Protocol Title

Impact of Fever Prevention in Brain Injured Patients (INTREPID)

Sponsor

Bard Medical Division
C. R. Bard, Inc.
8195 Industrial Blvd.
Covington, GA 30014

SAP Version/Date

Version 2.0 / March 25, 2021

Clinical Trial Registration

NCT02996266

STATISTICAL ANALYSIS PLAN

Title: Impact of Fever Prevention in Brain Injured Patients (INTREPID)

Protocol No.: BMD-1111

Study Device: Arctic Sun® 5000 Temperature Management System

Sponsor:
Bard Medical Division
C. R. Bard, Inc.
8195 Industrial Blvd.
Covington, GA 30014

Prepared By:
Shu Zhang
Becton Dickinson and Company
1 Becton Drive
Franklin Lakes, NJ 07417

Revised By:

Version & Date: Version 2.0 – March 25, 2021

Approval

Individual or Function	Name	Signature/Date
Statistician (Author)	Shu Zhang	This document is signed electronically in the eTMF system
Project Manager	Peter Fayek	This document is signed electronically in the eTMF system
Business Representative	Stuart Hart	This document is signed electronically in the eTMF system
Statistics Functional Manager	Ying Wan	This document is signed electronically in the eTMF system
Clinical Scientist	Jaime Ritter	This document is signed electronically in the eTMF system

Table of Contents

1 Introduction 6
1.1 Administration of the Interim Analysis 6
2 Study Objective and Endpoints 7
2.1 Study Objective 7
2.2 Study Endpoints 7
2.2.1 Primary Endpoint 7
2.2.2 Key Secondary Efficacy Endpoint 7
2.2.3 Secondary Efficacy Endpoints 7
2.2.4 Safety Endpoints 8
3 Study Design 8
3.1 Overview 8
3.2 Study Population 8
3.3 Randomization 8
3.4 Sample Size Consideration 9
3.5 Interim Analysis 10
3.6 Study Procedure 12
4 Analysis Set 16
5 Study-wise Type I Error control 16
6 Interim Analysis 17
6.1 Interim Analysis for DMC Decision Making 17
6.1.1 Primary Endpoint 17
6.1.2 Key Secondary Endpoint 17
7 Statistical Analysis of the Primary Endpoint – Fever Burden 19
7.1 Primary Endpoint Definition 19
7.2 Hypothesis Testing and Primary Analysis 20
7.3 Supportive Analysis 21
7.3.1 Binary Analysis 21
7.3.2 Total Fever Burden 21
7.3.3 Total Fever Duration 21
7.4 Assessment of Poolability of Sites 21
7.5 Handling of Missing Data and Sensitivity Analysis 22
7.5.1 Multiple Imputation Method 22
7.5.2 Tipping Point Analysis 22

7.6 Subgroup Analysis 22
8 Statistical Analysis of the Key Secondary Endpoint – 3-month mRS 23
8.1 Key Secondary Endpoint Definition 23
8.2 Hypothesis Testing and Primary Analysis: van Elteren test 23
8.3 Supportive Analysis 23
8.3.1 Proportional Odds Model 23
8.3.2 Utility-weighted 24
8.3.3 Binary Outcome 24
8.4 Assessment of Poolability of Sites 24
8.5 Handling of Missing Data and Sensitivity Analysis 24
8.5.1 Multiple Imputation Method 25
8.5.2 Tipping Point Analysis 25
8.5.3 Worst Case Analysis 25
8.6 Subgroup Analysis 25
8.7 Exploratory Analysis 26
8.7.1 Analysis of fever/fever potential subjects 26
9 Statistical Analysis of the Secondary Efficacy Endpoint 26
9.1 Modified Rankin Scale measured at 6- and 12-months 26
9.2 National Institutes of Health Stroke Scale measured at 3- and 6-months 27
9.3 Barthel Index measured at 3- and 6-months 27
9.4 Glasgow Outcome Scale Extended measured at 3- and 6-months 27
9.5 Montreal Cognitive Assessment measured at 3- and 6-months 27
9.6 Intensive Care Unit Length of stay 27
9.7 ICU Delirium 28
9.8 Use of Mechanical Ventilation 28
9.9 Hospital Length of Stay 28
9.10 Mortality 28
10 Statistical Analyses of Safety Endpoints 29
10.1 AE 29
10.2 MAE 29
10.3 Incidence of Infection during Acute Phase 30
10.4 Incidence of Shivering 30
11 Statistical Analyses of Other Endpoints 31
11.1 Subjects Dispositions 31

11.2 Protocol Deviations 31
11.3 Demographics and Background Disease Characteristics 31
11.4 Relevant medication and therapy 32
11.5 Follow-up Period 32
11.6 Acute Phase Treatment Characteristics 32
11.7 Device failure, malfunctions and defects 32
12 Reference 32
13 Appendix 33
Appendix 1.1 Tables/Listing/Figures Shell for DMC close session output 33
Appendix 1.2 Tables/Listing/Figures Shell for 3 Month Full Report 33
Appendix 1.3 Tables/Listing/Figures Shell for 12 Month Report 33
Appendix 2 Derived data specification 33
Appendix 3 Protocol Deviation Major/Minor classification 33

1 Introduction

This document provides details of the statistical analysis plan (SAP) for the C. R. Bard, Inc. protocol BMD-1111. There is a statistical interim analysis planned when the first 588 subjects (half of the planned sample size) have completed the 3-month visit (including subjects who were discontinued early from the study), at which time the primary analysis will be performed to evaluate the treatment effect in fever burden. If the result of the primary endpoint at interim is a success, the key secondary endpoint of modified Rankin Scale (mRS) will be evaluated for futility and sample size re-estimation. If the study is not futile at the interim, then the study will continue and be reported when all randomized subjects finish the 3-month visit. The 6- and 12-month follow up data will be reported after all subjects have finished the study and a supplementary report will be provided.

The tables to be presented in the interim and the final analyses will be listed separately in the appendix. The statistical methods described here are based on the analyses proposed in the Protocol issued on March 23, 2021, Version 7.06 DRAFT.

1.1 Administration of the Interim Analysis

An independent data monitoring committee (DMC) will be assembled which will be composed of medical experts with experience in treating subjects with stroke and a statistician, all of whom are not otherwise involved in any other aspects of the conduct of the study. The DMC will review safety data on an ongoing basis including adverse events. During the course of the study, the DMC will review the efficacy data in accordance with the guidelines for the pre-planned interim analyses. An independent statistician will provide data analyses support to DMC.

The interim analysis will be performed when half of the planned sample size subjects have completed the 3-month visit. A two-step decision will be made at the interim based on the following decision criteria:

1) futility decision based on the primary endpoint; and,
2) futility or sample size increase decision based on the key secondary endpoint.

The study can only stop because of futility; there will be no stopping of the study due to strong result in effectiveness. The study will continue unless futility is shown for the key secondary endpoint.

All of the interim results will be blinded to any personnel at the investigational sites and any BD/Bard personnel who are involved with the conduct of the study.

All data processing, summarization, and analyses will be performed using Statistical Analysis System (SAS), Version 9.4 software package or later. Simulations may be run in R version 3.4.4 or later.

2 Study Objective and Endpoints

2.1 Study Objective

The objective of this study is to assess fever burden and the impact on outcomes of fever prevention using the Arctic Sun 5000 Temperature Management System as compared to standard fever care in brain-injured patients.

