STATISTICAL ANALYSIS PLAN

MARVEL： Methylprednisolone for acute large vessel occlusion: a randomized double-blind, placebo-controlled trial in revascularization patients

Principal Investigators

Wenjie Zi, MD and Qingwu Yang, PhD., MD

Xinqiao Hospital, Army Medical University, Chongqing, China

Supplement 2 – Statistical Analysis Plan

This supplement contains the following items:

STATISTICAL ANALYSIS PLAN

MARVEL： Methylprednisolone for acute large vessel occlusion: a randomized double-blind, placebo-controlled trial in revascularization patients

Principal Investigators

Wenjie Zi, MD and Qingwu Yang, PhD., MD

Xinqiao Hospital, Army Medical University, Chongqing, China

SAP Authors

Wenjie Zi, MD, PhD; Changwei Guo, MD; Xinqiao Hospital, Army Medical University, Chongqing, China

Duolao Wang, PhD, Global Health Trials Unit, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

Mengjie Lu, PhD, School of Public Health, Shanghai JiaoTong University School of Medicine, Shanghai, China

CONFIDENTIAL

SAP version history		
Version Date	SAP Version #	Details of Changes
2021/09/29	*1.0*	

TABLE OF CONTENTS

Abbreviations and Definitions of Terms

AE	Adverse Event
AIS	Acute ischemic stroke
ASPECTS	Alberta Stroke Program Early CT Score
CI	Confidence Interval
CRF	Case Report Form
CT	Computed Tomography
CTA	CT angiography
DSA	digital subtraction angiography
DSMB	Data Safety Monitoring Board
EQ-5D-5L	European Quality of Life Five-Dimension Five-Level scale
EVT	Endovascular Therapy
GCP	Good Clinical Practice
GEE	Generalized Estimating Equation
GLM	Generalized Linear Model
ICH	Intracranial Haemorrhage
ICH-E9	International Conference on Harmonization - Statistical principles for clinical trials
ITT	Intention-to-Treat
IVT	Intravenous Thrombolysis
MR	Magnetic Resonance
MRA	MR angiography
mRS	Modified Rankin Scale
NIHSS	National Institute of Health Stroke Scale
NMPA	National Medical Products Administration
PI	Principal Investigator
PP	Per-protocol
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan

SICH	Symptomatic Intracranial Haemorrhage
SOC	System Organ Class
TIA	Transient Ischemic Attack
tPA	Tissue Plasminogen Activator

1. PREFACE

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for protocol ChiCTR2100051729 entitled “Methylprednisolone for acute large vessel occlusion: a randomized double-blind, placebo-controlled trial in revascularization patients” (Version 1.0, 2022/9/29).

This study is being completed to assess the efficacy and safety of adjunctive intravenous methylprednisolone compared with placebo for the treatment of acute ischemic stroke (AIS) in patients with large vessel occlusion (LVO) underwent endovascular treatment (EVT).

The structure and content of this SAP provides sufficient details to meet the requirements determined by the National Medical Products Administration (NMPA) and International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH-E9): Guidance on Statistical Principles in Clinical Trials[1]. All work planned and reported for this SAP will follow internationally recognized guidelines, published by the American Statistical Association[2] and the Royal Statistical Society[3], for statistical practice.

In preparing this SAP, the following documents were reviewed:

- Clinical Study Protocol Version
- Case report forms (CRFs).
- ICH-E9 Guidance on Statistical Principles for Clinical Trials.
- Guidelines for the Content of Statistical Analysis Plans in Clinical Trials

Readers of this SAP are encouraged to also read the clinical trial protocol to understand the implementation details of this study, the operational aspects of clinical evaluation, and the schedule of patients completing this study.

2. PURPOSE OF SAP

The purpose of this SAP is to outline the planned analyses to be completed to support the completion of the clinical study report. The planned analyses identified in this SAP will be included in regulatory submissions and/or future manuscripts. In addition, exploratory analyses not necessarily identified in this SAP may be conducted to support the clinical development plan. Any post-hoc, or unplanned, analyses outside of this SAP will be clearly identified in the respective clinical study report.

3. VALIDATION OF SAP AND AMENDMENTS

The SAP should be validated and signed by:

Principal investigator(s)

Chair of Steering Committee

Chair of Data Safety and Monitoring Committee

Senior statistician

After validation, the SAP may be amended, and all changes will be clearly tracked. Any amendments can occur only before the database is locked and the treatment code is unblinded.

4. STUDY OBJECTIVES AND ENDPOINTS

4.1 Study Objectives

The primary objective is to test the efficacy and safety of adjunctive intravenous methylprednisolone compared with placebo for the treatment of acute ischemic stroke in patients with large vessel occlusion underwent endovascular treatment.

4.2 Study Endpoints (Target Variables)

4.2.1 Primary Efficacy and Safety endpoints

4.2.1.1 Primary Efficacy Endpoint

The primary efficacy outcome measure for this study is: The distribution of the global disability defined as a modified Rankin Scale (mRS) score at Day 90 (±7) after randomization.

4.2.1.2 Primary Safety Endpoints

The primary safety outcome measure for this study are as follows:

- Mortality due to any cause within 90 (±7) days after randomization
- Proportion of patients with symptomatic intracranial haemorrhage within 48 h after treatment. Symptomatic intracranial haemorrhage will be adjudicated by an independent Imaging Core Laboratory according to the modified Heidelberg Bleeding Classification, as specified in the trial protocol. SICH defined as a new ICH seen on imaging will be later confirmed on independent reading by blinded central study radiologists

4.2.2 Secondary Efficacy and safety endpoints

4.2.2.1 Secondary Efficacy Endpoints

The secondary efficacy outcomes are as follows:

The secondary endpoints include:

1) proportion of mRS score 0 to 4 at 90 days;

2) proportion of mRS score 0 to 3 at 90 days

3) proportion of mRS score 0 to 2 at 90 days;

4) proportion of mRS score 0 to 1 or return to pre-morbid mRS score at 90 days (for patients with mRS >1) ;

5) NIHSS score at 5-7 days after EVT or at early discharge;

6) European Quality Five-Dimension scale score at 90 days.

4.2.2.2 Secondary Safety Endpoints

The secondary safety outcomes measure for this study are as follows:

- Proportion of patients with any radiologic intracranial haemorrhage (based on central neuroimaging reading) within 48 hours after receiving intravenous study drug;
- Proportion of patients with pulmonary infection;
- Proportion of patients with gastrointestinal bleeding within 7 days after EVT;
- Proportion of patients with any adverse events;
- Proportion of patients with serious adverse events;

5. STUDY METHODS

5.1 Overall Study Design and Plan

MARVEL is a multi-center, randomised, double-blinded, clinical trial. Patients will be randomized 1:1 to either intravenous methylprednisolone or intravenous placebo. It is planned to recruit at least 1672 patients in about 80 sites in China. The follow-up period covers the period from the subacute stage (day 5-7 or hospital discharge) to the final follow-up examination 83-97 days after randomization which defines the primary endpoint. The study patient flow outline following Consort diagram is shown in Figure 1.

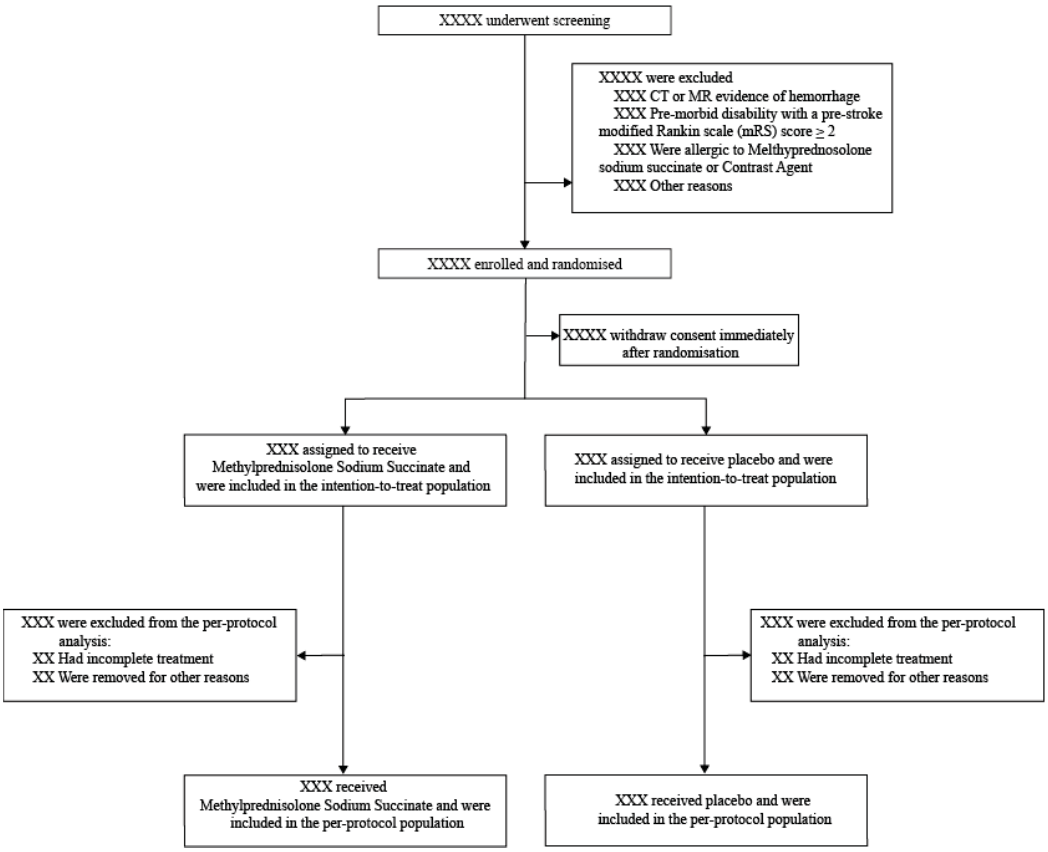


Figure 1. The study patient flow outline following Consort diagram.

5.2 Selection of Study Population

5.2.1 Inclusion Criteria

- Aged 18 years or older;
- The time from onset to randomization was within 24 hours;
- Baseline National Institutes of Health Stroke Scale (NIHSS) ⩾ 6
- Anterior circulation ischemic stroke was preliminarily determined according to clinical symptoms or imaging examination;
- Baseline Alberta Stroke Program Early CT Score (ASPECTS) ⩾ 5;

- Computed tomography angiography (CTA) /magnetic resonance angiography MRA) /digital subtraction angiography (DSA) confirmed occlusion of intracranial segment of internal carotid artery and middle cerebral artery, basilar artery and endovascular therapy was decided;
- Written informed consent signed by patients or their family members.

5.2.2 Exclusion Criteria

Patients meeting any of the following criteria will be excluded from study enrolment.

- Intracranial hemorrhage confirmed by cranial computed tomography (CT) or magnetic resonance imaging (MRI);
- mRS score ⩾ 2 before onset;
- Pregnant or lactating women;
- Allergic to contrast agents;
- Allergic to glucocorticoids;
- Participating in other clinical trials;
- Systolic blood pressure > 185 mmHg or diastolic pressure > 110 mmHg, and oral antihypertensive drugs can not control;
- Genetic or acquired bleeding constitution, lack of anticoagulant factors; Or oral anticoagulants and INR > 1.7;
- Blood sugar < 2.8 mmol/L (50 mg/dl) or > 22.2 mmol/L (400 mg/dl), platelet < 90 x 10^9/L;
- The artery is tortuous so that the thrombectomy device cannot reach the target vessel;
- Bleeding history (gastrointestinal and urinary tract bleeding) in recent 1 month;

- Chronic hemodialysis and severe renal insufficiency (glomerular filtration rate < 30 ml/min or serum creatinine > 220 umol/L [2.5 mg/ dL]);
- Life expectancy due to any advanced disease < 6 months;
- Follow-up is not expected to be completed;
- Intracranial aneurysm and arteriovenous malformation;
- Brain tumors with imaging mass effect;
- Severe systemic infectious diseases.

Inclusion and exclusion criteria will be assessed based on information available during the screening period.

5.3 Method of Treatment Assignment and Randomization

Randomization will be done by a web-based APP (Jinlingshu) on mobile phone or computer (https://jinlingshu.com/). The automated system will assign an appropriate set of study medication to each patient. Eligible patients are randomized to treatment with either intravenous methylprednisolone sodium succinate or intravenous placebo with a ratio of 1:1. Randomization will be stratified by participating centre with permutation block size of 4. Randomization will be completely concealed by having both web-based real-time allocation and identical appearance of methylprednisolone sodium succinate and placebo bottles. (All bottles will have a unique number. Subjects will be assigned a random serial number according to the time they were enrolled, and corresponding masked medications will be provided).

