Statistical Analysis Plan

A Randomized, Three-Arm, Open-Label Phase 3b Clinical Trial of Aumolertinib, versus Aumolertinib with Chemotherapy, versus Osimertinib for Patients with Metastatic NSCLC and an EGFR Mutation (TREBLE)

Prepared by: [redacted]

Date of Plan: 16 December 2022

Protocol Version: 4.0

SAP version: 1.0

NCT Number: NCT05493501

Prepared for:
EQRx International, Inc.
50 Hampshire Street
Cambridge, MA 02139
USA

SIGNATURE PAGE

This document has been prepared by:

[redacted]

[Electronic Signature Page Appended]

Signature Date

[redacted]
EQRx International, Inc.

SPONSOR'S SIGNATORIES

As representative(s) of the Sponsor, EQRx International, Inc., I (we) confirm the review and approval of the planned statistical analyses described herein. This statistical analysis plan was subjected to critical review. The information contained herein is consistent with study protocol(s) and applicable regulatory guidance and guidelines.

This document has been reviewed and approved by:

[redacted]

[Electronic Signature Page Appended]

Signature Date

[redacted]
EQRx International, Inc.

[redacted]

[Electronic Signature Page Appended]

Signature Date

[redacted]
EQRx International, Inc.

Table of Contents

Version History

SAP Version	Date	Change	Rationale	Protocol Version
1.0	12/16/2022	Not applicable	Original version	Protocol V4.0

List of Abbreviations

Abbreviation	Definition
AE	Adverse event
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
BICR	Blinded independent central review
CI	Confidence interval
CMH	Cochran-Mantel-Haenszel
CNS	Central nervous system
CPK	Creatinine phosphokinase
CR	Complete response
CTCAE	Common Terminology Criteria for Adverse Events
ctDNA	Circulating tumor deoxyribonucleic acid
DCR	Disease control rate
DepOR	Depth of response
DoR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal growth factor receptor
HR	Hazard ratio
IA	Interim analysis
ICF	Informed consent form
IVRS	Interactive voice response system
IWRS	Interactive Web response system
MedDRA	Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
NCI PRO-CTCAE	National Cancer Institute Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events
NSCLC	Non-small cell lung cancer
ORR	Overall response rate

OS	Overall survival
PD	Progressive disease
PEG	Percutaneous endoscopic gastrostomy
PFS	Progression-free survival
PFS2	Time to second objective disease progression
PK	Pharmacokinetic(s)
PR	Partial response
PRO	Patient-reported outcome
PS	Performance status
QoL	Quality of life
QTcF	QT interval with Fridericia's correction
RANO	Response Assessment in Neuro-Oncology
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SD	Stable disease
TEAE	Treatment-emergent adverse event(s)
TFST	Time to first subsequent anticancer therapy
TGR	Tumor growth rate
TSST	Time to second subsequent anticancer therapy
US	United States
WT	Wild-type

1. Introduction

This SAP focuses on efficacy and safety analyses. PK analysis will be documented in a separate PK SAP.