2.2 Study Endpoints

2.2.1 Primary Endpoint

The primary endpoint is the daily average fever burden (°C-hour). Each patient's daily average fever burden will be calculated as the patient's total fever burden divided by the total number of hours the patient was in the acute observation period and then multiplied by 24 hours (Total fever burden/(total hour) ×24 hour). The acute observation period is 336 hours (i.e., 14 days) or total hours from randomization to discharge/deemed medically ready for discharge from the intensive care unit, whichever comes first. Total fever burden is defined as the area under the temperature curve (AUC) that is above 37.9°C during the acute observation period. Although the Arctic Sun System records the patient's temperature every minute, to compare the treatment groups consistently, calculation of fever burden will be based on hourly temperature data recorded by the clinician.

2.2.2 Key Secondary Efficacy Endpoint

A key objective of this study is to determine whether fever prevention with the Arctic Sun 5000 Temperature Management System improves outcomes in moderate/severe brain-injured patients compared with standard care. Improved outcomes will be assessed using the endpoint of modified Rankin Scale (mRS) assessed at month 3 (6-category mRS outcome scores; mRS score 5 and 6 combined together). Note that mRS score 0-5 are directly collected at month 3; mRS will be scored 6 if subject died before the 3-month visit.

2.2.3 Secondary Efficacy Endpoints

Other secondary endpoints will include, but not be limited to:

- Neurologic outcome measures
    - Modified Rankin Scale measured at 6- and 12-months
    - National Institutes of Health Stroke Scale measured at 3- and 6-months
    - Barthel Index measured at 3- and 6-months
    - Glasgow Outcome Scale Extended measured at 3- and 6-months
    - Montreal Cognitive Assessment measured at 3- and 6-months
- Intensive Care Unit (ICU) length of stay
- ICU Delirium
- Use of mechanical ventilation

- Hospital length of stay
- Mortality [7-day (or hospital discharge), 3-, 6-, and 12-month]

2.2.4 Safety Endpoints

The safety endpoints will include major adverse events (MAEs), the overall incidence of adverse events (AEs), the incidence of infection, and the incidence of shivering.

3 Study Design

3.1 Overview

This is a randomized, controlled multicenter clinical investigation designed to assess fever burden and early, short- and long-term clinical outcomes of fever prevention (FP) using the Arctic Sun 5000 Temperature Management System (test device) compared to standard fever care (control device(s)) in the treatment of moderate-to-severe brain injured patients. The target temperature in the FP group is 37.0^0C through 14 days (336 hours) or until discharge from the ICU/deemed medically ready for discharge from the ICU.

The study will follow brain-injured patients throughout their hospital stay and at 3-, 6-, and 12-months post injury for neurologic assessment. After a pre-specified stabilization period, patients meeting inclusion/exclusion criteria will be consented and randomized to either the fever prevention (FP) or the standard care (SC) group. Randomization will be stratified by study site, the type of brain injury (i.e., acute ischemic stroke (IS), intracerebral hemorrhage (ICH), or subarachnoid hemorrhage (SAH)), subject age, and injury-specific severity score.

3.2 Study Population

Adult subjects with a primary stroke diagnosis (i.e., acute ischemic stroke (IS), intracerebral hemorrhage (ICH), or subarachnoid hemorrhage (SAH)) will be screened for potential eligibility against the study protocol inclusion and exclusion criteria. Medical records documenting tests, examinations and evaluations may be used as the source documents for the initial screening criteria. Specific screening tests and/or evaluations, typically performed as standard of care in neurocritical patients, will be recorded and used to verify inclusion/exclusion criteria. Only those patients having gone through the prescribed disease-specific stabilization period will be considered for enrollment.

3.3 Randomization

Randomization will occur after informed consent is obtained, after study personnel ensure the subject meets all general and disease-specific inclusion and none of the exclusion criteria. Using a minimization algorithm, enrolled subjects will be stratified by four factors: neurological diagnosis (IS, ICH or SAH), investigational site, age (<70 years vs. ≥70 years), and baseline severity and randomized into either the fever prevention group or the standard care group at a 1:1 ratio. Baseline severity for IS subjects will be according to the NIHSS score (<17 vs. ≥17, on a scale of 0-42, with higher scores indicating greater severity). For ICH subjects, baseline severity will be based on the ICH score (≤2 vs. ≥3, with higher scores indicating greater severity). For SAH subjects, baseline severity will be according to the World

Federation of Neurological Societies Grading System for Subarachnoid Hemorrhage (≤III vs. ≥IV, on a scale of I-V, with higher scores indicating greater severity).

The randomization process will be handled via use of an interactive web/voice-response system (IRS).

3.4 Sample Size Consideration

The study sample size is primarily driven by the key secondary endpoint, and the study is powered for both the primary endpoint and the key secondary endpoint. The test of the key secondary endpoint is conditional on the success of the test of the primary endpoint.

The key secondary endpoint is neurologic outcomes assessed by the modified Rankin Scale (mRS). In addition to its standard scale, the 6-category mRS outcome score will also be dichotomized as a binary outcome (scores of 0 to 3 for success and scores of 4 to 6 for failure). Sample size is estimated based on the dichotomized binary outcome (as a more conservative approach) with the assumed FP success rate of 50% and the SC success rate of 40%. A sample size of 1000 subjects (500 in each group) will provide 88% power based on a two-group chi-square test with two-sided α=0.05 (nQuery 7.0). Given such a sample size, the analysis based on the standard scale may be slightly more powerful. Both analyses, either using the original standard scale or dichotomized scale, will be performed for mRS in this study.

Assuming a 15% attrition rate, the total sample size of randomized subjects is 1,176, with a 1:1 ratio to the FP and SC group (588 subjects in each of the group).

The primary endpoint of fever burden will be analyzed at the interim when 588 subjects (half of the planned sample size of 1,176 subjects) have completed the 3-month visit. Since fever burden does not follow a normal distribution, Wilcoxon rank-sum tests will be used to analyze the data and the sample size estimation is based on such test as well. In the SC group, the proportion of patients who develop fever and therefore show positive fever burden is estimated at 40%. In the FP group, it is expected that few will have positive fever burden; using a conservative assumption, we estimate 10% of subjects will show a positive fever burden. Under such assumptions, the probability that fever burden in the SC group is higher than in the FP group is 65% (nQuery 7.0 Assistant). Consequently, the test of the primary endpoint will have more than 99% power to show a statistically significant difference between the two groups with 500 subjects (nQuery 7.0).

Sample size re-estimation:

An interim analysis will be conducted when the first 588 subjects (half of the planned sample size) have completed the 3-month visit (including subjects who were discontinued early from the study), at which time the primary analysis will be performed to evaluate the treatment effect in fever burden. The key secondary endpoint of modified Rankin Scale will be evaluated as well for futility and sample size re-estimation. If recommended by the Data Monitoring Committee (DMC) and deemed appropriate, the sample size may increase up to double the planned sample size (up to 2,000 subjects). The overall study characteristics will be displayed by simulation as further described below.

