A multicenter, randomized, double-blind, placebo-controlled phase III trial of camrelizumab plus chemotherapy with or without radiation therapy (RT) as first-line treatment for patients with brain metastatic non-small cell lung cancer (CTONG2003)

Statistical Analysis Plan

(SAP)

Author: Fengzhuo Cheng, Clinical Statistician

Company: Jiangsu Hengrui Medicine Co., Ltd.

Version: V1.0

Date: Feb 18, 2024

TABLE OF CONTENTS

1. Revision

Version Number	Version Date	Main Revisions	Reason for Revision
V1.0	Feb 18, 2024	First draft	

2. Introduction

This statistical analysis plan (SAP) is prepared for the " A multicenter, randomized, double-blind, placebo-controlled phase III trial of camrelizumab plus chemotherapy with or without radiation therapy (RT) as first-line treatment for patients with brain metastatic non-small cell lung cancer ", providing specific statistical analysis rules and strategies for its efficacy and safety analyses. The data to be aggregated and analyzed, as well as the details of the statistical analysis, will be described in this document.

This SAP is written according to the study protocol (Version number: 2.0, Version date: May 05, 2023). It will be finalized before the database lock and submitted to various functional departments for signature.

2.1. Study Design

This study is a multicenter, randomized, double blinded, controlled clinical study to compare the efficacy and safety of camrelizumab and radiochemotherapy (Group A) versus placebo and radiochemotherapy (Group B) in patients with advanced Non-Small-Cell Lung Cancer with brain metastases. Patients will be randomly assigned to either study group A or control group B in a 1:1 ratio stratified with squamous cell carcinoma/non-squamous carcinoma, number of intracranial metastases(1-5, >5) and with/without planned radiotherapy therapy at baseline. The primary endpoint are intracranial progression-free survival (iPFS) and overall progression-free survival (PFS), both are assessed by investigators according to the Response Evaluation Criteria in Solid Tumours version 1.1 (RECIST 1.1). A total of 200 patients are planned to be enrolled.

The study includes a screening period (no more than 28 days from when the subject signs the informed consent form (ICF) to the first treatment), a treatment period and a follow-up period (including both safety follow-up and survival follow-up).

2.2. Study Objectives

2.2.1. Primary Objective

To compare the iPFS and PFS assessed by investigators according to RECIST 1.1 in patients with advanced Non-Small-Cell Lung Cancer with brain metastases followed by camrelizumab and radiochemotherapy (Group A) versus placebo and radiochemotherapy (Group B).

2.2.2. Secondary Objectives

- To compare the iPFS and PFS assessed by central site investigators according to RECIST 1.1 in the study group (Group A) versus the control group (Group B);
- To compare the overall survival (OS) in the study group (Group A) versus the control group (Group B);
- To compare other efficacy endpoints including intracranial objective response rate (iORR), overall objective response rate (ORR), disease control rate (DCR) and duration of response (DoR) assessed by investigators according to RECIST 1.1, and safety of treatment regimens and health-related quality of life (HRQoL) to assess neurocognitive function changes(MMSE, HVLT-R) in the study group (Group A) versus the control group (Group B).

2.2.3. Exploratory Objective

To evaluate the patients-reported outcomes (PROs): EORTC QLQ-C30 and QLQ-BN20.

3. Analysis Plan

3.1. Final Analysis

After all subjects have been enrolled and followed up for 38 months, at least 153 iPFS events are expected to be collected in the control group and the study group, at which time the final analysis will be conducted.

4. Methods

4.1. Statistical Hypotheses

Primary endpoint I:

Null hypothesis (H_0): There is no difference in median iPFS between the study group and the control group.

Alternative hypothesis (H_1): There is a difference in median iPFS between the study group and the control group.

Primary endpoint II:

Null hypothesis (H_0): There is no difference in median PFS between the study group and the control group.

Alternative hypothesis (H_1): There is a difference in median PFS between the study group and the control group.

4.2. Multiple Comparisons/Multiplicity

The primary endpoints of this study are iPFS and PFS, with a test level of two-sided α= 0.05. A sequential design is used to control the overall Type I error level. iPFS will be first tested for Group A versus Group B, and subsequent PFS tests will be continued only if result of the previous test shows a significant difference.

4.3. Sample Size Determination and Statistical Decision Rules

This is a randomized controlled study. According to BRAIN study reports, the median iPFS of treatment regimen in the control group B is 5 months. According to the Keynote-021/189/407 data of the regimen under investigation of this study, it is expected in this study that the median iPFS in study group A for patients with advanced Non-Small-Cell Lung Cancer with brain metastases can reach 8 months. At a significance level α of 0.05 (two-sided), approximately 153 iPFS events need to be observed to provide a power of no less than 80% to detect the difference in iPFS between treatment regimen of the study group A and that of the control group B. An enrollment period of 14 months and a follow-up period of 38 months are planned for this study. A total of 200 subjects needs to be enrolled.

The primary endpoints of this study are iPFS and PFS, with a test level of two-sided α= 0.05. A sequential design is used to control the overall Type I error level. iPFS will be first tested for Group A versus Group B, and subsequent PFS tests will be continued only if result of the previous test shows a significant difference. 200 subjects would meet 92.8% power to detect a difference in PFS under assumptions that median PFS would increase from 4 months to 7 months (HR=0.57).

