Statistical Analysis Plan

For CLP 11899

MIND: A Prospective, Multicenter Study of Artemis a **M**inimally **I**nvasive **N**euro Evacuation **D**evice, in the Removal of Intracerebral Hemorrhage

Prepared by: Penumbra Inc.
NCT number: NCT03342664
Date: January 02, 2020

1. Overview

This is a multi-center randomized controlled study designed to compare the safety and efficacy of minimally invasive hematoma evacuation with the Artemis Neuro Evacuation Device to best medical management for the treatment of intracerebral hemorrhage (ICH).

Up to 500 patients will be enrolled from up to 50 global centers (US and OUS). Each site will be limited to a maximum enrollment of up to 120 patients (~20% of total enrollment).

This Statistical Analysis Plan (SAP) will provide details to further elaborate statistical methods as outlined in the protocol and will describe analysis conventions. The SAP will be signed off prior to unblinding.

2. Sample Size

A sample size of up to 500 patients was selected for this study. Based on simulations, this trial has 81% power for a cumulative odds ratio of 1.7 with a one-sided alpha of 0.025. The minimum sample size is 200 patients (approximately 133 MIS and 67 MM).

The Day 180 mRS distribution used for sample size estimation was based on control arm data from the following published studies: MISTIE II (Hanley 2016), STICH (Mendelow 2005), and SICHPA (Teernstra 2003). The distribution of mRS functional outcomes utilized in the sample size estimation is provided in Table 1.

Table 1. Distribution of MRS functional outcomes

mRS Distribution	**0**	**1**	**2**	**3**	**4**	**5/6**
MM	3%	10%	7%	7%	14%	59%
MIS	5%	15%	10%	9%	16%	46%

3. Randomization

Randomization takes place centrally through a commercially available Interactive Web Response System (IWRS). Randomization will occur in a 2:1 ratio to either MIS or MM. The treatment allocation will be balanced by Hemphill Score (0-2, 3-4) and hemorrhage location (primarily lobar, primarily deep).

If a lack of equipoise for randomization between treatment and control occurs during trial execution, the trial randomization assignment will be revised to randomize all patients to the MIS. The minimum sample size for patients assigned to MIS treatment is 133 patients.

4. Interim Analysis

Interim data analysis of the primary efficacy and safety endpoints is planned after 200 subjects have been enrolled. Additional interim analyses will be conducted after every 50 subjects. The interim analyses will include imputing the 180-day mRS score for subjects who do not have 180 day scores. Based on the predictive probability that the study would be a success, the study may be stopped early for futility or enrollment may be stopped based on expected success. The specific mathematical details of the planned analyses are described in the adaptive design report.

If a lack of equipoise for randomization between treatment and control occurs during trial execution, the supplemental adaptive design report provides the details to evaluate the superiority of MIS as compared to the combined MM arms of the MIND trial and any publicly available trial results with randomized MM arms.

5. Decision Rules for Adjusting the Sample Size

The maximum study sample size is 500 subjects and the minimum sample size is 200 subjects. The final study sample size is estimated to be low if the treatment effect is small and the study could be stopped for futility or if the treatment effect is large and the study could be stopped for expected success. The final study sample size will be high if the evidence for the treatment effect is inconclusive. The specific decision rules for sample size adjustments are described in the adaptive design report.

6. Analysis Populations

All primary and secondary effectiveness endpoints will be performed for both the intent-to-treat (ITT) population and per-protocol (PP) population.

6.1. Target Population: The target population is patients 18-80 years of age who have a diagnosis of spontaneous, non-traumatic, intracerebral hemorrhage (ICH) ranging in volume between and including 20 and 80 cc, with an associated significant neurological deficit (NIHSS > 6) who do not require emergent open surgical decompression related to uncontrolled intracranial pressure or mass effect.

6.2. Intent to Treat Sample: As the primary analysis, all efficacy and safety outcome measures will be analyzed under the intent-to-treat (ITT) principle. Under this principle, the evaluable sample includes all subjects who are randomized. Each subject will be analyzed according to the treatment group to which they were randomly assigned at the time of randomization. This population is the primary population for all efficacy parameters.

6.3. Per Protocol Sample: In addition to the defined ITT analysis sample, a per-protocol (PP) sample is defined as a subset of the ITT sample. The per-protocol sample will include all randomized subjects that do not have significant protocol deviations (e.g. eligibility violation, crossover).