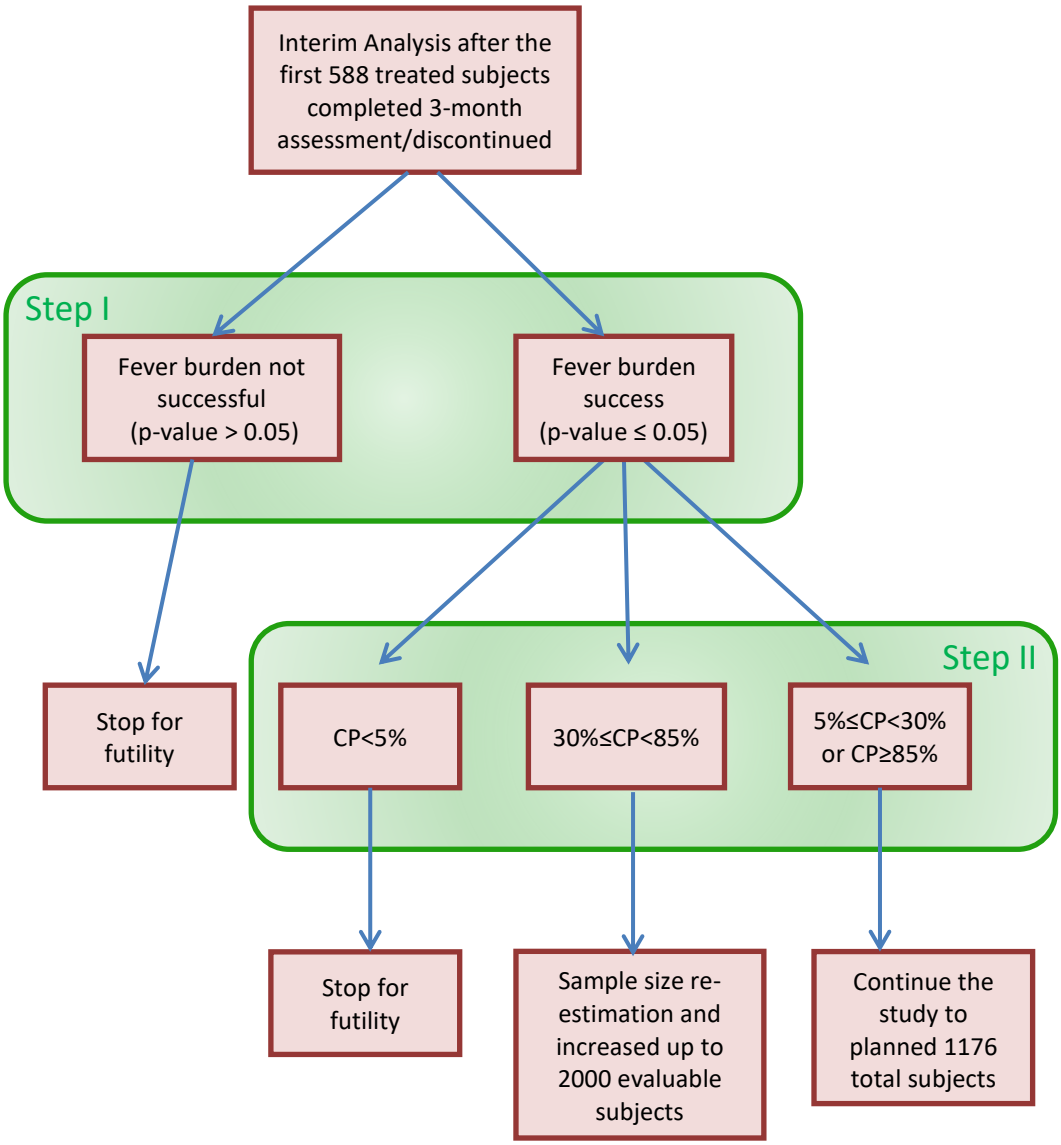
3.5 Interim Analysis

It is well known that adaptive design with sample size re-estimation can increase the power of the study but may cause inflation of type 1 error rate. On the other hand, futility analysis may decrease the power of the study and reduce the type 1 error rate. The proposed study design includes both features in the interim analysis for the key secondary endpoint of mRS. It is important to evaluate the impact on study power and type 1 error rate.

The rules of the two-step decision to be made at the interim are described below (See Figure 1 for illustration):

1. If the primary analysis of fever burden is not successful (p-value > 0.05), then the study will stop for futility; otherwise, the study will continue or stop based on the step 2 decision.
2. If the primary analysis of fever burden is successful, the futility and sample size re-estimation decision will be evaluated based on the dichotomized binary mRS outcome (scores of 0 to 3 for success and scores of 4 to 6 for failure) using the conditional power (CP) calculated under the current trend: If CP is less than 5%, the study will stop for futility; if the CP is between 5% and 30% or greater than or equal to 85%, the study will be finished at planned sample size (1000 evaluable subjects); if the CP is between 30% and 85%, the sample size will be re-estimated based on the observed effect at IA (current trend). If the re-estimated sample size is greater than 2000, the final sample size will be capped at 2000 evaluable subjects.

Figure 1. Proposed Interim Analysis Procedure



Note: CP is the conditional power under current trend on the key secondary binary endpoint (mRS scores of 0 to 3 for success and scores of 3 to 6 for failure).

$$CP_{\hat{\theta}} = 1 - \Phi\left(\frac{z_{\alpha/2} - E_{\hat{\theta}}\{B(1)|B(t)=b\}}{\sqrt{1-t}}\right) = 1 - \Phi\left(\frac{z_{\alpha/2} - \hat{\theta}}{\sqrt{1-t}}\right)$$

Where the B-value and $\hat{\theta}$ are following Section 2 in Statistical monitoring of clinical trials 2006: assume the success rate for the FP is p_T, and for the SC is p_C. At the interim, treatment difference $\hat{\delta}_1 = \hat{p}_T - \hat{p}_C$, with the standard error $SE_{\hat{p}_T - \hat{p}_C} = \{\hat{p}_T(1-\hat{p}_T)/N_T + \hat{p}_C(1-\hat{p}_C)/N_C\}^{1/2}$. Then test statistics $z(t) = \frac{\hat{p}_T - \hat{p}_C}{SE_{\hat{p}_T - \hat{p}_C}}$, and the empirical estimate $\hat{\theta} = \frac{B(t)}{t} = \frac{Z(t)}{\sqrt{t}}$ (t=1/2).

A simulation study was performed to evaluate the operating characteristics. The study was simulated under null hypothesis ($p_C = 40\%$, and $p_T = 40\%$) and under alternative hypothesis ($p_C = 40\%$, and $p_T = 50\%$), respectively, with 1000 subjects (500 in treatment arm and 500 in control arm). At the interim, the CP was calculated and the decision was made based on Figure 1. For each simulation, at the end of the study, the two treatment arms were compared by a chi-square test at two-sided α=0.05. Based on 10000 simulations, the overall type 1 error for this study is 2.1%, which is well under control (<5%), and the overall study power is 90.8%.

Table 1: Trial operating characteristics

	Futility at IA	Probability of Sample Size increased	Reject Null
Under Null Hypothesis ($p_C = 40\%$, and $p_T = 40\%$)	73.3%	9.6%	2.1%
Under Alternative Hypothesis ($p_C = 40\%$, and $p_T = 50\%$)	4.8%	24.4%	90.8%

3.6 Study Procedure

All subjects who sign the informed consent and meet eligibility criteria will be randomized to either the FP or SC group and will be followed as per the protocol defined study procedures and follow-up visit schedule. A confirmatory temperature check must be performed no more than one hour prior to randomization to ensure that the subject does not become febrile. If the subject becomes febrile between enrollment and randomization, they will be considered a screen failure. Randomization and initiation of study procedures (e.g., index procedure, hourly temperature measurements) should take place as close to enrollment as possible.

The study procedures and visit schedules for enrolled subjects is shown in Table 2. With the exception of the post-acute phase visit, all follow-up visits must be scheduled based on the date of symptom onset. If a subject is unable to attend a scheduled study visit, this must be documented in the subject's file and the site should request that the subject return as close to the scheduled visit date as possible.

Note: For the purposes of this study, the definition of "deemed medically ready for discharge" is defined as follows: patient is medically stable and no longer requires critical/intensive level care or observation. For example, a patient may be medically stable and no longer requires critical level care but cannot be discharged from the ICU due to a lack of beds in a standard care unit. In this case, the acute phase of the study would be discontinued and the date the patient was deemed ready for discharge would be recorded in the appropriate case report form.

Should a subject in the Fever Prevention arm prematurely discontinue the Arctic Sun 5000 System OR a subject in the Standard Care arm have a targeted temperature management system

initiated for fever control, the subject does not switch arms of the study. All study procedures/assessments should continue as scheduled for their assigned arm.

