Protocol Title: *Simple, Home-use neurostimulAtion tReatment for Parkinson's disease dEmeNtia* **(**Brief Title: SHARPEN)

STATISTICAL ANALYSIS PLAN v1.0

for **SHARPEN Study (ID: SNS-PD-004)**

Protocol Version and Date: Version 3.0 (February 12, 2024)
Version 2.0 (November 6, 2023)
Version 1.0 (January 10, 2023)

List of authors and reviewers

Authors

Martin Slade, MPH, PhD, Lead Trial Statistician, Lecturer in Medicine, Lecturer in Public Health, Director of Research, Yale Occupational & Environmental Medicine

Kristen Ade, PhD, Vice President and Chief Scientific Officer, Scion NeuroStim, Inc.

1. INTRODUCTION

This document describes the statistical analysis plan (SAP) for SNS-PD-004, the multicenter single arm safety and feasibility trial described in the protocol **Simple, Home-use neurostimulAtion tReatment for Parkinson's disease dEmeNtia** (Brief title: SHARPEN). This analysis plan is meant to supplement the study protocol. Any deviations from this analysis plan will be described in the clinical study report.

2. STUDY DESIGN

This is a single arm study designed to evaluate the safety and feasibility of twice daily time-varying caloric vestibular stimulation (tvCVS) treatments using a solid-state Device developed by Scion NeuroStim, Inc. (SNS), also known as ThermoNeuroModulation (TNM™), in patients with mild/moderate Parkinson's disease dementia (PDD) and to collect initial evidence to support proof of concept for the treatment of mild/moderate dementia in Parkinson's disease (PD). Participants will self-administer (with assistance, if necessary) ~19-minute *BID* treatments of tvCVS treatment over a period of 12 weeks (84 days). The study will be conducted at 2 centers located in the United States. This single-arm exploratory clinical trial will continue with competitive enrollment until 12 participants receive at least 1 treatment with the study Device at which point competitive enrollment will be closed.

2.1. STUDY OBJECTIVES

Primary Objectives: The primary objective will be to evaluate the safety and feasibility of time varying caloric vestibular stimulation (tvCVS), delivered with a solid-state TNM™ Device, in people with mild/moderate Parkinson's disease dementia (PDD).

Secondary Objectives: This study will also seek to assess the effects of tvCVS on cognition in people with mild/moderate PDD.

Exploratory Objectives: This study will also seek to further assess the effects of tvCVS on cognition and other signs and symptoms in people with mild/moderate PDD.

2.2. STUDY ENDPOINTS

Primary Endpoints

- Safety will be evaluated by determining whether tvCVS is associated with a worsening of PD signs and symptoms according to The International Parkinson and Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) combined measure of Parts I, II, and III (Goetz, Tilley et al. 2008). This endpoint will evaluate the change in the summed score of MDS-UPDRS Parts I, II and III at the end of the treatment period (day 84), which is defined based on the change relative to baseline (day 0) score.

Adverse events will also be monitored.

- Feasibility will be defined by evaluating both retention rate and treatment adherence rate. Specifically,
    - Retention rate will be defined by the percentage of participants that complete assessments required for analysis of the secondary endpoint:
        - High retention: >90%
        - Moderate retention: 60-90%
        - Low retention: < 60%
    - Treatment adherence rate for each participant will be defined as the percentage of twice-daily treatments with the neuromodulation device completed during the 12-week treatment period:
        - High adherence: >84%
        - Moderate adherence: 55-84%
        - Low adherence: < 55%

Secondary endpoint: The secondary endpoint will evaluate change in the Montreal Cognitive Assessment (MoCA).

Exploratory endpoints: The following outcome measures will be used:

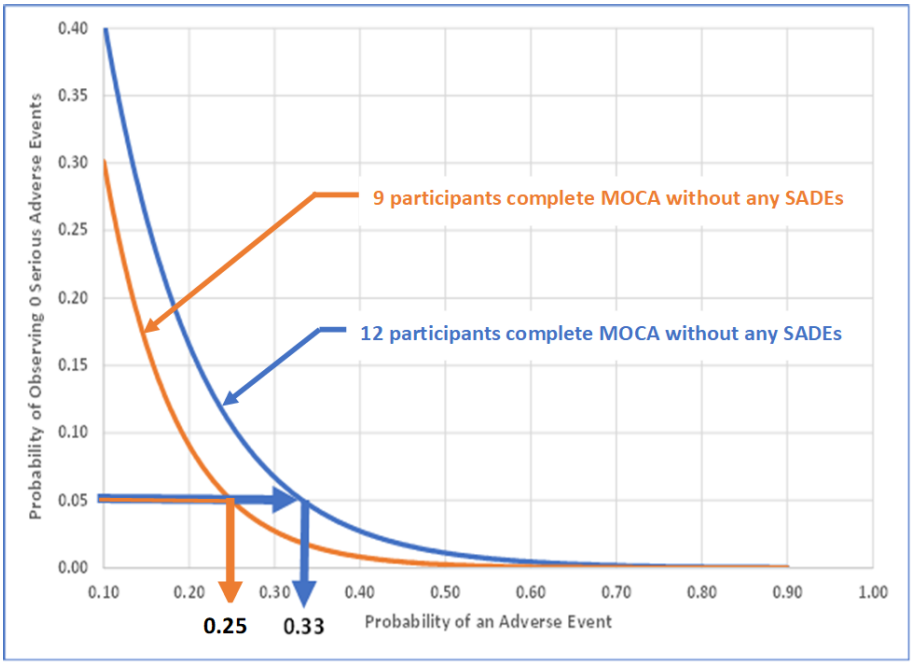
- The Zarit Burden Interview of Caregivers (ZBI)
- The Patient Global Impression of Improvement (PGI-I)
- The Neuropsychiatric Inventory - Questionnaire (NPI-Q)
- The Modified Schwab and England Activities of Daily Living Scale (S & E)
- A customized cognitive battery (CCB) which will assess the following functions:
    - Memory
        - Hopkins Verbal Learning Test-Revised (total learning)
        - Hopkins Verbal Learning Test-Revised (delayed recall)
        - Hopkins Verbal Learning Test-Revised (recognition discrimination)
    - Visuospatial
        - Neuropsychological Assessment Battery - Visual Discrimination
    - Language
        - Delis-Kaplan Executive Function System - Semantic Fluency
    - Attention
        - Neuropsychological Assessment Battery – Digits Forward
        - Hopkins Verbal Learning Test-Revised (Trial 1)
    - Processing speed
        - Oral Symbol Digit Modality
    - Executive Function
        - Delis-Kaplan Executive Function System - Semantic Fluency Switching
        - Neuropsychological Assessment Battery – Digits Backward

3. TREATMENT

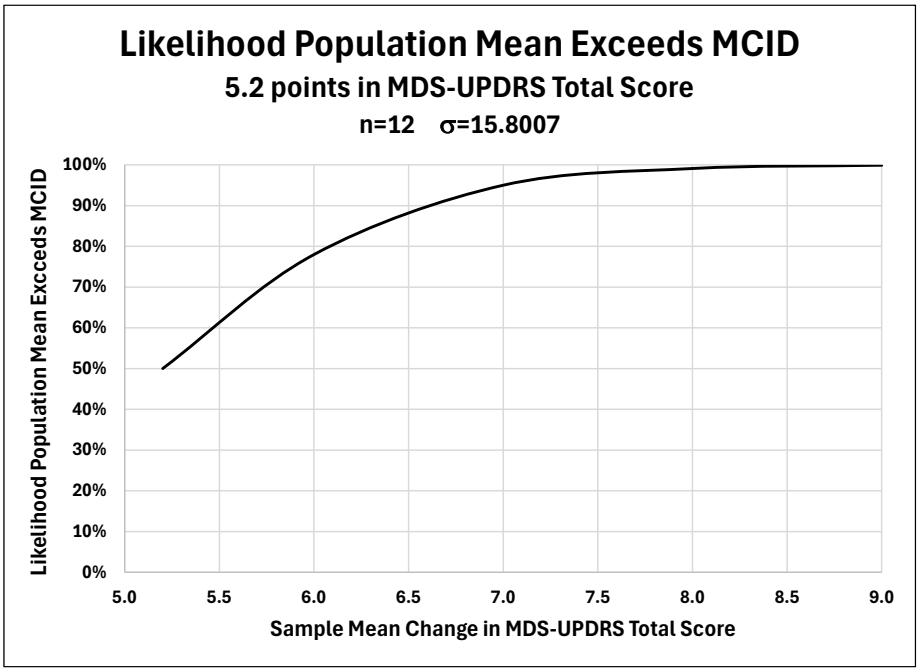
The study is single arm with all study participants receiving time-varying caloric vestibular stimulation treatments.