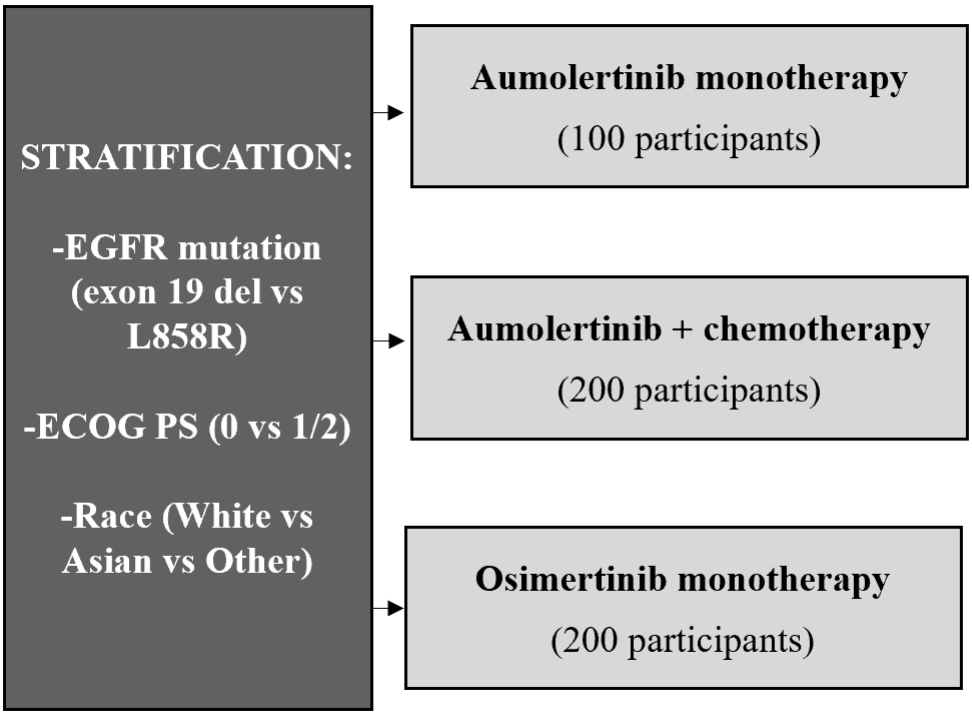
1.1. Objectives and Endpoints

Table 1-1 Objectives and Endpoints

Objectives	Endpoints
Primary	
• To assess the efficacy of aumolertinib in combination with chemotherapy compared with osimertinib monotherapy	• PFS, as assessed by blinded independent central review (BICR) per RECIST v1.1
Key Secondary	
• To further assess the efficacy of aumolertinib in combination with chemotherapy compared with osimertinib monotherapy	• OS
Other Secondary	
• To assess other efficacy measures of aumolertinib in combination with chemotherapy compared with osimertinib monotherapy	• TGR, ORR, DCR, DepOR, and DoR, each assessed by BICR per RECIST v1.1 • PFS, TGR, ORR, DCR, DepOR, and DoR, each assessed by Investigator per RECIST v1.1 • ctDNA clearance
• To assess the efficacy of aumolertinib monotherapy compared with osimertinib monotherapy	• PFS, TGR, ORR, DCR, DepOR, and DoR, each assessed by BICR per RECIST v1.1 • PFS, TGR, ORR, DCR, DepOR, and DoR, each assessed by Investigator per RECIST v1.1 • OS • ctDNA clearance
• To assess safety, including toxicities mediated by EGFR WT inhibition, in the study treatment arms	• TEAEs, vital signs, ECGs, and laboratory test results (including chemistry, hematology, and urinalysis) by treatment arm
• To evaluate selected PROs and QoL, including toxicities mediated by EGFR WT inhibition, in the study population	• NCI PRO-CTCAE questionnaire

Objectives	Endpoints
• To characterize the PK profile of aumolertinib alone and in combination with chemotherapy	• Aumolertinib concentration for population PK analysis
Exploratory	
• To characterize mechanisms of acquired resistance in the treatment population	• Exploratory biomarker analysis of plasma samples banked for the purpose of characterizing resistance mechanisms in participants treated with aumolertinib in combination with chemotherapy, aumolertinib monotherapy, and osimertinib monotherapy
• To evaluate CNS efficacy for participants in the study treatment arms with measurable CNS lesions at baseline	• Description of CNS efficacy as measured by PFS; OS; and ORR, DCR, DepOR, and DoR (each assessed by the Investigator per RANO criteria)
• To evaluate additional efficacy measures in the study treatment arms	• PFS2, as applicable • TFST and TSST, as applicable
• To evaluate the correlation between TGR and ctDNA clearance with PFS and OS in the study treatment arms	• TGR, ctDNA, PFS assessed by the Investigator per RECIST v1.1, and OS
• To evaluate the exposure-response relationship of aumolertinib alone and in combination with chemotherapy	• Exposure-response analysis of safety and efficacy endpoints

1.2. Study Design



Abbreviations: del = deletion; EGFR = epidermal growth factor receptor; ECOG PS = Eastern Cooperative Oncology Group performance status

EQ143-301 (TREBLE) is a Phase 3b multicenter, randomized, 3-arm, open-label study to evaluate aumolertinib + chemotherapy combination therapy compared with osimertinib monotherapy, and to describe aumolertinib monotherapy in comparison with osimertinib monotherapy.