Should a subject's medical care be changed to comfort-measures only, the acute phase of the study will be ended.

An overview of the study visit schedules is shown in Table 2, which summarizes the study procedures to be performed at each study visit.

Table 2: Time and Events Schedule

	Screening / Enrollment	Acute Phase	Post-Acute Phase Visit	3 Month F/U Visit	6 Month F/U Visit	12 Month F/U Visit
Procedure/Test/Follow-Up Windows	Hospital / ICU Admission	336 Hours (14 Days) [i]	5 Days Post-Acute Phase OR Hospital Discharge[ii] ±1 day	3 mo ±30 days from LKN/Ictus	6 mo±30 days From LKN/Ictus	12 mo ±30 days from LKN/Ictus
Inclusion/Exclusion Criteria	X					
Describe study to potential subject	X					
Informed Consent/Enrollment	X					
Randomization*	X					
Demographics and medical history	X					
Modified Rankin Scale	X (Pre-Morbid)		X	X	X	X[vi]
Glasgow Coma Scale	X					
Intracerebral Hemorrhage Score (ICH Score)	X[iv]					
Barthel Index			X	X	X	
Glasgow Outcome Scale – Extended			X	X	X	
Montreal Cognitive Assessment			X	X	X	
Diagnostic Imaging (CT Scan or MRI)	X[iii]					
NIH Stroke Scale	X[v]	X[vii]	X	X	X	
Daily Intensive Care Delirium Screening Checklist		X				
World Federation of Neurological Surgeons Grading System (SAH subjects only)	X[iii]					
Fever Prevention Index Procedure		X				
Hourly Body Temperature*		X				
Urinalysis (Subjects with Foley catheter)	X	X[viii]				
Daily White Blood Cell Count		X				
Shiver monitoring and control protocol		X				
Antipyretic use / supplemental shiver control		X				
End of full case data download from Arctic Sun		X				

Adverse Events		X	X[ix]	X[ix]	X[ix]	X[ix]

* A confirmatory temperature check must be performed no more than one hour prior to randomization; should they be found to be febrile they will be considered a screen failure

i – Acute Phase will continue through 336 Hours (14 days) for ALL subjects OR until the patient is discharged/deemed medically ready for discharge from the ICU OR the subject goes on comfort-measures only, whichever occurs first; neurologic assessments performed during this phase are not blinded

ii - Window will be 5 days after completion of Acute Phase OR at hospital discharge, whichever comes first

iii - Testing performed at admission may be used for screening and as study baseline measures; if not previously performed, tests should be completed for eligibility screen and baseline

iv – For ICH subjects only at screening

v - Only for IS and ICH subjects at screening and enrollment

vi – May be assessed by phone unless an actual visit is otherwise performed within the follow-up window

vii – NIH Stoke Scale will be performed at day 2 and day 5 of the study for IS and ICH subjects when not fully sedated; not blinded assessments

viii – Urinalysis performed at study enrollment, Day 7 (if still in the acute phase) and at the end of the acute phase/time of catheter removal for all catheterized subjects

ix – Limited to MAEs only

4 Analysis Set

Enrolled: The enrolled subjects consist of all subjects who sign the informed consent.

Intent-to-Treat (ITT): This is the population of all subjects who have been consented and randomized.

Interim Intent-to-Treat (ITT): This is the first 588 randomized ITT subjects.

As Treated (AT): This is the same as the ITT population but based on the actual treatment received, not based on randomization, if there are patients who received the treatment other than that which they were randomized.

Per Protocol (PP): This population is defined as all subjects in the ITT who do not have any major protocol deviation or the following conditions. Major protocol deviations include major inclusion/exclusion deviations or other major protocol deviations, as will be defined in the statistical analysis plan Appendix 3. For subjects in the FP arm, if the Artic Sun was suspended for a long period (greater than 12 hours in one incidence) or was taken off early during the acute phase (greater than 24 hour before the end of the acute phase), the subject will be excluded from the PP. For subjects in the SC arm, if the subject used continuous temperature management (e.g., tier 4 fever treatment) within 24 hours of randomization and remained on continuous temperature management for more than half of the acute phase (cumulative), then the subject will be excluded from the PP. Randomized but not treated subjects will all be excluded from PP.

All efficacy analyses including the primary and key secondary endpoints will be primarily based on the ITT population. Analysis of the AT and PP populations will also be performed on the primary and key secondary endpoints as sensitivity analyses. The safety analyses will be performed on the AT population using the as-treated principle, i.e., subjects will be analyzed based on the actual treatment they receive instead of the randomized treatment.

5 Study-wise Type I Error control

A fixed sequence of testing procedures will be applied to the testing of the primary and key secondary endpoint hypotheses.

The primary endpoint (the daily average fever burden) will be tested at the interim analysis time on the first 588 subjects at type I error level of 0.05.

If the hypothesis on the primary endpoint is statistically significant, the hypothesis on the key secondary endpoint will be tested when all subjects (sample size will be determined after interim analysis) have all complete the 3-month visit (including subjects who were discontinued early from the study) at type I error level of 0.05. Otherwise, the study will be stopped at the interim and no further testing will be provided.

For all other secondary endpoints, no hypotheses will be tested, and descriptive statistics will be provided.

6 Interim Analysis

Interim analysis will be performed on ITT subjects when the first 588 subjects (half of the planned sample size) have completed the 3-month visit (including subjects who were discontinued early from the study). The Interim ITT only include the first 588 subjects.

6.1 Interim Analysis for DMC Decision Making

6.1.1 Primary Endpoint

The primary endpoint (the daily average fever burden) is defined in Section 7.1. If the primary endpoint based on Interim ITT demonstrated success (p value $\leq$0.05 see Section 7.2 for details), then the interim decision can move to next step for decision based on key secondary endpoint.

6.1.2 Key Secondary Endpoint

If the primary analysis of fever burden is successful, the futility and sample size re-estimation decision will be evaluated based on the dichotomized binary mRS outcome (scores of 0 to 3 for success and scores of 4 to 6 for failure) using the conditional power (CP) calculated under the current trend. The following R code will be used to calculate the CP:

```
CPcurrent.binary<-function(n, nx, ny, x1, y1,alpha=0.05) {
 # n: original planned sample size per group
 # nx: actual sample size in treatment arm
 # ny: actual sample size in control arm
 # x1: treatment arm data or number of event
 # y1: control arm data or number of event

 t1<-(nx+ny)/2/n

 pt1<-sum(x1)/nx
 pc1<-sum(y1)/ny
 delta.hat1<- (pt1-pc1)
 I.n1 <- 1/(pt1*(1-pt1)/nx+pc1*(1-pc1)/ny)

 z.n1=sqrt(I.n1)*delta.hat1   #Z(t)
 b.n1<-sqrt(t1)*z.n1       #b(t)
 theta.hat1<-b.n1/t1     #theta(t)

 CP.n1.theta.hat1<-1-pnorm((abs(qnorm(alpha/2))-theta.hat1)/sqrt(1-t1))

 return(c(CP.n1.theta.hat1,z.n1))
 # CP.n1.theta.hat1: Conditional power under current trend
 # z.n1: Z(t) at n1
}
```

```
CPower<-CPcurrent.binary(n=500, nx=nx, ny=ny, x1=x1, y1=y1,alpha=0.05)
CPower
```

As illustrated in Section 3.5, if CP is less than 5%, the study will stop for futility; if the CP is between 5% and 30% or greater than or equal to 85%, the study will be finished at planned sample size (1000 evaluable subjects); if the CP is between 30% and 85%, the sample size will be re-estimated based on the observed effect at IA (current trend). If the re-estimated sample size is greater than 2000, the final sample size will be capped at 2000 evaluable subjects.