6.4. Safety Analysis (As Treated) Sample: In the case of cross-overs, a safety sample that is the same as the ITT sample will be examined in which subjects will be analyzed according to the actual treatment received. Subjects who receive Artemis device-based therapy are included in the MIS arm and subjects who receive only medical therapy are included in the MM arm.

6.5. The following additional population definitions apply to this study:

6.5.1. Screened: All patients considered for participation in the study, whether or not they sign an informed consent.

6.5.2. Screen Failure: All patients considered for participation in the study, who failed to meet inclusion or met exclusion criteria. Patients can be screen failed based on general or imaging criteria. These patients may or may not have signed an informed consent.

6.5.3. Enrolled (Randomized): All subjects who have been randomized based on the result of the baseline imaging and other inclusion/exclusion criteria. Informed consent must be obtained prior to randomization.

6.5.4. Completed: All subjects who were enrolled (randomized) and completed the study follow-up or were known to have died prior to the follow-up timepoint. The completed subject metric will be provided for Day 180 and Day 365 follow-up.

6.5.5. Early Termination: Subjects who were enrolled (randomized) but did not complete follow-up and were not known to have died. The early termination subject metric will be provided for Day 180 and Day 365 follow-up.

7. Statistical Methods

Summary tables (descriptive statistics and/or frequency tables) will be provided for all baseline variables, efficacy variables, and safety variables, as appropriate. Continuous variables will be summarized with descriptive statistics (n, mean, standard deviation, range, and median). Frequency counts and percentage of subjects within each category will be provided for categorical data. Based on the variable distribution, parametric or non-parametric tests will be used for comparisons between groups.

8. Baseline Characteristics

Baseline data will be analyzed to assess the comparability of treatment groups. Baseline data including, but not limited to demographics, clinical characteristics, and baseline ICH characteristics will be summarized using descriptive statistics. Statistical testing will be performed as appropriate.

9. Subject Disposition

The number of subjects for each of the following categories will be summarized.

- Screened patients

- Screen failures
- Randomized (Enrolled) subjects
- Subjects randomized that did not receive assigned treatment
- Subjects completing the study; subjects not completing the study
- Subjects included in the intent to treat population
- Subjects included in the per protocol population
- Subjects included in the safety population

10. Efficacy Analysis

10.1. Primary Efficacy Analysis

The primary endpoint is the Day 180 global disability assessed via the ordinal modified Rankin score (mRS).

The null hypothesis is that the cumulative odds ratio for mRS at 180 days in the MIS group compared to MM group is less than or equal to 1. The alternative hypothesis is that the cumulative odds ratio for mRS at 180 days is greater than 1. Formally, the null and alternative hypotheses to be tested are as follows:

H0: OR ≤ 1

HA:OR > 1,

where OR is the cumulative odds ratio for the mRS at the 180-day follow-up visit, with higher values indicating better outcomes in the MIS group.

Statistical analysis of the primary endpoint will be conducted using a logistic regression analysis of the 180-day mRS scores. The primary efficacy endpoint is met if the overall treatment effect is positive at a one-sided alpha of 0.02. The odds ratio and corresponding 95% confidence interval will be estimated from the proportional odds model. The proportional odds assumption will be assessed visually for the cumulative proportions and odds ratio estimates. The primary analysis will be unadjusted. A secondary analysis model will include the minimization variables of Hemphill Score and hemorrhage location. The mRS scores of 5 and 6 will be combined into a single group for the purposes of endpoint evaluation. Subjects deceased during study follow-up will be scored as mRS 6. As a sensitivity analysis, the mRS scores of 5 and 6 will be evaluated as separate groups.

10.2. Secondary Efficacy Analysis

The secondary endpoints of Day 180 and Day 365 mRS average improvement in mRS will be analyzed using a generalized linear model. The mRS scores will be weighted as the following: 1.0 for mRS level 0; 0.91 for mRS level 1; 0.76 for mRS level 2; 0.65 for mRS level 3; 0.33 for mRS level 4; 0 for mRS level 5; and 0 for mRS level 6. Statistical analysis of the dichotomized Rankin outcome scores of 0 to 2 and 0 to 3 will be conducted with a logistic regression model. Day 365 mRS analysis will also include ordinal and dichotomous (0 to 2; 0 to 3) outcomes. Group differences will be analyzed for the following: SIS-ADL, SIS-mobility, EQ-5D-5L, length of stay, length of procedure.