In summary, to provide a power of at least 80% to test the difference in iPFS between the study group A and the control group B, and to provide a power of at least 90% for the sequential test of the difference in PFS between the study group A and the control group B, a total of 200 subjects are required for this study.

4.4. Randomization

This is a randomized, double-blinded, controlled study. A randomization coding table will be generated using stratified blocked randomization, and the enrolled subjects will be randomly assigned to either study group A or control group B in a 1:1 ratio. Randomization will be stratified according to squamous cell carcinoma/non-squamous carcinoma, number of intracranial metastases(1-5, >5) and with/without planned radiotherapy at baseline.

4.5. Power analysis

Recruitment was terminated early due to therapeutic paradigm shifts globally, totally 60 patients were randomized. Conditional power (CP) and predicted power (PPOS) were computed for this study. Conditional power is the probability that the final result will be significant, given the data obtained up to the time of the early termination look. And predictive power (a Bayesian concept) is the result of averaging the conditional power over the posterior distribution of effect size.

5. Analysis Set

5.1. Intention-to-Treat Set

All subjects who have been successfully randomized and enrolled will constitute the intention-to-treat (ITT) set for this study. The ITT set is the primary data set used to evaluate the efficacy endpoints and describe the characteristics of the population.

5.2. Full Analysis Set

All randomized subjects who have at least one dosing record will constitute the full analysis set (FAS) of this study.

5.3. Per Protocol Set

The per protocol set (PPS) is a subset of the ITT population, which includes all subjects in the ITT population who have completed the treatments and visits specified in the protocol and have not had any major protocol deviations. The criteria for determination of major protocol deviations need to be finalized before the database lock. The list of subjects to be included or excluded from the PPS needs to be determined by the principal investigator, statistician, and sponsor at the data review meeting before the database lock. The PPS is the secondary analysis set for the efficacy endpoint of this study, and will be analyzed according to the randomly assigned treatment groups.

5.4. Safety Set

Safety set (SS) includes all randomized subjects who have received at least one dose of the study drug. The SS will be used for all safety analyses, and will be analyzed according to the actual treatment received by the subject.

5.5. Other Analysis Sets

5.5.1. Patient-Reported Outcomes Analysis Set

As a subset of the ITT set, the patient-reported outcomes (PROs) analysis set includes all subjects in the ITT population who have completed the baseline and at least one post-baseline PROs scale.

6. Endpoints and Estimands

6.1. Efficacy Endpoints

6.1.1. Primary Endpoint and Primary Estimand

- Intracranial progression-free survival

Population	Patients with advanced Non-Small-Cell Lung Cancer with brain metastases who have not received systemic treatment
Treatment	Refer to protocol 4.7.2
Endpoint	iPFS assessed by investigators according to RECIST 1.1
Summary at Population Level	Median iPFS and hazard ratio (HR)
Intercurrent Events (ICEs) and Management Strategies	ICE1: In case of discontinuation of study treatment (treatment method strategy), data collection will be continued after the discontinuation of study treatment for analysis. ICE2: In case of use of subsequent anti-tumor treatment (in-treatment strategy), only efficacy data prior to the start of subsequent anti-tumor treatment will be used. If the use of subsequent anti-tumor treatment is due to extracranial progression (treatment method strategy), efficacy data after the start of subsequent anti-tumor treatment will be collected for analysis.

- Progression-free survival

Population	Same as that of the primary endpoint iPFS
Treatment	Same as that of the primary endpoint iPFS
Endpoint	PFS assessed by investigators according to RECIST 1.1
Summary at Population Level	Median PFS and hazard ratio (HR)
Intercurrent Events (ICEs) and Management Strategies	ICE1: In case of discontinuation of study treatment (treatment method strategy), data collection will be continued after the discontinuation of study treatment for analysis. ICE2: In case of use of subsequent anti-tumor treatment (in-treatment strategy), only efficacy data prior to the start of subsequent anti-tumor treatment will be used.

6.1.2. Secondary Endpoints and Secondary Estimands

- Overall Survival

Population	Same as that of the primary endpoint iPFS
Treatment	Same as that of the primary endpoint iPFS
Endpoint	OS
Summary at Population Level	HR
ICEs and Management Strategies	ICE1: In case of discontinuation of study treatment (treatment method strategy), data collection will be continued after the discontinuation of study treatment for analysis. ICE2: In case of use of subsequent anti-tumor treatment (treatment method strategy), survival information after the start of subsequent anti-tumor treatment will be collected for analysis.

