STATISTICAL ANALYSIS PLAN
Protocol: ICH02
NCT01827046

Minimally Invasive Surgery plus rt-PA for Intracerebral Hemorrhage Evacuation Phase III
(MISTIE III)

July 25, 2018

Prepared for:

BIOS Coordinating Center
Johns Hopkins University

Prepared by:

MISTIE III Data Coordinating Center
Johns Hopkins University
Bloomberg School of Public Health
Department of Biostatistics
615 N. Wolfe Street
Baltimore, MD 21205

Approved by:	Signature	Date
Daniel F. Hanley, MD Trial PI	Daniel Hanley	7/25/18
Richard Thompson, PhD Statistical PI	Richard Thompson	7/25/18

Chapter 22.0: STATISTICAL ANALYSIS PLAN

Table of Contents

List of Abbreviations:

AE	Adverse event
AIC	Akaike Information Criterion
AUC	Area under the curve
BCa	Accelerated bootstrap
BP	Blood pressure
CCC	Clinical Coordinating Center
CES-D	Center for Epidemiological Studies - Depression
CI	Confidence interval
CLEAR III	Clot Lysis: Evaluating Accelerated Resolution of Intraventricular Hemorrhage Phase 3
CT	Computed tomography
DCC	Data Coordinating Center
DSMB	Data and Safety Monitoring Board
EDA	Exploratory data analyses
EDC	Electronic data capture
eGOS	Extended Glasgow Outcome Scale
EOT	End of treatment
EQ-5D	EuroQol five dimension
EQ-VAS	EuroQol Visual Analogue Scale
EVD	Extraventricular drainage
GCS	Glasgow Coma Scale
GEE	Generalized estimating equations
GOS	Glasgow Outcome Scale
ICH	Intracerebral hemorrhage
ICP	Intracranial pressure
ICU	Intensive care unit
ITT	Intent-to-treat
IVH	Intraventricular hemorrhage
LGR	Lower good recovery (eGOS)
LM	Lower moderate disability (eGOS)
LS	Lower severe disability (eGOS)
LTF	Loss to follow-up
MD	Moderate disability (eGOS)
MedDRA	Medical Dictionary for Regulatory Activities
MIS	Minimally invasive surgery
MISTIE III	Minimally Invasive Surgery plus rt-PA for Intracerebral Hemorrhage Evacuation Phase 3
MMSE	Mini-Mental State Examination
mRS	Modified Rankin Scale
NIH	National Institutes of Health
NIHSS	National Institutes of Health Scale
PBSI	Preference Based Stroke Index
QIC	Independence model criterion
QOL	Quality of life
ROC	Receiver Operating Characteristic
rt-PA	Recombinant tissue plasminogen activator
SAE	Serious adverse event
SIS	Stroke Impact Scale
UGR	Upper good recovery (eGOS)
UK	United Kingdom
US	Upper severe disability (eGOS)
VS	Vegetative state (eGOS)

1.0. Statistical Analysis Plan and Statistical Reports

This chapter provides the details of statistical considerations, analyses, and reports planned for the MISTIE III Trial, including proposed regression models, sample size estimates, power considerations, interim analyses, and analysis of adverse events. The proposed analyses will be conducted on the whole patient sample and on subsets of patients defined by gender, race, and ethnicity as well as patient-specific clinical factors/categories and treatment categories. In addition, this chapter discusses statistical issues relevant to these analyses (e.g., sample data to be used, missing data, adjustments for multiplicity, etc.).

Initially, the Data Coordinating Center (DCC) will examine data quality, create secondary variables, and perform exploratory data analyses (EDA). All variables will be evaluated to detect gaps, patterns, and inconsistencies in the data. These analyses will emphasize examination of the nature and extent of variability for all variables. Visual techniques to explore continuous variables will include stem-and-leaf plots, box plots, and quantile-quantile plots. Outliers will additionally be examined for data entry errors. Summary statistics for continuous variables will include the number of patients, median, mean, standard deviation, and range, while statistics for categorical variables will include the frequency and percentage of patients in each category.

Statistical analyses will be performed on data that have been quality-assured through DCC protocols and monitoring reviews, and that have been exported by the DCC directly from the MISTIE III database. The following analysis procedures may be applied to blinded (no treatment arm designation, for study investigators and staff), partially unblinded (treatment designation only as “A” and “B,” for DSMB and at interim analyses), and unblinded (full treatment designation, for final analyses) data, by DCC staff having the appropriate role permissions.

The MISTIE III Executive Committee receives the open DSMB reports.

2.0. Objectives of the Study

2.1. Efficacy

The primary aim of this multicenter, international, randomized, open-label, phase 3 trial is to determine the efficacy of intra-clot catheter placement and aspiration of hematoma contents, followed by gentle clot irrigation with low dose recombinant tissue plasminogen activator (rt-PA), referred to as Minimally Invasive Surgery plus rt-PA (MIS+rt-PA). The primary hypothesis of this trial is that the MIS+rt-PA approach to intracerebral hemorrhage (ICH) management will result in more patients having overall better functional outcome at 365 days, as defined by the proportion of subjects with dichotomized, adjudicated modified Rankin Scale (mRS) score 0-3 vs. 4-6, when compared to the mRS scores for patients managed with standard aggressive medical treatment. Although the mRS includes a category for death, and hence the primary analysis incorporates mortality outcomes, in-hospital and 30-day mortality will also be examined as safety outcomes for the trial. Furthermore, 365-day mortality will be evaluated separately as a secondary efficacy outcome. An intent-to-treat (ITT) paradigm for the analyses will be incorporated. The MISTIE III trial will enroll a total of 500 eligible patients, 250 randomized each treatment arm (MIS+rt-PA or medical management). The power to detect (88%) is for an estimated effect size of 13% across all clot locations with this sample size (500).

Additional analyses of efficacy outcomes will evaluate the therapeutic benefit of MIS+rt-PA relevant to medical management in improving subjects’ function and quality of life at 365 days post symptom onset as measured by extended Glasgow Outcome Scale (eGOS), days in various locations such as home

during the 365-day period, patient disposition at 180 and 365 days including proportion of days in long-term care facilities, National Institutes of Health Stroke Scale (NIHSS), Barthel score, days in hospital, days in intensive care unit (ICU), Stroke Impact Scale (SIS), Preference-Based Stroke Index (PBSI), EuroQol-5 Dimension (EQ-5D), Mini-Mental State Examination (MMSE), Personal Health Utility Interview, and Center for Epidemiological Studies – Depression (CES-D).