10.3. Health Economics Information

The study site will complete CRFs containing healthcare utilization information (e.g. ICU days). This information may be used for analyses to compare overall healthcare costs and resource utilization between MIS and MM.

10.4. Handling of Multiplicity

There will be no adjustment on the comparison between MIS and MM on the primary effectiveness variable since the primary comparison is specified in the protocol. The following 8 secondary efficacy analyses will be adjusted using the Bonferroni correction and may be used for labeling claims:

- Stroke Impact Scale – Mobility at 180 days
- Stroke Impact Scale – ADLs at 180 days
- Stroke Impact Scale – Mobility at 365 days
- Stroke Impact Scale – ADLs at 365 days
- EQ-5D-5L at 180 days
- EQ-5D-5L at 365 days
- Length of hospital stay
- Length of ICU stay

10.5. Efficacy Subgroup Analysis

To evaluate the impact of baseline conditions on treatment effect on functional outcome, subgroup analyses will be performed for the primary efficacy variable, 180-day ordinal mRS. Subgroup comparisons will be considered secondary analyses and will not be adjusted or used for labeling claims. The subgroups below will be used for these analyses:

- **Age** (< 65, or >= 65)
- **Gender** (Male, or Female)
- **Race/Ethnicity**
- **Baseline Hemphill (**0-2, or 3-4)
- **Site of hemorrhage** (primarily lobar, primarily deep)
- **Geographic Location** (US, OUS)
- **Time from stroke symptom onset to randomization** (below the median, above the median)
- **Time from stroke symptom onset to sheath placement** (below the median, above the median)
- **Baseline GCS for ICH volume 20-30 cc (A x B x C/2)** (GCS >= 9, GCS <=8)

The subgroup analysis will be conducted using regression with terms of treatment group and treatment-by-subgroup interaction. In addition, Hemphill Score and hemorrhage location will be included in the model. The primary statistical inference is the treatment-by-subgroup interaction, which is tested at the significance level of 0.100. These analyses will be performed on the Intent-To-Treat population. When the treatment-by-subgroup interaction is statistically significant ($p <= 0.100$) for a specific subgroup, the treatment group differences will be evaluated within each stratum of that subgroup. Analyses stratified by geographic location will be examined for consistency with the overall trial result. A stepwise multivariate regression model with a subset of best baseline predictors for the primary endpoint of mRS will be used for the stratified analyses for consistency with the overall study result with respect to mRS. The time from stroke symptom onset to sheath placement will be evaluated for the MIS group only.

11. Safety Analysis

11.1. Primary Safety Analysis

The primary safety endpoint is 30 day mortality. The data will be analyzed as a binary variable with each subject counted only once. The primary safety analysis is an analysis of all patients according to treatment received. The treatment group difference and 95% confidence interval will be employed to assess the primary safety endpoint.

11.2. Analysis of Adverse Events

Adverse events will be coded using the MedDRA dictionary. The number and percentage of subjects with AEs and SAEs will be summarized by body system and preferred term. Each subject will be counted only once within a category. The specific categories analyzed will be those that are reported by at least three (3) percent of the subjects. The Clinical Events Committee adjudicated data supersedes the investigator reported data in all analyses of adverse events.

11.3. Handling of Multiplicity

There will be no adjustment for multiple comparisons between MIS and MM on the primary safety variable since the primary comparison is specified in the protocol; all other safety comparisons will be considered secondary analyses and will not be adjusted or used in labeling claims.

11.4. Analysis of Deaths

The Kaplan-Meier product-limit method will be the primary method utilized to assess the mortality rate. With the date of randomization set at Day 0, any death occurring on or before calendar day 30 will be counted as a death. If clinical assessment is missing for a patient who has not died, the patient will be censored at the last follow-up date. Patients who are alive at day 30 will be censored at day 30. The log-rank test will be used to compare the groups. This comparison weights earlier and later differences equally. The time to death will be plotted with confidence intervals at monthly intervals.

Additionally, the death data will be presented as binary deaths. The number of deaths will be presented for each group. A proportional hazards model may be used to account for confounders and center effects, if applicable.