- Intracranial objective response rate

Population	Same as that of the primary endpoint iPFS
Treatment	Same as that of the primary endpoint iPFS
Endpoint	iORR assessed by investigators according to RECIST 1.1
Summary at Population Level	The proportion of patients that respond either partially or fully to therapy, and the risk difference and odds ratio in iORR between the study group and the control group
ICEs and Management Strategies	ICE1: preoperative dropout due to any reason (hypothetical strategy) ICE2: preoperative use of subsequent anti-tumor treatment (hypothetical strategy)

- Objective response rate

Population	Same as that of the primary endpoint iPFS
Treatment	Same as that of the primary endpoint iPFS
Endpoint	ORR assessed by investigators according to RECIST 1.1
Summary at Population Level	The proportion of patients that respond either partially or fully to therapy, and the risk difference and odds ratio in ORR between the study group and the control group
ICEs and Management Strategies	ICE1: preoperative dropout due to any reason (hypothetical strategy) ICE2: preoperative use of subsequent anti-tumor treatment (hypothetical strategy)

Other secondary endpoints

- Duration of response is defined as the time from randomization to disease progression or death in patients who achieve complete or partial response;
- Disease control rate (DCR) describes the percentage of patients with advanced cancer whose therapeutic intervention has led to a complete response, partial response, or stable disease;
- patient reported HRQoL is an evaluation of patient's quality of life with respect to health status over time, HRQoL is assessed using MMSE, HVLT-R to find the differences in toxicity of neurocognitive function and using EORTC QLQ-C30 and QLQ-BN20 to detect the general changes in quality of life between group A and group B.

6.2. Safety Endpoints

6.2.1. Adverse Events

AEs will be coded according to the MedDRA dictionary, and summarized by preferred term (PT), the correlation to the study drug, and severity grading according to NCI-CTCAE v5.0. Events, the number of subjects with each event in each treatment group and the incidence, the start and end time of the event, whether the event is a serious adverse event (SAE), the correlation, and the outcome will be listed. And the results for each cycle will be listed.

Treatment emergent adverse events (TEAEs)

An AE that meets any of the following conditions will be determined as a TEAE:

- Any untoward medical event that occurs after the first dose of the study drug and before the end of the safety follow-up period (90 days after the last dose of camrelizumab for the study group, and within 28 days after the last dose of the study drug for the control group);
- Any medical condition/disease with the onset date before the first dose of the study drug but worsens during the treatment with the increased severity according to the NCT-

CTCAE version 5.0;

- Progressive disease leading to death occurring before the safety follow-up period.

Treatment-related adverse events

TEAEs that, as judged by the investigator, have a certain, possible, or indeterminable correlation with the study drug (including radiochemotherapy).

6.2.2. Other Safety Data

6.2.2.1. Laboratory Tests

Laboratory test data such as hematology, blood chemistry/electrolytes, coagulation, urinalysis and stool routine, thyroid function, myocardial zymogram, CA-125, pregnancy test, and viral investigations will be collected at the visit points specified in the protocol.

6.2.2.2. 12-lead Electrocardiogram

Heart rate, QT interval, and P-R interval

Bazett's Correction (msec)

$$QTcB\ (\text{msec}) = \frac{QT\ (ms)}{\sqrt{RR\ (ms)/1000}}$$

Fridericia's Correction (msec)

$$QTcF\ (\text{msec}) = \frac{QT\ (ms)}{\sqrt[3]{RR\ (ms)/1000}}$$

RR interval - If RR interval is not available it will be derived from HR as follows, for the derivation of the QTc corrections

$$RR\ (\text{msec}) = \ 1000 * \frac{60}{HR\ (bpm)}$$

6.2.2.3. Physical Examination and Vital Signs

Vital signs include body temperature, blood pressure, pulse, and respiratory rate.

The physical examination includes general condition, head, face, eyes, ears, nose, throat, oral cavity, skin, lymph nodes, respiratory system, cardiovascular system, gastrointestinal system, genitourinary system, musculoskeletal system, nervous system, mental status, and others.

The specific time points can be found in the visit flow chart.

6.2.2.4. Echocardiogram

Left ventricular ejection fraction (LVEF, %)

6.2.2.5. ECOG Performance Status

The ECOG performance status (PS) will be scored at the pre-set visits, with the judgment rules as follows:

ECOG Score	Criteria
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work

2	Ambulatory and capable of all self-care but unable to carry out any work activities, up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
5	Dead

7. General Considerations

7.1. Reference Start Date and Study Day

The study day corresponding to the investigation or event will be calculated using the following formula with the day of first treatment as the study start date:

If the date of investigation/event is before the study start date, then the study day = investigation/event date - study start date;

If the date of investigation/event is on or after the study start date, then the study day = investigation/event date - study start date + 1.

7.2. Baseline

Unless otherwise specified, the "baseline" in this study is defined as the last non-missing measurement before the first dose of the study drug, including measurements taken on the day of the first dose and before the first dose. If the examination at screening is completed within 3 days before the first dose, there is no need for duplication before the first dose.

If multiple measurements are made on the same day, then:

- If both scheduled and unscheduled visits exist, then the scheduled visit is considered the last visit;
- If more than one scheduled visit exists, the visit with the largest visit number is considered the last visit;
- If there is no scheduled visit, but multiple unscheduled visits exist, the last effective measurement before dosing from the unscheduled visits is taken as the baseline.

7.3. Unscheduled Visits, Retests

Typically, data summarization by visit only includes data from scheduled visits. Data from unscheduled visits are not included in the aforementioned summarization, but are taken into account during the summarization for transposed tables.

If duplicate measurements are obtained at the same visit, the last measurement of that visit is taken as the analysis value and used for analysis.