2.2. Efficacy Measures Summary: All Analyses Adjusted Similarly to the Primary Analysis

- **Primary Efficacy Analysis:** Dichotomized, adjudicated, cross-sectional mRS 0-3 vs. 4-6 at 365 days post-ictus, adjusting for baseline (pre-randomization) variables used in covariate adaptive randomization as well as the clinically established severity variables IVH size and ICH location (lobar or deep).

- **Additional, Exploratory Analyses of mRS at 365 days:**
  a. Dichotomized, adjudicated, cross-sectional mRS 0-3 vs. 4-6 at 365 days post ictus, adjusting for baseline (pre-randomization) variables used in covariate adaptive randomization as well as additional baseline variables determined to be associated with the outcome.
  b. Ordinal, adjudicated mRS (0-6) at 365 days.
  c. mRS 0-3 vs. 4-6 at 365 days based on a full longitudinal analysis of the mRS data at 90, 180, and 270 days.
  d. Severity subgroup analyses (including, but not limited to, IVH size, location of ICH, presentation co-morbidities such as diabetes, CVD, and white matter disease, etc.)
  e. Cross-sectional of mRS 0-3 vs. 4-6 for demographic subgroups of patients, including for race and gender. (For d and e, see section 5.5.)
  f. Analysis of cross-sectional mRS 0-3 vs. 4-6 at 365 days that models random effects for hospital sites.

- **Secondary Analyses of other outcomes**

  1. Dichotomized eGOS UGR-US vs. LS-Death at 365 days.
     - Subordinate: Ordinal eGOS UGR, LGR, MD, LM, US vs. LS, VS, Death at 365 days.
  2. All-cause mortality longitudinally from ictus to 365 days.
     - Subordinate: Predictors of mortality.
  3. Relationship between clot removal as an AUC clot-assessment that estimates the time-weighted clot volume from ictus to end of treatment (EOT; i.e., 24 hours after last dose) as AUC clot exposure and functional outcome (proportion 0-3 mRS).
  4. Patient disposition:
     a. Home days over 365 days' time from ictus.
     b. Patient location at 365 days post ictus (i.e., good vs. bad location. See Tertiary 2g for definition.)
  5. 180-day efficacy:
     a. Dichotomized, adjudicated, cross-sectional mRS 0-3 vs. 4-6.
        - Subordinate (1): Ordinal, adjudicated mRS (0-6).
     b. Dichotomized, adjudicated eGOS UGR, LGR, MD, LM, US vs. LS, VS, Death.
        - Subordinate (2): Ordinal, eGOS UGR, LGR, MD, LM, US vs. LS, VS, Death.
        - Subordinate analyses 2-5 as defined under the Primary but for eGOS.
  6. Type and intensity of ICU management:
     a. ICU days

b. ICU management (ICP, Ventilation, BP support, osmotherapies)
c. Hospital days

7. EQ-VAS and EQ-5D

- **Tertiary (exploratory/sensitivity):**

1. Functional status at 180 and 365 days post-stroke:
   a. GOS
   b. Barthel Index
   c. NIHSS
   d. MMSE
2. Quality of life at 180 and 365 days post-stroke:
   a. SIS
   b. PBSI
   c. Personal Health Utility Assessment Interview
   d. CES-D
   e. EQ dimensions individually assessed
   f. Home days (180 only)
   g. Patient location including proportion of days in home, rehabilitation, acute care, long-term care facilities. These will be evaluated by individual location type and dichotomized as good vs bad outcome defined as home and rehabilitation vs acute care, long-term care and death.
   h. Proportion of subjects in vegetative condition
3. Cost-benefit model

In addition to the primary, secondary, and tertiary (exploratory) measures, there are a number of related and intermediate surrogate outcomes of interest, including time to treatment, surgical task descriptors, clot locations, Glasgow Coma Scale (GCS), Graeb Scale, clot lysis rate and final clot size reduction.

2.3. Safety

The safety of MIS+rt-PA relative to medical management is assessed by comparing the proportion of subjects between the two groups with treatment-related SAEs within the 30-day time frame: mortality, rebleeding, infection, and first-week (designated as operative mortality for the surgical group) mortality. Additionally, the number and proportion of all SAEs across the entire time frame in the respective randomized groups will be evaluated starting with the follow-up time frames of 180 and 365.

Interim safety analyses will be prepared for the external Data Safety and Monitoring Board (DSMB) on a pre-arranged schedule (such as semi-annual or after enrollment of a fixed number of subjects) to evaluate efficacy and safety. Safety measures include: monitoring the incidence of symptomatic and asymptomatic intracranial bleeding events (i.e., hemorrhage extension, new hemorrhage, catheter-tract hemorrhage) through daily computed tomography (CT) scans for the first four days after randomization and then repeated one day (approximately 24 hours) after ICH catheter removal, and monitoring the incidence of confirmed and suspected infection and in-hospital mortality.

2.3.1. Safety Measures Summary

- Mortality and safety events at 30 days post-randomization:
  a. First-week (operative) mortality
  b. All-cause mortality

c. Symptomatic bleeding rate
d. Infection rate
e. Total SAEs at 30 days
f. Summary of AEs and SAEs by MedDRA code and grouped by organ system

2.3.2. Safety Monitoring and Recruitment Suspension Rules

Safety events above pre-specified thresholds will trigger a "suspend recruitment and review" by the DSMB to investigate the presumed cause and impact of these events. Such events are initially reviewed by DCC staff to determine if an event threshold has been reached at which time the study investigators and the DSMB will be notified. The thresholds triggering such a review are: 30-day mortality >60%, a MIS+rt-PA–related symptomatic bleeding rate >25% (during active treatment or during the 72 hours after treatment will be monitored), first-week operative death rate >10%, and a procedure-related infection rate >15% (over the initial 30 days). If any of these SAEs are attributable to the intervention (catheter insertion or manipulation, or rt-PA injection), the study will be suspended for a complete safety review.

Interim analyses for efficacy will occur after 375 subjects are enrolled and 180 days outcomes are available. These analyses will be based on O'Brien-Fleming stopping boundaries for efficacy and more aggressive (i.e., more likely to stop if there is no early signal of benefit) stopping boundaries for efficacy and safety analyses will be every six months or at other intervals as determined by the external DSMB.

3.0. Study Design

This study is a phase 3, 1:1 randomized, open-label, multicenter evaluation of MIS and ICH lysis with rt-PA versus medical care. The study (n=500) will evaluate the efficacy and safety of MIS plus 1 mg of rt-PA administered every eight hours for up to nine doses as compared to subjects treated with conventional medical management.