The intended study population is systemic treatment-naïve participants with metastatic EGFR-mutant NSCLC.

The primary study endpoint is efficacy of aumolertinib in combination with chemotherapy compared with osimertinib monotherapy, as measured by PFS (assessed by BICR per RECIST v1.1). PFS is defined as the time from date of participant randomization to either the first date of disease progression (per RECIST v1.1) or the date of death by any cause, whichever occurs first.

Approximately 500 study participants will be randomized in a 2:2:1 ratio (by the stratified permuted block randomization method) into one of 3 treatment arms, as follows:

1. **Aumolertinib with chemotherapy** (platinum-based doublet), as follows ***(approximately 200 participants)***:
    - For adenocarcinoma, either:
        - Aumolertinib + cisplatin with pemetrexed, **or**
        - Aumolertinib + carboplatin with pemetrexed
    - For squamous cell carcinoma, one of the following:
        - Aumolertinib + cisplatin or carboplatin with paclitaxel;
        - Aumolertinib + cisplatin or carboplatin with albumin-bound paclitaxel; **or**
        - Aumolertinib + cisplatin or carboplatin with gemcitabine
2. **Osimertinib alone *(approximately 200 participants)***
3. **Aumolertinib alone *(approximately 100 participants)***

A cycle of treatment is defined as 21 days for all 3 treatment arms; Cycles 2+ have a window of 2 days prior to, or 7 days following, the cycle due date. All imaging assessments are to be performed as scheduled, regardless of any delays in the participant's treatment.

Participants can continue to receive treatment in the study as long as they have not withdrawn consent and are judged by the Investigator to continue to receive clinical benefit, in the absence of meeting treatment and/or study discontinuation criteria as detailed in the protocol.

All Participants will be stratified by the following factors:

- EGFR mutation (ex19del versus L858R)
- ECOG PS (0 versus 1/2)
- Race (White versus Asian versus Other Races Combined)

In this study, aumolertinib will be orally administered as two 55-mg tablets for a total dose of 110 mg once daily.

Participants randomized to receive aumolertinib with chemotherapy will receive up to 4 cycles of a platinum-based doublet with aumolertinib at 110 mg once daily; thereafter, at the Investigator's discretion, based on assessment of continued clinical benefit for the individual participant, the participant may go on to receive maintenance pemetrexed with once-daily aumolertinib for the remainder of their study treatment, provided that they do not meet treatment and/or study discontinuation criteria.

Osimertinib (tablet) will be administered daily at 80 mg once orally.

2. Statistical Hypothesis

The primary analysis is BICR-assessed PFS comparison between aumolertinib + chemotherapy vs Osimertinib. The null hypothesis that the PFS is the same for the two arms will be tested at 2-sided 0.05 significance level.

The hypotheses for the primary endpoint of BICR-assessed PFS are as follows:

- H_0: $S_{aumo\ chemo}(t) = S_{osi}(t)$
- H_A: $S_{aumo\ chemo}(t) > S_{osi}(t)$

where $S_{aumo\ chemo}(t)$ is the function of PFS for the aumolertinib with chemotherapy arm and $S_{osi}(t)$ is the function of PFS of the osimertinib monotherapy arm.

Comparisons on the secondary (except the key secondary) and exploratory endpoints between arms (aumolertinib + chemotherapy vs Osimertinib, aumolertinib monotherapy vs. osimertinib) will be at nominal significance level of 0.05 based on 2-sided tests.

2.1.1. Multiplicity Adjustment

The key secondary endpoint of OS (aumolertinib with chemotherapy vs osimertinib monotherapy) would be tested sequentially at a 2-sided 0.05 Type I error if the result of the primary endpoint analysis (PFS) is statistically significant.

The Lan DeMets approach, which approximates the O'Brien-Fleming spending function, will be used in IAs and final PFS analyses to maintain an overall 2-sided 0.05 Type I error for the primary PFS comparison. After the result of the primary endpoint analysis (PFS) is statistically significant, the Lan DeMets approach will be used in IAs and final OS analyses to maintain an overall 2-sided 0.05 Type I error for the OS comparison.