7.4. Analysis Window

For visits after the baseline, data will be summarized according to the visit time points recorded in the Electronic Case Report Form (eCRF), without considering whether the examination is performed beyond the time window specified by the protocol.

For tabular summarization and analysis by visit, statistical analysis will be conducted according to the scheduled visit time points specified by the protocol, that is, unscheduled time points not specified in the protocol will not be considered. All visits will be listed when the listing is

created.

7.5. Common Calculations

Change from baseline = Visit measurement value - Baseline measurement value

Percentage change from baseline (%) = Change from baseline * 100 / Baseline measurement value

8. Statistical Considerations

8.1. Derived Variables

For detailed explanations of other derived variables, please refer to the corresponding programming specification documents.

8.2. Missing Data

8.2.1. Missing Dates

8.2.1.1. Medical History Date

Handling of missing or incomplete medical history (date of initial diagnosis, local progression or metastasis):

- If only the day is missing, the 15th day will be imputed if it does not contradict with other dates.
- If both the day and month are missing, and the year is prior to that of the study treatment, July 1st will be imputed if it does not contradict with other dates. If the year is the same as the year of first dose, January 1 will be imputed.
- If the year, month, and day are all missing, they will not be imputed and will be treated as missing values.

8.2.1.2. Date of Adverse Event/Concomitant Medication

Handling of missing AE/CM onset date:

- If the year, month, and day of the event are all missing, then imputation should be performed using the data collection date.
- If only the day is missing, but the year and month of onset are consistent with the year and month of the study treatment, then the day of the start of the study treatment will be used for imputation. In other cases, the first day of the month will be used for imputation.
- If both the month and day of the event are missing, but the year of onset is consistent with the year of the study treatment, then the month and day of the start of the study treatment will be imputed. In other cases, January 1 will be used for imputation.

Handling of missing AE/CM end date:

- If the year, month, and day of the end date of event are missing, and the AE is ongoing, then no imputation will be performed.
- If the year, month, and day of the end date of event are missing, and the AE is not ongoing, then imputation will be performed using the data collection date.
- If the day of the end date is missing, then the last day of that month will be used for imputation. If the imputed date is later than the study end date or the death date, then the earliest date among them will be taken as the imputed date.

- If the month of the end date is missing, then December 31 will be imputed. If the imputed date is later than the study end date or the death date, then the earliest date among them will be taken as the imputed date.

If the imputed end date is later than the start date, then the end date will be imputed using the start date.

8.2.1.3. Date of Subsequent Anti-Tumor Treatment

Handling of missing start date of subsequent anti-tumor treatment:

- If only the day is missing, the 1st day will be imputed.
- If both the day and month are missing, January 1 will be imputed.
- If the year, month, and day are all missing, no imputation will be made.

The imputed new date of the anti-tumor treatment will be compared with the date of the last dose of anti-tumor treatment of the subject, and the latest date of "imputed new date of the anti-tumor treatment ", "date of the last treatment dosing +1", and "date of disease progression used to determine the subject's withdrawal" will be taken as the final date of the anti-tumor treatment.

8.2.1.4. Drug Exposure Date

The utmost effort should be made to obtain the actual start and end dates of dosing. If it is impossible to obtain the information, imputation will be made according to the following rules.

Handling of missing or incomplete date of first dose:

- If only day is missing, the 1st day will be imputed. If the imputed start time of dosing is earlier than the date of randomization, then the date of randomization will be used for this imputation.
- If both the day and month are missing, January 1 will be imputed. If the imputed start time of dosing is earlier than the date of randomization, then the date of randomization will be used for this imputation.
- If the year, month, and day are all missing, no imputation will be made.

If the imputed date of the first dose is later than the date of the last dose, then the date of the last dose will be the date of the first dose.

Handling of missing or incomplete date of last dose:

If only day is missing

- If the date of the last dose, as indicated by year and month, is earlier than the end of treatment date, then imputation will be performed using the last day of the month. The limitations of the length of the dosing cycle should also be considered.
- If the year and month are the same as those of the end of treatment date, then the min(end of treatment date, death date) will be used for imputation. The limitations of the length of the dosing cycle should also be considered.

If both the day and month are missing

- If the date of the last dose, as indicated by the year, is earlier than the end of treatment

date, then imputation will be performed using December 31.

- If the year is the same as that of the end of treatment date, then min(end of treatment date, death date) will be used for imputation.

In other cases, min(end of treatment date, death date) will be used for imputation.

8.2.1.5. Protocol Deviation Date

The protocol deviation date will not be subjected to imputation.

8.2.1.6. Date of Death

Handling of missing or incomplete date of death:

- If only the day is missing, the 1st day will be imputed.
- If both the day and month are missing, January 1 will be imputed.
- If the year, month, and day are all missing, the date of death will be imputed with "the last known date of survival +1".

The imputed date of death will be compared with the date the subject was last known to be alive, and the later of the "date of death after imputation" and "last known date of survival +1" will be taken as the final date of death.

8.2.2. Missing Value in Patient Reported Outcomes

If more than one answer is collected for a question, the more severe answer will be retained. If multiple item functions or at least half of the items in the scale are answered, the score of the scale can be proportionally rated based on the items that are not missing. Questions with missing answers are not subjected to imputation.