4.0. Sample Size Determination for the Primary Efficacy Analysis

The MISTIE phase 2 results, for all available data (non-randomized trial-standardized "pilot" and randomized participants), demonstrated an estimated absolute increase in the probability of achieving an mRS 0-3 of 13.0% [95% CI: (0.4% to 26.4%)] comparing MIS+rt-PA treatment versus medical management 180 days post-stroke.[1] For the intention-to-treat (ITT) analysis, when only randomized patients are considered (n=96), the estimate for the unadjusted treatment effect is 10.9% [95% CI: (-8.8% to 29.4%)], and 16.2% [95% CI: (0.3% to 32.3%)] after adjusting for baseline clinical characteristics. This range of estimates provides a strong motivation for conducting MISTIE III to confirm these clinically beneficial findings. The analogous unadjusted and adjusted effect sizes based on non-lobar ICH patients are very similar to these effect sizes. The second stage of MISTIE II provided evidence of increasing frequency of mRS scores in the 0-2 categories between 180 and 365 days with an unadjusted proportion of mRS 0-3 increasing to 14%. The sample size was selected based on 180-day outcomes in MISTIE II.

It was decided to change the primary outcome to be based on 365-day outcomes because of the lack of a plateau in good functional outcome at 180 days.

5.0. Definition of Study Samples

5.1. Target Population

The target population to which the MISTIE intervention or conventional medical management may be applied are patients with site-determined ICH ≥30 mL and not requiring extraventricular drainage (EVD) for the management of acute obstructive hydrocephalus.

5.2. Intent-to-Treat (ITT) Sample

As the primary analysis, all efficacy and safety outcome measures are analyzed under the ITT. Under this principle, each eligible subject is analyzed according to the treatment group to which they were assigned at the time of randomization. A total of 500 eligible subjects will be randomized. ***Exposure to the MISTIE procedure will be considered evidence of intention to treat.***

5.3. Safety Analysis Sample

All randomized subjects are included in the safety analysis sample. Further safety analyses may be determined by clinical events and subgroups of the surgical task.

5.4. As-Treated Sample for MIS+rt-PA

The potential for cross-overs to the MIS+rt-PA treatment in this study is minimal; none occurred in our prior experience. However, in the case of cross-overs, a per-protocol sample will be constructed and examined in which treatment-as-received is analyzed.

Given that amount of clot resolution is a potentially treatment-related post-randomization variable, this covariate will be examined as an important potential mediator of final outcomes as part of a per-protocol analysis. This treatment performance variable will be explored at the protocol thresholds of 15 mL, and 10 mL and the observed level of 5 mL. Other thresholds may be explored as determined by inspection of the data and clinical findings. Surgical performance will be evaluated by chosen trajectory, surgical experience level and compliance with Surgical Center recommendations, and by actual task performance versus planned performance.

5.4.1. As-Treated Sample for Any Form of Surgery

Given the lack of guidelines to indicate whether any form of surgery is beneficial, different combinations of surgical versus medical therapy will be explored, starting with all subjects receiving any surgical procedure versus all subjects receiving only medical therapy. Important groups that represent actual practices and the timing of such practices occurring will be explored. Given that clot resolution is a potentially treatment-related post-randomization variable, this covariate will be examined as an important potential mediator of final outcomes as part of the as-treated analysis.

Given the absence of clinical trial demonstration of benefit of craniotomy, we will assess the benefit of craniotomy or not by comparing all subjects receiving craniotomy versus those receiving only MIS+rt-PA and those receiving only medical therapy. Craniotomies will be divided into subgroups by time performed (in equal size groups and by day and half day). Finally, we will consider all craniotomies as salvage and assign them to death, then determine the frequency at which they are “life saving.”

5.5. Demographic and Clinically Specific Subgroups

All the primary, secondary, and safety analyses will be performed by gender, race, and ethnicity to determine if the treatment effect or safety differs by these demographics. Patient-specific clinical or "severity" factors/categories and treatment categories will be explored. Planned explorations include age, GCS (or NIHSS), ICH location, time to treatment, and initial ICH size.

Subgroups utilized for adaptive cut points are: age <56, age 56 through 67, and age ≥67; GCS 3–8, 9–12 and 13–15; initial ICH size (0–<25; 25–<35; 35–<45; 45–<55; 55–<65; and ≥65). We will examine the following subgroups: stability ICH size <50 and ≥50; lobar vs. deep; stability IVH size; time to surgery at <24 hours, 24-48 hours, and >48 hours from ictus; GCS 3–8, 9–12 and 13–15; and age <65 and ≥65.

Additional analysis of groupings of these clinical factors will be dictated by group size and balance between groups. We will also explore time as a continuum and at other clinically meaningful cut points. We will treat age, ICH size, and IVH size as a continuum and collapse them to other meaningful subgroups, as described above. NIHSS and GCS will also be treated as a continuum of scores. Based on post hoc analysis of the STICH trials, we will evaluate GCS 9–12 versus all other score levels.

5.6. Definition of Treatment Failures

Treatment failures are defined as: rt-PA administrations discontinued due to clinically significant rebleeding (prolonged GCS motor deterioration >2 motor points), the Investigator's judgment to not treat, and failure to achieve the target volume of 15 mL or less.

6.0. Randomization

Randomization will occur after completed screening for vascular malformation and central validation of inclusion and exclusion criteria. Subjects will be randomized to MIS+rt-PA or medical management using a covariate-adaptive design similar to that used for the CLEAR III trial.[2] The goal of this randomization scheme is to obtain an improved balance across study arms in the number of subjects with certain pre-randomization variables that are most strongly predictive of the functional outcomes (mRS) at 180 and 365 days post-stroke. These variables include pre-randomization clot volume, age, and baseline severity of impairment as measured by GCS. Under this scenario, a randomization scheme will be used which increases a newly enrolled subject's probability of being assigned to the study arm that improves the overall balance in these important prognostic factors. Briefly, prior to randomization for each incoming patient, the imbalance between treatment arms will be determined using the accumulating available data on clot volume, age and GCS. The patient will then be randomized with high probability (e.g., "weighted coin toss") to the treatment arm that lowers the imbalance. Patients at any given site will not be considered for covariate-adaptive randomization until the site has enrolled four or six (blinded; sites are randomized to four or six to maintain blinding) patients into each treatment arm by within-site block randomization. This randomization scheme will be implemented using software included in the MISTIE III EDC (electronic data capture) system similar to what has been successfully implemented in the CLEAR III trial.