The randomization list will be prepared by the independent statistical centre using SAS 9.4. It will only be sent to the data coordination centre responsible for central randomization and data management, the pharmacy centre responsible for treatment packaging, and the centre responsible for blinding emergency patients.

5.4 Treatment Masking (Blinding)

All trial personnel (investigators, their clinical staff, and the data management group) and patients will be blinded to treatment assignment. The Data Safety Monitoring Board (DSMB) will have access to masked data.

In addition, the person in charge of the data management group responsible for managing the programming of the randomized system will be unblinded. If it is necessary to unblind, this person will become the contact person. This person will not participate in data management and will only communicate to unblind when contacted by the medical monitor.

If the treatment code needs to be broken for the safety of the patient, the site principal investigator (PI) will call the medical monitor for all unblinding queries. The medical monitor will discuss with the researcher whether it is necessary to unblind the patient. If the medical monitor determines that the PI should be unblinded for the patient, the medical monitor will contact the person in charge of the data management group by e-mail or phone, instructing the unblinding of the specific patient. Members of the unblinding data management team will provide only to the site PI with allocation information, the unblinding date, site number, PI name, and patient number via email. Any cases unblinded in this way will be recorded in the central file. Only the physician who requests the unblinding will receive the unblinding message. Study drugs will be stopped afterwards. It is not expected that there will be any clinical cases that need to be unblinded. The randomization data will be strictly confidential, and only authorized personnel can access it until the database is locked.

5.5 Contents of Investigational Product Kit

Methylprednisolone sodium succinate and its placebo are manufactured and provided by Lummy Pharmaceutical Group Co., Ltd., Chongqing, China. The medication packages are visually identical (including labelling, dosage form, size, and colour), except for a unique number. Methylprednisolone sodium succinate and its placebo will be packed in sterile, disposable, individually labelled bottles. Each kit has a unique six-digit identification number and will be stored in a safe location at room temperature on

the clinical site with restricted access. All study medication will be manufactured, tested, released, packaged, labelled, and shipped in accordance with Good Manufacturing Practice, Good Clinical Practice (GCP) guidelines, and any national regulatory requirements.

6. FORMAL ANALYSES AND REPORTING

All formal planned analyses identified in the protocol and in this SAP will be performed only after the formal database lock. Before database lock and completion of the final analyses, a blinded data review meeting will be held. Also, no database may be locked, random code unblinded, or analyses completed until this SAP has been approved and signed.

Any post-hoc, exploratory analyses completed to support planned study analyses, which were not identified in this SAP, will be documented and reported in appendices. Any results of these unplanned analyses will also be clearly identified in appendices.

7. SAMPLE SIZE DETERMINATION

The current trial is designed detect a shift on the modified Rankin Scale score, which represents the global disability, to a lower score. The expected distribution of the modified Rankin Scale score, derived from the EVT for acute anterior circulation ischemic stroke (ACTUAL) registry[4] and the Highly Effective Reperfusion evaluated in Multiple Endovascular Stroke Trials (HERMES) collaboration[5], is as follows: 7%, 20%, 16%, 11%, 13%, 8% and 25%.

The sample size was calculated using the cumulative proportion of patients with mRS 0-2. We assume a moderate effect of 7% absolute increase in the cumulative proportion of patients with mRS 0-2 in the intervention group, compared with controls, indicating an odds ratio (OR) of 1.33, which have substantially exceeds the minimal clinically important difference.[6,7] In order to demonstrate the expected

treatment effect with a type-1 error alpha =0.05 (two-tailed) and a power of 80% (beta = 20%). A sample size of n= 1588 patients (n=794 per treatment group) is required.

The intention-to-treat principle will be applied to the primary analysis and, therefore, to safeguard against dilution of the treatment effect associated with an approximate 5% non-adherence rate (due to loss to follow-up, consent withdrawal and other reasons), we initially planned to enrol n=1672 patients (n=836 per treatment group) for this study. This estimation was performed based on PASS (NCSS, LLC. Kaysville, Utah, USA) version 15.0.

8. ANALYSIS POPULATIONS

The following analysis populations are planned for the studies:

8.1 Intention-to-Treat (ITT) Population

The ITT population includes all patients randomized into the trial and who were recorded as receiving any amount of study drug, even if the subject does not receive the correct treatment, or does not follow the protocol until completion[8]. The ITT population will be the primary analysis population for the efficacy endpoints and subjects will be analysed according to the treatment group to which they were assigned at randomization.

The ITT analysis strategy[8] for MARVEL is defined as follows:

- Is based on an ITT design that aims to collect all outcome data on all randomized subjects;
- Includes a main analysis that keeps subjects in their randomized groups, analyses all available outcome data, and is valid under a named plausible assumption about the missing data;
- Includes sensitivity analyses that consider a range of plausible alternative assumptions about the missing data;

- All randomised individuals are included in sensitivity analyses.

8.2 Per-Protocol (PP) Population

The PP population is defined as the subset of the ITT population excluding major protocol violators deemed to have the potential to affect patient outcome in terms of efficacy.

The PP population includes patients who actually received the assigned treatment and do not have major protocol violations or deviations. Major protocol violations or deviations will be identified in a blinded fashion prior to database lock. More specifically, patients with anyone of the following criteria will be excluded from the PP population. These deviations will be determined based on the medical monitors' records, as well as programmatically, using the following criteria at a minimum:

- Received but did not complete treatment with study drug, or dose of study drug administered outside recommended dose.
- Violate inclusion or exclusion criteria.
- Missing essential information.

A list of patients to be excluded from the randomized patients to create the PP-Efficacy analysis will be established and validated by the Steering Committee prior to unblinding.

8.3 Safety Population

The Safety Population includes all patients who received any amount of study drug. In case of violation of the randomization scheme, patients will be classified according to the treatment they actually received. Patients will be assigned to the different populations prior to unblinding of the database. Patients who withdraw informed consent immediately after randomization and are not to receive any treatment should be excluded from Safety Populations.

9. GENERAL ISSUES FOR STATISTICAL ANALYSIS

9.1 Analysis Software

All statistical analyses will be performed using SAS® Software version 9.4 in a Windows environment[9] and R version 4.1.1.

9.2 Methods for Withdrawals and Missing Data

9.2.1 Withdrawals

If any subject withdraws from the study prematurely (before the final visit day 90±7 for evaluation), they are required to complete the withdrawal visit in the Case Report Form. The reasons for withdrawal will be listed in a summary table. Subjects who withdrew before the subject's last follow up (90 ± 7 days) will be included in the analysis. Subjects who do not receive any study drugs after randomization will be excluded from any end-point analysis because no post-baseline data are available.

9.2.2 Missing data

The handling of missing data will follow the principles specified in the ICH-E9[1] and the CPMP/EWP/1776/99Rev1. Guideline on Missing Data in confirmatory trials Guidelines.[10] All efforts will be made to minimize the amount of missing data, particularly the 90-day outcome assessment. After the patients are randomized, the patients or their families and doctors will keep WeChat (a Chinese social networking software similar to Facebook) and phone numbers for each other to avoid loss to follow-up.

9.2.2.1 Baseline covariates

Missing baseline covariates will be imputed using simple imputation methods in the covariate adjusted analysis based on the covariate distributions, should the missing values for a particular covariate be less than 5%. For a continuous variable, missing

values will be imputed from random values from a normal distribution with mean and SD calculated from the available sample. For a categorical variable, missing values will be imputed from random values from a uniform distribution with probabilities P_1, P_2, …, and P_k from the sample.

If the missing values for a covariate are ⩾5% then they will be imputed using Markov chain Monte Carlo (MCMC) methods using SAS PROC MI. A total of 10 multiply imputed dataset will be generated. The seed for the imputations will be 128.

9.2.2.2 Efficacy outcomes

Sensitivity analyses based on different hypotheses about the missingness pattern of the primary outcome will be performed to test for the robustness of the primary analysis results[11]. Primary analysis of the primary outcome will be conducted in the complete case of the ITT population (in which patients with missing data will be censored).

9.3 Data Transformations

Data transformation is not relevant for most of the primary and secondary endpoints, which are categorical variables. The analysis of continuous variables may require transformation to normalize the distributions, but these variables will be mainly analysed using nonparametric statistical tests. Decisions to transform the distribution of some variables for analysis may be taken after the blind review of the data, and before unblinding the treatment code.

9.4 Multicenter Studies

No adjustment on centre will be performed for the primary analysis, on the basis of the following rationale:

The MARVEL trial aims to recruit 1672 patients from about 80 stroke centres in China, about 24 patients per centre. However, it is likely that many centres will recruit far fewer patients, while a few large centres will recruit far more patients. This situation is

the argument for not adjusted the primary analysis on centres. As the International Conference on Harmonization 9 guidelines quotes:

"In some trials, for example some large mortality trials with very few subjects per centre, there may be no reason to expect the centres to have any influence on the primary or secondary variables because they are unlikely to represent influences of clinical importance."

However, the central effect will be considered in the sensitivity analysis together with other important prognostic covariates.

9.5 Multiple Comparisons and Multiplicity

Analyses of secondary safety outcomes and additional analyses for the primary outcome are regarded as exploratory in nature, therefore, multiplicity adjustment will not apply to the primary and secondary outcome analyses.

9.6 Covariates

In the covariates-adjusted analysis, the following variables will be adjusted using the IPTW method.

- Age
- Baseline NIHSS score
- Pre-stroke modified Rankin Scale score
- Baseline aspects score
- Intravenous Thrombolysis
- Time from stroke onset to randomization
- Occlusion location

The same IPTW weight will be used for all adjusted analyses.

9.7 Planned Subgroups

Subgroup analyses could be carried out irrespective of whether there is a significant

treatment effect on the outcome. Their purpose is to supplement evidence from the primary analysis to help to fully characterize the treatment effect. Results from subgroup analyses will be interpreted in this context.

Subgroup analyses will be performed on a few important patient characteristics and other variables of interest. Pre-planned subgroup analyses will explore the effects of:

- Age
- Sex
- Baseline NIHSS score
- Pre-stroke mRS
- Baseline ASPECTS score
- Intravenous thrombolysis
- Time from last known well to randomization. (mins)
- Occlusion location
- Complete Reperfusion defined as Extended Thrombolysis In Cerebral Infarction grade 2c or 3
- Patients with any radiological haemorrhages
- Patients with symptomatic haemorrhages

Subgroup analyses will be conducted irrespective of whether there is a significant treatment effect on the outcome. The purpose is to supplement evidence from the primary analysis to help to fully characterize the treatment effect. These analyses are exploratory and will not lead to formal inferential conclusions about the treatment effect in any particular subgroup.

A forest plot will be used to present the treatment effect of methylprednisolone for each of the individual subgroups. The plot will show the cOR described above with two-sided 95%CIs, corresponding to each level of these subgroups. Separately, the adjusted

estimates for each of the patient characteristics defining a subgroup will be presented from the multivariate model.

9.8 Derived and Computed Variables

Some derived and computed variables have been initially identified and described as appropriate later in the document. It is expected that additional variables will be required. The SAP will not be amended for additional variables that are not related to the primary target or key secondary target variables.

9.9 Presentation of results

Description of continuous variables includes the mean, standard deviation, median, first and third quartiles, and range. The description of categorical variables includes the frequency and percentage n (%). The description of ordered variables includes the frequency, percentage and cumulative percentage.

10. DISPOSITION OF SUBJECTS AND WITHDRAWALS

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

- Total number of assessed patients
- Number of randomized patients
- Number of patients completing the study and not completing the study (grouped by treatment and main reason)
- Number of patients included in the ITT population
- Number of patients included in the PP population

11. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Clinically important demographic and disease characteristics at screening will be summarized with descriptive statistics for each treatment group. Summaries will be produced for the ITT population and PP population.

If the missing value exceeds 5%, the denominator will be added in the footnote of the corresponding summary table.

The variables to be summarized include but are not restricted to the following:

Age

Sex

Medical History (hypertension, coronary heart disease, smoking, diabetes mellitus, hyperlipidaemia, and prior stroke)

Pre stroke onset mRS score

Baseline NIHSS score

Baseline ASPECTS Score

Systolic and Diastolic Blood Pressure

Stroke Aetiology

Serum Glucose

Occlusion Sites

Intravenous Thrombolysis

ASITN/SIR

Time From Stroke Onset (Last Known Well) to Randomization.

Time From Stroke Onset (Last Known Well) to Puncture

Time From Puncture to Reperfusion or the End of the Operation

Time From Randomization to Initiation of study drugs.