8.3. Covariates and Subgroups

8.3.1. Covariates

The baseline characteristics of the subjects, such as age, gender, ECOG score, PD-L1 expression, clinical T stage, smoking status, intracranial hypertension status, tumor burden and stratification factors (see section 4.3 for stratification factors) will be used as covariates to analyze their impact on the primary endpoint iPFS and PFS, and secondary endpoints OS, iORR and ORR.

8.3.2. Subgroups

The subgroup settings for the subgroup analysis are as follows:

- Age (< 65 years, ≥ 65 years)
- ECOG score (0, 1)
- Gender (male, female)
- Clinical T stage (T1-T2, T3, T4)
- PD-L1 CPS (< 1 vs. ≥ 1) and TPS (< 1% vs. ≥ 1%)
- Intracranial hypertension status (Yes, No)
- Smoking status (Never, Current or ever)
- Tumor burden (single organ, multiple organ)

- Pathology (squamous cell carcinoma, non-squamous carcinoma)
- Number of intracranial metastases (1-5, >5)
- Planned radiotherapy (Yes, No)

If the number of subjects in a subgroup is insufficient (<=5), then the subgroup should be merged with the nearest subgroup in classification, or subgroup analysis will not be conducted.

The detailed subgroup analysis plan can be found in section 9.2.5.

9. Statistical Analysis

9.1. Summary of Study Data

9.1.1. Disposition of Subjects

In all analysis sets, statistical analysis will be performed by Group A, Group B, and the total number of subjects. For all subjects who have signed the ICF and participated in screening, only the total number of subjects will be used for statistical analysis. All subjects who have signed the ICF and participated in the screening will be summarized by the screening situation: including the total number and percentage of screening failures, and the reasons for screening failures. The reasons for screening failure are recorded on the Enrollment Information page of the eCRF.

The information on subject screening, randomization, and analysis population, as well as the reasons for subject termination of treatment/study will all be tabulated.

9.1.2. Protocol Deviations

The number of subjects with protocol deviations will be summarized by treatment group based on the ITT set, and the levels and categories of protocol deviations will be further summarized.

Before the database lock, the relevant study personnel from the sponsor's project team and investigators will review and discuss all protocol deviations, and identify subjects that need to be removed from the PPS due to major protocol deviations.

Criteria for judging protocol deviations:(Needs further discussion)

The final protocol deviation list (including categories, severity, and whether it is a major protocol deviation) and the inclusion and exclusion list for each analysis set will be finalized at the data review meeting.

9.1.3. Demographic and Baseline Characteristics

Demographic characteristics is recorded on the Demographic Information page of the eCRF. Based on the ITT set, the following demographic baseline characteristics will be summarized by treatment group:

- Age (years) according to measurement data
- Age (< 65 years vs. ≥ 65 years)
- Body mass index (BMI) (mg/kg^2) according to measurement data
- ECOG score (0, 1)
- Gender (male, female)
- Clinical T stage (T1-T2, T3, T4)
- PD-L1 CPS (< 1 vs. ≥ 1) and TPS (< 1% vs. ≥ 1%)

- Intracranial hypertension status (Yes, No)
- Smoking status (Never, Current or ever)
- Tumor burden (single organ, multiple organ)
- Pathology (squamous cell carcinoma, non-squamous carcinoma)
- Number of intracranial metastases (1-5, >5)
- Sum of diameters in intra target lesions at baseline according to measurement data
- Planned radiotherapy (Yes, No)
- Radiotherapy (WBRT, SRT, others)
-

Based on the ITT set, the measurement data such as age, and BMI and sum of diameters in intra target lesions at baseline will be summarized by the treatment group using the number of evaluable subjects (N), mean (Mean) and standard deviation (Stand Deviation), median (Median), minimum (Min), maximum (Max) and other descriptive statistics.

- Age will be calculated as the complete number of years between the date of birth and the date of informed consent (the date of informed consent - date of birth / 365.25 will be rounded down to the nearest whole number +1);
- BMI will be calculated as the weight (kg) / body height squared (m^2), and BSA (m^2) will be calculated using the formula 0.0061 * height (cm) + 0.0128 * weight (kg) - 0.1529 (Stevenson's formula).

Detailed demographic and baseline characteristics of the subjects will be tabulated.

9.2. Efficacy Analysis

All efficacy analyses in this study will be based on the ITT population. Sensitivity analysis of efficacy will be conducted in the FAS and PPS. The software to be used for analysis is SAS version 9.4 or higher.

All study endpoints hypothesis tests will be conducted at a significance level of α=0.05 (two-sided). Hierarchical test procedure are considered to control overall type I error in this study.