Now assume that the final sample size per arm n* is not known, the sample size at the interim is n_1 per arm. To achieve a pre-set conditional power ($=1-\beta$) under current trend, we have

$$1-\beta = CP_{\hat{\theta}} = 1-\Phi\left(\frac{z_{\alpha/2}-\hat{\theta}}{\sqrt{1-t}}\right) = 1-\Phi\left(\frac{z_{\alpha/2}-z_1/\sqrt{t}}{\sqrt{1-t}}\right)$$

where $\hat{\theta} = \frac{B(t)}{t} = \frac{b}{t} = \frac{\sqrt{t}z_1}{t} = \frac{z_1}{\sqrt{t}}$, and $z_1 = \frac{\hat{p}_T-\hat{p}_C}{SE_{\hat{p}_T-\hat{p}_C}}$ and the unpooled standard error

$$SE_{\hat{p}_T-\hat{p}_C} = \{\hat{p}_T(1-\hat{p}_T)/N_T + \hat{p}_C(1-\hat{p}_C)/N_C\}^{1/2}$$

Therefore

$$z_\beta\sqrt{1-t}\sqrt{t} + z_{\alpha/2}\sqrt{t} - z_1 = 0$$

Solve the *t* based on the above function (using uniroot in R), and as $t=n_1/N_{final}$, then $N_{final}=n_1/t$ to get the re-estimated sample size.

The following R code will be used to do sample size re-calculation:

```
### function to determine the needed info fraction t in order to re-estimate sample size
### z1 from CP calculation z.n1
library(rootSolve)
samplesz<-function(t,beta=0.1,alpha=0.05,z1){
  abs(qnorm(beta))*sqrt((1-t)*t)+abs(qnorm(alpha/2))*sqrt(t)-z1
}

#if CPower>0.3 and <0.85 then
tpower=0.15
tt<-uniroot(samplesz, c(0,0.95), beta=tpower, alpha=0.05, z1=CPower[2])
nn<-ceiling( (nx+ny)/2/tt$root )
nn
# nn is the reestimated sample size per arm
# if nn >1000, then set nn to 1000
```

The interim output shell for the DMC can be found in Appendix 1.1.

7 Statistical Analysis of the Primary Endpoint – Fever Burden

7.1 Primary Endpoint Definition

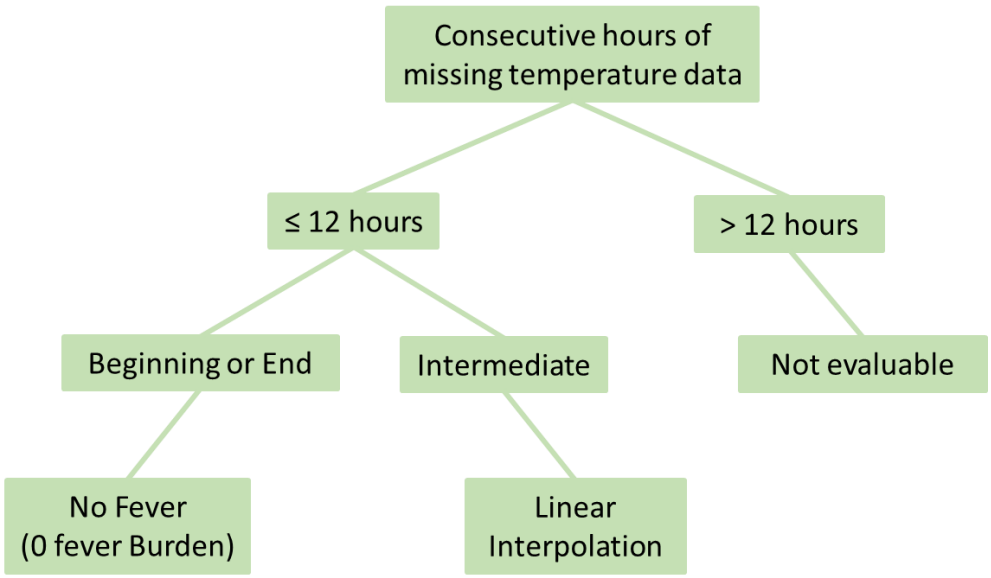
The primary endpoint is the daily average fever burden (°C*hour). Each patient's daily average fever burden will be calculated as the patient's total fever burden divided by the total number of hours the patient was in the acute phase and then multiplied by 24 hours (Total fever burden/(total hours) ×24 hour).

Acute Phase is defined as from randomization through 14 days (336 hours) for all subjects OR until the patient is discharged/deemed medically ready for discharge from the ICU OR the subject goes on comfort-measures only, whichever occurs first. The acute phase duration is 336 hours or calculated as (Date/time of the ICU discharge/ready to discharge/comfort-measure – date/time of the randomization), whichever is less.

If the subject is discontinued before ICU discharge/deemed medically ready for discharge and the period is less than 336 hours, the acute phase should end at the discontinuation day, and the discontinuation event will be included in the acute phase. For the acute phase duration for fever burden calculation, it will be ended at the last temperature collection time.

Although the Arctic Sun System records the patient's temperature every minute, to compare the treatment groups consistently, calculation of fever burden will be based on hourly temperature data collected on the CRF by the clinician.

Total fever burden is defined as the area under the temperature curve (AUC) during the acute phase that is above 37.9°C. Here the AUC will be the sum of the hourly bars that are above 37.9°C. If more than 12 consecutive hours of data during the acute phase are missing, the subject will be considered as not evaluable. If there are missing temperature collections at the beginning or at the end of the acute phase, normal temperature (0 fever burden) will be assumed as the subject should be at normal temperature at the time of randomization and when they completed the acute phase. If there are missing intermediate temperature(s) for a subject, a linear interpolation of the available adjunctive two temperatures will be used to impute the missing intermediate temperature(s). Below figure illustrates the missing temperature handling.



7.2 Hypothesis Testing and Primary Analysis

The primary endpoint is daily average fever burden and will be evaluated by the following hypothesis:

H_0: The distribution of the daily average fever burden of FP treatment group is the same as that of SC treatment group.

H_1: The distribution of the daily average fever burden of FP treatment group is different from that of SC treatment group.

The primary analysis will be based on the first 588 randomized subjects (the interim ITT analysis population). Only when statistical significance is achieved at 0.05 level (i.e. two-sided p-value $\leq$ 0.05) and the mean value of the FP group is less than the SC group will the study be considered successful on the primary endpoint. At that time the study may continue to its full sample size to evaluate other study endpoints, depending on the interim analysis decision of futility and sample size re-estimation based on the key secondary endpoint. Otherwise, the study will stop for futility.

Since fever burden is not normally distributed, the daily average fever burden in the two treatment groups will be compared using a van Elteren test, which is a direct extension of the Wilcoxon rank-sum tests, with stratification on patient diagnosis (ICH, AIS or SAH). Superiority of FP will be demonstrated if the p-value from the two-sided test is less than or equal to 0.05, and the mean value of the FP group is less than SC group.

The same analysis will be performed for the AT and PP interim population as sensitivity analysis.

When all subjects (sample size will be determined after interim analysis) have complete the study, this analysis will be performed for the full ITT, AT and PP population.

7.3 Supportive Analysis

Temperature data will be analyzed in other ways than daily average fever burden, and below are the alternative definitions. These endpoints will be analyzed for interim ITT population.

7.3.1 Binary Analysis

A binary endpoint will be derived including all subjects with an evaluable primary endpoint. If the daily average fever burden is greater than 0, then the subject will be considered as having positive fever burden; otherwise the subject will be considered as no fever burden. A binary variable (positive fever burden versus no fever burden) will be analyzed by Cochran-Mental-Haenszel (CMH) with patient diagnosis (ICH, AIS or SAH) as stratification factor.