7.0. Blinding

The examiner performing the 30-, 180-, and 365-day follow-up assessments will video record the mRS interview assessments and upload the video to the Outcome CCC at the Western Infirmary in Glasgow, UK, where trained reviewers will classify the interview objectively without knowledge of the examiner's

score, patient name, or the treatment details including therapy intensity levels.

8.0. Multiplicity

The primary analysis of efficacy is tested at the two-sided alpha level of 0.05. In addition, secondary analyses will consider the alpha=0.05 two-sided type I error. Secondary analyses will be examined in the order specified and will not be corrected for multiplicity of testing.

Secondaries 2-7 are defined. The other analyses without specific hypotheses are considered exploratory. We will present descriptive statistics and CIs.

For safety outcomes, we do not account for multiplicity as this study was not powered to observe safety. Thus, a cautionary approach is justified.

9.0. Missing Data

Based on previous MISTIE and CLEAR studies performed in this group of stroke patients, the CCC will make substantial efforts to ensure complete collection of data for all patients, and to accrue minimal loss to follow-up (LTF) to optimize evaluation of the primary outcome of 365-day adjudicated mRS post-stroke. The CLEAR III trial experienced an overall LTF rate of 2%. Since we expect missingness will be more prevalent for the 365-day mRS (the primary efficacy end point) than at early time points, we will use the targeted maximum likelihood estimator (TMLE)[3] to adjust for censoring, as specified in the description of the model for the primary efficacy model given in Section 11.1 below. The effects of incompleteness/noncompliance will be quantified through sensitivity analyses, the gold standard in the field.[4]

10.0. Baseline Characteristics

10.1. Demographics

Summary statistics will be presented for age (years), gender (male or female), race and ethnicity, presentation center, pertinent medical history including illicit drug use, primary diagnosis (ICH vs. ICH with IVH), hypertension history, diabetes, anticoagulation, tobacco and cocaine use, and antiplatelet drug use.

10.2. Medical History

Summary statistics will also be presented for medical history, as well as presentation variables: temperature, pulse, systolic blood pressure, diastolic blood pressure, presenting NIHSS, GCS, ICP, ICH volume, IVH volume, extent of white matter disease (van Sweiten index), and location of bleed.

11.0. Efficacy Analyses

11.1. Primary Efficacy Analysis: Dichotomized, Adjudicated mRS 0-3 vs. 4-6 at 365 Days Post Ictus

This primary analysis will produce an estimate and 95% confidence interval for the average treatment effect comparing assignment to MIS+rt-PA treatment (MISTIE) versus control (standard of care, i.e., medical management), using the ITT principle. Specifically, we will estimate the difference between the probability of an adjudicated mRS score of 3 or less at 365 days post-stroke comparing assignment to

treatment versus control; this difference is called the average treatment effect. Below is a description of the procedure for computing estimators of the average treatment effect. We also will test the null hypothesis of zero treatment effect at level 0.05.

To estimate the average treatment effect, we will use the targeted maximum likelihood estimator (TMLE) of van der Laan and Gruber.[3] This estimator adjusts for baseline (i.e., pre-randomization) variables in order to improve precision. The baseline variables, denoted by the vector B, include the three variables used in the covariate-adaptive randomization as given in Section 6.0 (i.e., pre-randomization clot volume, age, and baseline severity of impairment as measured by GCS) as well as the following additional variables: IVH volume and location of ICH (indicator of lobar vs. deep). All baseline variables are treated as categorical and included as main terms in the models below.

The TMLE estimator also involves adjustment with the goal of reducing bias due to informative censoring, that is, participant dropout that may be correlated with the primary outcome. The TMLE adjusts for missing outcomes through a combination of inverse probability of censoring weights and outcome prediction models. Specifically, the weights and prediction models use the aforementioned baseline variables, study arm, and the following post-randomization variables: mRS scores measured at 30, 180, 365 days from ictus, denoted M_{30}, M_{180}, M_{365} respectively. The motivation is that if mRS score at a given visit is correlated with the mRS score at the next visit and also with dropout, then adjusting for observed differences in the mRS history of those who drop out versus those who stay on study may reduce bias (under the assumption that censoring is missing at random). As we describe below, this adjustment is implemented by fitting models for dropout and models for the primary outcome based on mRS measured at the previous visit, study arm, and baseline variables.

Let A denote the study arm, where A=1 for the MISTIE treatment arm and A=0 for the standard of care (medical management) arm. Let C_{30}, C_{180}, C_{365} denote the indicator of being uncensored at the 30-, 180-, 365-day visits, respectively. E.g., C_{180}=1 means the participant attended the 180-day visit and had his or her mRS M_{180} recorded.

The TMLE involves first constructing inverse probability of censoring weights W_{30}, W_{180}, W_{365}, defined as follows:

$W_{30}=1/[Pr(C_{30}=1 \mid B, A)\ P(A \mid B)]$
$W_{180}= W_{30}/Pr(C_{180}=1 \mid C_{30}=1, B, A, M_{30})$
$W_{365}= W_{180}/Pr(C_{365}=1 \mid C_{180}=1, B, A, M_{180})$

where the probabilities are estimated by fitting logistic regression models. Specifically, we fit logistic regression models for the probability of censoring at each visit given mRS at the previous visit, study arm, and baseline variables, among those who attended the previous visit. Each model includes main terms and an intercept, e.g., the model

$$Pr(C_{365}=1 \mid C_{180}=1, B, A, M_{180}) = \text{logit}^{-1}(\beta_0+\beta_1 B+\beta_2 A+\beta_3 M_{180})$$

for the probability of remaining uncensored at 365 days given 180-day mRS, study arm, and baseline variables, is fit among those uncensored at 180 days (C_{180}=1).

We also fit a logistic regression model for the probability of A=1 given baseline variables B, used in W_{30}. When the mRS score at an interim visit (30, 180 days) is included as a covariate in a regression model, it is coded as a categorical variable. We also truncate each weight at 20 (i.e., for weights above 20 we replace them by 20). Below, there is no connection between the coefficients in different models—we

use β to denote generic coefficients in each model, for simplicity.

The TMLE involves a sequence of weighted regression model fits using the inverse probability of censoring weights constructed above. First, the weighted logistic regression model:
$\Pr(M_{365} \leq 3 \mid C_{365}=1, B, A, M_{180}) = \text{logit}^{-1}(\beta_0+\beta_1 B+\beta_2 A+\beta_3 M_{180})$ **(1)**
using weights W_{365} is fit using all participants who are uncensored at 365 days ($C_{365}=1$).
Next, for each participant who is uncensored at 180 days (the previous visit), this model fit is used to generate the predicted probability, denoted by S_{180}, of Y=1 given their 180-day mRS M_{180}, study arm A, and baseline variables B.