9.2.1. Primary Efficacy Analysis

The analysis of the primary efficacy endpoint iPFS is as follows:

Analysis Set	ITT set
Endpoint indicator	iPFS is defined as the time from randomization to the first occurrence of any of the following events: intracranial tumor progression, symptoms related to intracranial hypertension in patients with asymptomatic BM at baseline, worsening of previously controlled intracranial hypertension, or death from any cause; iPFS (months) = [Date of intracranial disease progression/ date of death from any causes/ date of censoring – date of randomization + 1]/30.4375
Analytical procedure	The Kaplan-Meier (KM) method will be used to calculate the 6-month, 12-month iPFS%, and the log(-log) method will be used to calculate the 95% CI of the iPFS. The KM survival curves will be plotted. The stratified log-rank test will be used to compare the difference in iPFS between the study group and the control group, and the stratified Cox proportional hazard model will be used to estimate the HR between the study

	group and the control group and its 95% CI estimated by Brookmeyer and Crowley method. If there is a discrepancy between the results of the log-rank test and the estimation of the Cox proportional hazard model, the results of the log-rank test should be taken as the standard. The score method will be used to process the 95% CI of the HR estimated by the Cox proportional hazard model. Sensitivity analysis: If the Cox model does not meet the proportional hazards assumption, the relative magnitude of the effect between groups will be measured by the difference in restricted mean survival time (RMST) between the study group and the control group.
Handling of missing data, censoring rules	The censoring rules can be found in Appendix 10.2

The analysis of the primary endpoint PFS is as follows:

Analysis Set	ITT set
Endpoint indicator	PFS is defined as the time from randomization to the first occurrence of any of the following events: document overall disease progression or death from any causes; PFS (months) = [Date of overall disease progression/ date of death from any causes/ date of censoring – date of randomization + 1]/30.4375
Analytical procedure	Same as iPFS endpoints
Handling of missing data, censoring rules	The censoring rules can be found in Appendix 10.2

9.2.2. Secondary Efficacy Analysis

Secondary efficacy endpoint: OS

Analysis Set	ITT set
Endpoint indicator	OS refers to the time from the date of randomization of a subject to death from any cause. OS (in months) = [Date of death/date of censoring - date of randomization + 1]/30.4375.
Analytical procedure	Same as iPFS
Handling of missing data, censoring rules	If no death event occurs, the survival time will be censored to the last known date of survival. If there is no data after randomization, the survival time will be censored to the date of randomization.

9.2.3. Other Secondary Efficacy Analyses

For the binary data in the secondary endpoints, such as iORR, ORR and DCR the stratified Cochran-Mantel-Haenszel test will be used to summarize the risk differences (RD) and odds ratio (OR) between groups, the corresponding two-sided 95% CIs, and *p* values.

For the time-event data of DoR in secondary endpoint data, the KM method will be used to calculate the median DoR based on the ITT population, and the 95% CI for the survival time will be calculated using Brookmeyer and Crowley method.

Reserve Kaplan-Meier in OS, which refers reversing events such that loss-to-follow-up are treated as “events” while the outcome events are treated as “censored”, to estimate median follow-up time.

9.2.4. Subgroup Analysis

The following subgroup analysis in primary and secondary endpoints will be conducted on the subgroups listed in Section 8.3.2: the KM method will be used to estimate the difference in median survival time (with its 95% CI estimated using the Brookmeyer Crowley method). The hazard ratios of the study group versus the treatment group within each subgroup will be calculated using the unstratified Cox proportional hazard model. If the number of events in a certain subgroup is less than 10, the subgroup will be combined with the adjacent subgroup for further analysis. Subgroup analyses of iPFS, PFS and OS will be conducted within the ITT population.

9.3. Safety Analysis

All safety analyses will be based on the SS.

For continuous indicators, measurements and changes from baseline will be summarized at each scheduled visit point. For categorical indicators, the number and percentage of subjects in each category will be provided at each scheduled visit point.

For the results of laboratory tests, electrocardiogram status, and physical examinations, the normal and abnormal changes in the clinical significance of each indicator at baseline and after treatment will be summarized in a shift table. The most severe clinical significance judgment result after treatment will be selected for summary, including data from unscheduled visits.

In addition, for laboratory test indicators, the severity will be graded according to CTCAE 5.0, and the grading changes before and after treatment will be summarized in a shift table. The most severe grading after treatment will be selected for summary, including data from unscheduled visits.

All abnormal and clinically significant laboratory test results, vital signs, electrocardiogram results, and physical examination findings will be tabulated.

9.3.1. Adverse Events

AE information is recorded on the Adverse Event page of the eCRF. All AEs will be coded using the MedDRA (v22.0) version, and graded according to the NCI-CTCAE v5.0.

TEAE in this study is defined as: any AE that occurs or worsens from the first dose of radiotherapy to 90 days after the last dose of camrelizumab in the study group, and any AE that occurs or worsens from the first dose of radiotherapy to 28 days after the last dose of the study drug in the control group. All summary tables will be based on TEAEs, and AEs during the screening phase will only be provided in a listing.

If a subject experiences multiple AEs with the same SOC and/or PT coding, the subject will be counted only once at the corresponding SOC/PT level in the statistics. AEs will be sorted in descending order according to the incidence in the study group by SOC/PT. Within the same SOC, PTs will be sorted in descending order according to the incidence in the study group. If at least two PTs are equally proportional, they will be sorted in alphabetical order. If no AEs occur within a certain SOC or PT, then no analysis will be conducted.

If one subject reports multiple AEs of different CTCAE grades, the events with the highest grade will be used for counting. AEs with missing CTCAE grades will be analyzed as Grade 3 AEs.