12. EFFICACY ANALYSES

All efficacy analyses will provide the point estimate of treatment effect with associated two-sided 95% confidence intervals (CIs). Statistical analyses of the primary and

secondary efficacy endpoints will primarily be based on the ITT population. Secondarily, the same analyses will be repeated on the PP population.

Assessment of efficacy will be conducted based on endpoints collected during the Day 90 follow-up. The protocol allows this visit to be scheduled within 7 days of the target 90 days after randomization.

12.1 Primary Efficacy Endpoint

The mRS score is an assessment of disability with values from 0 (no symptoms at all) to 5 (sever disability) to 6 (death). The mRS score is collected at 90 days after randomization. If death occurs prior to an assessment day (Day 90), the mRS score will be considered available and will be set to 6.

As the primary endpoint, mRS score at Day 90 will be further explored on the range of the values (0 to 6). The distribution of full range of mRS scores at Day 90 will be summarized by treatment group. Both tabular displays and stacked bar charts will be prepared. Win ratio method will be used to calculate the win ratio statistic as a measurement of treatment effect in mRS score.[12] Primary analysis of the primary outcome will be conducted in the complete case of the ITT population (in which patients with missing data will be deleted) using the proportional odds model.

12.1.1 Hypothesis

The null and alternative hypotheses to be tested appear below:

The null hypothesis for proportional odds logistic regression is that the common OR is equal to one, which needs to be disproved. The null hypothesis for WMW test is the equality of ranks when ties are split evenly. This can be rephrased as both probabilities Prob ($Y_M > Y_P$) + 0.5Prob ($Y_M = Y_P$) and Prob ($Y_M < Y_P$) + 0.5Prob ($Y_M = Y_P$) being equal to 0.5, resulting in WMW GenOR being equal to 1. Therefore, the null hypothesis for the WMW test states that there's an equal probability of treatment observation being worse or better than control observation (with equally splitting ties), meaning that

WMW GenOR isn't different from one.

12.1.2 Crude Analysis

The categorical shift in the distribution of mRS scores between the two treatment groups will be analysed fitting a proportional-odds logistic regression model, assuming a common odds ratio across all cut points of the modified Rankin scale.

a) If the proportional odds assumptions are satisfied (both Brant test and Approximate likelihood-ratio test of proportionality of odds are not significant), common OR and 95%CI will be calculate from an GLM with multinomial distribution and cumlogit-link(ordinal logistic regression model) [13,14];

b) If the proportional odds assumptions are not satisfied for the crude analysis (significant Brant test or Approximate likelihood-ratio test of proportionality of odds), assumption-free method will be used.[15]

12.1.3 Covariates Adjusted Analysis

a) If the proportional odds assumptions are satisfied for crude analysis, and still satisfied when adding the covariates in the ordinal logistic regression model, then the common odds ratio from an ordinal logistic regression model, IPTW method will be employed in the co-variate adjusted analysis. [13,14,16]

b) If the proportional odds assumptions are not satisfied for the IPTW adjusted analysis (significant Brant test or Approximate likelihood-ratio test of proportionality of odds), assumption-free method will be used.[15]

Covariate adjusted analysis will be performed on both ITT and PP populations.

Main conclusion will be drawn from the adjusted ITT analysis of the primary endpoint.

12.1.4 Subgroup analysis

Subgroup analysis of the primary endpoint will be performed using modified Poisson regression model on the following subgroup variables:

- age
- baseline NIHSS score
- pre-stroke mRS
- baseline ASPECTS score
- intravenous thrombolysis (yes vs. no)
- time from last known well to randomization. (mins)
- occlusion location
- sex
- pre-stroke mRS
- Complete Reperfusion defined as Extended Thrombolysis in Cerebral Infarction grade 2c or 3
- Patients with any radiological haemorrhages
- Patients with symptomatic haemorrhages

Assessment of the homogeneity of treatment effect by a subgroup variable will be conducted by a GLM with the treatment, subgroup variable, and their interaction term as predictors, adjusted for age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location, and the *P*-value from the interaction term will be used to assess exploratorily homogeneity of treatment effect by a subgroup variable.

Subgroup analysis will be performed on both ITT and PP populations.

12.1.5 Sensitivity analysis

To assess the influence of the missing primary endpoints on the treatment effect estimate, sensitivity analyses will be performed using the same statistical methods as described in Section 12.1.1 and 12.1.2, under the different assumptions of the missing

outcome below:

(1) Worst-case scenario:

All patients with a missing primary endpoint will be considered as having the worst mRS score of 6.

(2) Best-case scenario:

All patients with a missing primary endpoint will be considered as having the best mRS score of 0.

(3) Multiple Imputation

Missing mRS score will be imputed using multiple imputation method via SAS PROC MI. Imputation procedure will be performed under the missing-at-random (MAR) assumption. Fully conditional method (FCS) will be used to predict missing mRS score by fitting an ordered logistic regression model with following predictors: treatment group, age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location, and last observed NIHSS if available.[17] The seed will be 128. A total of 10 imputed datasets will be generated. cOR will be estimated for each dataset separately, and the pooled together to calculate the final pooled cOR and 95% confidence interval using the Robin's role.

(4) Mixed-effect model: A general linear mixed effect model will be fitted to estimate the common odds ratio between the methylprednisolone sodium succinate and placebo treatment with treatment as fixed effect, age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location as covariates, and the centre as a random effect. The results will be calculated for each dataset after multiple imputation and combined following the Rubin's Rule.

12.2 Secondary Efficacy Analysis

12.2.1 Dicotomized modified Rankin Scale score.

Secondary efficacy outcomes include mRS 0-4, mRS 0-3, mRS0-2, mRS 0-1(or return to pre-morbid mRS score) and EQ-5D at Day 90. These results will further describe the treatment effect.

The treatment effect will be estimated in the ITT population using a GLM with Poisson distribution and log-link function to determine the risk ratio (RR) and its 95%CI.[18] In case of nonconvergence of this model, the following generalized linear models (GLMs) will be fitted sequentially until a model is converged:

- GLM with Binomial distribution and log link function;
- GLM with Negative Binomial distribution and log link function;
- GLM with Binomial distribution and logit link function, from which odds ratio (OR) will be converted into RR.[16,19]

In addition, risk difference and OR will be estimated and presented using two GLMs (identity-binomial regression and logit-binomial regression) as supportive measurements of treatment effect.

Adjusted analyses will be carried out on the primary efficacy outcome to determine whether the treatment effect estimate is affected with the inclusion of covariates at baseline. The covariates that will be included in the adjusted analyses are listed and described in Section 9.6. Imputation for baseline missing covariates (see description in the Section 9.2

12.2.2 NIHSS score at 5-7 days or at early discharge

The NIHSS is a standardized neurological examination score, with scores ranging from 0 to 42, with higher scores indicating increasing severity.

The mean difference and 95%CI of NIHSS score between two groups will be estimated using a GLM with treatment group as the sole predictor and the NIHSS score at 5-7 days or at early discharge as the dependent variable. Subjects with missing NIHSS score at 5-7 days or at early discharge will be excluded from the analysis. IPTW method will

be employed in calculating the co-variates adjusted mean difference (Co-variates will be defined as Section 9.6). Normality and variance homogeneity of residuals will be accessed graphically. The win ratio method will be used if normality assumption is seriously violated.

12.2.3 European Quality Five-Dimension (EQ-5D)

The EQ-5D is a tool for describing and assessing health-related quality of life 90±7 days after randomization. The tool is designed for self-completion.-The EQ-5D will be completed on the 90 (±7) days after randomization. Reasons for not performing the assessment of EQ-5D at Day 90 will be summarized.

The mean difference and 95%CI of EQ-5D score between two groups will be estimated using a GLM with treatment group as the sole predictor and the 90 days EQ-5D score as the dependent variable. IPTW method will be employed in calculating the co-variates adjusted mean difference and corresponding 95%CI (Co-variates will be defined as Section 9.6).

Subjects with missing EQ-5D score will be excluded from the analysis. Normality and variance homogeneity of residuals will be accessed graphically. The win ratio method will be used if normality assumption is seriously violated.

13. SAFETY ANALYSES

Safety analyses will be conducted on the Safety population only.

13.1 Mortality at 90 days

13.1.1 Defined of endpoint

Mortality due to any cause is the primary safety endpoint. The outcome is measured on a binary scale and is defined as positive in the presence of death of any cause, and as negative otherwise. With the date of randomization at day 0, any death occurring on or before calendar day 90 will be counted as a death. Patients who are alive at day 90 will

be censored at day 90. Mortality rates are defined as the number of deaths observed divided by the number of subjects observed over the 90-day study period.

13.1.2 Statistical Analysis

The modified Poisson regression with a treatment group as an independent variable and the cumulative rate death due to any cause at 90 days as the dependent variable. The treatment effect will be presented as RR with the corresponding 95%CI. IPTW will be used in calculating the adjusted RR and its 95%CI.

In addition, proportional hazards regression model will be employed to calculate the hazard ratio with 95%CI. IPTW method will be used to calculate the adjusted hazard ratio. In addition, the Kaplan-Meier method will be used to display the survival curves by treatment arm. The log-rank test will be applied to compare survival curves between the two treatment groups.

13.2 SICH rate within 48 hours

13.2.1 Defined of endpoint

Proportion of patients with symptomatic intracranial haemorrhage (SICH) within 48 hours post-treatment is considered another primary safety outcome. SICH will be evaluated according to the Heidelberg Bleeding Classification[20].

13.2.2 Statistical Analysis

Percentage of subjects with SICH within 48 hours after treatment will be presented for each definition by treatment group. Frequency counts and percentage of patients within each category will be provided for categorical data. Subject rates will be compared between treatment groups with Chi-square test or Fisher's exact test.

The modified Poisson Regression model will be used to estimates the RR and corresponding 95%CI with treatment group as an independent variable and the presence or absence of SICH as the dependent variable. The treatment effect will be presented as RR with the corresponding 95%CI. IPTW methods will be employed to calculate the adjusted RR and 95%CI.

13.3 Any ICH within 48 hours

The statistical analysis of differences in proportions of patients with any radiologic ICH within 48 hours between methylprednisolone sodium succinate and placebo groups will be performed as per the Section 13.2. SICH rate within 48 hours.

13.4 Proportion of patients with pneumonia

Number of patients and percentage of subjects with pulmonary infection within 48 hours after treatment will be presented by treatment group. The between group difference will be tested using the modified Poisson Model. And the Adjusted RR and corresponding 95% CI will be estimated using the IPTW method as other binary outcomes. Both unadjusted and adjusted results will be presented.

13.5 Proportion of patients with gastrointestinal bleeding within 7 days after EVT

Number of patients and percentage of subjects' gastrointestinal bleeding within 7 days after EVT will be presented by treatment group. The between group difference will be

tested using the modified Poisson Model. The Adjusted RR and corresponding 95%CI will be estimated using the IPTW method as other binary outcomes. Both unadjusted and adjusted results will be presented.

13.6. ADVERSE EVENTS

Adverse events (AEs) and diseases will recorded in this study.

13.6.1 Adverse Events

Any adverse change in health or the appearance of or worsening of any undesirable sign, symptom, or medical condition occurring after enrollment into the trial will be recorded as Adverse Event (AE) whether or not it is considered to be related to the study drug. An AE also includes a new illness; aggravated in severity of frequency from the baseline condition, abnormal results of diagnostic procedures, or a combination of the above. Pre-existing medical conditions are not to be reported as AEs.

13.6.2 Serious Adverse Events

A serious adverse event (SAE) is any untoward medical occurrence that at any dose: Result in death; Are life-threatening or fatal; Require or prolong hospitalization; Result in persistent or significant disability/incapacity, Constitutes a congenital anomaly or birth defect, or Significant medical event.

A SAE can also be an important medical event that may not result in death, be life-threatening, or require hospitalization, but may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. For example, any new diagnosis of cancer (made after study enrollment) is considered an important medical event. Because our primary safety outcomes for the trial are also SAEs, they will be reported dually as SAEs and as outcomes. SAEs should be managed according to the best current standard of care. All deaths occurring during the follow-up to Day 90 will be reported as an SAE. When reporting a death, the event or condition that caused or contributed to the fatal outcome should be reported as a single medical

concept.