The above two-step process (model fitting then generating predictions) is iterated by replacing ($M_{365} \leq 3$) by S_{180} in (1) and regressing on previous visit mRS, study arm, and baseline variables.

Specifically, fit $E(S_{180} \mid C_{180}=1, B, A, M_{30}) = \text{logit}^{-1}(\beta_0+\beta_1 B+\beta_2 A+\beta_3 M_{30})$ with weights W_{180} **(2)**
using all participants who are uncensored at 180 days ($C_{180}=1$).
and use the model fit to generate the predicted mean S_{30} of S_{180} for each participant with $C_{30}=1$;
next, fit $E(S_{30} \mid C_{30}=1,B, A) = \text{logit}^{-1}(\beta_0+\beta_1 B+\beta_2 A)$ **(3)**
with weights W_{30} using all participants who are uncensored at 30 days ($C_{30}=1$).

Lastly, the probability of Y=1 under assignment to study arm A=1 (MISTIE treatment) is estimated by generating for every participant (in both arms) a prediction by substituting their vector of baseline variables B and setting A=1 in the model fit (3). Similarly, the probability of Y=1 under assignment to study arm A=0 (medical management) is estimated by generating for each participant (in both arms) a prediction by substituting their vector of baseline variables B and setting A=0 in the model fit (3). The primary analysis estimator of the average treatment effect as a risk difference is defined as the difference between the former and the latter.

The 95% confidence interval will be computed by the non-parametric bootstrap (resampling participants with replacement and recomputing the above estimator), and using the bias corrected and accelerated (BCa) method.[5] The null hypothesis of zero average treatment effect will be rejected if this confidence interval excludes 0. This completes the description of the primary efficacy analysis, except for how missing baseline variables and intermittent (non-monotone) missingness are handled, as described next.

Missing baseline variables are imputed using the median for continuous-valued variables and mode for categorical variables. This and the following paragraph are done before any of the steps above.

We next describe how intermittent (non-monotone) missing data are handled, such as when a participant misses the 30-day visit only. We consider each interim visit at v=30 and 180, in turn. First, for visit v=30 days, we fit a multinomial logistic regression for M_{30} given A, B among those with $C_{30}=1$; for those with $C_{30}=0$, we then set their M_{30} to be the value with highest predicted probability from the model fit. Next, for visit v=180 days, we fit a multinomial logistic regression for M_{180} given A, B, C_{30}, among those with $C_{180}=1$; for those with $C_{180}=0$, we set their M_{180} to be the value with highest predicted probability from the model fit. We do not apply this process to the final outcome at 365 days. All models include intercepts and main terms for the variables listed. Next, we modify model (1) to include the additional main term $C_{30}C_{180}$ and its interaction with A. Similarly, in model (2) we include the additional main term C_{30} and its interaction with A. The idea is that each model (1)-(2) now contains an indicator of whether the mRS score on the right side was directly measured or was predicted by the above process based on previous mRS measurements, study arm, and baseline variables.

We will conduct sensitivity analyses to evaluate the impact (beyond what is explained by variables used in the analysis above) of informative censoring on our estimators, following recommendations from the U.S. National Research Council Panel on Handling Missing Data in Clinical Trials.[6]

11.2 Additional (Non-Primary) Efficacy Analyses

Models as detailed in equations (1) to (3) above an analogous approach will be used to assess 365-day differences with other corrections described next. Estimates from these models will account for design aspects such as potential within-site clustering that might otherwise lead to biased estimates in the standard errors associated with the regression coefficients. Additionally, we will control for clinically important variables, including pre-randomization ICH volume, IVH volume (or presence, depending on the variability in IVH volumes), baseline stroke severity, age, co-morbidity (diagnosis of diabetes or history of cardiovascular disease), and extent of white matter disease.

11.2.1. Additional Efficacy Analysis: Ordinal, Adjudicated mRS (0-6) at 365 Days

When we consider analysis of the ordinal mRS, we will estimate a proportional odds model of the form:

$$log\frac{\gamma_k}{1-\gamma_k} = \alpha_k + \beta X + \theta Trt \quad (6)$$

where α_k = Prob(mRS < k), k=0, 1,...,5.

This proportional odds model is very similar to the models described under 11.1. above, with the difference being that a treatment effect leads to an increase in the likelihood of a patient being in any subsequently lower mRS category. The probabilities, α_k, are cumulative probabilities which incorporate all outcomes below any given level. The set of intercepts, α_k k=0,1,…,5, define the initial probabilities of these cumulative levels, covariates are again included through β and X, and the treatment difference θ now describes the log-odds of moving from any cumulatively higher (worse) category, into any cumulatively lower (better) category. For instance, here θ describes both how the MIS+rt-PA changes the probability of ICH patients away from a potential mRS category of “Dead” (6) to “No Symptoms through Severe Disability” (0-5), as well as how the MIS+rt-PA changes patient probabilities away from “Severe Disability or Dead” (5,6) to “No Symptoms through Moderate Disability” (0-4) and so forth. This model is potentially more efficient than a simple dichotomous approach since it uses the available information across the scale of the measure. To examine the proportional odds assumption, we will additionally examine generalized ordered logit models that replace β and θ with β_k and θ_k, allowing non-parallel effects between successive cumulative categories. In the event that the proportional odds assumption is rejected, we will present results from the generalized ordered logit models.

For the ordinal outcome, the average treatment effect is defined as the average of the cumulative log-odds:

$$\beta = \frac{1}{K-1}\sum_{i=1}^{K-1} log\frac{P(Y_1 \le i)/P(Y_1 > i)}{P(Y_0 \le i)/P(Y_0 > i)} \quad (7)$$

where Y_r is the mRS score at 365 days and r=1 for MIS+rt-PA and 0 for medical management.