7.3.2 Total Fever Burden

The total fever burden in the two treatment groups will be compared using a van Elteren test, which is a direct extension of the Wilcoxon rank-sum tests, with stratification on patient diagnosis (ICH, AIS or SAH).

7.3.3 Total Fever Duration

The total hours with fever and the percent of acute phase with fever may be summarized by treatment group.

7.4 Assessment of Poolability of Sites

The poolability of site analysis for the primary endpoint is for interim ITT. The sites with less than 10 randomized subjects will be sorted by site number and pooled by order within geographic regions to form one or more combined site(s) with at least 10 randomized subjects.

The poolability of the investigational sites on the primary endpoint will be tested using a ranked ANOVA model. The daily average fever burden will be ranked first and the ranks will be fitted with an ANOVA model with treatment, sites, and treatment and site interaction as factors. If the p-value for the interaction test is <0.15, it will be considered evidence of a statistically significant interaction effect, and additional analyses will be performed to explore the differences among the investigational sites to assess potential causes and whether or not they may be clinically meaningful.

The daily average fever burden will be summarized by site and the p-value of the site effect will be calculated using a van Elteren test with stratification on patient diagnosis (ICH, AIS or SAH).

7.5 Handling of Missing Data and Sensitivity Analysis

7.5.1 Multiple Imputation Method

If more than 12 consecutive hours of hourly temperature data is missing during the acute phase, the subject will be considered as not evaluable and become missing data for the primary endpoint. In order to assess the robustness if the primary efficacy endpoint, multiple imputation (MI) for the interim ITT population will be performed, and tipping point analysis will be produced to evaluate the missing data impact.

SAS PROC MI will be used to generate 100 imputations sets of the study data using the following variables to model the primary endpoint (the daily average fever burden): gender, age, patient diagnosis, injury severity, study site and method of thermometry for the MI analysis. If a negative value was generated, then a value 0 will be assigned. For each imputed data set, the primary analysis will be analyzed by the same van Elteren test with stratification on patient diagnosis (ICH, AIS or SAH) as the primary analysis. These will be provided to SAS PROC MIANALZYE in order to complete the final analysis. The amount of the variability associated with imputations and study analysis will be included in the summary.

7.5.2 Tipping Point Analysis

Tipping point approach is performing the sensitivity analysis under the missing at not random (MNAR) assumption. In other words, the tipping point approach is like a progressive stress-testing to assess how severe departures from missing at random (MAR) must be in order to overturn conclusions from the primary analysis.

Tipping point approach can be seen as a special application of the multiple imputation. It can also be considered as a special case of controlled imputation method (i.e., applying the shift parameter only to the active treatment group, not to the placebo group), that is to generate multiple imputed data sets, with a set of specified shift parameters that adjust the imputed values for observations in the treatment group, not the placebo group.

7.6 Subgroup Analysis

The primary endpoint will be explored in the following subgroups:

- Gender
- Age (<70, ≥70)
- Patient primary diagnosis (ICH, AIS or SAH)
- Injury severity (high or low, where high score includes NIHSS ≥ 17 for IS subjects, or baseline ICH score ≥ 3 for ICH subjects, or WFNS ≥ IV for SAH subjects, and low score includes NIHSS < 17 for IS subjects, or baseline ICH score ≤ 2 for ICH subjects, or WFNS ≤ III for SAH subjects).
- Acetaminophen/Paracetamol use: yes/no is for indication of any use of a drug from the identified list of antipyretic drugs (including Acetaminophen) during the acute phase.

- Sedation drug use: yes/no is for indication of any use of sedation drugs during the acute phase.
- Method of Thermometry (Core body temperature (Bladder, Rectal) versus others) to be used for hourly temperature collection
- Geographic Region (US, EU, Asia, other (Australia, etc))

8 Statistical Analysis of the Key Secondary Endpoint – 3-month mRS

8.1 Key Secondary Endpoint Definition

The key secondary endpoint is the mRS outcome score at 3-months. mRS score 0-5 are directly collected at 3-month visit; mRS will be scored 6 if the subject missed the 3-month mRS evaluation due to death or discontinuation due to comfort measures before the end of the 3-month visit window (Day 120). mRS score 5 and 6 will be combined together as score 5 to form a 6-category mRS outcome score. Otherwise, if the subject missed the 3-month visit assessment or discontinued before having 3-month visit, the subject will be considered not evaluable.

8.2 Hypothesis Testing and Primary Analysis: van Elteren test

The key secondary endpoint is the 6-category mRS outcome score and will be evaluated by the following hypothesis:

H0: The distribution of mRS at month 3 in the FP treatment group is the same as that in the SC treatment group.

H1: The distribution of mRS at month 3 in the FP treatment group is different from that in the SC treatment group.

A “shift analysis” will be employed to analyze the 6-category mRS outcome scales for the ITT population. This is essentially a van Elteren test (a particular form of Cochran-Mantel-Haenszel (CMH) test), with stratification on patient diagnosis (ICH, IS or SAH). When statistical significance is achieved at 0.05 level (i.e. p-value $\leq$ 0.05) and the median value of the FP group is less than SC group, the study will be considered successful on the key secondary endpoint.

The same analysis will be performed for the AT and PP population as sensitivity analysis.

8.3 Supportive Analysis

8.3.1 Proportional Odds Model

Additionally, the 6-category mRS outcome scores will be analyzed using a proportional odds model that includes treatment and patient diagnosis (ICH, AIS or SAH) as factors in the model. This is a straightforward generalization of logistic regression where the odds ratio is calculated for each cut-point across the mRS (for example 0 versus 1 to 5, then 0 to 1 versus 2 to 5, and

so on), and then a summary odds ratio is calculated from the individual odds ratios under the assumption that the individual odds ratios are the same. This approach has the advantage of providing a clinically interpretable parameter (the estimated summary odds ratio), but the disadvantage of requiring the assumption that the individual odds ratios are the same (the "proportional odds assumption"). The overall odds ratio and its 95% confidence interval will be estimated and reported.

8.3.2 Utility-weighted

A utility weighted mRS (UW-mRS) [Chaisinanunkul 2015] will be derived with utility values as shown in the following table:

mRS	0	1	2	3	4	5/6
Weight	1.0	0.91	0.76	0.65	0.33	0

The two treatments will be compared by an ANOVA model with treatment and diagnosis categories as factors.

8.3.3 Binary Outcome

The mRS will also be dichotomized as binary outcomes, one with scores of 0 to 2 for success and scores of 3 to 6 for failure; the other one with scores of 0 to 3 for success and scores of 4 to 6 for failure. The mRS success rate at month 3 will be analyzed by the Cochran-Mantel-Haenszel (CMH) test for the two treatment groups with stratification on patient diagnosis (ICH, IS or SAH).

8.4 Assessment of Poolability of Sites

The sites with less than 10 randomized subjects will be sorted by site number and pooled by order within geographic regions to form one or more combined site(s) with at least 10 randomized subjects.

The poolability of the investigational sites on the key secondary endpoints will be tested using a logistic regression model using the full 6-category range of the mRS (the proportional odds model) as response with treatment, sites, and treatment and site interaction as factors. If the p-value for the interaction test is <0.15, it will be considered evidence of a statistically significant interaction effect, and additional analyses will be performed to explore the differences among the investigation sites to assess potential causes and whether or not they are clinically meaningful.

8.5 Handling of Missing Data and Sensitivity Analysis

Study endpoints may be missing due to withdrawal of consent, investigator's decision, or lost to follow-up. As long as the missing data is unrelated to the study intervention and the observed and unobserved data, limiting the analysis to those subjects who contribute endpoints produces unbiased estimates of the event rates.

