deviation of 0.29 m/s, at the first interim look conditional power will be <20% if the interim Z-statistic is < -1.73 (indicating a strong trend of a benefit in favor of the placebo arm). Under the null hypothesis of no treatment effect, with n=33 observations the probability of observing such a trend (and thus the estimated chance of stopping for futility) is 4.2%. At the second interim look with n=67, conditional power is <20% if the Z-statistic is < 0.46; if there is truly no effect of active shunting, this corresponds to an estimated chance of stopping for futility by the second look of 67.7%.

5.1.3 Blinding in the Interim Analysis

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 "Arm A" and "Arm 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.

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

The target sample size for the PENS Trial was determined based on available preliminary data for the primary efficacy outcome of change in gait velocity. An NPH database from three centers that will participate in the trial (n=97) showed mean gait velocity improvement of 0.25 ± 0.29 *m/s* post-shunting among iNPH patients with baseline velocities <1 *m/s*. The AHCRN Registry (n=50) found similar mean post-shunting improvement of 0.32 ± 0.26 *m/s*. The AHCRN network's pilot randomized trial found a strong trend for active shunting benefit on 3-month gait velocity (active arm change: 0.28 ± 0.28, placebo arm: 0.06 ± 0.15, improvement vs. placebo of 0.21 *m/s*, 95% CI = [−0.03,0.45]).

Conservatively assuming a standard deviation of 0.29 *m/s* for 3-month change in gait velocity and estimating power based on a two-sided *t*-test with Type I error of 0.05, 46 evaluable participants in each arm yield 90% power to detect a significant treatment effect, if the true between-arm difference in change velocities is at least 0.2 *m/s*. Based on these calculations, target recruitment has been set to 100 randomized subjects to maintain power assuming conservative interim efficacy monitoring, as well as in case of slight loss to follow-up at 3 months or if variability in velocity change scores is larger than expected.

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.
- **Modified Intention to Treat Population (mITT):** The modified Intention-to-Treat Population includes all subjects who provide informed consent, who are randomized into the trial, and in whom a shunt is successfully surgically implanted, regardless of adherence to assigned shunt setting or any other aspects of the protocol. The mITT population (which may be identical to the ITT population) will be used for the primary efficacy analyses in the study.
- **Safety Population (SAFETY):** The Safety Population includes all randomized patients in whom a shunt implantation procedure has been initiated, 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 categorized age to assure approximate treatment balance within each age stratum at all times. As the number of centers is large relative to the total sample size and therefore some centers may enroll a very small number of patients into the trial, randomization is not additionally stratified by center, nor is it planned to adjust for the effect of center in the primary analysis. It is planned to examine variability in treatment effect by clinical center as much as possible in the study data, although the small sample size 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 three secondary efficacy endpoints, a Bonferroni-Holm stepdown test[6] will be used for assessing significance, with a total alpha level of 0.05 for these three outcomes.

Tertiary efficacy outcomes will be reported using unadjusted *p*-values for each individual comparison. However, all reports of these outcomes will explicitly note that multiple tertiary efficacy outcomes have been evaluated, and note numbers of relevant assessments as appropriate.

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.

8.5 Planned Subgroups, Interactions, and Covariates

The PENS Trial will assess outcomes separately for the three age randomization strata, as well as by gender, race, and ethnicity.

More specifically, the outcomes will be re-analyzed using the same outcome-

specific linear model applied to the overall trial population, with the addition of a main effect of the subgroup and an interaction between the subgroup and treatment arm. If the interaction p-value is significant at the 0.05 level, this indicates some level of evidence that the effect of arm differs by subgroup level when predicting the outcome.

Multiplicity will not be adjusted for in the presented subgroup analyses, and this will be clearly noted in any results published from this trial. The key motivation for eschewing formal multiplicity adjustment is that we do not *a priori* expect extremely strong modifications of treatment effect by any of the above factors (although we certainly cannot rule out that such interactions exist). The PENS Trial has therefore been designed and "powered" to detect a treatment effect in the overall population, rather than allowing definitive specification of differential strategies for subgroups of enrolled patients. Any significant subgroup effects will be strictly reported as exploratory for potential confirmation in future studies. Reports of any such effects will also note the number of subgroups being examined, and discuss the strength of evidence for such an effect using clinical criteria as well as within-trial assessment of the criteria developed in the Predictive Approaches to Treatment Effect (PATH) statement[7].