Survival models will be constructed to examine the mortality for the MIS+rt-PA and medically managed groups. Standard Cox proportional hazard models will initially be examined. These models are commonly written as:

$$h_i(t) = \lambda_0(t)\exp(\gamma X + \phi \cdot TRT) \qquad (8)$$

where $\lambda_0(t)$ is a non-parametric baseline hazard function, γ is a vector of regression coefficients related to the X covariates, and ϕ is the parameter for the treatment effect of MIS+rt-PA on the hazard of death (mortality). This model may be extended by including time-dependent covariates and non-proportional hazards as deemed necessary through diagnostic checks. Observed mortality will also be compared with predicted mortality based on clinical presentation, overall, and by treatment group. Predicted mortality will be estimated. We will develop a severity model that predicts mortality based on the Tuhrim model.[2,7,8]

11.2.2. Additional Efficacy Analysis: Longitudinal Analysis

Longitudinal analysis will examine the secondary outcomes that are longitudinally collected at baseline and the follow-up time points (e.g., at 30, 90, 180, 270, and 365 days post-stroke). These data will have within-patient correlations from one time point to the next, making it necessary to use appropriate analytical methods that properly estimate the coefficients and their associated standard errors in the analysis models. Since the overall goal of the trial is to compare subjects receiving the MIS+rt-PA intervention with subjects receiving medical standard of care, a marginal model estimated using a generalized estimating equations (GEE) technique is appropriate. The GEE technique requires specification of the standard "mean" model, as well as specification of a working "association" model. One strength of GEE is that the fixed-effect parameters in the mean model (i.e., the intervention effect) are consistent regardless of whether the association model is specified correctly under mild regularity conditions. The fixed-effects will include treatment arm (coded as 1 for MIS+rt-PA and 0 for medical management), follow-up time represented with four indicator variables for 90, 180, 270, and 365 days post ictus), and their interactions. The interaction terms will test (at 0.05 level of statistical significance) whether the longitudinal change in the outcome is statistically different between study arms. Additional variables will include ICH volume, baseline GCS, age, IVH, and comorbidities, such as history of diabetes or cardiovascular disease. Models will be compared using quasi-likelihood under the independence model criterion (QIC) for GEE.

11.2.3. Additional Efficacy Analysis: Severity Subgroup Analysis

Regression of the treatment by subject severity. In subsequent analyses, the model in equation (1) will be modified to include multiple baseline severity variables, such as NIHSS, IVH, extent of white matter disease, co-morbidity, etc. The models will be compared using Akaike Information Criterion (AIC) to assess the best combination of predictors. The discriminatory ability of the models will be compared using the cross-validated area under the Receiver Operating Characteristic (ROC) curve.

11.2.4. Additional Efficacy Analysis: Subgroup Analysis

Summary statistics will be presented for the analysis sets and subgroups of gender, race, and ethnicity. In addition, summary statistics will be examined for the subgroups of the patients who completed the study and by patients who discontinued the study early from the data available for these patients. The reasons for early discontinuation, including bleeding, loss to follow-up, and withdrawal of care, will be tabulated and recorded to help improve future recruitment and retention.

In addition, since data from the MISTIE phase 2 trial neither supports nor negates any significant difference in treatment efficacy by gender, race, or ethnicity, the analyses detailed in equations (1) to (4) above will be performed separately for males and females, by race (e.g., Caucasian, African

American, Asian, Native American, Pacific Islanders, etc.), and by ethnicity (e.g., Hispanic, non-Hispanic), where adequate data are available to determine if the treatment effect varies by these demographic categories, which would indicate the presence of effect modification by treatment. These sub-analyses are required by the NIH Policy and Guidelines on The Inclusion of Women and Minorities as Subjects in Clinical Research, amended October 2001. Note that the trial is not powered to see treatment effect modifications by these demographic categories.

11.2.5. Additional Efficacy Analysis: Random Effects (Site Differences) Analyses

We will conduct analyses based on mixed models that assume random effects and their variance-covariance matrix structure, including models without random intercepts, and possibly random slopes, for hospital sites.

12.0. Interim Efficacy Analyses

At the discretion of the DSMB, a formal interim efficacy analysis will occur after 375 subjects reach 180 days of follow-up. These analyses will be based on approximate O'Brien-Fleming stopping boundaries for efficacy (defined below).

In the MISTIE III design, subjects will be randomized to MIS+rt-PA versus medical management using a covariate-adaptive design described previously. The trial will initially enroll ICH patients until either a maximum of 500 eligible patients is accrued or the trial is stopped early for efficacy or futility using group sequential boundaries described below.

Analyses will be conducted after 375 and 500 patients, respectively, have 180-day post-stroke measurements. Efficacy stopping boundaries will be based on the error-spending (also called alpha-spending) approach using the power family with exponent rho=3, which approximates the O'Brien-Fleming boundaries. The (one-sided) alpha allocated to the analysis (called the interim analysis) at 375 participants is defined to be $alpha_1 = 0.025(375/500)^3 = 0.0105$. All remaining alpha, i.e., $(0.025\text{-}alpha_1)$ is spent at the analysis with 500 participants (called the final analysis).

13.0. Secondary Efficacy Outcomes Analyses

13.1. Secondary Efficacy Analysis 1: Dichotomized eGOS UGR-US vs. LS-Death at 365 days

Analyses will follow the methods as described in equations 1, 2, and 3 above (11.2.1–11.2.5). Subordinate analyses will be adjusted for other clinically meaningful variables, in addition to the ones used in the covariate-adaptive randomization.

13.1.1. Additional, Subordinate Efficacy Analysis 1: 365-Day Ordinal eGOS

Clot removal with MIS+rt-PA produces improvements in functional outcome assessed by alternative outcome measures such as the extended Glasgow Outcomes Scale (eGOS). Analyses: Ordinal, linear, logistic, and polytomous data models to examine other levels of the eGOS as part of Secondary Efficacy Analysis 1 will be similar to those presented above.

13.2. Secondary Efficacy Analysis 2: All-Cause Mortality Longitudinally from Ictus to 365 Days

Mortality at 365 days post ictus for MIS+rt-PA–treated subjects will be improved as compared to that of medical managed subjects. Comparison of treatment groups will be examined using logistic regression analysis.

Mortality by group will be graphically displayed with Kaplan-Meier curves, and the statistical association between treatment assignment and mortality will be assessed with the log-rank test. Multivariable Cox regression models will then be created to assess for treatment differences in the hazards of survival after adjusting for potential group imbalances in the covariates listed above. Finally, we will account for the effect of the decision to withdraw patient care on treatment differences in mortality.

As a secondary assessment, time to death within the one-year follow-up period will be compared between the MIS+rt-PA and medically managed groups adjusting for appropriate baseline covariates. Provided the model assumptions are met, a proportional hazards regression model may be used for the analysis as described above.