References

1. CPMP/ICH/363/96. ICH E9 Statistical Principles for Clinical Trials.
2. Association AS. Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics.
3. Society RS. The Royal Statistical Society: Code of Conduct.
4. Zi W, Wang H, Yang D, et al. Clinical Effectiveness and Safety Outcomes of Endovascular Treatment for Acute Anterior Circulation Ischemic Stroke in China. Cerebrovasc Dis 2017;44:248-58.
5. Goyal M, Menon BK, van Zwam WH, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. LANCET 2016;387:1723-31.
6. Cranston JS, Kaplan BD, Saver JL. Minimal Clinically Important Difference for Safe and Simple Novel Acute Ischemic Stroke Therapies. Stroke 2017;48:2946-51.
7. Berkhemer OA, Fransen PSS, Beumer D, et al. A Randomized Trial of Intraarterial Treatment for Acute Ischemic Stroke. New England Journal of Medicine 2015;372:11-20.
8. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. Bmj 2011;342:d40.
9. SAS version 9.4 software, SAS Institute Inc., Cary, NC,.
10. Rev1. CE. Guideline on Missing Data in confirmatory trials.
11. (EMA/CHMP/ICH/436221/2017). ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Draft.
12. Wang D, Pocock S. A win ratio approach to comparing continuous non-normal outcomes in clinical trials. Pharm Stat 2016;15:238-45.
13. Agresti A. Categorical Data Analysis, Second Edition.
14. Saver JL. Novel end point analytic techniques and interpreting shifts across the entire range of outcome scales in acute stroke trials. Stroke 2007;38:3055-62.
15. Churilov L, Arnup S, Johns H, et al. An improved method for simple, assumption-free ordinal analysis of the modified Rankin Scale using generalized odds ratios. International Journal of Stroke 2014;9:999-1005.
16. Morris TP, Walker AS, Williamson EJ, White IR. Planning a method for covariate adjustment in individually randomised trials: a practical guide. Trials 2022;23:328.
17. Young-Saver DF, Gornbein J, Starkman S, Saver JL. Handling of Missing Outcome Data in Acute Stroke Trials: Advantages of Multiple Imputation Using Baseline and Postbaseline Variables. J Stroke Cerebrovasc Dis 2018;27:3662-9.
18. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159:702-6.
19. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34:3661-79.
20. von Kummer R, Broderick JP, Campbell BCV, et al. The Heidelberg Bleeding Classification Classification of Bleeding Events After Ischemic Stroke and Reperfusion Therapy. Stroke 2015;46:2981-6.

STATISTICAL ANALYSIS PLAN

MARVEL： Methylprednisolone for acute large vessel occlusion: a randomized double-blind, placebo-controlled trial in revascularization patients

Principal Investigators

Wenjie Zi, MD and Qingwu Yang, PhD., MD

Xinqiao Hospital, Army Medical University, Chongqing, China

SAP Authors

Wenjie Zi, MD, PhD; Changwei Guo, MD; Xinqiao Hospital, Army Medical University, Chongqing, China

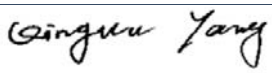
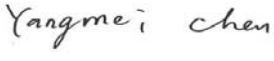
Duolao Wang, PhD, Global Health Trials Unit, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

Mengjie Lu, PhD, School of Public Health, Shanghai JiaoTong University School of

Medicine, Shanghai, China

CONFIDENTIAL

SAP version history		
Version Date	SAP Version #	Details of Changes
2021/09/29	*1.0*	
2023/10/01	*2.0*	*Changed the primary analysis of the primary outcome from inverse probability treatment weighting common odds ratio to generalized odds ratio.*

	Signature	Date
Dr. Changwei Guo (Trial Statistician)		*2023/10/01*
Prof. Duolao Wang (Senior Statistician)		
Prof. Qingwu Yang (Chief Investigator)		
Prof. Yangmei Chen (Chief of the DSMB)		

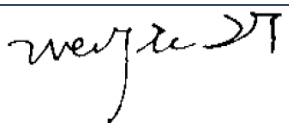
Prof. Wenjie Zi (Chief Investigator & SC Chair)		

TABLE OF CONTENTS

Abbreviations and Definitions of Terms

AE	Adverse Event
AIS	Acute ischemic stroke
ASPECTS	Alberta Stroke Program Early CT Score
CI	Confidence Interval
CRF	Case Report Form
CT	Computed Tomography
CTA	CT angiography
DSA	digital subtraction angiography
DSMB	Data Safety Monitoring Board
EQ-5D-5L	European Quality of Life Five-Dimension Five-Level scale
EVT	Endovascular Therapy
GCP	Good Clinical Practice
GEE	Generalized Estimating Equation
GLM	Generalized Linear Model
ICH	Intracranial Haemorrhage
ICH-E9	International Conference on Harmonization - Statistical principles for clinical trials
ITT	Intention-to-Treat
IVT	Intravenous Thrombolysis
MR	Magnetic Resonance
MRA	MR angiography
mRS	Modified Rankin Scale
NIHSS	National Institute of Health Stroke Scale
NMPA	National Medical Products Administration
PI	Principal Investigator
PP	Per-protocol
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan

SICH	Symptomatic Intracranial Haemorrhage
SOC	System Organ Class
TIA	Transient Ischemic Attack
tPA	Tissue Plasminogen Activator

1. PREFACE

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for protocol ChiCTR2100051729 entitled “Methylprednisolone for acute large vessel occlusion: a randomized double-blind, placebo-controlled trial in revascularization patients” (Version 2.0, 2022/8/18).

This study is being completed to assess the efficacy and safety of adjunctive intravenous methylprednisolone compared with placebo for the treatment of acute ischemic stroke (AIS) in patients with large vessel occlusion (LVO) underwent endovascular treatment (EVT).

The structure and content of this SAP provides sufficient details to meet the requirements determined by the National Medical Products Administration (NMPA) and International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH-E9): Guidance on Statistical Principles in Clinical Trials[1]. All work planned and reported for this SAP will follow internationally recognized guidelines, published by the American Statistical Association[2] and the Royal Statistical Society[3], for statistical practice.

In preparing this SAP, the following documents were reviewed:

- Clinical Study Protocol Version
- Case report forms (CRFs).
- ICH-E9 Guidance on Statistical Principles for Clinical Trials.
- Guidelines for the Content of Statistical Analysis Plans in Clinical Trials

Readers of this SAP are encouraged to also read the clinical trial protocol to understand the implementation details of this study, the operational aspects of clinical evaluation, and the schedule of patients completing this study.

2. PURPOSE OF SAP

The purpose of this SAP is to outline the planned analyses to be completed to support the completion of the clinical study report. The planned analyses identified in this SAP will be included in regulatory submissions and/or future manuscripts. In addition, exploratory analyses not necessarily identified in this SAP may be conducted to support the clinical development plan. Any post-hoc, or unplanned, analyses outside of this SAP will be clearly identified in the respective clinical study report.

3. VALIDATION OF SAP AND AMENDMENTS

The SAP should be validated and signed by:

Principal investigator(s)

Chair of Steering Committee

Chair of Data Safety and Monitoring Committee

Senior statistician

After validation, the SAP may be amended, and all changes will be clearly tracked. Any amendments can occur only before the database is locked and the treatment code is unblinded.

4. STUDY OBJECTIVES AND ENDPOINTS

4.1 Study Objectives

The primary objective is to test the efficacy and safety of adjunctive intravenous methylprednisolone compared with placebo for the treatment of acute ischemic stroke in patients with large vessel occlusion underwent endovascular treatment.

4.2 Study Endpoints (Target Variables)

4.2.1 Primary efficacy and safety endpoints

4.2.1.1 Primary Efficacy Endpoint

The primary efficacy outcome measure for this study is: The distribution of the global disability defined as a modified Rankin Scale (mRS) score at Day 90 (±7) after randomization (Ordinal data).

4.2.1.2 Primary Safety Endpoints

The primary safety outcome measure for this study are as follows:

- Mortality due to any cause within 90 (±7) days after randomization (Binary data)
- Proportion of patients with symptomatic intracranial haemorrhage within 48 h after treatment. Symptomatic intracranial haemorrhage will be adjudicated by an independent Imaging Core Laboratory according to the modified Heidelberg Bleeding Classification, as specified in the trial protocol. SICH defined as a new ICH seen on imaging will be later confirmed on independent reading by blinded central study radiologists (Binary data).

4.2.2 Secondary efficacy and safety endpoints

4.2.2.1 Secondary Efficacy Endpoints

The secondary efficacy outcomes are as follows:

The secondary endpoints include:

1) proportion of mRS score 0 to 4 at 90 days (Binary data);

2) proportion of mRS score 0 to 3 at 90 days (Binary data)'

3) proportion of mRS score 0 to 2 at 90 days (Binary data);

4) proportion of mRS score 0 to 1 or return to pre-morbid mRS score at 90 days (for patients with mRS >1) (Binary data);

5) NIHSS score at 5-7 days after EVT or at early discharge; (Continuous data)

6) European Quality Five-Dimension scale score at 90 days (Continuous data).

4.2.2.2 Secondary Safety Endpoints

The secondary safety outcomes measure for this study are as follows:

- Proportion of patients with any radiologic intracranial haemorrhage (based on central neuroimaging reading) within 48 hours after receiving intravenous study drug (Binary data);
- Proportion of patients with pneumonia (Binary data);
- Proportion of patients with gastrointestinal bleeding within 7 days after EVT (Binary data);
- Proportion of patients with any adverse events;
- Proportion of patients with serious adverse events;

5. STUDY METHODS

5.1 Overall Study Design and Plan

MARVEL is a multi-center, randomised, double-blinded, clinical trial. Patients will be randomized 1:1 to either intravenous methylprednisolone or intravenous placebo. It is planned to recruit at least 1672 patients in about 80 sites in China. The follow-up period covers the period from the subacute stage (day 5-7 or hospital discharge) to the final follow-up examination 83-97 days after randomization which defines the primary endpoint. The study patient flow outline following Consort diagram is shown in Figure 1.

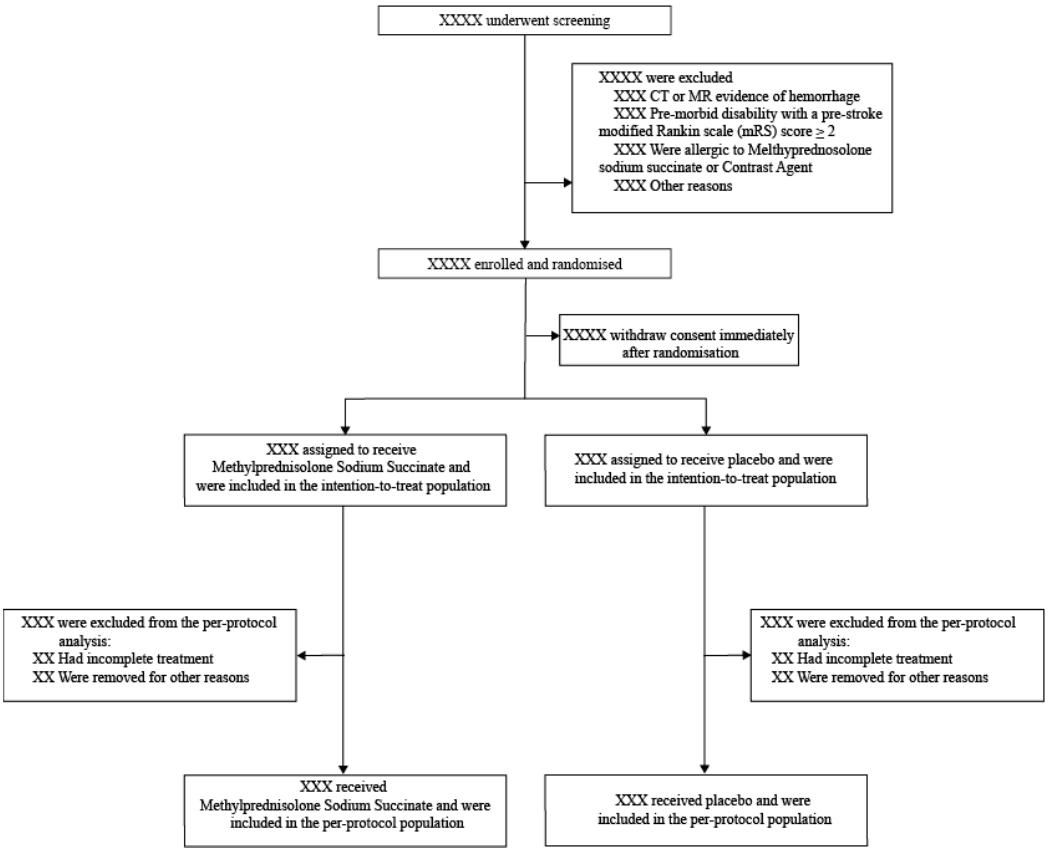


Figure 1. The study patient flow outline following Consort diagram.