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 or did not receive a shunt, 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 during the presecreening and final screening phases 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 3 months after initial shunting, will include the following test of hypothesis:

H_0: the average change in gait velocity rate at 3 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 3 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 model will be fit with three-month gait velocity as the dependent outcome variable, and assigned treatment arm as a binary predictor along with continuous baseline gait velocity and the patient's age at the time of randomization categorized into three levels.

As the PENS pilot study indicated possible heterogeneity of variances with respect to 3-month outcome change between the two treatment groups, the linear model for this outcome will fit separate error variance terms for the two groups.

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 arms 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 rapid 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 calculated over up to three valid trials at each evaluation timepoint.

10.2 Secondary Efficacy Variable Analysis

The three prespecified secondary efficacy endpoints of this study are:

1. Tinetti test score
2. Montreal Cognitive Assessment Test (MoCA) total score
3. Overactive Bladder Questionnaire Composite (OAB-q) Symptom Score

10.2.1 Change in Tinetti score

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

H_0: the average change in Tinetti Score at 3 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 Tinetti Score at 3 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 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 MoCA total score

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

H_0: the average change in MoCA Total Score at 3 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 Total Score at 3 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. However, for the the analysis of this outcome, age will be included as a continuous covariate rather than categorized, and years of education will be included as an additional continuous covariate. This analysis will be carried out with the same attention to robustness of results from any departures from assumptions of the regression model.

10.2.3 Change in OAB-q Symptom Score

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

H_0: the average change in OAB-q Symptom Score at 3 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-q Symptom Score at 3 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 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.

10.2.4 Type I Error for Secondary Efficacy Analyses

It is desired to have a total Type I error of 0.05 for the three secondary efficacy analyses described in the above sections. To achieve this overall error a Bonferroni-Holm step-down procedure[6] will be implemented. The three *p*-values from the analyses discussed in the above two sections will be considered, as detailed below.

If the smallest of the three *p*-values is less than 0.05/3=0.0167, then the null hypothesis corresponding to that *p*-value will be rejected. In this instance, the remaining two *p*-values will be examined, and if the smaller of these two *p*-value is less than 0.05/2=0.025, then the null hypothesis corresponding to that *p*-value will be rejected as well. Only in this instance, the remaining p-value will also be evaluated, and if this p-value is less than 0.05, then then the null hypothesis corresponding to that *p*-value will be rejected as well.

If the smallest of the three *p*-values above is greater than or equal to 0.05/3=0.0167, then none of the three null hypotheses for the secondary analyses will be rejected.

10.3 Tertiary Study Efficacy Outcomes

The various tertiary efficacy outcomes identified in this protocol will be treated as tertiary/exploratory 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 3-month study timepoint (within the time interval 14 days prior and 45 days after), which is valid in the judgment of the examiner, will be eligible for analysis of that outcome.

10.3.2 Statistical approaches for tertiary analyses of efficacy

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

Covariates included for adjustment, in addition to baseline level of the relevant measurement, will be stratified age category for the analysis of Boon and modified Randkin Scale, and continuous age and years of education for the several core assessments of cognition and mood.

In the event that changes from baseline to three 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 relevant outcomes assessed, and note that results must be viewed with caution due to the substantial numbers of comparisons performed.

10.4 Handling of Missing Values

The primary analysis strategy for efficacy outcomes at 3 months specifies exclusion of patients with missing outcome data, as this proportion is expected to be very small at 3 months. However, analyses will be repeated using multiple imputation strategies to determine robustness to missing data.

When imputation is performed, we will use a multivariate sequential regression approach to model and impute item missing values and unit non-responses[8]. We will simultaneously model and impute values for primary, secondary, tertiary, and safety outcomes. We will use forward stepwise variable selection to reduce the computational burden of fully conditional specification, as the resulting simpler models have been shown to produce similar results[8].

We will stratify imputation models by treatment assigned. Clinical and demographic predictor variables, such as age, gender, Evans' index, baseline measures of efficacy outcomes, and other important predictors will be used as general model variables. Indicators of missingness will be included as predictors.

Multiple imputation will be performed using IVEware (University of Michigan, Ann Arbor, MI, USA). We will execute IVEware and perform all post-imputation analyses using SAS/STAT software (SAS Institute Inc. Cary, NC, USA). The most recent versions of the software will be used.