13.2.1. Subordinate, Secondary Efficacy Analysis 2: Predictors of Mortality Analysis

Additionally, survival models will be constructed to examine the mortality, morbidity, and ICU distributions for the MIS+rt-PA and medical management groups. Standard Cox proportional hazard models will initially be examined. These models are commonly written as:

$$h_i(t) = \lambda_0(t)\exp(\gamma X + \phi \cdot TRT)$$

where $\lambda_0(t)$ is a non-parametric baseline hazard function, γ is again a vector of regression coefficients related to the X covariates, and ϕ is the parameter for the treatment effect of EVD plus rt-PA on the hazard of death (mortality). This model may be extended by including time-dependent covariates and non-proportional hazards as deemed necessary through diagnostic checks. Observed mortality will also be compared with predicted mortality based on clinical presentation, overall, and by treatment group. Predicted mortality will be estimated. We will develop a severity model for predicting mortality based on the Tuhrim model.[2,7,8]

13.3. Secondary Efficacy Analysis 3: Relationship Between Clot Removal as an AUC that Estimates the Time-Weighted Clot Volume from Ictus to EOT as AUC Clot Exposure and Functional Outcome (Proportion 0-3 mRS)

EOT clot volume and/or percent EOT ICH reduction are related to mRS functional outcome regardless of treatment. Based on the experience of the CLEAR III trial, time-weighted average clot volumes will also be considered as a potential predictor of mRS outcome. The null hypothesis is that ICH clot reduction time course (time weighted clot volume, EOT clot volume, and/or percent EOT clot reduction) has no effect on functional outcome regardless of treatment.

In addition, since a “plateau” or “floor” effect of clot resolution over time is likely, subsequent models to statistically test for these fixed effects will consider clot resolution in non-linear terms, such as linear or restricted cubic splines and other non-linear functions. Placement of knots for the splines will be based on the visual inspection of the average ICH trajectory over time. In addition, when using restricted cubic splines, we will consider placing the knots at equally spaced percentiles of the “follow-up time” variable’s marginal distribution as recommended by Harrell.[9] Alternative methods that quantify the time-weighted clot volume will also be considered in this model.[7]

From the final, most parsimonious model, the empirical Bayes estimates of the random slopes will be computed to be used in the next phase. The second-phase analysis will be based on multivariable logistic regression models that include the estimated clot resolution trajectories over time, the covariates listed in equation (1) above, and the dichotomized mRS at 365 days post-stroke as the outcome.

13.4. Secondary Analysis 4: Patient Disposition

13.4.a. Secondary Efficacy Analysis 4a: Home Days Over 365 Days' Time from Ictus

Multivariable linear regression models will be built on this continuous outcome of time to home.

13.4.b. Secondary Efficacy Analysis 4b: Patient Location at 365 Days Post Ictus

Multivariable logistic regression models will be built on the dichotomized outcome of home plus rehabilitation locations versus other locations for patients 365 days post-ictus. Because variations in the more severe levels of mRS (4,5) and eGOS (US, LS, VS) do not fully account for disposition, it will be assessed by adjusting for all covariates. Where discrepant conclusions about severe disability might be drawn from mRS, eGOS, or final location, all three descriptors will be presented for the severe levels.

13.5. Secondary Analysis 5: 180-Day Efficacy

13.5.a. Secondary Analysis 5a: Dichotomized, Adjudicated, Cross-Sectional mRS 0-3 vs. 4-6

Analysis methods described in section 11.1 for the Primary Efficacy Analysis will be repeated using day 180 data.

13.5.b. Dichotomized, adjudicated eGOS UGR-US vs. LS-Death

Analysis methods described in section 13.1 for the Secondary Efficacy Analysis 1 will be repeated using day 180 data.

13.5.1. Additional, Secondary Efficacy Analysis 5: Ordinal, Adjudicated mRS (0-6)

Analysis methods described in section 11.1.1. for the Subordinate Analysis (1) for the Primary Efficacy Analysis will be repeated using day 180 data.

13.5.2. Additional, Secondary Efficacy Analysis 5: Ordinal, eGOS UGR-US vs. LS-Death

Analysis methods described in section 13.1.1. for the Subordinate Analysis for Secondary Efficacy Analysis 1 will be repeated using day 180 data.

13.5.3. Additional, Secondary Efficacy Analysis 5: Subordinate Analyses 2-5 as Defined under the Primary Efficacy Analysis

Analysis methods described in sections 11.2–11.5 for the Primary Efficacy Analysis will be repeated using day 180 data.

13.6. Secondary Efficacy Analysis 6: Type and Intensity of ICU Management

MIS+rt-PA treatment of ICH leads to decreased intensity of hospital care compared to medical management. This includes fewer hospital days, ICU days, decreased use or duration of ventilatory management, decreased intensity of ICP management, decreased BP support, decreased osmotherapy, shorter periods of CSF drainage, lower utilization of ventriculoperitoneal shunts, and fewer general critical care complications. Analyses: Linear, log-linear, and logistic data models will be similar to those presented above.

13.7. Secondary Efficacy Analysis 7: EQ-VAS and EQ-5D

Clot removal with MIS+rt-PA leads to improved health-related QOL and will be assessed by subject and surrogate QoL domains through the EQ-VAS and EQ-5D.

14.0. Safety Analyses

14.1. Mortality and Safety Events at 30 Days Post Randomization

The primary safety measures for this study are symptomatic intracranial bleeding, infection, and 30-day mortality. The safety thresholds for these measures have been set as: 30-day mortality >60%, a MIS+rt-PA–related symptomatic bleeding rate >25% (occurring during active treatment or during the 72 hours after treatment will be monitored), first-week operative death rate >10%, and a procedure-related infection rate >15% (over the initial 30 days). Bleeding will be further characterized as during the exposure to dosing of alteplase and the subsequent 72 hours. Comparisons to medical subjects in the same time frame will be made. Total SAEs at 30 and 180 days, as well as summary of AEs and SAEs by MedDRA code and grouped by organ system, will be compared by treatment groups.

Statistical Hypothesis: Early use of MIS+rt-PA for 3 days for the treatment of ICH is as safe as medically managed ICH, measured by rates of procedure-related mortality, rebleeding, and infection within 30 days post-randomization. The null hypothesis is: *Early use of MIS+rt-PA is worse than medical management for a specific safety measure.*

Analysis: The number of SAEs will be tabulated by SAE type post-stroke for the total population as well as by treatment group, with particular emphasis placed on group differences at 30 days post-randomization. Any treatment group differences in SAEs enumerated for interim analyses will remain blinded to investigators associated with this study. Unmasked results (e.g., group A versus B, where the labels “A” and “B” are randomly assigned to the surgical and medically managed groups) will be made available only during the closed sessions of DSMB meetings.