The correlation with the study treatment is categorized as "definitely related", "possibly related", " unlikely related", "not related", and "not determined". During the analysis, "definitely related", "possibly related", "not determined" or missing values will be considered

as related to the study drug, while "unlikely related" and "not related" will be considered as not related to the study drug.

Summary statistics will be conducted for the following types of AEs:

- All TEAEs, all TEAEs related to the study treatment, and surgical complications;
- TEAEs with CTCAE grade ≥3, TEAEs with CTCAE grade ≥3 and related to the study treatment, surgical complications with CTCAE grade ≥3;
- SAEs during the treatment period, SAEs related to the study treatment;
- TEAEs leading to permanent discontinuation, TEAEs related to study treatment leading to permanent discontinuation;
- TEAEs leading to dose reduction only, TEAEs related to the study treatment leading to dose reduction;
- TEAEs leading to dose interruption, TEAEs related to the study treatment leading to dose interruption;
- TEAEs leading to withdrawal from the study, TEAEs related to study treatment leading to withdrawal from the study;
- TEAEs of special interest, TEAE of special interest related to the study treatment;
- TEAEs leading to death, TEAEs related to study treatment leading to death.

Simultaneously, the above events will be statistically analyzed by the stages of neoadjuvant radiochemotherapy, neoadjuvant drug treatment, and adjuvant drug treatment.

The number and percentage of subjects who experienced the aforementioned types of TEAEs will be summarized by treatment group based on the SS population, and these TEAEs will be further summarized by MedDRA SOC and PT classification for the count and proportion. In addition, a further detailed summary will be conducted on the CTCAE grading of the aforementioned types of TEAEs.

A detailed listing of all AEs, all immune-related adverse events (irAEs), all surgical complications, all TEAEs, treatment-emergent serious adverse events (TESAEs), TEAEs related to the study drug, TEAEs leading to discontinuation of study drug, TEAEs leading to interruption or dose reduction of study drug, TEAEs leading to withdrawal from the study, TEAEs leading to death, and adverse events of special interest (AESIs) will be provided by treatment group and subject number, including the name, SOC, PT, start time, outcome time, outcome, action taken with the study drug, whether it is an SAE, whether it is a TEAE, the correlation with the study drug, whether it is related to medical history, whether concomitant treatment is used, and whether it is an RECCP. Among them, except for the listing of all AEs, all the other listings do not require the item "whether it is a TEAE", and except for the listing of TESAEs, all the other listings do not require the item "whether it is an SAE".

9.3.1.1. Death

The number and percentage of subjects who died during the study treatment and the entire study will be summarized by treatment group in the SS population, and the cause of death will be further subdivided and summarized.

A detailed listing of deaths will be provided by treatment group and subject number, which will include the date of death and cause of death.

9.3.1.2. Other Adverse Events of Special Interest

The number and percentage of subjects with TEAEs of special interest will be summarized by study group and by MedDRA SOC or SMQ, PT in the SS population, and a detailed listing will be provided with the content the same as that of listing of TEAE.

A listing of AESIs is provided in Appendix 11.5.

9.3.2. Clinical Laboratory Evaluations

The laboratory test data is recorded on the Laboratory Test Details page of the eCRF. Baseline values, measured values at scheduled postbaseline visits, and the changes from the baseline will be summarized for consecutively measured laboratory indicators, such as white blood cell count, urine protein, and red blood cells, by treatment group in the SS population.

Measurement values from the laboratory test beyond the normal range may potentially be clinically significant abnormalities. According to the investigator's judgment, the test results can be classified as normal, abnormal but not clinically significant, and abnormal and clinically significant. A comparison of baseline value and the worst investigator-determined post-baseline result will be conducted for each indicator by treatment group in the SS population (transposed table), the denominator for percentage calculation is the number of subjects who have undergone baseline and at least one post-baseline test in SS.

Based on the SS, AEs related to laboratory tests will be graded according to the NCI-CTCAE V5.0 criteria, and then, the baseline value and the worst CTCAE grade after baseline of each indicator will be compared by treatment group (transposed table).

All laboratory results will be listed in detail by subject number, abnormal values will be identified by H/L, and whether they have clinical significance will be indicated. Subjects with laboratory indicators that are abnormal and clinically significant after baseline will also be tabulated.

9.3.3. Vital Signs

The detailed information on the vital signs investigation is recorded on the Vital Signs page in the eCRF. The investigation of vital signs includes: body temperature, pulse, respiratory rate, systolic blood pressure, and diastolic blood pressure.

Details of vital signs and manifest abnormal values will be tabulated.

9.3.4. 12-lead ECG

The detailed information on ECGs is recorded on the 12-lead ECG page in the eCRF. The investigation items of ECG include: HR (beats/minute), PR interval (MS), QT interval (MS), and QTC (MS). If an abnormality (with clinical significance) is found, another two ECGs will be performed, and the mean of the three measurements will be taken for statistical analysis. A transposed table comparing the baseline and post-baseline results (normal, abnormal but not clinically significant, and abnormal and clinically significant) of subjects who have undergone baseline and at least one valid post-baseline examination by the scheduled visit point and treatment group will be provided.

Detailed ECG-related data will also be provided in a listing.