For all efficacy outcomes, analyses will be performed on each imputed dataset. We will use the MIANALYZE procedure in SAS/STAT software in order to combine the multiple results into a single result. This procedure uses standard methods which combine estimates and within- and between-imputation variance estimates to derive valid univariate and multivariate inference[9].

We will initially impute and analyze 10 datasets. We will increase that number if the relative efficiency for the primary outcome test statistic is less than 0.99.

In imputation analyses, any patients who have missing data due to being deceased at the 3-month timepoint will be excluded, but such deaths will be noted in all reports.

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 any patients with off-protocol valve adjustments. Specifically, if a patient randomized to a closed shunt

has an adjustment to an open position prior to the 3-month study timepoint (more specifically, prior to the time interval during which 3-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 3-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 3-month evaluation window. For example, any patients whose shunt was in an open position 14 days prior to the 3-month evaluation window would be treated as "open shunt" for the "as-treated" evaluation of treatment effect at 3 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 (overall and clinically significant)
3. Occurrence of other complications (subdural hematoma, subdural effusion, shunt infection, shunt obstruction, and wound infection, overall and classified as surgical versus non-surgical)

11.1 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.

11.2 Falls (All, and clinically significant)

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.

If, at any time during the trial, the observed number of patients with one or more significant falls in the placebo arm is at least double that in the active arm, and the number of patients with falls in the placebo arm exceeds that in the active shunting arm by at least two, the medical monitor will be notified of the trend and the trend will be investigated. For example, if 2/4 patients in the active arm and 0/4 in the placebo arm have clinically significant falls, recruitment would be suspended, while recruitment would continue if the within-arm numbers were 2/4 and 1/4 or 1/4 or 0/4. As an example of invoking this algorithm later in the trial, if 3/20 patients in the active arm have clinically significant falls, recruitment would be suspended if at least 6/20 in the placebo arm report significant falls. The DSMB will be notified of the trend and the results of the investigation. The DSMB will be consulted for guidance on whether enrollment should be suspended.

11.3 Incidence of complications

Other complications of interest including mortality will be treated as binary. Overall rates of surgical complications will also be calculated. If surgical complications are seen in greater than 30% of trial participants at any time, the trend will be investigated to ensure it is not related to the study. The DSMB will be notified of the trend and the results of the investigation.

11.4 Analysis of Binary Safety Outcomes

Statistical comparisons of ninary outcomes 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 5 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[10] 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.

11.5 Analysis of Safety Outcomes as Rate Data

Rates of falls and clinically significant 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.

11.6 Handling of Patients Changing Treatment Arms

While primary assessments of safety in the PENS Trial will occur during the first 3 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 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 12 months of active shunting in the entire PENS cohort

In the presentation of results for the entire trial cohort after open shunting, it will be critical to assess the "delay hypothesis" for each outcome, that is, whether the pattern of outcomes over time after open shunting differs for patients by their initially assigned treatment arm. Differences in gait velocity and other measures at these later follow-up timepoints might, for example, 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 3, 6 and 12 months after active shunting** making use of all available study data, will be implemented for this analysis. In this model, the baseline value of the outcome will be the study "baseline" for patients assigned to initial open shunting, and the value obtained at 3-month followup (i.e., immediately before the shunt is opened) will be used as the baseline value for those assigned to open shunt. Covariates appropriate for the outcome being examined (e.g., age stratum for gait assessment, continuous age and years of education for MoCA) will also be included in the model. Timepoint will be included as a categorical predictor. The outcomes fit will be measures at **3, 6, and 12 months** after active shunting.

This model will be fit with a timepoint-by-assigned treatment interaction. If this interaction is not significant, this is evidence that after open shunting, the two study arms have similar outcome patterns, and thus that it is appropriate to summarize post-open-shunting trends and change magnitudes for the entire cohort. If the interaction is significant, this will be evidence that post-open-shunting trends should be described separately for the two treatment arms, with the open-shunt-assigned group perhaps more representative of outcomes expected after initial shunting.

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 Neuropsychological Substudy Analyses

The additional neuropsychological outcomes collected among the approximately 60-80 participants at selected North American centers participating in the neuropsychological substudy will be analyzed and reported in the same fashion as

Core Battery outcomes discussed above, for assessment of treatment efficacy at 3 months, as well as for assessment of outcomes **after 12 months of active shunting** in the entire Substudy cohort. While treatment effect will be again assessed using analysis of covariance controlling for age and education as continuous covariates, due to the smaller sample size and resulting lower power to detect an assigned treatment effect, presentation will focus on trends observed over time in the entire substudy.