3. Analysis Sets

For the purposes of analysis, the following analysis sets are defined:

Table 3-1 Analysis Sets

Participant Analysis Set	Description
Full analysis set (FAS)	• The Full Analysis Set will include all randomized participants. Participants will be analyzed in the treatment arm to which they were randomized, and strata (with stratification by EGFR mutation, ECOG PS, and race) will be assigned at the time of randomization.
Safety analysis set (SAS)	• The Safety Analysis Set will include all participants who received at least one dose of study intervention. Safety data will be summarized based on the actual treatment received.
CNS analysis set (cFAS)	• A subset of FAS population that includes all participants at Screening/baseline who have been identified as having at least 1 measurable brain lesion at baseline based on Investigator assessment.

The full analysis set will be used to analyze endpoints related to the efficacy objectives and the safety analysis set will be used to analyze the endpoints and assessments related to safety.

4. Statistical Analyses

4.1. General Considerations

Summary statistics for continuous variables will include, at the very least, n, median, mean, standard deviation, minimum, and maximum. For categorical variables, frequencies and percentages will be presented. For time-to-event variables, the Kaplan-Meier method will be used for descriptive summaries. Graphical displays will be provided as appropriate.

All statistical analysis outputs except tumor growth rate analyses will be produced using SAS® version 9.4 or a later version in a secure and validated environment. Tumor growth rate analyses will be performed in R-4.2.1 or a later version.

4.1.1. Definitions and Conventions

The date of randomization is the date the participant was randomly assigned to aumolertinib or aumolertinib + chemo or osimertinib arm using an IVRS/IWRS.

The date of first dose is the date of the first dose of study treatment.

The baseline value of a safety assessment is the last non-missing value observed prior to the first dose of study treatment.

The baseline value of an efficacy assessment is the last non-missing value observed prior to the date of randomization.

The study day of a safety event or assessment will be calculated as:

- The difference between the date of the event or assessment and the date of first dose plus 1 for all events or assessments occurring on or after the day of first dose.
- The difference between the date of the event or assessment and the date of first dose for all events or assessments occurring before the day of first dose.

The study day of an efficacy event or assessment will be calculated as:

- The difference between the date of the event or assessment and the date of randomization plus 1 for all events or assessments occurring on or after the date of randomization.
- The difference between the date of the event or assessment and the date of randomization for all events or assessments occurring before the date of randomization.

One month is defined as 365/12 days.

4.1.2. Visit Window Rules

Safety summaries of vital signs, laboratory tests, ECG, ECOG PS score and LVEF will be presented by the analysis visits. Non-missing assessments from all scheduled and unscheduled visits will mapped to an appropriate analysis visit using a windowing scheme. Analysis visit windows are defined in **Table 4-1**

If there is more than one value in the same visit window as defined above, the closest value to the scheduled study day will be used in summaries, or the earlier in the case that values are equidistant from the nominal visit date. The other visits will not be used in the summary table. If multiple observations on the same day, then the average of those values will be used in summary of that visit. For non-numeric results, the best value will be taken as baseline and the worst value will be taken for a post-baseline visit as this is the most conservative. In derivation of extreme values such as minimum/maximum post-baseline value, all post baseline values including those collected at unscheduled visits will be considered.

Table 4-1 Visit Windows

Serial number	Analysis visit name	Scheduled study day	Lower limit of the window	Upper limit of the window
1	Baseline	1	-28	1
2	C2D1	22	2	32
3	C3D1	43	33	53
4	C4D1	64	54	74
5	C5D1	85	75	95
6	C6D1	106	96	116
7	C7D1	127	117	169
8	C11D1	211	170	253
9	C15D1	295	254	337
10	C19D1	379	338	421
	C*x*D1	(*x*-1)×21+1	Study Day- 41	Study Day+ 42

4.1.3. Handling of Dropouts or Missing Data

In general, missing baselines will not be imputed. On summary tables, the following approaches will be used:

- Categorical data will be summarized based on the denominator of the total number of participants in a corresponding treatment arm. Missing data may be presented as a separate category.
- Continuous data will be summarized based on observed data only.

4.1.3.1. Handling of Partial Dates for Concomitant Medications and Adverse Events.

In general, the imputation should be conservative. The start dates should be imputed to be as early as possible, and the end date should be imputed to be as late as possible. If the partial start date of medication or AE cannot indicate whether it starts prior to study treatment or after, the medication or AE will be classified as concomitant or treatment-emergent.