9.3.5. Physical Examination

The detailed information on the physical examination is recorded on the Physical Examination page of the eCRF. The physical examination includes: head, skin system, lymph nodes, eyes,

ears, nose, throat, oral cavity, respiratory system, cardiovascular system, abdomen, genitourinary system, musculoskeletal system, nervous system, mental status, etc. A transposed table comparing the examination results (normal, abnormal, not done) of subjects who have undergone baseline and at least one valid post-baseline examination by the scheduled visit point and treatment group in the SS will be provided.

Detailed information on physical examinations will be provided in a listing.

9.4. Quality-of-life Evaluation Analysis

9.4.1. Quality of Life Scale

The analysis will be performed according to the actual visit points collected and the corresponding content in the CRF.

- Scale compliance
- The LS estimated change from baseline, with a threshold of clinical significance of ±10
- Proportion of improvement in quality of life

10. Appendices

10.1. Schedule of Events

Refer to protocol

10.2. Censoring Rules

intracranial progression-free survival (iPFS) /progression-free survival (PFS) censoring rules:

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessment	Date of randomization	Censored
No on-study tumor assessments and no death	Date of randomization	Censored
Documented progression	Date of first documented progression per RECIST 1.1	Progressed
No progression and no death	Date of last evaluable tumor assessment	Censored
New anticancer therapy, New tumor-directed radiotherapy or tumor-directed surgery received without progression reported prior or on the same day	Date of the last evaluable tumor assessment before subsequent anti-tumor treatment	Censored
Death without progression	Date of Death	Progressed

overall survival (OS) censoring rules:

Situation	Date of Progression or Censoring	Outcome
Death from any causes	Date of death	Death
Alive	Date of last known alive	Censored

10.3. RECIST 1.1 Objective Response Evaluation Table

Table 1 Time Point Response: Subjects with Target (+/- Non-Target) Lesions

Target Lesions	Non-target Lesions	New Lesions	Overall Response
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
CR	NE	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR=complete response, PR=partial response, SD=stable disease, PD=progressive disease, NE=inevaluable

Table 2 Time Point Response: Subjects with Non-Target Disease Only

Non-target Lesions	New Lesions	Overall Response
CR	No	CR
Non-CR/Non-PD	No	Non-CR or Non-PD
Not all evaluated	No	NE
Unequivocal PD	Yes or No	PD
Any	Yes	PD

CR=complete response, PR=partial response, SD=stable disease, PD=progressive disease, NE=inevaluable

Note: “Non-CR/non-PD” is preferred over “SD” for non-target lesions. Since SD is increasingly used as the endpoint for assessment of efficacy in some trials, so non-CR/non-PD efficacy criteria have been developed so to assign this category when no lesions can be measured is not advised. For equivocal findings of progression (e.g., very small and undefined new lesions; cystic changes or necrosis in existing lesions), treatment may continue until the next scheduled assessment. If at the next scheduled assessment, progression is confirmed, the date of progression should be the earlier date when progression was suspected.

Table 3: Best Overall Response Requiring Confirmation for CR and PR Response

Overall Response at First Time Point	Overall Response at Subsequent Time Points	Best Overall Response
CR	CR	CR
CR	PR	SD, PD, or PR[a]
CR	SD	SD, if SD lasts long enough (6 weeks, 42 days), otherwise it should be PD
CR	PD	SD, if SD lasts long enough, otherwise it should be PD
CR	NE	SD, if SD lasts long enough, otherwise it should be NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD, if SD lasts long enough, otherwise it should be PD
PR	NE	SD, if SD lasts long enough, otherwise it

		should be NE
NE	NE	NE

CR=complete response, PR=partial response, SD=stable disease, PD=progressive disease, NE=inevaluable.

a: If a CR is truly present at the first time point and any disease occurs at a subsequent time point, the subject's response will be assessed as PD at a later time point, even if the subject's response meets the criteria for PR relative to baseline (since disease will reappeared after CR). Best response depends on whether minimum duration for SD is met (6 weeks, 42 days). However, sometimes CR may be evaluated at first assessment, but subsequent scans suggest that small lesions are likely still present, and in fact, the subject has PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR, and the best response is PR.

10.4. Creatinine Clearance Calculation

Cockroft-Gault formula:

Serum creatinine concentration in mg/dL:

$$Creatinine\ clearance\ for\ men\ (mL/min) = \frac{(140 - \text{age}) \times \text{weight}}{72 \times \text{serum Cr}}$$

$$Creatinine\ clearance\ for\ women\ (mL/min) = \frac{0.85 \times (140 - \text{age}) \times \text{weight}}{72 \times \text{serum Cr}}$$

Serum creatinine concentration in μmol/L:

$$Creatinine\ clearance\ for\ men\ (mL/min) = \frac{(140 - \text{age}) \times \text{weight}}{0.81 \times \text{serum Cr}}$$

$$Creatinine\ clearance\ for\ women\ (mL/min) = \frac{0.85 \times (140 - \text{age}) \times \text{weight}}{0.81 \times \text{serum Cr}}$$

Note: The unit for age is "years", and the unit for weight is "Kg".

10.5. Adverse Events of Special Interest (AESIs)

Adverse Events of Special Interest (AESIs)