STATISTICAL ANALYSIS PLAN

A randomized, concurrent controlled trial to assess the safety and effectiveness of the Separator 3D as a component of the Penumbra System in the revascularization of large vessel occlusion in acute ischemic stroke

---

Protocol CLP4853

2/5/2013

Contents

1 Overview

This is a 90-day multi-center, randomized, single blind study designed to evaluate the effectiveness and safety of the Penumbra System versus the Penumbra System with Separator 3D in the treatment of adult patients with acute ischemic stroke from large vessel occlusion in the cerebral circulation.

Approximately 230 adult patients will be enrolled from approximately 50 centers. Each site will be limited to a maximum enrollment of approximately 45 patients (~20% of total enrollment).

Patients will be randomized 1:1 to receive one of the following:

- Penumbra System (control);
- Penumbra System with Separator 3D (3D).

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 database lock.

2 Study Success Criteria

The study will be considered a success if the primary effectiveness non-inferiority criterion is met.

Primary Effectiveness Non-Inferiority: The proportion of patients in the Penumbra System with Separator 3D group who experience TICI 2/3 is no more than .15 below the proportion of patients in the Penumbra System group who experience TICI 2/3.

Additional analyses conducted for secondary endpoints and subgroups will be presented, but not utilized in evaluating the study success or Penumbra System product claims.

3 Sample Size

Approximately 230 patients are to be randomized in a 1:1 ratio to the Penumbra System group (115 patients) or the Penumbra System with Separator 3D group (115 patients).

3.1 Non-inferiority Margin Evaluation

The primary efficacy analysis will be performed by comparing the lower bound of a one-sided, 95.0% confidence interval, calculated using normal approximation, for the observed difference between the Penumbra System with Separator 3D and the control on the percentage of subjects that have a post procedure score of TICI 2-3 with a pre-specified non-inferiority margin (-15%). Should the lower limit of this difference be > -15%, it can be concluded that the Penumbra System with Separator 3D is non-inferior to the control with respect to the primary efficacy measure.

3.2 Effectiveness Sample Size Evaluation

The sample size calculations assume that 80% of the standard Penumbra System patients experience success (TICI of 2 to 3) and 80% of the Penumbra System with Separator 3D patients experience success. Based on a binomial non-inferiority analysis with a non-inferiority margin of 15%, a study of 103 patients per group will have 85% power with a one-sided alpha of 0.05. The sample size was adjusted to 115 patients per group to account for up to10% attrition rate.

Assuming that both the 3D arm and the control arm observe a rate of 80% (82/103), the 95% confidence interval for the difference between groups is (-11.0%, 11.0%).The sample size was calculated for a two-group, one-sided test for two proportions using the normal approximation to the binomial using SAS v 9.2 (SAS Institute, Cary, NC). The output of the proc power procedure for a Pearson Chi-square Test for Two Proportions is provided below.

Computed N Total	
Actual Power	**N Total**
0.852	206

The sample size calculations support enrolling a total of 230 patients (Table 1).

Table 1. Effectiveness Sample Size Parameters

Penumbra System Success Proportion	0.8
Penumbra System with Separator 3D Success Proportion	0.8
Enrollment Ratio (Control: 3D)	1:1
Power	85%
Alpha	0.05
Number of Sides	1

Null Proportion Difference	-0.15
Sample Size	206 (103 per group)
Estimated Attrition	10%
Total Sample Size	230 (115 per group)

3.3 Randomization Methods

Patients who are eligible based on inclusion and exclusion criteria and have had all pre-randomization screening procedures performed will be randomized by center in a 1:1 ratio to one of following two treatment groups:

- Penumbra System group (control)
- Penumbra System with Separator 3D group (3D)

The interactive voice response (IVRS) will be used to determine the randomization of patients. The IRVS vendor statistician will provide the randomization schedule. The Penumbra Inc. statistician will review the randomization schedule. Detailed instructions for using the IVRS will be provided to center personnel.

3.4 Blinding

The Penumbra Inc. clinical team responsible for the conduct of the study, the investigator, and center study personnel will not be blinded to each patient's randomized treatment group throughout the course of the study. The study subjects and the Core Laboratory will be blinded to treatment assignment.

4 Interim Analysis

No interim analyses are planned for the purpose of stopping the study early.