4. SAMPLE SIZE DETERMINATION

An enrollment number of 12 was set ensure adequate evaluation of the safety and the feasibility of the study design. Based on the Go/No Go Decision criteria for a follow-on Phase II randomized controlled trial (RCT) in PDD, this Phase I study would need to demonstrate that there were no Serious Adverse Device Effects (SADEs) and that at least 75% of the participants complete assessments required for analysis of the secondary endpoint. This second criterion would mean that at least 9 participants complete the MoCA at the end-of-treatment visit without any SADEs. Utilizing a Poisson distribution, fulfilling these two criteria would yield 95% probability that the true probability of a serious adverse event among the PDD population would be less than 0.33. If all 12 participants completed the MoCA at the end-of-treatment visit without any SADEs, then the study would yield 95% probability that the true probability of a SADEs among the PDD population would be less than 0.25. This relationship is depicted in the following figure:



The likelihood of the population mean being at least that of the outcome's MCID can be determined as a function of the average change in MDS-UPDRS Total score among the sample. The figure below depicts the relationship between the sample (n=12) mean change in MDS-UPDRS Total Score and the likelihood that the true population mean change is 5.20 points (MCID).



To illustrate the usefulness of the above figure, if the average change in MDS-UPDRS Total Score (i.e., the combined measure of Parts I, II and III) among the sample is 6.2 points, then there would be a 80% likelihood that the true population mean change is 5.20 points (i.e., the MCID value) indicating a clinically meaningful worsening of Parkinson's disease signs and symptoms.

5. INDEPENDENT DATA MONITORING

The TNM™ Device has been designated as a Non-Significant Risk Device for this study. However, the Sponsor has chosen to implement an Independent Medical Monitor (IMM) for the trial to provide independent safety monitoring and to confirm data integrity. The IMM is a physician with extensive experience in Parkinson's disease dementia and clinical trials and has no perceived or real conflict of interest with the Sponsor or in the study outcomes. The IMM will review the research protocol and ongoing study activities with emphasis on data integrity, protocol adherence and study participant safety issues while making recommendations to the continuation, modification or conclusion of the trial. The IMM will review data at least four times per year during the course of the RCT.

In addition to the IMM, the Sponsor will ensure critical safety and data points will be monitored per the SHARPEN Clinical Monitoring Plan (CMP) and the SHARPEN Data and Safety Management Plan (DSMP). A summary of key monitoring procedures is noted below:

Each site will be assigned a Clinical Research Associate (CRA) who will monitor the data to ensure the protection of rights and safety of human subjects, to verify the reported trial data are accurate, complete, and verifiable from source documents, and that the trial is conducted in compliance with the currently approved protocol/amendment(s), with Good Clinical Practice and with the applicable regulatory requirements. For all enrolled study participants, 100% source data review (SDR) and source data verification (SDV), where applicable, will occur of informed consent documents and process, inclusion and exclusion criteria, primary and secondary endpoints, all protocol deviations and safety reports, including adverse events and safety endpoints. Targeted source data verification will occur for all other endpoints and clinical data collected per the Targeted Source Data Verification (TSDV) portion of the CMP. Review of study-wide trends and key risk indicators metrics will be reviewed in a departmental meeting to occur at least once a month. If trends are identified, the Sponsor will review with the study-assigned medical monitor and may implement a Correction and Preventative Action (CAPA) Plan, as necessary. The study-specific Medical and Safety monitor, who has extensive experience with the device, will review and sign off on all adverse events (AEs) at least monthly, and all Unanticipated Adverse Device Effects (UADEs) within 2 days of acknowledgment of the event. All UADEs will be reported to the ethics committees and all site Investigators per the protocol and regulatory requirements.

No Serious Adverse Device Effects (SADEs) or UADEs are anticipated for this study, if two or more serious and related adverse events (SADE or UADE) are reported, the study enrollment will pause while universality and unblinded review occurs by the Medical and Safety monitor. Additionally, the IMM will review these events for adjudication and agreement and will provide recommendations to the Sponsor.

6. STATISTICAL METHODS

The primary objective of this study is to evaluate the safety and feasibility of tvCVS delivered using the TNM™ Device (Generation 4.0) as an adjuvant treatment for patients diagnosed with mild to moderate PDD.

6.1. ANALYSIS POPULATIONS

Intent-to-treat Population (ITT): All participants who start at least one treatment with the study device will be included in this primary analysis of efficacy and safety.

Per-protocol Population (PP): All eligible participants who (1) demonstrate at least 70% adherence of the randomized expected treatments with the Study Device and (2) have completed the regularly scheduled end of treatment visit (day 84), and (3) have not had changes to medications used to treat motor and/or non-motor symptoms of PD during the trial will be included in the PP population.

6.2. DEFINITIONS

Study day: The study day will be calculated in reference to the date of the baseline visit (Day 0).

No data collected at the study screen will be utilized in the evaluation of study outcomes.

6.3. ANALYSIS OF STUDY CONDUCT

The number of participants screened and the number of those randomized falling into the ITT and PP cohorts will be summarized.