Relationships of neurological outcomes (Core and substudy) to one another, and to the various outcomes observed among all trial participants, will be examined. A key exploratory analysis making use of these relationships will generate principal components (among substudy participants, using Core Battery as well as substudy measures at baseline) to create a set of cognitive composites. The subset of these composites that are associated with eigenvalues > 1 will be compared between treatment arms, and these composites will also be assessed over time using linear modeling approaches as for efficacy outcomes in the overall trial.

Additional analysis details for this substudy will be detailed in a separate Statistical Analysis Plan.

12.3 CSF Biomarker Study Analyses

For the CSF biomarker study, biomarker data is expected to be available for approximately 300 subjects undergoing prescreening who are potentially eligible for PENS (meeting all inclusion and exclusion criteria that are evaluable prior to drainage trial and subsequent testing). We will use baseline CSF biomarker levels from all 300 samples, alone and combined with available patient data such as age gender, Evans index, and baseline gait and MOCA to distinguish responders from the phenotypically similar non-responders. Biomarker levels will be reported and compared between these two populations using appropriate criteria including rank-based tests for biomarkers with highly skewed distributions. We note that this component of PENS is viewed as hypothesis-generating, and thus formal adjustment for multiplicity in these univariable comparisons will not be performed.

Univariable and multivariable logistic regression models, the latter incorporating the lasso approach for variable selection as appropriate, will be used for these analyses. As these biomarkers are typically used diagnostically using cutpoints and biomarker ratios, and between-marker interrelationships between markers are likely, Classification and Regression Trees (CART) will also be used as an exploratory analytic approach. Random forests and other machine learning approaches will also be implemented to gain further insight regarding relative importance of various markers in segregating responders. Past studies have found that “black box” machine learning approaches such as random forests have

markedly superior diagnostic performance, albeit without the explicit identification of specific biomarker cutpoints and interactions that can be obtained from CART and regression models.

Within the population of 100 patients who are enrolled in PENS, similar modeling approaches will be used to assess relationships of CSF biomarkers with 3-month and longer-term outcomes post shunting, i.e., to assess CSF biomarkers predictive of short-term and sustained response to active shunting. Magnitude of relationships (e.g., with change in gait velocity and MOCA) will be quantified using correlations and other appropriate measures.

12.4 Imaging Study Analyses

For the Imaging Study, MRI assessments of cerebral ventricular and brain volumetrics, brain morphology, CSF flow, white matter connectivity, and resting-state functional network alteration will be obtained and used as potential predictors of favorable outcome post-shunting.

Prespecified primary analysis parameters will include Evans ratio and volumes of the left lateral, right lateral, third, and fourth ventricles. Secondary parameters, evaluated on the hemisphere contralateral to the valve, include brain (parenchyma) volume and thickness of the subarachnoid space. Exploratory parameters include subcortical gray matter volumes (thalamus, putamen, caudate), cerebellar volume, cortical gray matter volume and thickness (on the contralateral side), and white matter volume (also on the contralateral side).

Modeling approaches analogous to those described for the efficacy analyses above will be used to compare change with respect to these MRI imaging parameters (i.e., assess morphological differences) from study baseline to 3-month evaluation in the active versus placebo arms. In addition, for the entire randomized cohort, morphological changes assessed at imaging 12 months after active shunting will be evaluated and assessed for association with improvement in efficacy outcomes at 12 months. Various modeling strategies will also be used to identify potential MRI biomarker predictors of 3-month and longer-term improvement in these efficacy outcomes after active shunting.

Additional analysis details for the imaging study will be detailed in a separate Statistical Analysis Plan.