5 Analysis Populations

5.1 Definitions

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

The ITT population will consist of all patients who signed the informed consent and are randomized in the study. The data from the ITT population will be analyzed by the treatment group assignment given at the time of randomization, even if the subject does not receive the correct treatment, or does not follow the protocol until completion. All randomized patients will

be followed and assessed for 90 days post-procedure, even when no neurothrombectomy is performed.

For the PP analysis, only patients who actually received the assigned treatment and do not have major protocol violations will be included.

For the safety analysis, patients will be analyzed according to the treatment they received.

Any treated patients who are not randomized through the IVRS system will be excluded from all analyses. The reason(s) for their exclusion from the study will be recorded. Listings will be provided for these patients, and they will be discussed in the clinical report.

Roll-in patients will be followed for 90 days and their safety will be summarized separately in the clinical report.

5.2 Randomization Variables

The randomization of the study will be stratified based on the following:

- Clinical study center

6 Statistical Methods

Summary tables (descriptive statistics and/or frequency tables) will be provided for all baseline variables, effectiveness 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 patients within each category will be provided for categorical data.

7 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 vessel characteristics will be summarized using descriptive statistics. Differences between the treatment groups and their 95% confidence intervals will be calculated. Statistical testing will be performed as appropriate.

8 Patient Disposition

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

- Randomized patients
- Patients completing the study; patients not completing the study
- Patients included in the intent to treat population
- Patients included in the per protocol population
- Patients included in the safety population

9 Effectiveness Analysis

9.1 Primary Effectiveness Analysis

The primary effectiveness variable is the proportion of patients with a post procedure TICI score of 2 to 3. The proportion of patients in each group who are successful based on this criterion will be calculated. The Core Laboratory data supersede the investigator-reported data in all analyses of TICI scoring.

For those patients who are treated with IA tPA or other non-FDA-approved treatments for the purpose of reducing the clot burden, their TICI score will be analyzed as a treatment failure (i.e. TICI 0 to 1) for the primary effectiveness analysis.

The primary effectiveness analysis will be the difference between the Penumbra System only group (control) and the Penumbra System with Separator 3D group. A binomial comparison will be used to test the one-sided null hypothesis that the difference in proportions is less than or equal to -0.15 (H_0: $P_{3D} - P_{Control} \leq -0.15$) vs. the alternative ($H_1$: $P_{3D} - P_{Control} > -0.15$). This is equivalent to evaluating that the lower bound of the 90% confidence interval for the difference is above 15%. The primary effectiveness analysis will be performed on the Per Protocol (PP) subject population. The analysis based on the Intent-To-Treat (ITT) population will be considered as supportive.

9.2 Secondary Effectiveness Analysis

The secondary effectiveness variables:

- Good clinical outcome at 30 days post-procedure. Patients meeting one of the following are evaluated as having good clinical outcome: 10 points or more improvement in the NIHSS at Discharge; a NIHSS score of 0-1 at Discharge; 30-day mRS score of 0-2.

- Good neurological outcome at 90 days post-procedure. Patients meeting one of the following are evaluated as having a good neurological outcome at 90 days:
    - mRS score ≤ 2 at 90 days
    - mRS score at 90 days equal to the pre-stroke mRS score (if the pre-stroke mRS score was higher than 2)
    - 10 or more points improvement on the NIHSS score at 90 days
- The proportion of patients with a modified Rankin Scale (mRS) of ≤ 2 at 90 days post procedure.

Frequency counts and percentage of patients within each category will be provided for categorical data. Estimates of the treatment differences and their 95% confidence intervals will be calculated. Subject rates will be compared between treatment groups with Fisher's exact test.

9.3 Handling of Multiplicity

There will be no adjustment for multiple comparisons between the Penumbra System versus the Penumbra System with Separator 3D on the primary effectiveness variable since the primary comparison is specified in the protocol. All other comparisons will be considered secondary analyses and will be adjusted using the Bonferroni correction.

9.4 Effectiveness Subgroup Analysis

To evaluate the impact of baseline conditions on treatment effect, subgroup analyses will be performed for the primary effectiveness variable, TICI 2-3, the secondary effectiveness variables (good clinical outcome at 30 days, 90-day mRS 0-2, good neurological outcome at 90 days) and the primary and secondary safety variables. The subgroups below will be used for these analyses:

- Age (< 65, or >= 65)
- Baseline NIHSS (< 20, or >= 20)
- Site of occlusion

The subgroup analysis will be conducted using a logistic regression with terms of treatment group and treatment-by-subgroup interaction. 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.