5.2 Selection of Study Population

5.2.1 Inclusion Criteria

- Aged 18 years or older;
- The time from onset to randomization was within 24 hours;
- Baseline National Institutes of Health Stroke Scale (NIHSS) ⩾ 6
- Anterior circulation ischemic stroke was preliminarily determined according to clinical symptoms or imaging examination;
- Baseline Alberta Stroke Program Early CT Score (ASPECTS) ⩾ 3;

- Computed tomography angiography (CTA) /magnetic resonance angiography MRA) /digital subtraction angiography (DSA) confirmed occlusion of intracranial segment of internal carotid artery and middle cerebral artery, basilar artery and endovascular therapy was decided;
- Written informed consent signed by patients or their family members.

5.2.2 Exclusion Criteria

Patients meeting any of the following criteria will be excluded from study enrolment.

- Intracranial hemorrhage confirmed by cranial computed tomography (CT) or magnetic resonance imaging (MRI);
- mRS score ⩾ 2 before onset;
- Pregnant or lactating women;
- Allergic to contrast agents;
- Allergic to glucocorticoids;
- Participating in other clinical trials;
- Systolic blood pressure > 185 mmHg or diastolic pressure > 110 mmHg, and oral antihypertensive drugs can not control;
- Genetic or acquired bleeding constitution, lack of anticoagulant factors; Or oral anticoagulants and INR > 1.7;
- Blood sugar < 2.8 mmol/L (50 mg/dl) or > 22.2 mmol/L (400 mg/dl), platelet < 90 x 10^9/L;
- The artery is tortuous so that the thrombectomy device cannot reach the target vessel;
- Bleeding history (gastrointestinal and urinary tract bleeding) in recent 1 month;

- Chronic hemodialysis and severe renal insufficiency (glomerular filtration rate < 30 ml/min or serum creatinine > 220 umol/L [2.5 mg/ dL]);
- Life expectancy due to any advanced disease < 6 months;
- Follow-up is not expected to be completed;
- Intracranial aneurysm and arteriovenous malformation;
- Brain tumors with imaging mass effect;
- Severe systemic infectious diseases.

Inclusion and exclusion criteria will be assessed based on information available during the screening period.

5.3 Method of Treatment Assignment and Randomization

Randomization will be done by a web-based APP (Jinlingshu) on mobile phone or computer (https://jinlingshu.com/). The automated system will assign an appropriate set of study medication to each patient. Eligible patients are randomized to treatment with either intravenous methylprednisolone sodium succinate or intravenous placebo with a ratio of 1:1. Randomization will be stratified by participating centre with permutation block size of 4. Randomization will be completely concealed by having both web-based real-time allocation and identical appearance of methylprednisolone sodium succinate and placebo bottles. (All bottles will have a unique number. Subjects will be assigned a random serial number according to the time they were enrolled, and corresponding masked medications will be provided).

The randomization list will be prepared by the independent statistical centre using SAS 9.4. It will only be sent to the data coordination centre responsible for central randomization and data management, the pharmacy centre responsible for treatment packaging, and the centre responsible for blinding emergency patients.

5.4 Treatment masking (Blinding)

All trial personnel (investigators, their clinical staff, and the data management group) and patients will be blinded to treatment assignment. The Data Safety Monitoring Board (DSMB) will have access to masked data.

Both the person responsible for investigational product labels and the independent statistical group that prepare reports for the DSMB will be unblinded. In addition, the person in charge of the data management group responsible for managing the programming of the randomized system will be unblinded. If it is necessary to unblind, this person will become the contact person. This person will not participate in data management and will only communicate to unblind when contacted by the medical monitor.

If the treatment code needs to be broken for the safety of the patient, the site principal investigator (PI) will call the medical monitor for all unblinding queries. The medical monitor will discuss with the researcher whether it is necessary to unblind the patient. If the medical monitor determines that the PI should be unblinded for the patient, the medical monitor will contact the person in charge of the data management group by e-mail or phone, instructing the unblinding of the specific patient. Members of the unblinding data management team will provide only to the site PI with allocation information, the unblinding date, site number, PI name, and patient number via email. Any cases unblinded in this way will be recorded in the central file. Only the physician who requests the unblinding will receive the unblinding message. Study drugs will be stopped afterwards. It is not expected that there will be any clinical cases that need to be unblinded. The randomization data will be strictly confidential, and only authorized personnel can access it until the database is locked.

5.5 Contents of Investigational Product Kit

Methylprednisolone sodium succinate and its placebo are manufactured and provided by Lummy Pharmaceutical Group Co., Ltd., Chongqing, China. The medication packages are visually identical (including labelling, dosage form, size, and colour), except for a unique number. Methylprednisolone sodium succinate and its placebo will

be packed in sterile, disposable, individually labelled bottles. Each kit has a unique six-digit identification number and will be stored in a safe location at room temperature on the clinical site with restricted access. All study medication will be manufactured, tested, released, packaged, labelled, and shipped in accordance with Good Manufacturing Practice, Good Clinical Practice (GCP) guidelines, and any national regulatory requirements.

6. FORMAL ANALYSES AND REPORTING

All formal planned analyses identified in the protocol and in this SAP will be performed only after the formal database lock. Before database lock and completion of the final analyses, a blinded data review meeting will be held. Also, no database may be locked, random code unblinded, or analyses completed until this SAP has been approved and signed.

Any post-hoc, exploratory analyses completed to support planned study analyses, which were not identified in this SAP, will be documented and reported in appendices. Any results of these unplanned analyses will also be clearly identified in appendices.

7. SAMPLE SIZE DETERMINATION

The current trial is designed detect a shift on the modified Rankin Scale score, which represents the global disability, to a lower score. The expected distribution of the modified Rankin Scale score, derived from the EVT for acute anterior circulation ischemic stroke (ACTUAL) registry[4] and the Highly Effective Reperfusion evaluated in Multiple Endovascular Stroke Trials (HERMES) collaboration[5], is as follows: 7%, 20%, 16%, 11%, 13%, 8% and 25%.

The sample size was calculated using the cumulative proportion of patients with mRS 0-2. We assume a moderate effect of 7% absolute increase in the cumulative proportion of patients with mRS 0-2 in the intervention group, compared with controls, indicating an odds ratio (OR) of 1.33, which have substantially exceeds the

minimal clinically important difference.[6,7] In order to demonstrate the expected treatment effect with a type-1 error alpha =0.05 (two-tailed) and a power of 80% (beta = 20%). A sample size of n= 1588 patients (n=794 per treatment group) is required.

The intention-to-treat principle will be applied to the primary analysis and, therefore, to safeguard against dilution of the treatment effect associated with an approximate 5% non-adherence rate (due to loss to follow-up, consent withdrawal and other reasons), we initially planned to enrol n=1672 patients (n=836 per treatment group) for this study. This estimation was performed based on PASS (NCSS, LLC. Kaysville, Utah, USA) version 15.0.

8. ANALYSIS POPULATIONS

The following analysis populations are planned for the studies:

8.1 Intention-to-Treat (ITT) population

The ITT population includes all patients randomized into the trial and who were recorded as receiving any amount of study drug, even if the subject does not receive the correct treatment, or does not follow the protocol until completion[8]. The ITT population will be the primary analysis population for the efficacy endpoints and subjects will be analysed according to the treatment group to which they were assigned at randomization.

The ITT analysis strategy[8] for MARVEL is defined as follows:

- Is based on an ITT design that aims to collect all outcome data on all randomized subjects;
- Includes a main analysis that keeps subjects in their randomized groups, analyses all available outcome data, and is valid under a named plausible assumption about the missing data;
- Includes sensitivity analyses that consider a range of plausible alternative assumptions about the missing data;

- All randomised individuals are included in sensitivity analyses.

8.2 Per-Protocol (PP) population

The PP population is defined as the subset of the ITT population excluding major protocol violators deemed to have the potential to affect patient outcome in terms of efficacy.

The PP population includes patients who actually received the assigned treatment and do not have major protocol violations or deviations. Major protocol violations or deviations will be identified in a blinded fashion prior to database lock. More specifically, patients with anyone of the following criteria will be excluded from the PP population. These deviations will be determined based on the medical monitors' records, as well as programmatically, using the following criteria at a minimum:

- Received but did not complete treatment with study drug, or dose of study drug administered outside recommended dose.
- Violate inclusion or exclusion criteria.
- Missing essential information.

A list of patients to be excluded from the randomized patients to create the PP-Efficacy analysis will be established and validated by the Steering Committee prior to unblinding.

8.3 Safety population

The Safety Population includes all patients who received any amount of study drug. In case of violation of the randomization scheme, patients will be classified according to the treatment they actually received. Patients will be assigned to the different populations prior to unblinding of the database. Patients who withdraw informed consent immediately after randomization and are not to receive any treatment should be excluded from Safety Populations.

9. GENERAL ISSUES FOR STATISTICAL ANALYSIS

9.1 Analysis Software

All statistical analyses will be performed using SAS® Software version 9.4 in a Windows environment[9] and R version 4.3.0.

9.2 Methods for Withdrawals and Missing Data

9.2.1 Withdrawals

If any subject withdraws from the study prematurely (before the final visit day 90±7 for evaluation), they are required to complete the withdrawal visit in the Case Report Form. The reasons for withdrawal will be listed in a summary table. Subjects who withdrew before the subject's last follow up (90 ± 7 days) will be included in the analysis. Subjects who do not receive any study drugs after randomization will be excluded from any end-point analysis because no post-baseline data are available.

9.2.2 Missing data

The handling of missing data will follow the principles specified in the ICH-E9[1] and the CPMP/EWP/1776/99Rev1. Guideline on Missing Data in confirmatory trials Guidelines.[10] All efforts will be made to minimize the amount of missing data, particularly the 90-day outcome assessment. After the patients are randomized, the patients or their families and doctors will keep WeChat (a Chinese social networking software similar to Facebook) and phone numbers for each other to avoid loss to follow-up.

9.2.2.1 Baseline covariates

Missing baseline covariates will be imputed using simple imputation methods in the covariate adjusted analysis based on the covariate distributions, should the missing values for a particular covariate be less than 5%. For a continuous variable, missing

values will be imputed from random values from a normal distribution with mean and SD calculated from the available sample. For a categorical variable, missing values will be imputed from random values from a uniform distribution with probabilities P_1, P_2, …, and P_k from the sample.

If the missing values for a covariate are ⩾5% then they will be imputed using Markov chain Monte Carlo (MCMC) methods using SAS PROC MI. A total of 10 multiply imputed dataset will be generated. The seed for the imputations will be 128.

9.2.2.2 Efficacy outcomes

Sensitivity analyses based on different hypotheses about the missingness pattern of the primary outcome will be performed to test for the robustness of the primary analysis results[11]. Primary analysis of the primary outcome will be conducted in the complete case of the ITT population (in which patients with missing data will be censored). About other missing pattern please refer to Section 12.2.4.

9.3 Data Transformations

Data transformation is not relevant for most of the primary and secondary endpoints, which are categorical variables. The analysis of continuous variables may require transformation to normalize the distributions, but these variables will be mainly analysed using nonparametric statistical tests. Decisions to transform the distribution of some variables for analysis may be taken after the blind review of the data, and before unblinding the treatment code.

9.4 Multicenter Studies

No adjustment on centre will be performed for the primary analysis, on the basis of the following rationale:

The MARVEL trial aims to recruit 1672 patients from about 80 stroke centres in China, about 24 patients per centre. However, it is likely that many centres will recruit far

fewer patients, while a few large centres will recruit far more patients. This situation is the argument for not adjusted the primary analysis on centres. As the International Conference on Harmonization 9 guidelines quotes:

"In some trials, for example some large mortality trials with very few subjects per centre, there may be no reason to expect the centres to have any influence on the primary or secondary variables because they are unlikely to represent influences of clinical importance."

However, the central effect will be considered in the sensitivity analysis together with other important prognostic covariates.

9.5 Multiple Comparisons and Multiplicity

Analyses of secondary safety outcomes and additional analyses for the primary outcome are regarded as exploratory in nature, therefore, multiplicity adjustment will not apply to the primary and secondary outcome analyses.

9.6 Covariates

In the covariates-adjusted analysis, the following variables will be adjusted using the IPTW method.

- Age (Continuous)
- Baseline NIHSS score (Continuous)
- Pre-stroke modified Rankin Scale score (Ordered)
- Baseline aspects score (Continuous)
- Intravenous Thrombolysis (Binary)
- Time from stroke onset to randomization (Continuous)
- Occlusion location (Categorical)

The same IPTW weight will be used for all adjusted analyses.

9.7 Planned Subgroups

Subgroup analyses could be carried out irrespective of whether there is a significant treatment effect on the outcome. Their purpose is to supplement evidence from the primary analysis to help to fully characterize the treatment effect. Results from subgroup analyses will be interpreted in this context.