Imputation of concomitant medication and adverse events start dates

- If year is missing (or completely missing), do not impute
- If year is present, and month and day are missing:
  - Impute as January 1st if the year is different to the first dose date or the end date is prior to the first dose date
  - Impute as the first dose date if the year is the same as the first dose date and the end date is on or after the first dose date.
- If year and month are present and day is missing:
  - Impute as the first day of the month, if the year or month are different to the first dose date or the end date is prior to the first dose date
  - Impute as the first dose date if the year and month are the same as the first dose date and the end date is on or after the first dose date.

Imputation of concomitant medication and adverse events end dates

- If year is missing (or completely missing), do not impute
- If year is present, and month and day are missing, impute as the earliest of:
  - December 31st
  - Date of Data cutoff
  - Date of Death
- If year and month are present and day is missing, impute as the earliest of:
  - The last day of the month
  - Date of Data cutoff
  - Date of Death

4.1.3.2. Handling of Missing Dates/Months/Years for Date of Death

For subjects who are confirmed dead, if the complete date of death cannot be known, the rules for imputation are as follows:

- If only the day is missing, impute as the 1st date of the month or the known last date of survival + 1, whichever is later;
- If only the month and day are missing, impute as the January 1 of the year or the known last date of survival + 1, whichever is later;
- If the year, month, and day are all missing, impute as the known last date of survival + 1.

4.1.3.3. Handling of Partial Dates for Subsequent Anti-Cancer Therapy

Imputation of start dates

- If year is missing (or completely missing), do not impute
- If year is present, and month and day are missing:
  - Impute as January 1st if the year after the last dose date
  - Impute as the last dose date + 1 if the year is the same as the last dose date
- If year and month are present and day is missing:
  - Impute as the first day of the month, if the year/month is after the last dose date
  - Impute as the last dose date + 1 if the year and month are the same as the last dose date

Imputation of end dates

- If year is missing (or completely missing), do not impute
- If year is present, and month and day are missing, impute as the earliest of
  - December 31st
  - Date of Data cutoff
  - Date of Death
- If year and month are present and day is missing, impute as the earliest of
  - The last day of the month
  - Date of Data cutoff
  - Date of Death

Imputation of progression dates of subsequent anti-cancer therapy

- If year is missing (or completely missing), do not impute
- If year is present, and month and day are missing, impute as the latest of
  - January 1st
  - Start date of the therapy
- If year and month are present and day is missing, impute as the latest of
  - The last day of the month
  - Start date of the therapy

4.1.4. Participant Disposition

A detailed description of participant disposition will be provided. It will include a summary of the number and percentage of participants randomized in the study, and treated in the study, reasons for discontinuation from study treatment (randomized population), and reasons for discontinuation from study (randomized population). Reason for discontinuation from both study treatment and the study will be summarized by pre-determined categories. Randomized population's disposition data (dates and reasons for both from study treatment and from study) will be listed.

4.1.5. Participant Characteristics

4.1.5.1. Demographics and Baseline Characteristics

Race, Ethnicity, Age, Sex, Height, Baseline Weight, Body mass index (BMI), and smoking status will be listed and summarized.

4.1.5.2. Baseline Disease Characteristics

Disease characteristics will include the following:

- EGFR gene mutation type (Ex19del or L858R)
- Histologic type
- Disease stage at study entry
- Brain metastasis
- Baseline ECOG PS

4.1.5.3. Medical History and Current Condition

Preexisting conditions are defined as adverse events that begin prior to the first dose of study drug.

Preexisting conditions coded by MedDRA and graded per the NCI CTCAE version 5.0 will be presented in tables and listings using MedDRA system organ class and preferred term.

4.1.5.4. Prior Cancer Surgeries/Radiotherapies

Prior cancer surgeries/radiotherapies will be listed.

4.1.5.5. Prior Cancer Therapies

Prior cancer therapies will be listed.

4.1.5.6. Subsequent Anti-Cancer Therapies

Therapies received following study treatment discontinuation will be summarized by arm. Therapies will be summarized overall and by category.