The reason for missing data for all subjects will be reported. In addition, the following analyses will be produced to evaluate the missing data impact:

8.5.1 Multiple Imputation Method

SAS PROC MI will be used to generate 100 imputations sets of the study data using the following variables to model the key secondary endpoint – 3-month mRS: Gender, age, patient diagnosis categories, injury severity, study site and Baseline Pre-morbid mRS score (0,1,2) for the MI analysis. Other variables may be determined to be included in the MI process. The imputed value will be rounded to the whole number and capped at 0 and 5 for the lowest and highest value. For each imputed data set, the primary analysis will be analyzed by the same van Elteren test with stratification on patient diagnosis (ICH, AIS or SAH) as in the primary analysis. These will be provided to SAS PROC MIANALZYE in order to complete the final analysis. The amount of the variability associated with imputations and study analysis will be included in the summary.

8.5.2 Tipping Point Analysis

Tipping point approach is performing the sensitivity analysis under the missing at not random (MNAR) assumption. In other words, the tipping point approach is like a progressive stress-testing to assess how severe departures from missing at random (MAR) must be in order to overturn conclusions from the primary analysis.

Tipping point approach can be seen as a special application of the multiple imputation. It can also be considered as a special case of controlled imputation method (i.e., applying the shift parameter only to the active treatment group, not to the placebo group), that is to generate multiple imputed data sets, with a set of specified shift parameters that adjust the imputed values for observations in the treatment group, not the placebo group.

8.5.3 Worst Case Analysis

For missing mRS, score of 5/6 will be assigned to the FP treatment group, and score of 0 will be assigned to the SC treatment group. Data will be analyzed by the same van Elteren test with stratification on patient diagnosis (ICH, AIS or SAH) as in the primary analysis.

8.6 Subgroup Analysis

The primary endpoint will be explored in the following subgroups:

- Gender
- Age (<70, ≥70)
- Patient primary diagnosis (ICH, AIS or SAH)
- Injury severity (high or low, where high score includes NIHSS ≥ 17 for IS subjects, or baseline ICH score ≥ 3 for ICH subjects, or WFNS ≥ IV for SAH patients, and low

score includes NIHSS < 17 for IS subjects, or baseline ICH score ≤ 2 for ICH subjects, or WFNS ≤ III for SAH patients).

- Acetaminophen/Paracetamol drug use: yes/no is for indication of any use of a drug from the identified list of antipyretic drugs (including Acetaminophen) during the acute phase.
- Sedation drug use: yes/no is for indication of any use of sedation drugs during the acute phase.
- Baseline Pre-morbid mRS score (0,1,2)
- Region (US, EU, Asia, other (Australia, etc))

8.7 Exploratory Analysis

8.7.1 Analysis of fever/fever potential subjects

The Arctic sun case data collected during the acute phase for the SC subjects will be analyzed by independent individuals to determine if the subject would have developed fever if not on Arctic sun (fever potential subject). All subjects with fever (in both SC and FP groups) or subjects in FP with potential fever will be included in this analysis. The key secondary endpoint, 3 month mRS, in the two treatment groups will be compared using a van Elteren test (a particular form of Cochran-Mantel-Haenszel (CMH) test), with stratification on patient diagnosis (ICH, IS or SAH) as in the primary analysis.

9 Statistical Analysis of the Secondary Efficacy Endpoint

The following secondary endpoints will be summarized using descriptive statistics for the two treatment groups on ITT population. The secondary endpoints are initially described in Section 2.2.3, and the details on endpoint definition and analysis will be added below.

9.1 Modified Rankin Scale measured at 6- and 12-months

The mRS scores will be collected at Screen/Baseline, Post-Acute Phase, 3 Month, 6 Month, 12 Month follow up visits. If there is missing data, the 6-category mRS will be scored 5/6 if subject died or discontinued due to comfort measure only before each of the visits' window end. Data will be summarized at each visit as change from Post-Acute Phase by treatment group.

The mRS scores at 6- and 12-months will be analyzed similarly to the mRS scores at 3-months for the 6-category scores and the dichotomized binary outcome. A "shift analysis" will be employed to analyze the 6-category mRS outcome scales with a van Elteren test with stratification on patient diagnosis (ICH, IS or SAH). The dichotomized binary mRS success rate will be analyzed by the Cochran-Mantel-Haenszel (CMH) test with stratification on patient diagnosis (ICH, IS or SAH).

9.2 National Institutes of Health Stroke Scale measured at 3- and 6-months

NIHSS score will be collected at screening, Day 2, Day 5, Post-Acute Phase, Month 3 and Month 6 follow up visits. Day 2 and Day 5 scores are for measurement of ICU stay progress, and won't otherwise be summarized. Data collected for other visits will be summarized by treatment group. Only observed data will be used. The total score of change from Baseline will be summarized by treatment group.

The total NIHSS score at 3 month and 6 month will be analyzed by an ANOVA model with treatment and diagnosis categories as factors.

9.3 Barthel Index measured at 3- and 6-months

The Barthel Index score will be collected at Post-Acute Phase, Month 3 and Month 6 follow up visits.

The total Barthel Index score at 3 month and 6 month will be analyzed by an ANOVA model with treatment and diagnosis categories as factors. The total score of change from Post-Acute Phase will be summarized by treatment group.

9.4 Glasgow Outcome Scale Extended measured at 3- and 6-months

The Glasgow Coma Scale will be collected at Baseline, and the Glasgow Outcome Scale Extended score will be collected at Post-Acute Phase, Month 3 and Month 6 follow up visits. If there is missing data, the total Glasgow Outcome Scale Extended score will be scored 1 if subject died or discontinued due to comfort measures only before each of the visits' window end.

The total Glasgow Outcome Scale Extended score at 3 month and 6 month will be analyzed by an ANOVA model with treatment and diagnosis categories as factors. The total score of change from Post-Acute Phase will be summarized by treatment group.

9.5 Montreal Cognitive Assessment measured at 3- and 6-months

The Montreal Cognitive Assessment score will be collected at Post-Acute Phase, Month 3 and Month 6 follow up visits.

The total Montreal Cognitive Assessment score at 3 month and 6 month will be analyzed by an ANOVA model with treatment and diagnosis categories as factors. The total score of change from Post-Acute Phase will be summarized by treatment group.

9.6 Intensive Care Unit Length of stay

The ICU length of stay (in hours) is calculated as:

(Date/time of the ICU discharge– date/time of the ICU admission).

If ICU discharge date/time is missing, then the ICU duration of the subject will be missing and won't be included in the summary. If a subject died before ICU discharge, the subject death date should be entered as ICU discharge date.

The ICU length of stay will be analyzed by an ANOVA model with treatment and diagnosis categories as factors.

9.7 ICU Delirium

Intensive Care Delirium Screening Checklist (ICDSC) will be collected daily during the acute phase. ICDSC between 4-8 will be classified as Delirium. Percent of subjects with any delirium classification on any day, number of days with delirium classification and percentage of the acute phase with delirium will be summarized by treatment groups with diagnosis categories as factors. Note that only the observed days with Delirium checklist data will be included in this analysis.

9.8 Use of Mechanical Ventilation

Whether the subject is on a ventilator at any time will be collected daily during the acute phase. Percent of subjects with any use of ventilator and number of days with ventilator during the acute phase will be summarized by treatment groups with diagnosis categories as factors.

In the case where ventilator data is missing on certain day(s), if the day before was checked on ventilator, then the missing day(s) will be considered as on ventilator; otherwise will be considered as not on ventilator.

9.9 Hospital Length of Stay

The hospital length of stay is calculated as:

(Date of the hospital discharge – date of the hospital admission) + 1.

If hospital discharge date is missing, then the length of hospital stay of the subject will be missing and won't be included in the summary. If a subject died, the subject death date should be entered as hospital discharge date.

The hospital length of stay will be analyzed by an ANOVA model with treatment and diagnosis categories as factors.