Subgroup analyses will be performed on a few important patient characteristics and other variables of interest. Pre-planned subgroup analyses will explore the effects of:

- Age (> median vs. ≤ median years old)
- Sex (Male vs. Female)
- Baseline NIHSS score (> median vs. ≤ median)
- Pre-stroke mRS (0 vs. 1 vs. 2 or more)
- Baseline ASPECTS score (> 5 vs. ≤ 5)
- Intravenous thrombolysis (yes vs. no)
- Time from last known well to randomization. (mins) (> median vs. ≤ median)
- Occlusion location (ICA vs. M1 vs. M2)
- Complete Reperfusion defined as Extended Thrombolysis In Cerebral Infarction grade 2c or 3 (yes or no)
- Patients with any radiological haemorrhages
- Patients with symptomatic haemorrhages

Subgroup analyses will be conducted irrespective of whether there is a significant treatment effect on the outcome. The purpose is to supplement evidence from the primary analysis to help to fully characterize the treatment effect. These analyses are exploratory and will not lead to formal inferential conclusions about the treatment effect in any particular subgroup.

A forest plot will be used to present the treatment effect of methylprednisolone for each of the individual subgroups. The plot will show the GenOR described above with two-

sided 95%CIs, corresponding to each level of these subgroups. Separately, the adjusted estimates for each of the patient characteristics defining a subgroup will be presented from the multivariate model.

Consistent with regulatory[12-14] to reduce the effect of the COVID-19 pandemic on the trial results, we prespecified the following subgroup analysis.

This subgroup included patients who completed 90-days follow up before 8th Jan 2023 when the Chinese Government officially downgraded the management of the disease from Class A to Class B in accordance with the country's law on prevention and treatment of infectious disease and remove it from quarantinable infectious disease management carried out in accordance with the Frontier Health and Quarantine Law of the People's Republic of China.

9.8 Derived and Computed Variables

Some derived and computed variables have been initially identified and described as appropriate later in the document. It is expected that additional variables will be required. The SAP will not be amended for additional variables that are not related to the primary target or key secondary target variables.

9.9 Presentation of results

Description of continuous variables includes the mean, standard deviation, median, first and third quartiles, and range. The description of categorical variables includes the frequency and percentage n (%). The description of ordered variables includes the frequency, percentage and cumulative percentage.

10. DISPOSITION OF SUBJECTS AND WITHDRAWALS

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

- Total number of assessed patients
- Number of randomized patients

- Number of patients completing the study and not completing the study (grouped by treatment and main reason)
- Number of patients included in the ITT population
- Number of patients included in the PP population

11. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Clinically important demographic and disease characteristics at screening will be summarized with descriptive statistics for each treatment group. Summaries will be produced for the ITT population and PP population.

If the missing value exceeds 5%, the denominator will be added in the footnote of the corresponding summary table.

The variables to be summarized include but are not restricted to the following:

Age

Sex

Medical History (hypertension, coronary heart disease, smoking, diabetes mellitus, hyperlipidaemia, and prior stroke)

Pre stroke onset mRS score

Baseline NIHSS score

Baseline ASPECTS Score

Systolic and Diastolic Blood Pressure

Stroke Aetiology

Serum Glucose

Occlusion Sites

Intravenous Thrombolysis

American Society of Interventional and Therapeutic Neuroradiology/Society of Interventional Radiology (ASITN/SIR)

Time From Stroke Onset (Last Known Well) to Randomization.

Time From Stroke Onset (Last Known Well) to Puncture

Time From Puncture to Reperfusion or the End of the Operation

Time From Randomization to Initiation of study drugs.

12. EFFICACY ANALYSES

All efficacy analyses will provide the point estimate of treatment effect with associated two-sided 95% confidence intervals (CIs). Statistical analyses of the primary and secondary efficacy endpoints will primarily be based on the ITT population. Secondarily, the same analyses will be repeated on the PP population.

Assessment of efficacy will be conducted based on endpoints collected during the Day 90 follow-up. The protocol allows this visit to be scheduled within 7 days of the target 90 days after randomization.

12.1 Primary Efficacy Endpoint (Ordinal data)

The mRS score is an assessment of disability with values from 0 (no symptoms at all) to 5 (sever disability) to 6 (death). The mRS score is collected at 90 days after randomization. If death occurs prior to an assessment day (Day 90), the mRS score will be considered available and will be set to 6.

For the analysis of the ordinal outcomes, Agresti proposed a generalized odds ratio (GenOR) for evaluating the difference between two groups.[15]Agresti's generalised odds ratio was defined as the ratio between the proportion of all pairs with better outcomes in the treatment group and the proportion of all pairs with better outcomes in the control group.[16]GenOR can be calculated using the following formula:

$$GenOR = \frac{P(Y_T < Y_C)}{P(Y_T > Y_C)}$$

Where Y_T and Y_C are the mRS scores in the treatment arm and control arm, respectively. $P(Y_T < Y_C)$ is the probability that among all possible pairwise comparisons between treatment and control a subject in treatment arm has a smaller mRS score than that in control arm, and $P(Y_T > Y_C)$ is the probability that among all possible pairwise comparisons between treatment and control a subject in treatment arm has a larger mRS score than that in control arm. [13]

The GenOR is actually identical to the win ratio statistic proposed by Pocock et al in 2012.[15,17] We will therefore use the WINS package in R to perform GenOR analysis, in which the variance of win ratio (GenOR) is estimated using a method proposed by Dong et al.[18] In case of binary outcome, genearalized odds ratio is actually the conventional odds ratio. Since the generalised odds ratio was proposed in 1980 by Agresti)[15] well before win ratio statistic proposed in 2012 by Pocock et al[17], we will use the term generalised odds ratio or genOR in the statistical reporting.

The distribution of full range of primary endpoint, mRS scores at Day 90, will be summarized by treatment group using a summary table and stacked bar chart. The primary endpoint will also be descriptively summarised using proportion of wins in treatment arm over control arm ($P(Y_T < Y_C)$ and proportion of wins in control arm over treatment arm ($P(Y_T > Y_C)$. The treatment effect will be measured using genOR together with its 95%CI and *P*-value.

The primary analysis of the primary endpoint will be conducted based on the complete case of the ITT population (patients with missing outcome data will be excluded) Sensitivity analysis based on different missing pattern and statistical method will be employed to test to robustness of the results.

12.1.1 Crude analysis

The null and alternative hypotheses to be tested appear below:

The null hypothesis for the generalized OR (GenOR) is equal to one:

$$GenOR = 1$$

The above hypothesis can be rephrased as the probability Prob ($Y_T < Y_C$) is equal to the probability Prob ($Y_T > Y_C$). Alternative hypothesis is
$GenOR \neq 1$. The point estimate and 95%CI of the crude GenOR will be estimated using the method by Dong et al[18] using the WINS package in R. Null hypothesis will be rejected if the 95%CI for the GenOR does not include 1.

12.1.2 Covariates adjusted analysis

The covariates adjusted generalized odds ratio will be calculated using the IPTW combined using the WINS package.

We will first calculate a propensity score with treatment as the dependent variable (1 for active treatment group and 0 for control) and all covariates listed above as independent variables through a logistic regression model, and then calculate the inverse probability of treatment weighting (IPTW) for each subject. In the next step, we will calculate the adjusted generalized odds ratio with the IPTW method. The application of these weights to the study population creates a pseudopopulation in which confounders are equally distributed across treatment and control groups.

12.1.3 Sensitivity analysis

To assess the influence of the missing primary endpoints on the treatment effect estimate, sensitivity analyses will be performed using the same statistical methods as described in Section 12.1.1 and 12.1.2, under the different assumptions of the missing outcome below:

(1) Worst-case scenario:

All patients with a missing primary endpoint will be considered as having the worst mRS score of 6.

(2) Best-case scenario:

All patients with a missing primary endpoint will be considered as having the best mRS

score of 0.

(3) Multiple Imputation

Missing mRS score will be imputed using multiple imputation method via SAS PROC MI. Imputation procedure will be performed under the missing-at-random (MAR) assumption. Fully conditional method (FCS) will be used to predict missing mRS score by fitting an ordered logistic regression model with following predictors: treatment group, age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location, and last observed NIHSS if available.[19] The seed will be 128. A total of 10 imputed datasets will be generated. GenOR will be estimated for each dataset separately, and the pooled together to calculate the final pooled genOR and 95% confidence interval using the Robin's role.

The above analysis will be performed on the imputed primary outcome for both ITT and PP populations as described in Section 12.1.2.

12.1.4 Supplementary analysis

To access the robustness of the primary analysis, the following supportive analysis will be employed as complement to the primary analysis. The analysis will be conducted in the complete case population of the primary outcomes.

(1) Proportional odds model. A proportional odds model will be employed in the sensitivity analysis. Both the unadjusted and adjusted results will be presented. The adjusted model will include age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, the use of intravenous thrombolysis, time from onset to randomization, and occlusion location as covariates.

(2) Patial Proportional Odds Model. If the proportional odds assumption holds true for unadjusted analysis in the proportional odds model, but fails when adjusted for age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, the use of intravenous thrombolysis, time from onset to randomization, and occlusion location,

then a partial proportional odds model will be fitted to estimate the common odds ratio of treatment.[20,21]

To access the influence of centre on the treatment effect, the following methods will be used:

(3) Proportional odds model with random effect: A proportional odds model with random effect will be fitted to estimate the common odds ratio between the methylprednisolone sodium succinate and placebo treatment with treatment as fixed effect, age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location as covariates, and the centre as a random effect.

12.1.5 Subgroup analysis

Subgroup analysis of the primary endpoint will be performed on the following subgroup variables, and the GenOR and its 95% CI will be estimated for each category of a subgroup variable and displayed in a forest plot to assess exploratorily homogeneity of treatment effect by a subgroup variable.

- age (> median vs. ≤ median years old)
- sex (Male vs. Female)
- baseline NIHSS score (> median vs. ≤ median)
- pre-stroke mRS (0 vs. 1 vs. 2 or more)
- baseline ASPECTS score (> 5 vs. ≤ 5)
- intravenous thrombolysis (yes vs. no)
- time from last known well to randomization. (mins) (> median vs. ≤ median)
- occlusion location (ICA vs. M1 vs. M2)
- Complete Reperfusion defined as Extended Thrombolysis in Cerebral Infarction grade 2c or 3 (yes or no)

- Patients with any radiological haemorrhages (with vs without)
- Patients with symptomatic haemorrhages (with vs without)

Subgroup analysis will be performed on both ITT and PP populations.

12.2 Secondary Efficacy Analysis

12.2.1 Dichotomized modified Rankin Scale score. (Binary data)

The treatment effect for binary secondary outcomes will be estimated in the ITT population using a modified Poisson regression as the risk ratio (RR) and its 95%CI.[22]

In the adjusted analysis, IPTW method will be used: we will first calculate a propensity score with treatment as the dependent variable (1 for active treatment group and 0 for control) and all covariates listed below as independent variables through a logistic regression model, and then perform an inverse probability of treatment weighting (IPTW) analysis and calculate the Adjusted RR with 95% CI.[23,24] Covariates including age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location to randomization.

In addition, RD (risk difference) and OR will be estimated and presented using two GLMs (identity-binomial regression and logit-binomial regression) as supportive measurements of treatment effect.[25] IPTW method will be calculate the covariate adjusted RD and OR.

12.2.2 NIHSS score at 5-7 days or at early discharge (Continuous data)

The NIHSS is a standardized neurological examination score, with scores ranging from 0 to 42, with higher scores indicating increasing severity.

The mean difference and 95%CI of NIHSS score between two groups will be estimated using a GLM with treatment group as the sole predictor and the NIHSS score at 5-7 days or discharge as the dependent variable. Subjects with missing NIHSS score at 5-7

days or discharge will be excluded from the analysis. IPTW method will be employed in calculating the co-variates adjusted mean difference (Co-variates will be defined as Section 9.6). Normality and variance homogeneity of residuals will be accessed graphically. The win ratio method will be used if normality assumption is seriously violated.

12.2.3 European Quality Five-Dimension (EQ-5D) (Continuous data)

The EQ-5D is a tool for describing and assessing health-related quality of life 90±7 days after randomization. The tool is designed for self-completion.-The EQ-5D will be completed on the 90 (±7) days after randomization. Reasons for not performing the assessment of EQ-5D at Day 90 will be summarized.

The mean difference and 95%CI of EQ-5D score between two groups will be estimated using a GLM with treatment group as the sole predictor and the 90 days EQ-5D score as the dependent variable. IPTW method will be employed in calculating the co-variates adjusted mean difference and corresponding 95%CI (Co-variates will be defined as Section 9.6).