4.1.5.7. Concomitant Therapy

All medications will be coded to the generic preferred name according to the current World Health Organization (WHO) drug dictionary. All concomitant medications will be summarized for the FAS population based on the Anatomical Therapeutic Chemical (ATC) classification the first level (ATC1), the third level (ATC 3) and preferred name.

4.2. Efficacy Analysis

Participant response will be assessed by BICR and the Investigator per RECIST v1.1. CNS efficacy will be evaluated by investigator per RANO-brain metastases criteria (Chukwueke & Wen, 2019).

All efficacy analyses will be performed on the FAS, based on the treatment arm assigned in the randomization. Additionally, CNS efficacy will be based on cFAS.

All stratified models and tests will be based on the stratification information used for randomization. If randomization strata have few participants/events due to the three race categories, race might be combined to two categories in the stratified models.

Primary efficacy endpoint will be analyzed when 251 PFS events as assessed by BICR are reached on the aumolertinib with chemotherapy and Osimertinib monotherapy arms.

Please refer to **Table 4-2** below for efficacy endpoints and the corresponding definitions.

Table 4-2 Efficacy Endpoints and Definitions

Endpoint	Definition
Primary Endpoint	

Progression-free survival (PFS)	The duration from date of participant randomization to either the first date of disease progression or the date of death by any cause, whichever occurs first. Duration is calculated as end date – start date +1. For instance, if a PFS event occurs, then PFS time (in days) is defined as event date – date of randomization +1. If a censoring event occurs, then PFS time is defined as the censoring date – date of randomization +1. Please refer to **Table 4-3** for details on PFS outcome status (PFS event vs. censored) and censoring definitions.
Secondary Endpoints	
Overall survival (OS)	Duration from randomization until the date of death due to any cause. If death event did not occur during the follow-up period, the participant is censored at the last contact date.
Objective response rate (ORR)	Proportion of participants who achieve a CR or PR at any time before PD or initiating a subsequent anticancer therapy.
Disease control rate (DCR)	Proportion of participants who have a best overall response of CR or PR or SD.
Duration of response (DoR)	Duration from the date of first response (CR or PR) until the date of PD or death due to any cause, whichever occurs first. DoR will follow the same outcome and censoring definition used for the primary PFS.
Time to Response (TTR)	Duration from randomization until to the date of first response (CR or PR).
Depth of Response (DepOR)	Defined as the relative change in target lesion tumor size (calculated as the sum of the longest diameters of the target lesions, in the absence of new lesions or progression of nontarget lesions) as compared to baseline.

Tumor growth rate	A constant exponential growth rate estimated using the Sum of Longest Diameters, see Section 4.2.4.2 for details.
Circulating tumor DNA (ctDNA) clearance at 6 weeks	Defined as when a detectable EGFR mutation in ctDNA at baseline becomes undetectable or drops below a predetermined threshold at 6 weeks.
Exploratory Endpoints	
Time to second objective disease progression (PFS2)	Duration from randomization until the date of the first progression event subsequent to that used in the primary analysis of PFS, or death due to any reason, whichever occurs first. Please refer to **Table 4-4** for details on the outcome and censoring definitions.
Time to first subsequent anticancer therapies (TFST)	Duration from randomization to the start date of the first anticancer therapy subsequent to the discontinuation of the study intervention, or death due to any reason, whichever occurs first. Any participant not known to have received a first subsequent anti-cancer therapy, or to have not died at the time of the analysis, will be censored at the last contact date.
Time to second subsequent anticancer therapies (TSST)	Defined as the time from randomization to the start date of the second subsequent anticancer therapy after the discontinuation of the study intervention, or death due to any reason, whichever occurs first. Any participant not known to have received a second subsequent anti-cancer therapy, or to have not died at the time of the analysis, was to be censored at the last contact date.

Table 4-3 Primary PFS outcome and censoring definition

Situation	Date of event or censoring	Outcome

No baseline or post-baseline tumor assessment	Randomization	Censored
Death or PD after a gap of 2 or more consecutively missed tumor assessments	Date of last adequate tumor assessment* before the gap	Censored
Death or PD without a gap of 2 or more consecutively missed tumor assessments	Date of death or PD, whichever occurs first	PFS event
No PD or death on or before a. database cut, b. discontinuation of study treatment plus 14 days, c. start of new anti-cancer treatment plus 14 days, whichever occurs first	Date of last adequate tumor assessment on or prior to the earliest occurrence of the events listed on the left column	Censored

* Adequate tumor assessment will have an overall response of CR, PR, SD, PD.