9.10 Mortality

The mortality data will be summarized for 7-day, hospital discharge, 3-, 6-, and 12-month by treatment. If a subject discontinued before the specified timepoint (7-day, hospital discharge, 90-day, 180-day and 365-day, respectively), the subject will be treated as not evaluable and will not be included in the denominator. If a subject was determined to discontinue due to undergo comfort measure only, the subject should be considered as death on the day of discontinuation.

A Kaplan-Meier analysis will be used to estimate the mortality rate by treatment groups for overall population and the three diagnosis categories for 7-day, 3-, 6-, and 12-month.

10 Statistical Analyses of Safety Endpoints

10.1 AE

Each subject will be monitored for the occurrence of both adverse events (AEs) and major adverse events (MAEs) according to definitions in Protocol Section 7. All AEs will be monitored from study randomization through the end of the acute phase of the study (336 hours or discharge/ready to discharge from the ICU, whichever comes first) and MAEs will be monitored from randomization through the end of study participation.

Events with an onset prior to randomization should be reported in the subject's medical history. AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA).

To ensure consistency in determination of relatedness of adverse events, a Clinical Events Committee (CEC) will adjudicate all AEs determined by the site investigator to be related or possibly related to a study device or procedure. Additionally, the CEC will adjudicate all SAEs and MAEs.

All AEs from study randomization through Acute Phase (including MAE) will be summarized for:

- The number (n) and percentage (%) of patients with at least 1 event by System Organ Class (SOC) and Preferred Terms (PT)
- Event by intensity (severe, moderate, mild), presented by SOC and PT
- Event by CEC adjudicated relationship to treatment or procedure (related, not related), presented by SOC and PT
- CEC adjudicated Serious adverse event presented by SOC and PT

Post Acute Phase MAE will be summarized by System Organ Class (SOC) and Preferred Terms (PT).

A listing of all AEs, as well as of the SAEs, UADEs will be provided.

The above summaries will be performed on AT analysis set if ITT and AT are not the same.

10.2 MAE

In this study, MAEs will be monitored from randomization through the end of study participation, and an MAE will be defined as one of the following:

- Pneumonia;
- Sepsis;
- Malignant cerebral edema; or

- Death

The CEC adjudicated MAEs will be classified into the above categories by preferred terms of AEs for Pneumonia, Sepsis or Malignant cerebral edema. Death will be identified in the AE outcome. Rate of subject with MAE during the follow-up period of 3 month, 6 month and 12 month will be summarized by treatment, and the rate difference along with 95% CI will be estimated by Wald asymptotic method. The Preferred terms will be summarized under each category.

For 3 month report, data will be cut by Day 90.

10.3 Incidence of Infection during Acute Phase

Incidence of infection will be identified from AE data. Rate of subjects with infection during the acute phase will be summarized by treatment, and the rate difference along with 95% CI will be estimated by Wald asymptotic method. The total number of infection incidence and number of infections per subject (each AE reported counted as one incidence) will be summarized by treatment groups.

10.4 Incidence of Shivering

Shivering will be identified using “Shiver Assessment/Control” CRF page.

Rate of subjects with any shivering will be summarized by treatment, and the rate difference along with 95% CI will be estimated by Wald asymptotic method.

Number of days with shivering and percentage of days during the acute phase with shivering will be analyzed by subject, if there is any shivering incidence on the day, that day will be counted one of the days for the subject with shivering.

Significant shivering is defined as BSAS score of 2 or greater, all other incidences will be regarded as not significant. The rate of subjects with significant shivering will be summarized by treatment. Number of days with significant shivering and percentage of days during the acute phase with significant shivering will be analyzed by subject, if there is any significant shivering incidence on the day, that day will be counted one of the days for the subject with significant shivering.

Severe shivering is defined as BSAS score of 3, all other incidences will be regarded as not severe. The rate of subjects with severe shivering will be summarized by treatment. Number of days with severe shivering and percentage of days during the acute phase with severe shivering will be analyzed by subject, if there is any severe shivering incidence on the day, that day will be counted one of the days for the subject with severe shivering.

11 Statistical Analyses of Other Endpoints

11.1 Subjects Dispositions

The summary of the number of subjects enrolled, intent to treated (ITT), discharged from ICU, discharged from hospital, completed the study, and discontinued from the study by reason of discontinuation will be provided. Screen failures will be summarized for each inclusion/exclusion criteria that were not met.

11.2 Protocol Deviations

The number of subjects with protocol deviations will be summarized with descriptive statistics by nature of the deviation. Protocol deviations will be listed with date of occurrence and the nature of deviation. This summary will be reported based on the ITT population.

11.3 Demographics and Background Disease Characteristics

Demographics and background disease characteristics will be summarized with descriptive statistics using the ITT analysis set. Summary statistics for categorical variables will include frequency counts and percentages and for continuous variables will include mean, standard deviation, minimum, median, and maximum.

Demographics and baseline characteristics variables include:

- Age at screening (year)
- Sex (Male, Female)
- Race (American Indian or Alaskan Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White and Other)
    - Asian will be further categorized to Japanese, Chinese, Korean and Other Asian Race
- Ethnicity (Hispanic/Latino)
- Baseline Weight
- Baseline Height
- Baseline Body mass index (BMI) calculated from weight and height.
- Baseline Temperature
- Baseline Skin Condition
- Baseline Active Infection
- Baseline Glasgow Outcome Scale
- Primary neurological diagnosis
- Testing Used to Confirm Diagnosis and Interventions or Treatments Given
- Subject Receive a Foley Catheter
- Method of Thermometry to be Used During the Study

Medical history (in 6 categories: risk factors, nervous system disorder, cardiovascular disorder, respiratory illness, metabolic disease, and other significant medical history; and in sub-terms) will be summarized.

11.4 Relevant medication and therapy

Relevant medications include antipyretics, antishivering medications, anticonvulsants, sedatives, and antimicrobials during the acute phase of the study. Only medications after randomization will be summarized by treatment group including rate by subject, indication and duration if available. Indication and duration of use will also be summarized across medication by treatment group by subject level data. The antipyretic medications will be summarized by treatment groups including acetaminophen and other similar drugs.

Fever control therapy and counter warming measure will be summarized by treatment group.

11.5 Follow-up Period

The duration of follow-up period after randomization during the study is calculated as:

(Date of the last study visit – date of the randomization) + 1.

Durations will be summarized using the ITT analysis set.

11.6 Acute Phase Treatment Characteristics

For the Fever Prevention group, the Arctic Sun 5000 therapy duration (in hours) is calculated as:

(Date/time of the therapy end– date/time of the therapy start).

The therapy suspension information will be summarized for fever prevention group.

11.7 Device failure, malfunctions and defects

Device failure, malfunctions or defects will be tabulated by the failure code.

12 Reference

Chaisinanunkul N, etc., Adopting a Patient-Centered Approach to Primary Outcome Analysis of Acute Stroke Trials Using a Utility-Weighted Modified Rankin Scale; Stroke. 2015; 46:2238-2243

Howard G, etc., A simple, assumption-free and clinically interpretable approach for analysis of modified Rankin outcomes; Stroke. 2012 March ; 43(3)

Proschan MA, Lan KK, Wittes JT. Statistical monitoring of clinical Trials: A unified approach. Springer 2006

Yuan Y, Sensitivity Analysis in Multiple Imputation for Missing Data, Paper SAS270-2014, SAS Institute Inc.

13 Appendix

Appendix 1.1 Tables/Listing/Figures Shell for DMC close session output

Appendix 1.2 Tables/Listing/Figures Shell for 3 Month Full Report

Appendix 1.3 Tables/Listing/Figures Shell for 12 Month Report

Appendix 2 Derived data specification

Appendix 3 Protocol Deviation Major/Minor classification