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. Group IGS. INPH Guidelines Study Group. Neurosurgery. September 2005 2005;57(3):S2vi.
5. Marmarou A, Bergsneider M, Relkin N, Klinge P, Black PM. Development of guidelines for idiopathic normal-pressure hydrocephalus: introduction. Neurosurgery. Sep 2005;57(3 Suppl):S1-3; discussion ii-v. doi:10.1227/01.neu.0000168188.25559.0e
6. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979;6:65–70.
7. Kent DM, Paulus JK, van Klaveren D, D'Agostino R, Goodman S, Hayward R, Ioannidis JPA, Patrick-Lake B, Morton S, Pencina M, Raman G, Ross JS, Selker HP, Varadhan R, Vickers A, Wong JB, Steyerberg EW. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement. Ann Intern Med. 2020 Jan 7;172:35-45.
8. Raghunathan T, Lepkowski J, Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology, 27, 11 2000.
9. Rubin D. B. Multiple imputation for nonresponse in surveys. Wiley, 1987.
10. Lancaster HO. The combination of probabilities arising from data in discrete distributions. Biometrika 1949; 36, 370-382.

DocuSign

Certificate Of Completion

Envelope Id: DCB86D9301D34B4DBF1F535F91A40F9D
Status: Completed
Subject: SIGNATURE REQUEST: PENS NIH TRIAL SAP VERSION 1.1 JAN 4.pdf
Source Envelope:
Document Pages: 37
Signatures: 2
Certificate Pages: 5
Initials: 0
AutoNav: Enabled
EnvelopeId Stamping: Disabled
Time Zone: (UTC-08:00) Pacific Time (US & Canada)

Envelope Originator:
CYNTHIA Jones
102 South 200 East, Suite 110
Salt Lake City, UT 84111
u0019811@utah.edu
IP Address: 98.167.161.154

Record Tracking

Status: Original
1/16/2024 11:05:58 AM
Holder: CYNTHIA Jones
u0019811@utah.edu
Location: DocuSign

Signer Events	Signature	Timestamp
RICHARD HOLUBKOV u0337124@utah.edu Security Level: Email, Account Authentication (None)	DocuSigned by:  577A811E45874DE...	Sent: 1/16/2024 11:11:11 AM Viewed: 1/16/2024 11:12:53 AM Signed: 1/16/2024 11:15:37 AM
	Signature Adoption: Drawn on Device Using IP Address: 155.100.47.7	
Electronic Record and Signature Disclosure: Not Offered via DocuSign		
Mark Luciano markluciano@jhu.edu Security Level: Email, Account Authentication (None)	DocuSigned by:  AEFA78C9944F4C6...	Sent: 1/16/2024 11:15:39 AM Viewed: 1/18/2024 3:22:39 AM Signed: 1/18/2024 3:23:09 AM
	Signature Adoption: Drawn on Device Using IP Address: 73.129.29.91 Signed using mobile	
Electronic Record and Signature Disclosure: Accepted: 1/18/2024 3:22:39 AM ID: b6634a08-7945-4e65-b4b2-9db8f321792f		

In Person Signer Events	Signature	Timestamp

Editor Delivery Events	Status	Timestamp

Agent Delivery Events	Status	Timestamp

Intermediary Delivery Events	Status	Timestamp

Certified Delivery Events	Status	Timestamp

Carbon Copy Events	Status	Timestamp
Jess Wollet Jwollet1@jhmi.edu Security Level: Email, Account Authentication (None)	COPIED	Sent: 1/18/2024 3:23:11 AM
Electronic Record and Signature Disclosure: Not Offered via DocuSign		

Witness Events	Signature	Timestamp

Notary Events	Signature	Timestamp

Envelope Summary Events	Status	Timestamps
Envelope Sent	Hashed/Encrypted	1/16/2024 11:11:11 AM
Certified Delivered	Security Checked	1/18/2024 3:22:39 AM
Signing Complete	Security Checked	1/18/2024 3:23:09 AM
Completed	Security Checked	1/18/2024 3:23:11 AM

Payment Events	Status	Timestamps

Electronic Record and Signature Disclosure

ELECTRONIC RECORD AND SIGNATURE DISCLOSURE

From time to time, University of Utah (we, us or Company) may be required by law to provide to you certain written notices or disclosures. Described below are the terms and conditions for providing to you such notices and disclosures electronically through the DocuSign system. Please read the information below carefully and thoroughly, and if you can access this information electronically to your satisfaction and agree to this Electronic Record and Signature Disclosure (ERSD), please confirm your agreement by selecting the check-box next to 'I agree to use electronic records and signatures' before clicking 'CONTINUE' within the DocuSign system.