Table 4-4 PFS2 outcome and censoring definition

Situation	**Date of event or censoring**	**Outcome**
Alive and no subsequent anti-cancer therapy	Last contact date	Censored
Death (by any cause) prior to starting any subsequent anti-cancer therapy	Date of death	Event
No progression or death during the first subsequent anti-cancer therapy and have started second subsequent anti-cancer therapy	Start date of the second subsequent anti-cancer therapy	Censored
No progression or death during the first subsequent anti-cancer therapy and have not started second subsequent anti-cancer therapy	The end date of the first subsequent anti-cancer therapy, or the last contact date if the end date is missing	Censored
Progression or death (by any cause) during the first subsequent anti-cancer therapy	The earlier date of disease progression date or date of death	Event

4.2.1. Analysis of Primary Endpoint

The primary endpoint of PFS as determined by BICR assessment per RECIST v1.1 will be analyzed when 251 PFS events are reached in the aumolertinib with chemotherapy and Osimertinib treatment arms. The primary analysis on PFS will be based on BICR assessment and using the stratified log-rank test to compare aumolertinib with chemotherapy versus Osimertinib monotherapy (2-sided 0.05 level).

The median PFS will be evaluated with a 95% CI for each treatment arm, using the Kaplan-Meier method. HR with a 95% CI will be estimated by a stratified Cox proportional hazard model, using the Breslow method of tie handling, with treatment as the factor. The strata will be the same stratification factors (EGFR mutation, ECOG PS, and race) used in the randomization.

4.2.2. Sensitivity Analysis of PFS

The following sensitivity analyses of PFS will be conducted in addition to the primary PFS analysis.

4.2.2.1. Analysis based on the PFS outcome and censoring definition for Sensitivity Analysis

PFS analysis will be repeated based on the following PFS outcome and censoring definition for sensitivity analysis as listed on the **Table 4-5**.

Table 4-5 PFS outcome and censoring definition for Sensitivity Analysis

Situation	Date of event or censoring	Outcome
No baseline or post-baseline tumor assessment	Randomization	Censored
No PD or death before database cut	Date of last adequate tumor assessment* on or prior to the database cut	Censored
Death or PD before database cut	Date of PD or death, whichever occurs first	PFS event

* Adequate tumor assessment will have an overall response of CR, PR, SD, PD.

4.2.2.2. Analysis based on the actual stratification information collected during the screening period

PFS analysis will be repeated based the actual stratification information collected during the screening period if more than 10 patients have screening stratification factors different than those entered in IRT. When the central retrospective EGFR testing data is available, the PFS analysis will be repeated using the central EGFR mutation as one of the stratification factors if more than 10 patients have central EGFR mutation status different than those entered in IRT.

4.2.3. Analysis of the Key Secondary Endpoints

The final OS analysis will be performed when 200 OS events are reached in the aumolertinib with chemotherapy and osimertinib monotherapy treatment arms. The OS comparison between aumolertinib with chemotherapy vs osimertinib monotherapy would be tested sequentially at a 2-sided 0.05 Type I error if the result of the primary endpoint analysis (PFS) is statistically significant. The OS comparison will use the same methods as the one used for primary PFS analyses.

4.2.4. Analysis of the Other Secondary Endpoints

Each other secondary endpoints will be analyzed by the following comparisons

- aumolertinib with chemotherapy vs. osimertinib
- aumolertinib monotherapy vs. osimertinib

Secondary tumor-based efficacy endpoints will include both BICR and Investigator assessments per RECIST v1.1.

Statistical comparison will be at nominal significance level of 0.05 based on 2-sided tests.

4.2.4.1. Analysis of Binary Endpoints

Secondary binary endpoints including ORR and DCR will be analyzed using the CMH test stratified by the randomization stratification factors. The CMH estimate of the odds ratio with a 95% CI will be reported. A 2-sided p-value will be provided to compare aumolertinib with chemotherapy vs Osimertinib, and aumolertinib monotherapy vs. osimertinib. For each arm, the point estimate for ORR and DCR with its exact 95% CI will be provided.