Study treatment administration, duration of follow-up, discontinuation from study treatment and the reasons for discontinuation will be summarized by treatment group for all randomized participants. Coded protocol deviations and protocol violations will also be summarized as either Major/Important, Minor or No Impact, where:

- Major/Important are protocol deviations that might significantly impact completeness, accuracy, and/or reliability of study data OR could significantly affect a participant's rights, safety or well-being
- Minor are protocol deviations with minor impact on data quality or patient safety
- No impact are protocol deviations that do not impact data quality or patient safety

6.4. ANALYSIS OF TREATMENT GROUP

6.4.1. Demographics and Baseline Characteristics

The following characteristics will be summarized for participants in each cohort (i.e., ITT and PP): age, sex, years since PD diagnosis, levodopa equivalent daily dose, race and ethnicity. Sex, race and ethnicity will be described with frequency and percentages. Age, years since PD diagnosis and levodopa equivalent daily dose will be described with minimum, 1Q, median, 3Q, and maximum values.

6.4.2. Disease characteristics

Baseline Hoehn & Yahr scores will be summarized. Additionally, motor phenotypes (i.e., tremor-dominant, postural instability gait difficulty and intermediate) classified based on the Baseline visit MDS-UPDRS part II and part III scores and previously-established methodology(Stebbins, Goetz et al. 2013). Specifically, a Tremor score will be calculated by summing all items from 2.10, 3.15, 3.16, 3.17 and 3.18 from the MDS-UPDRS. A postural instability and gait difficulty (PIGD) score will be calculated by summing all items from 2.12, 2.13, 3.10, 3.11 and 3.12 from the MDS-UPDRS. Ratios of Tremor score/ PIGD score will be used to classify motor phenotypes. Ratios ≥ 1.15 = Tremor dominant. Ratios ≤ 0.90 = PIGD dominant. Scores > 0.9 and < 1.15 = intermediate. If data required to calculate the scores is missing from the Baseline 2 visit, data from the Baseline 1 visit will be analyzed. Post randomization data will not be considered. Classification will only be

performed on complete case data. Disease characteristics will be described with minimum, 1Q, median, 3Q, and maximum values.

6.4.3. Pre-study and Concomitant Medications

The number and percentage of each treatment cohort on concomitant therapies at randomization will be reported categorically including the following:

- adenosine A2A antagonists
- amantadine
- anticonvulsants
- anti-depressant/antianxiety medications
- antiemetics
- antihistamines
- antihypertensives
- antipsychotics
- central anticholinergics
- cognitive enhancers
- COMT Inhibitors
- contraceptives
- corticosteroids
- dystonia treatments
- incontinence treatments
- inhaled levodopa
- inhaled or ingested cannabinoids
- laxatives/stool softeners
- MAO-B inhibitors
- narcotics
- oral or transdermal dopamine agonists
- oral levodopa-based therapies
- oral levodopa-based therapies/COMT inhibitors
- orthostatic hypotension treatments
- pain/cramping treatments
- pump therapies
- sexual dysfunction treatments
- sialorrhea treatments
- sleep aids
- stimulants

Changes in medications to treat symptoms associated with PD during the trial are a protocol deviation. Changes in these medications would exclude participants from the PP analysis. Changes in medications in the ITT cohorts will be summarized descriptively.

6.5. ENDPOINT ANALYSIS

All efficacy analyses will assume non-normality given the small sample size of the study. Paired comparisons (i.e., pre- and post-treatment measurements) will utilize Wilcoxon signed-rank tests unless otherwise noted.

6.6.1 PRIMARY ENDPOINTS

(Safety) MDS-UPDRS Parts I, II & III summed score: Changes will be considered to be clinically meaningful worsening if the median difference meets or exceeds the previously established MCID (+5.2 points for clinical worsening) (Makkos, Kovacs et al. 2018). Two-sided, 90% confidence intervals for the combined MDS-UPDRS Parts I, II, and III scores at baseline and the end of treatment period will be constructed. If the median difference (baseline- end of treatment) is less than 5.2 and the confidence interval of the difference (baseline – end of treatment) contains 5.2, then the no clinically meaningful worsening.