Subjects with missing EQ-5D score will be excluded from the analysis. Normality and variance homogeneity of residuals will be accessed graphically. The win ratio method will be used if normality assumption is seriously violated.

13. SAFETY ANALYSES

Safety analyses will be conducted on the Safety population only.

13.1 Mortality at 90 days (Binary data)

13.1.1 Defined of endpoint

Mortality due to any cause is the primary safety endpoint. The outcome is measured on a binary scale and is defined as positive in the presence of death of any cause, and as

negative otherwise. With the date of randomization at day 0, any death occurring on or before calendar day 90 will be counted as a death. Patients who are alive at day 90 will be censored at day 90. Mortality rates are defined as the number of deaths observed divided by the number of subjects observed over the 90-day study period.

13.1.2 Statistical Analysis

The modified Poisson regression with a treatment group as an independent variable and the cumulative rate death due to any cause at 90 days as the dependent variable. The treatment effect will be presented as RR with the corresponding 95%CI. IPTW will be used in calculating the adjusted RR and its 95%CI.

In addition, proportional hazards regression model will be employed to calculate the hazard ratio with 95%CI. IPTW method will be used to calculate the adjusted hazard ratio. In addition, the Kaplan-Meier method will be used to display the survival curves by treatment arm. The log-rank test will be applied to compare survival curves between the two treatment groups.

13.2 SICH rate within 48 hours (Binary data)

13.2.1 Defined of endpoint

Proportion of patients with symptomatic intracranial haemorrhage (SICH) within 48 hours post-treatment is considered another primary safety outcome. SICH will be evaluated according to the Heidelberg Bleeding Classification[26].

13.2.2 Statistical Analysis

Percentage of subjects with SICH within 48 hours after treatment will be presented for each definition by treatment group. Frequency counts and percentage of patients within each category will be provided for categorical data. Subject rates will be compared between treatment groups with Chi-square test or Fisher's exact test.

The modified Poisson Regression model will be used to estimates the RR and corresponding 95%CI with treatment group as an independent variable and the presence or absence of SICH as the dependent variable. The treatment effect will be presented as RR with the corresponding 95%CI. IPTW methods will be employed to calculate the adjusted RR and 95%CI.

13.3 Any ICH within 48 hours (Binary data)

The statistical analysis of differences in proportions of patients with any radiologic ICH within 48 hours between methylprednisolone sodium succinate and placebo groups will be performed as per the Section 13.2. SICH rate within 48 hours.

13.4 Proportion of patients with pneumonia (Binary data)

Number of patients and percentage of subjects with pulmonary infection within 48 hours after treatment will be presented by treatment group. The between group difference will be tested using the modified Poisson Model. And the Adjusted RR and corresponding 95% CI will be estimated using the IPTW method as other binary outcomes. Both unadjusted and adjusted results will be presented.

13.5 Proportion of patients with gastrointestinal bleeding within 7 days after EVT (Binary data)

Number of patients and percentage of subjects' gastrointestinal bleeding within 7 days after EVT will be presented by treatment group. The between group difference will be

tested using the modified Poisson Model. The Adjusted RR and corresponding 95%CI will be estimated using the IPTW method as other binary outcomes. Both unadjusted and adjusted results will be presented.

13.6. ADVERSE EVENTS

Adverse events (AEs) and diseases will recorded in this study.

13.6.1 Adverse Events

Any adverse change in health or the appearance of or worsening of any undesirable sign, symptom, or medical condition occurring after enrollment into the trial will be recorded as Adverse Event (AE) whether or not it is considered to be related to the study drug. An AE also includes a new illness; aggravated in severity of frequency from the baseline condition, abnormal results of diagnostic procedures, or a combination of the above. Pre-existing medical conditions are not to be reported as AEs.

13.6.2 Serious Adverse Events

A serious adverse event (SAE) is any untoward medical occurrence that at any dose: Result in death; Are life-threatening or fatal; Require or prolong hospitalization; Result in persistent or significant disability/incapacity, Constitutes a congenital anomaly or birth defect, or Significant medical event.

A SAE can also be an important medical event that may not result in death, be life-threatening, or require hospitalization, but may jeopardize the subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition. For example, any new diagnosis of cancer (made after study enrollment) is considered an important medical event. Because our primary safety outcomes for the trial are also SAEs, they will be reported dually as SAEs and as outcomes. SAEs should be managed according to the best current standard of care. All deaths occurring during the follow-up to Day 90 will be reported as an SAE. When reporting a death, the event or condition that caused or contributed to the fatal outcome should be reported as a single medical

concept.

References

1. CPMP/ICH/363/96. ICH E9 Statistical Principles for Clinical Trials.
2. Association AS. Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics.
3. Society RS. The Royal Statistical Society: Code of Conduct.
4. Zi W, Wang H, Yang D, et al. Clinical Effectiveness and Safety Outcomes of Endovascular Treatment for Acute Anterior Circulation Ischemic Stroke in China. Cerebrovasc Dis 2017;44:248-58.
5. Goyal M, Menon BK, van Zwam WH, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. LANCET 2016;387:1723-31.
6. Cranston JS, Kaplan BD, Saver JL. Minimal Clinically Important Difference for Safe and Simple Novel Acute Ischemic Stroke Therapies. Stroke 2017;48:2946-51.
7. Berkhemer OA, Fransen PSS, Beumer D, et al. A Randomized Trial of Intraarterial Treatment for Acute Ischemic Stroke. New England Journal of Medicine 2015;372:11-20.
8. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. Bmj 2011;342:d40.
9. SAS version 9.4 software, SAS Institute Inc., Cary, NC,.
10. Rev1. CE. Guideline on Missing Data in confirmatory trials.
11. (EMA/CHMP/ICH/436221/2017). ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Draft.
12. Administration. UFaD. Conduct of clinical trials of medical products during the COVID-19 public health emergency guidance for industry, investigators, and institutional review boards. March, 2020, updated on Aug, 30, 2021. https://www.fda.gov/ media/136238/download (accessed Oct 8, 2022).
13. Agency. EM. Guidance on the management of clinical trials during the COVID-19 (coronavirus) pandemic. Version 5. https://health.ec.europa.eu/system/files/2022-02/ guidanceclinicaltrials_covid19_en_1.pdf (accessed Oct 8, 2022).
14. 国家药品监督管理局. NMPA. Guiding Principles for the Management of Drug Clinical Trials During the COVID-19 Pandemic.《新冠肺炎疫情期间药物临床试验管理指导原则（试行）》.
15. A. A. Generalized odds ratios for ordinal data. . Biometrics 1980;36:59–67.
16. Churilov L, Arnup S, Johns H, et al. An improved method for simple, assumption-free ordinal analysis of the modified Rankin Scale using generalized odds ratios. International Journal of Stroke 2014;9:999-1005.
17. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J 2012;33:176-82.
18. Dong G, Li D, Ballerstedt S, Vandemeulebroecke M. A generalized analytic solution to the win ratio to analyze a composite endpoint considering the clinical importance order among components. Pharm Stat 2016;15:430-7.
19. Young-Saver DF, Gornbein J, Starkman S, Saver JL. Handling of Missing Outcome Data in Acute Stroke Trials: Advantages of Multiple Imputation Using Baseline and Postbaseline Variables. J Stroke Cerebrovasc Dis 2018;27:3662-9.
20. Schildcrout JS, Harrell FE, Jr., Heagerty PJ, et al. Model-assisted analyses of longitudinal, ordinal outcomes with absorbing states. Stat Med 2022;41:2497-512.
21. Desantis SM, Lazaridis C, Palesch Y, Ramakrishnan V. Regression analysis of ordinal stroke clinical trial outcomes: an application to the NINDS t-PA trial. Int J Stroke 2014;9:226-31.
22. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159:702-6.
23. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting

(IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34:3661-79.

24. Morris TP, Walker AS, Williamson EJ, White IR. Planning a method for covariate adjustment in individually randomised trials: a practical guide. Trials 2022;23:328.

25. Naimi AI, Whitcomb BW. Estimating Risk Ratios and Risk Differences Using Regression. Am J Epidemiol 2020;189:508-10.

26. von Kummer R, Broderick JP, Campbell BCV, et al. The Heidelberg Bleeding Classification Classification of Bleeding Events After Ischemic Stroke and Reperfusion Therapy. Stroke 2015;46:2981-6.

27. National Institute of Neurological D, Stroke rt PAS, Study G. Tissue plasminogen activator for acute ischemic stroke. The New England journal of medicine 1995;333:1581-7.

28. Hacke W, Kaste M, Fieschi C, et al. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Second European-Australasian Acute Stroke Study Investigators. Lancet (London, England) 1998;352:1245-51.

29. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med 2008;359:1317-29.

30. Wahlgren N, Ahmed N, Eriksson N, et al. Multivariable analysis of outcome predictors and adjustment of main outcome results to baseline data profile in randomized controlled trials: Safe Implementation of Thrombolysis in Stroke-MOnitoring STudy (SITS-MOST). Stroke 2008;39:3316-22.

Summary of Changes - Statistical Analysis Plan MARVEL Version 1.0 to Version 2.0

Below is the table of changes. Deleted items are identified with ~~Strikethrough~~ font. Additional wording is in ***bold, italics, and underlining*** font

Section(s)	SAP Version 1.0 Change From:	SAP Version 2.0 Change To:	Rationale
5.2.1 Inclusion Criteria	Baseline Alberta Stroke Program Early CT Score (ASPECTS)>= 5;	***Added:*** Baseline Alberta Stroke Program Early CT Score (ASPECTS)>= ~~5~~*3*;	Addition To be consistent with the Change of Inclusion Criteria of Protocol.
9.1 Analysis Software	All statistical analyses will be performed using SAS® Software version 9.4 in a Windows environment9 and R version 4.1.1.	All statistical analyses will be performed using SAS® Software version 9.4 in a Windows environment9 and R version ~~4.1.1.~~***4.3.0.***	Update the R version
9.2.2.2 Efficacy outcomes		***Added:*** About other missing pattern please refer to Section 12.2.4.	Added to refer other sensitivity analysis under different missing assumption.
Section 9.6 Planned Subgroups, Interactions, and Covariates	Age Baseline NIHSS score Pre-stroke modified Rankin Scale score Baseline aspects score Intravenous Thrombolysis Time from stroke onset to randomization Occlusion location (Categorical)	***Changed:*** Age ***(Continuous)*** Baseline NIHSS score ***(Continuous)*** Pre-stroke modified Rankin Scale score ***(Ordered)*** Baseline aspects score ***(Continuous)*** Intravenous Thrombolysis ***(Binary)*** Time from stroke onset to randomization ***(Continuous)*** Occlusion location ***(Categorical)***	Addition To make the covariates more clear.

9.7 Planned Subgroups	A forest plot will be used to present the treatment effect of methylprednisolone for each of the individual subgroups. The plot will show the cOR described above with two-sided 95%CIs, corresponding to each level of these subgroups. Separately, the adjusted estimates for each of the patient characteristics defining a subgroup will be presented from the multivariate model.	A forest plot will be used to present the treatment effect of methylprednisolone for each of the individual subgroups. The plot will show the ~~cOR~~***GenOR*** described above with two-sided 95%CIs, corresponding to each level of these subgroups. Separately, the adjusted estimates for each of the patient characteristics defining a subgroup will be presented from the multivariate model.	Addition To be consistent with the Change of analysis method of the primary analysis.
9.7 Planned Subgroups		***Consistent with regulatory12-14 to reduce the effect of the COVID-19 pandemic on the trial results, we prespecified the following subgroup analysis.*** ***This subgroup included patients who completed 90-days follow up before 8th Jan 2023 when the Chinese Government officially downgraded the management of the disease from Class A to Class B in accordance with the country's law on prevention and treatment of infectious disease and remove it from quarantinable infectious disease management carried out in accordance with the Frontier Health and Quarantine Law of the People's Republic of China.***	As COVID-19 is an important variable that will influence the effect of corticosteroid, a sensitivity analysis was determined by the steering committee.
12.1 Primary Efficacy Endpoint	12.1 Primary Efficacy Endpoint	12.1 Primary Efficacy Endpoint (***Ordinal data)***	Addition, To make the definition