Methods of non-inferiority will be used to statistically determine if the proportions of a particular type of SAE, where there are adequate data for a particular type, are not substantially different among both treatment groups in order to test the hypothesis that MIS+rt-PA is as safe as standard medical management. Non-inferiority testing deems the rates of any SAE event in the MIS+rt-PA group to not be substantially greater than the SAE event rate in the standard medical management group as long as the group difference in this rate is statistically within some pre-specified tolerance. For this trial, the tolerance value will be defined as 10% for the primary adverse events of 30-day mortality, rebleeding, and infection, and 5% for first-week operative mortality. Thus, the rates of adverse events seen in the MIS group can be higher than those in the medically managed group, and yet still be considered not substantially different as long as the upper one-sided 95% CIs on rate differences are ≤10% or ≤5%, depending on the safety measure.

Since data from the MISTIE phase 2 trial neither supports nor negates any significant difference in SAE event rates by gender, race, or ethnicity, the non-inferiority analyses described above will be performed separately for subpopulations within each of these demographics where there are adequate data available.

14.2. Total SAEs at 30 Days

Mortality attributed to MIS+rt-PA treatment plus disease-associated adverse events are similar to the morbidity and mortality attributed to medical management in the first 30 days. Analyses: For initial evaluations, differences between the MIS+rt-PA intervention and medical management groups will be examined using a logistic regression model.

14.3. Summary of Adverse Events and Serious Adverse Events by MedDRA Code and Grouped by Organ System

All AEs and SAEs are summarized by type and by treatment group in terms of frequency of the event, number of subjects having the event, timing relative to randomization, severity (mild, moderate, severe), and relatedness to the study treatment (definitely, probably, possibly, definitely not). At the end of the study, the cumulative incidences of these events will be compared between the two treatment groups using Fisher's exact test. Statistically significant differences between the treatment arms will be assessed after adjustments for multiplicity in the case of individual SAE/AE diagnoses and MedDRA codes. Additionally, generalized linear mixed models for binary data will be used to examine AE and SAE probabilities between treatment groups while accounting for potential confounders and site-clustering effects.

14.4. Proposed Interval Safety Monitoring

The DCC will generate periodic DSMB reports. (We assume the total number of subjects enrolled per six months in a single treatment group will be approximately 30; thus, eight semi-annual looks at the data will be suggested to the DSMB. As usual, though, periods for reporting will be discussed and finalized at the initial DSMB meeting.) Each report will provide cumulative summary statistics on enrollment; subject status in the study (e.g., number completed at 3-month, 6-month, 9-month, and 12-month assessments); baseline characteristics; protocol violations; safety data, including AEs and SAEs by AE code, severity, and relatedness to the study medication; outcomes data (if coinciding with an interim analysis—also to be discussed and finalized at the initial DSMB meeting); and data management/quality information (e.g., timeliness and completeness of data entry by the clinical sites via the MISTIE III electronic data capture (EDC) system; number of queries generated and resolved). Two reports will be generated: an open report with combined data from both treatment groups, and a closed report for the DSMB voting committee members with data provided by blinded treatment group (noted as A and B). The outline of these proposed DSMB reports will be included in the Manual of Operations and Procedures prepared by the CCC. In an interval as yet to be determined (monthly, quarterly, or semi-annually), the CCC will also generate a Safety Monitoring Report to be distributed to the DSMB. This report contains only enrollment, subject study status, safety, and data quality information. The trial's Executive Committee will also receive the Safety Monitoring Reports. The DSMB will have discretion to recommend stopping the trial early if safety concerns become substantial. Safety stopping rules for each of the primary safety outcomes will be developed and used to help the DSMB make its safety assessments.

15.0. Tertiary (Exploratory/Sensitivity) Analyses

15.1. Functional Status at 180 and 365 Days Post Ictus

Clot removal with MIS+rt-PA treatment produces improvements in functional outcome will be assessed by alternative outcome measures such as:

15.1.a. Glasgow Outcome Scale
15.1.b. Barthel Index
15.1.c. NIHSS
15.1.d. MMSE

Analysis methods will include linear, logistic, and polytomous data models similar to those presented above.

15.2. Quality of Life at 180 and 365 Days Post Ictus

Clot removal with MIS+rt-PA leads to improved health-related QoL and will be assessed by subject and surrogate QoL domains through the:

15.2.a. Stroke Impact Scale
15.2.b. PBSI
15.2.c. Personal Health Utility Assessment Interview
15.2.d. CES-D
15.2.e. EQ dimensions individually assessed
15.2.f. Home days
15.2.g. Patient location including proportion of days in acute care, home, rehabilitation, long-term care facilities, and in vegetative condition.

Analyses methods will include linear, logistic, and polytomous data models similar to those presented above.

15.3 Cost-Benefit Model

Cost-benefit analyses will be performed using data on median hospital stay, patient disposition categories (such as home, long-term ambulatory and nursing care, death), time to return home, and cost of procedures and hospital stay. These analyses will be performed with respect to both 180- and 365-day periods post-stroke. For cost-benefit analyses, the cost of ICU and the rest of the hospital stay, including the cost of MIS+rt-PA, will be considered. Cost data will be obtained from the literature. The overall amount for both treatments will be assessed, and the difference in dollars will be calculated.

16.0. References

1. Hanley DF, Thompson RE, Muschelli J, et al. Safety and efficacy of minimally invasive surgery plus alteplase in intracerebral haemorrhage evacuation (MISTIE): a randomised, controlled, open-label, phase 2 trial. Lancet Neurol 2016;15(12):1228-1237.
2. Hanley DF, Lane K, McBee N, et al. Thrombolytic removal of intraventricular haemorrhage in treatment of severe stroke: results of the randomised, multicentre, multiregion, placebo-controlled CLEAR III trial. The Lancet 2017;389:603-11.
3. van der Laan MJ, Gruber S. Targeted minimum loss based estimation of causal effects of multiple time point interventions. Int J Biostat 2012;8(1):1370.
4. Permutt T. Sensitivity analysis for missing data in regulatory submissions. Statistics in medicine 2016;35:2876-9.
5. Diciccio TJ, Efron B, Hall P, et al. Bootstrap confidence intervals. Statistical Science 1996;11(3):189-228.
6. National Research Council Panel on Handling Missing Data in Clinical Trials. The prevention and treatment of missing data in clinical trials. Washington, DC: National Academies Press (US); 2010.
7. Tuhrim S. Intracerebral hemorrhage—improving outcome by reducing volume? N Engl J Med 2008;358(20):2174–6.
8. Tuhrim S, Horowitz DR, Sacher M, Godbold JH. Volume of ventricular blood is an important determinant of outcome in supratentorial intracerebral hemorrhage. Crit Care Med 1999;27:617-621.
9. Harrell F. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2001.