Getting paper copies

At any time, you may request from us a paper copy of any record provided or made available electronically to you by us. You will have the ability to download and print documents we send to you through the DocuSign system during and immediately after the signing session and, if you elect to create a DocuSign account, you may access the documents for a limited period of time (usually 30 days) after such documents are first sent to you. After such time, if you wish for us to send you paper copies of any such documents from our office to you, you will be charged a $0.00 per-page fee. You may request delivery of such paper copies from us by following the procedure described below.

Withdrawing your consent

If you decide to receive notices and disclosures from us electronically, you may at any time change your mind and tell us that thereafter you want to receive required notices and disclosures only in paper format. How you must inform us of your decision to receive future notices and disclosure in paper format and withdraw your consent to receive notices and disclosures electronically is described below.

Consequences of changing your mind

If you elect to receive required notices and disclosures only in paper format, it will slow the speed at which we can complete certain steps in transactions with you and delivering services to you because we will need first to send the required notices or disclosures to you in paper format, and then wait until we receive back from you your acknowledgment of your receipt of such paper notices or disclosures. Further, you will no longer be able to use the DocuSign system to receive required notices and consents electronically from us or to sign electronically documents from us.

All notices and disclosures will be sent to you electronically

Unless you tell us otherwise in accordance with the procedures described herein, we will provide electronically to you through the DocuSign system all required notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you during the course of our relationship with you. To reduce the chance of you inadvertently not receiving any notice or disclosure, we prefer to provide all of the required notices and disclosures to you by the same method and to the same address that you have given us. Thus, you can receive all the disclosures and notices electronically or in paper format through the paper mail delivery system. If you do not agree with this process, please let us know as described below. Please also see the paragraph immediately above that describes the consequences of your electing not to receive delivery of the notices and disclosures electronically from us.

How to contact University of Utah :

You may contact us to let us know of your changes as to how we may contact you electronically, to request paper copies of certain information from us, and to withdraw your prior consent to receive notices and disclosures electronically as follows:
To contact us by email send messages to: DocuSignAdmin@utah.edu

To advise University of Utah of your new email address

To let us know of a change in your email address where we should send notices and disclosures electronically to you, you must send an email message to us at DocuSignAdmin@utah.edu and in the body of such request you must state: your previous email address, your new email address. We do not require any other information from you to change your email address.

If you created a DocuSign account, you may update it with your new email address through your account preferences.

To request paper copies from University of Utah

To request delivery from us of paper copies of the notices and disclosures previously provided by us to you electronically, you must send us an email to DocuSignAdmin@utah.edu and in the body of such request you must state your email address, full name, mailing address, and telephone number. We will bill you for any fees at that time, if any.

To withdraw your consent with University of Utah

To inform us that you no longer wish to receive future notices and disclosures in electronic format you may:

i. decline to sign a document from within your signing session, and on the subsequent page, select the check-box indicating you wish to withdraw your consent, or you may;

ii. send us an email to DocuSignAdmin@utah.edu and in the body of such request you must state your email, full name, mailing address, and telephone number. We do not need any other information from you to withdraw consent.. The consequences of your withdrawing consent for online documents will be that transactions may take a longer time to process..

Required hardware and software

The minimum system requirements for using the DocuSign system may change over time. The current system requirements are found here: https://support.docusign.com/guides/signer-guide-signing-system-requirements.

Acknowledging your access and consent to receive and sign documents electronically

To confirm to us that you can access this information electronically, which will be similar to other electronic notices and disclosures that we will provide to you, please confirm that you have read this ERSD, and (i) that you are able to print on paper or electronically save this ERSD for your future reference and access; or (ii) that you are able to email this ERSD to an email address where you will be able to print on paper or save it for your future reference and access. Further, if you consent to receiving notices and disclosures exclusively in electronic format as described herein, then select the check-box next to 'I agree to use electronic records and signatures' before clicking 'CONTINUE' within the DocuSign system.

By selecting the check-box next to 'I agree to use electronic records and signatures', you confirm that:

- You can access and read this Electronic Record and Signature Disclosure; and
- You can print on paper this Electronic Record and Signature Disclosure, or save or send this Electronic Record and Disclosure to a location where you can print it, for future reference and access; and
- Until or unless you notify University of Utah as described above, you consent to receive exclusively through electronic means all notices, disclosures, authorizations, acknowledgements, and other documents that are required to be provided or made available to you by University of Utah during the course of your relationship with University of Utah .