			more clear.
12.1 Primary Efficacy Endpoint (Ordinal data)	As the primary endpoint, mRS score at Day 90 will be further explored on the range of the values (0 to 6). The distribution of full range of mRS scores at Day 90 will be summarized by treatment group. Both tabular displays and stacked bar charts will be prepared. Win ratio method will be used to calculate the win ratio statistic as a measurement of treatment effect in mRS score.12 Primary analysis of the primary outcome will be conducted in the complete case of the ITT population (in which patients with missing data will be deleted) using the proportional odds model. 12.1.1 Hypothesis The null and alternative hypotheses to be tested appear below: The null hypothesis for proportional odds logistic regression is that the common OR is equal to one, which needs to be disproved. The null hypothesis for WMW test is the equality of ranks when ties are split evenly. This can be rephrased as both probabilities Prob (YM > YP) + 0.5Prob (YM = YP)	~~As the primary endpoint, mRS score at Day 90 will be further explored on the range of the values (0 to 6). The distribution of full range of mRS scores at Day 90 will be summarized by treatment group. Both tabular displays and stacked bar charts will be prepared. Win ratio method will be used to calculate the win ratio statistic as a measurement of treatment effect in mRS score.12 Primary analysis of the primary outcome will be conducted in the complete case of the ITT population (in which patients with missing data will be deleted) using the proportional odds model.~~ ~~12.1.1 Hypothesis~~ ~~The null and alternative hypotheses to be tested appear below:~~ ~~The null hypothesis for proportional odds logistic regression is that the common OR is equal to one, which needs to be disproved. The null hypothesis for WMW test is the equality of ranks when ties are split evenly. This can be rephrased as both probabilities Prob (YM > YP) + 0.5Prob (YM = YP) and Prob (YM < YP) + 0.5Prob (YM = YP) being equal to 0.5, resulting in WMW GenOR being equal to 1. Therefore, the null hypothesis for the WMW test states that there's an equal probability of treatment observation being worse or better than control observation (with equally splitting ties),~~	Changed the primary outcome analysis. The initial statistical method was the proportional odds regression adjusted for covariates using the Inverse probability treatment weighting method. The analysis method was change with the concern that the proportional odds model assumption can rarely held after adjusting for covariates. An assumption-free method was considered to be the primary analysis consisted with two previous studies published in the New England journal of medicine. (the ANGEL-ASPECTS study and the SELECT2 study). And the proportion odds assumption was

	and Prob (YM < YP) + 0.5Prob (YM = YP) being equal to 0.5, resulting in WMW GenOR being equal to 1. Therefore, the null hypothesis for the WMW test states that there's an equal probability of treatment observation being worse or better than control observation (with equally splitting ties), meaning that WMW GenOR isn't different from one.	~~meaning that WMW GenOR isn't different from one.~~ ***For the analysis of the ordinal outcomes, Agresti proposed a generalized odds ratio (GenOR) for evaluating the difference between two groups.15Agresti's generalised odds ratio was defined as the ratio between the proportion of all pairs with better outcomes in the treatment group and the proportion of all pairs with better outcomes in the control group.16GenOR can be calculated using the following formula:*** ***GenOR=(P(Y_T<Y_C))/(P(Y_T>Y_C))*** ***Where YT and YC are the mRS scores in the treatment arm and control arm, respectively. P(Y_T<Y_C) is the probability that among all possible pairwise comparisons between treatment and control a subject in treatment arm has a smaller mRS score than that in control arm, and P(Y_T>Y_C) is the probability that among all possible pairwise comparisons between treatment and control a subject in treatment arm has a larger mRS score than that in control arm. 13*** ***The GenOR is actually identical to the win ratio statistic proposed by Pocock et al in 2012.15,17 We will therefore use the WINS package in R to perform GenOR analysis, in which the variance of win ratio (GenOR) is estimated using a method proposed by Dong et al.18 In case of***	considered as an supportive analysis.

		binary outcome, genearalized odds ratio is actually the conventional odds ratio. Since the generalised odds ratio was proposed in 1980 by Agresti)15 well before win ratio statistic proposed in 2012 by Pocock et al17, we will use the term generalised odds ratio or genOR in the statistical reporting. ***The distribution of full range of primary endpoint, mRS scores at Day 90, will be summarized by treatment group using a summary table and stacked bar chart. The primary endpoint will also be descriptively summarised using proportion of wins in treatment arm over control arm (P(Y_T<Y_C) and proportion of wins in control arm over treatment arm (P(Y_T>Y_C). The treatment effect will be measured using genOR together with its 95%CI and P-value.*** ***The primary analysis of the primary endpoint will be conducted based on the complete case of the ITT population (patients with missing outcome data will be excluded) Sensitivity analysis based on different missing pattern and statistical method will be employed to test to robustness of the results.***	
12.1.1 Crude analysis	The categorical shift in the distribution of mRS scores between the two	~~The categorical shift in the distribution of mRS scores between the two treatment groups will be~~	

	treatment groups will be analysed fitting a proportional-odds logistic regression model, assuming a common odds ratio across all cut points of the modified Rankin scale. a) If the proportional odds assumptions are satisfied (both Brant test and Approximate likelihood-ratio test of proportionality of odds are not significant), common OR and 95%CI will be calculate from an GLM with multinomial distribution and cumlogit-link(ordinal logistic regression model) 13,14; b) If the proportional odds assumptions are not satisfied for the crude analysis (significant Brant test or Approximate likelihood-ratio test of proportionality of odds), assumption-free method will be used.15	~~analysed fitting a proportional-odds logistic regression model, assuming a common odds ratio across all cut points of the modified Rankin scale.~~ ~~a) If the proportional odds assumptions are satisfied (both Brant test and Approximate likelihood-ratio test of proportionality of odds are not significant), common OR and 95%CI will be calculate from an GLM with multinomial distribution and cumlogit-link(ordinal logistic regression model) 13,14;~~ ~~b) If the proportional odds assumptions are not satisfied for the crude analysis (significant Brant test or Approximate likelihood-ratio test of proportionality of odds), assumption-free method will be used.15~~ ***The null and alternative hypotheses to be tested appear below:*** ***The null hypothesis for the generalized OR (GenOR) is equal to one:*** ***GenOR=1*** ***The above hypothesis can be rephrased as the probability Prob (YT < YC) is equal to the probability Prob (YT > YC). Alternative hypothesis is*** ***GenOR≠1 . The point estimate and 95%CI of the crude GenOR will be estimated using the method by Dong et al18 using the WINS package in R. Null hypothesis will be rejected if the 95%CI for the GenOR does not include 1.***	

12.1.2.2 Covariates Adjusted Analysis	a) If the proportional odds assumptions are satisfied for crude analysis, and still satisfied when adding the covariates in the ordinal logistic regression model, then the common odds ratio from an ordinal logistic regression model, IPTW method will be employed in the co-variate adjusted analysis. 13,14,16 b) If the proportional odds assumptions are not satisfied for the IPTW adjusted analysis (significant Brant test or Approximate likelihood-ratio test of proportionality of odds), assumption-free method will be used.15 Covariate adjusted analysis will be performed on both ITT and PP populations. Main conclusion will be drawn from the adjusted ITT analysis of the primary endpoint.	~~a) If the proportional odds assumptions are satisfied for crude analysis, and still satisfied when adding the covariates in the ordinal logistic regression model, then the common odds ratio from an ordinal logistic regression model, IPTW method will be employed in the co-variate adjusted analysis. 13,14,16~~ ~~b) If the proportional odds assumptions are not satisfied for the IPTW adjusted analysis (significant Brant test or Approximate likelihood-ratio test of proportionality of odds), assumption-free method will be used.15~~ ~~Covariate adjusted analysis will be performed on both ITT and PP populations.~~ ~~Main conclusion will be drawn from the adjusted ITT analysis of the primary endpoint.~~***12.1.3 Covariates Adjusted Analysis*** ***a) If the proportional odds assumptions are satisfied for crude analysis, and still satisfied when adding the covariates in the ordinal logistic regression model, then the common odds ratio from an ordinal logistic regression model, IPTW method will be employed in the co-variate adjusted analysis. 13,14,16*** ***b) If the proportional odds assumptions are not satisfied for the crude analysis (significant Brant test or Approximate likelihood-ratio test of proportionality of odds), assumption-free method will be used.15***	

		Covariate adjusted analysis will be performed on both ITT and PP populations. ***Main conclusion will be drawn from the adjusted ITT analysis of the primary endpoint.***	
12.1.4 Sensitivity analysis	(3) Multiple Imputation Missing mRS score will be imputed using multiple imputation method via SAS PROC MI. Imputation procedure will be performed under the missing-at-random (MAR) assumption. Fully conditional method (FCS) will be used to predict missing mRS score by fitting an ordered logistic regression model with following predictors: treatment group, age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location, and last observed NIHSS if available.[19] The seed will be 128. A total of 10 imputed datasets will be generated. cOR will be estimated for each dataset separately, and the pooled together to calculate the final pooled cOR and 95% confidence interval using the Robin's role.	Missing mRS score will be imputed using multiple imputation method via SAS PROC MI. Imputation procedure will be performed under the missing-at-random (MAR) assumption. Fully conditional method (FCS) will be used to predict missing mRS score by fitting an ordered logistic regression model with following predictors: treatment group, age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location, and last observed NIHSS if available.19 The seed will be 128. A total of 10 imputed datasets will be generated. ~~cOR~~***GenOR*** will be estimated for each dataset separately, and the pooled together to calculate the final pooled ~~cOR~~***genOR*** and 95% confidence interval using the Robin's role.	Change from cOR to genOR to be consistent to the change of the primary outcome.
12.1.4 Supplementary		***Added: To access the robustness of the primary***	Added the cOR method

analysis		***analysis, the following supportive analysis will be employed as complement to the primary analysis. The analysis will be conducted in the complete case population of the primary outcomes.*** ***(2) Proportional odds model. A proportional odds model will be employed in the sensitivity analysis. Both the unadjusted and adjusted results will be presented. The adjusted model will include age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, the use of intravenous thrombolysis, time from onset to randomization, and occlusion location as covariates.*** ***(2) Patial Proportional Odds Model. If the proportional odds assumption holds true for unadjusted analysis in the proportional odds model, but fails when adjusted for age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, the use of intravenous thrombolysis, time from onset to randomization, and occlusion location, then a partial proportional odds model will be fitted to estimate the common odds ratio of treatment.20,21*** ***To access the influence of centre on the***	as supportive analysis method in consist to the change of the primary analysis method.

		treatment effect, the following methods will be used: ***(4) Proportional odds model with random effect: A proportional odds model with random effect will be fitted to estimate the common odds ratio between the methylprednisolone sodium succinate and placebo treatment with treatment as fixed effect, age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location as covariates, and the centre as a random effect.***	
12.1.5 Subgroup analysis	Assessment of the homogeneity of treatment effect by a subgroup variable will be conducted by a GLM with the treatment, subgroup variable, and their interaction term as predictors, adjusted for age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location, and the P-value from the interaction term will be used to assess exploratorily homogeneity of treatment effect by a subgroup	~~Assessment of the homogeneity of treatment effect by a subgroup variable will be conducted by a GLM with the treatment, subgroup variable, and their interaction term as predictors, adjusted for age, baseline NIHSS score, pre-stroke mRS, baseline ASPECTS score, use of intravenous thrombolysis, time from onset to randomization, and occlusion location, and the P-value from the interaction term will be used to assess exploratorily homogeneity of treatment effect by a subgroup variable.~~ ***Subgroup analysis of the primary endpoint will be performed on the following subgroup***	Rephrase the sentence, to be consistent with the change of primary outcome. And currently there was no robust method to calculate formal p value for interaction for the generalized odds ratio,

	variable.	***variables, and the GenOR and its 95% CI will be estimated for each category of a subgroup variable and displayed in a forest plot to assess exploratorily homogeneity of treatment effect by a subgroup variable.***	
12.2.1 Dichotomized modified Rankin Scale score. (Binary data)	12.2.1 Dichotomized modified Rankin Scale score.	12.2.1 Dichotomized modified Rankin Scale score. ***(Binary data)***	Added the variable type to define the variable more clearly.
12.2.2 NIHSS score at 5-7 days or at early discharge	12.2.2 NIHSS score at 5-7 days or at early discharge	12.2.2 NIHSS score at 5-7 days or at early discharge ***(Continuous data)***	Added the variable type to define the variable more clearly.
12.2.3 European Quality Five-Dimension (EQ-5D) (Continuous data)	12.2.3 European Quality Five-Dimension (EQ-5D)	12.2.3 European Quality Five-Dimension (EQ-5D) ***(Continuous data)***	Added the variable type to define the variable more clearly.
13.1 Mortality at 90 days (Binary data)	13.1 Mortality at 90 days	13.1 Mortality at 90 days ***(Binary data)***	Added the variable type to define the variable more clearly.
13.2 SICH rate within 48 hours (Binary data)	13.2 SICH rate within 48 hours	13.2 SICH rate within 48 hours ***(Binary data)***	Added the variable type to define the variable more clearly.