(Feasibility)

- **Retention rate** will be defined by the percentage of participants that complete assessments required for analysis of the secondary endpoint:
    - High retention: >90%
    - Moderate retention: 60-90 %
    - Low retention: < 60 %
- **Treatment adherence** rate for each participant will be defined as the percentage of twice-daily treatments with the neuromodulation device completed during the 12-week treatment period:
    - High adherence: >84%
    - Moderate adherence: 55-84 %
    - Low adherence: < 55 %
    - *Treatment adherence is captured by the study device and will be reported summarized with minimum, 1Q, median, 3Q, and maximum values. Treatment adherence of ≥70% is required for inclusion in the Per Protocol cohort.*

Adverse Events (AEs): Participants will be assessed for AEs every two weeks at minimum (either at study visits or during phone calls). For each group, the AEs will be summarized with frequency and percentage by preferred (PT) term, with all participants in that treatment group as the denominator. Classification will utilize MedDRA System Organ

Class (SOC) based on MedDRA® Version 26.0 or other terminology/classification common to Parkinson's disease (e.g., falls, freezing of gait, etc.). AE incidence will be summarized by both severity and causal relationship to Device treatment (as determined by the blinded Principal Investigator or designated study personnel). AEs deemed to be of "possible" or "probable" relationship or "related" to the Device will also be considered as device related. Additionally, the number of events and number of device-related events per person will be reported as a distribution.

The AE summary tables will provide an overall summary of AEs including the number and percentage of participants who experienced any AE, any SAE, any ADE, any SADE, any UADE, and any discontinuations in study participation due to an AE.

6.6.2 SECONDARY ENDPOINTS

The secondary endpoint will be the change at the end of treatment or early termination visit from baseline in the Montreal Cognitive Assessment (MoCA) score.

6.6.3 EXPLORATORY ENDPOINTS

These exploratory endpoints have been added to identify outcome measures to utilize in future trials. Mean or median change with corresponding 95% confidence intervals and p values will be provided. No adjustments will be made to address a multiplicity of endpoints. Only participants who have baseline and post-treatment data collected for the exploratory endpoints will be included in these efficacy analyses.

Exploratory outcomes will evaluate the change in the following measures:

- The Zarit Burden Interview of Caregivers (ZBI)
- The Patient Global Impression of Improvement (PGI-I): The Patient Global Impression-Improvement - a patient determined scale to assess how much their illness has improved or worsened relative to a baseline state at the beginning of the intervention (starting with the RCT). Scoring: 1 = very much better, 2 = much better, 3 = a little better, 4 = no change, 5 = a little worse, 6 = much worse, 7 = very much worse. For ease of interpretation, the PGI-I score will be converted by the following formula: Converted PGI-I = 4 - PGI-I. Thus a 0 corresponds to no change, higher magnitude positive values correspond to greater improvements and higher magnitude negative scores correspond to increased worsening. If the lower limit of the 95% CI is positive, this would indicate improvement.
- The Neuropsychiatric Inventory - Questionnaire (NPI-Q)
- The Modified Schwab and England Activities of Daily Living Scale (S & E)
- A customized cognitive battery (CCB) which will assess the following functions
    - Memory
        - Hopkins Verbal Learning Test-Revised (total learning)
        - Hopkins Verbal Learning Test-Revised (delayed recall)

        - Hopkins Verbal Learning Test-Revised (recognition discrimination)
    - Visuospatial
        - Neuropsychological Assessment Battery - Visual Discrimination
    - Language
        - Delis-Kaplan Executive Function System - Semantic Fluency
    - Attention
        - Neuropsychological Assessment Battery – Digits Forward
        - Hopkins Verbal Learning Test-Revised (Trial 1)
    - Processing speed
        - Oral Symbol Digit Modality
    - Executive Function
        - Delis-Kaplan Executive Function System - Semantic Fluency Switching
        - Neuropsychological Assessment Battery – Digits Backward
- *Analysis of the customized cognitive battery will be described in a supplemental statistical analysis plan and will not be included in the main statistical analyses for this study.*

7. REFERENCES

Goetz, C. G., B. C. Tilley, S. R. Shaftman, G. T. Stebbins, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, M. B. Stern, R. Dodel, B. Dubois, R. Holloway, J. Jankovic, J. Kulisevsky, A. E. Lang, A. Lees, S. Leurgans, P. A. LeWitt, D. Nyenhuis, C. W. Olanow, O. Rascol, A. Schrag, J. A. Teresi, J. J. van Hilten, N. LaPelle and U. R. T. F. Movement Disorder Society (2008). "Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results." Mov Disord **23**(15): 2129-2170.

Makkos, A., M. Kovacs, Z. Aschermann, M. Harmat, J. Janszky, K. Karadi and N. Kovacs (2018). "Are the MDS-UPDRS-Based Composite Scores Clinically Applicable?" Mov Disord **33**(5): 835-839.

Stebbins, G. T., C. G. Goetz, D. J. Burn, J. Jankovic, T. K. Khoo and B. C. Tilley (2013). "How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale." Mov Disord **28**(5): 668-670.