12. Pooling Across Centers

Analyses will be presented by treatment group using data pooled across site. The site analysis will be conducted using an ordinal logistic regression with terms of treatment group and treatment-by-site interaction. This analysis will be performed on the Intent-To-Treat population. The primary statistical inference is the treatment-by-site interaction, which is tested at the significance level of 0.15. When the treatment-by-site interaction is statistically significant ($p \leq 0.15$), the treatment group differences will be evaluated within each site. Adjusted analysis on the primary outcome using key baseline variables will be used for the site analyses for consistency with the overall study result. If the odds ratio of the treatment effect is found to vary by site, then a random-effects model analysis will be performed to assess whether there was significant variance in the primary endpoint according to study site.

13. Lost to Follow-up and Missing Data

Subjects not completing the 180-day follow-up mRS will be imputed for the primary endpoint using the mRS as of the last available follow-up visit (i.e. Day 30, Day 90) and the Bayesian longitudinal model. The median mRS score will be imputed for each patient corresponding to the last mRS value and the longitudinal model posterior distributions. The imputed mRS scores will be utilized in both the interim and final primary analyses.

As a sensitivity analysis, we will also perform the following analyses:

- include only patients with complete 180-day mRS evaluation
- impute patients with missing 180 day mRS as mRS 5/6
- impute all missing MIS patients as mRS 5/6 and all missing MM patients as mRS 0
- utilize multiple imputation from the longitudinal model to calculate the Bayesian posterior probability of superiority
- estimate 180-day mRS by extending the longitudinal multiple imputation models to include strata for baseline Hemphill Score and hemorrhage location

14. Blinding

The protocol is designed to have open label treatment assignment. The Penumbra, Inc. clinical team, or delegate is responsible for the conduct of the study, the investigator, site study personnel, and the subject will not be blinded to each subject's randomized treatment group throughout the course of the study. The evaluator who performs the 180-day mRS assessment will be blinded to the treatment received by the patient and will be instructed to follow a scripted interview to control for potential bias.

While the CEC review of adverse events specific to the interventional procedure will unblind the members, all members of the CEC will be blinded to the overall primary results of the study.

15. Control of Systematic Error and Bias

Randomization takes place centrally through a commercially available Interactive Web Response System (IWRS). The clinical study will be conducted under a common protocol for each investigational site with the intention of pooling the data for analysis. Every effort will be made to promote consistency in study execution at each investigational site.

The interim analysis of the primary endpoint will be conducted by an independent statistician.

16. Committees

16.1. Clinical Events Committee (CEC)

A Clinical Events Committee will review and adjudicate clinical events for causality and attribution.

16.2. Core Lab

An independent Core Lab will review and score all imaging scans for at a minimum hematoma volume.

16.3. Data Safety Monitoring Board (DSMB)

A Data Safety Monitoring Board (DSMB) will monitor the overall safety during the clinical study.

17. Changes to Planned Analyses

All changes to the statistical analysis plan (SAP) will be documented in a revised SAP or the clinical study report.

18. References

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Agresti, A. (2012). Categorical Data Analysis (Third Edition). *John Wiley & Sons.*

Chaisinanunkul et al (2015) Adopting a Patient-Centered Approach to Primary Outcome Analysis of Acute Stroke Trials Using a Utility-Weighted Modified Rankin Scale. *Stroke* 46(8);2238-43.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013) Bayesian Data Analysis, Third Edition. *Chapman & Hall/CRC Texts in Statistical Science*

Hanley, Daniel F., et al. "Safety and efficacy of minimally invasive surgery plus alteplase in intracerebral haemorrhage evacuation (MISTIE): a randomised, controlled, open-label, phase 2 trial." *The Lancet Neurology* 15.12 (2016): 1228-1237.

Hollander, M., Wolfe, D. A., & Chicken, E. (2013). Nonparametric Statistical Methods (Second Edition). *John Wiley & Sons.*

Kalbfleisch, J. D., & Prentice, R. L. (2011). The Statistical Analysis of Failure Time Data (Second Edition). *John Wiley & Sons.*

Mendelow AD, Gregson BA, Fernandes HM, Murray GD, Teasdale GM, Hope DT, et al. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the international surgical trial in intracerebral haemorrhage (STICH): A randomised trial. *Lancet*. 2005;365:387-397

Stokes, M. E., Davis, C. S., & Koch, G. G. (2012). Categorical Data Analysis Using SAS® (Third Edition). *Cary, NC, SAS Institute.*

Teernstra, O. P. M., Evers, S. M. A. A., Lodder, J., Leffers, P., Franke, C. L., & Blaauw, G. (2003). Stereotactic treatment of intracerebral hematoma by means of a plasminogen activator. *Stroke*, 34(4), 968-974.