10 Safety Analysis

10.1 Primary Safety Analysis

The primary safety endpoint is the proportion of patients with a device or procedure related SAE within 24 hours post-procedure. The proportion of patients in each group who meet the safety endpoint based on this criterion will be calculated.

The primary null hypothesis for safety in this study is that there is no difference in device-related and procedure-related serious adverse event rates. This null hypothesis will be tested against the alternative hypothesis that there is a difference between the two treatment groups. The null hypothesis will be rejected at the two-sided significance level of alpha = 0.05. The primary analysis is an analysis of all patients according to treatment received. The Fisher's Exact test will be employed to assess the primary safety endpoint.

10.2 Secondary Safety Analysis

The secondary safety variables:

- All cause mortality.
- Rate of symptomatic ICH

The proportion of patients in each group who experience a safety event based on these criteria will be calculated. The CEC/DSMB data supersede the investigator-reported data in all analyses Adverse Events. Rates of symptomatic and asymptomatic hemorrhage will additionally be provided. Frequency counts and percentage of patients within each category will be provided for categorical data. Estimates of the treatment differences and their 95% confidence intervals will be calculated. Subject rates will be compared between treatment groups with Fisher's exact test.

10.3 Analysis of Adverse Events

Tabulations of adverse events will be presented with descriptive statistics at baseline hospitalization and follow-up visits.

Adverse events will be categorized. Adverse event incidence rates will be summarized by category and severity of the adverse event. Each subject will be counted only once within a category by using the adverse event with the highest severity within each category.

All information pertaining to adverse events noted during the study will be listed by subject, detailing verbatim given by the investigator, category, date of onset, date of resolution, causality and severity. The onset of adverse events will also be shown relative (in number of days) to the day of procedure.

A tabulation of Serious Adverse Events (SAEs) will be provided by subject.

The specific categories analyzed will be those that are reported by at least three (3) percent of the patients.

The CEC/DSMB adjudicated data supersedes the investigator reported data in all analyses of adverse events.

10.4 Handling of Multiplicity

All safety variable comparisons between the Penumbra System versus the Penumbra System with Separator 3D will be considered secondary analyses and will be adjusted using the Bonferroni correction.

10.5 Analysis of Deaths

The Kaplan-Meier product-limit method will be the primary method utilized to assess the mortality rate. With the date of procedure set at Day 0, any death occurring on or before calendar day 90 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 90 will be censored at day 90. 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 90 Day binary deaths. The number of deaths will be presented for each group.

11 Pooling Across Centers

The clinical study will be conducted under a common protocol for each investigational center with the intention of pooling the data for analysis. Every effort will be made to promote

consistency in study execution at each investigational center. Analyses will be presented using data pooled across centers. Appropriate stratification and multivariate techniques, including contingency tables and logistic regression for binary outcomes, analysis of variance for continuous measures, and proportional hazards regression for time-to-event outcomes, will be used to assess differences between study centers to justify pooling data across centers. Centers with fewer than 10 patients will be combined and treated as one site in the pooling analysis.

12 Lost to Follow-Up and Missing Data

For sensitivity purposes, the following additional analyses will be conducted:

- Analyze only patients with complete primary endpoint data.
- For those patients without any observations after randomization, their final observations will be imputed by baseline scores and included in an analysis.
- A tipping point analysis will be conducted in which the missing data will be replaced with values such that the p-value is equal to the significance level.

13 Committees

13.1 Clinical Events Committee/Data Safety Monitoring Board (CEC/DSMB)

A CEC/DSMB will adjudicate serious adverse events for causality and attribution.

13.2 Core Lab

An independent Core Lab will review and score all imaging scans for TIMI/TICI scores and intracranial hemorrhage (ICH). The Core Laboratory will be blinded to treatment allocation.

14 Changed to Planned Analyses

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

15 References

Agresti A (2002). Categorical Data Analysis (Second Edition). Wiley.

Chow S, Shao J, Wang H. (2003). Sample Size Calculations in Clinical Research, CRC Press.

Kalbfleisch JD, Prentice RL. (2002). The Statistical Analysis of Failure Time Data (Second Edition). Wiley.

Laster LL, Johson MF, Kotler ML. Non-inferiority trials: the "at least as good as" criterion with dichotomous data. *Statistics in Medicine* (2006) 25:1115-1130.

(2009). SAS/STAT(R) 9.2 User's Guide, Second Edition. Cary, SAS Institute.