ORR and DCR will be calculated among FAS and among the participants with baseline measurable diseases in the FAS.

4.2.4.1. Analysis of Time to Event Endpoints

Other time-to-event secondary endpoints, including PFS, DoR and OS, will be analyzed using the stratified log-rank test and the stratified Cox proportional hazard model, as described above for the primary endpoint analysis. Specifically, these secondary endpoint analyses will be performed to compare aumolertinib with chemotherapy versus osimertinib monotherapy and aumolertinib monotherapy versus osimertinib monotherapy. The median event time will be evaluated with a 95% CI using the Kaplan-Meier method.

TTR will be summarized descriptively for each arm.

4.2.4.1. Analysis of ctDNA clearance

The rate of ctDNA clearance at 6 weeks will be evaluated as follows:

$$100 \times \frac{\text{Number of Participants who are ctDNA}-\text{negative at 6 weeks}}{\text{Number of Participants who are ctDNA}-\text{positive at baseline}}$$

ctDNA clearance at C1D1 (predose) is defined as the baseline.

The rate of ctDNA clearance will be analyzed using the same method as described for binary endpoints.

4.2.4.2. Analysis of Tumor Growth Rate

The following regression-growth models based on the constant exponential growth rate (***g***) and constant exponential decay (***d***) were proposed by Stein et al in 2011 (Stein, et al., 2011) and were retrospectively studied (Claret, et al., 2018; Gong, et al., 2020; Bruno, Mercier, & Claret, 2014),

(1) $f(t) = \exp(-\boldsymbol{d} \bullet t) + \exp(\boldsymbol{g} \bullet t) - 1$

(2) $f(t) = \exp(\boldsymbol{g} \bullet t)$

(3) $f(t) = \exp(-\boldsymbol{d} \bullet t)$

(4) $f(t) = \boldsymbol{f} \exp(-\boldsymbol{d} \bullet t) + (1-\boldsymbol{f}) \exp(\boldsymbol{g} \bullet t)$

Here, *f(t)* represents tumor burden (the sum of longest diameters of target lesions) at time t (in days), relative to a quantity of 1, at t = 0. ***g*** (in $days^{-1}$) is the tumor growth rate (TGR). ***d*** (in $days^{-1}$) is the rate of ***d***ecay. ***f*** is the treatment-sensitive fraction of tumor and (1– ***f***) the fraction absolutely or relatively resistant to treatment.

The basic starting model ***gd*** is described by **Equation (1)**. In cases where the data demonstrate only an increase in tumor burden from the beginning of therapy (i.e., only ***g*** differs significantly from 0 with $p<0.1$), ***d*** is eliminated, and tumor growth rate is estimated using the ***gx*** equation [**Equation (2**)]. Similarly, in cases where the data demonstrate only a reduction in tumor burden from the beginning of therapy (i.e., only ***d*** differs significantly from 0 with $p< 0.1$), ***g*** is eliminated, and the rate of tumor decay is estimated using the ***dx*** equation [**Equation (3**)]. ***f*** represents the proportion of tumor that is sensitive to therapy. In cases where the data allow the estimation of three parameters, ***f*** can be estimated using the ***gdf*** equation [**Equation (4)]**. Incorporation of time (*t*) in the equations renders the analysis indifferent to time, or intervals of assessment used by the study.

The *tumgr* package (https://cran.r-project.org/web/packages/tumgr/index.html) in R sequentially applies **Equations 1-4** and selects the equation that best fits the tumor's data. Values for ***g***, ***d***, or ***f*** were estimated only if the fit of the data had a p-value of <0.1. In cases where all parameters were significant predictors of tumor quantity in more than one model, the model that best minimized the Akaike Information Criterion (AIC) was selected.

In this study, we will analyze TGR based on the sum of longest diameters over time using the R package *tumgr*. The participants who had measurable diseases at baseline and at least one post-baseline tumor assessment will be included in the tumor growth rate analyses. The estimated participants' tumor growth rates will be summarized by treatment arm. Box plots of the tumor growth rates will be generated side by side to visualize the tumor growth rates of each arm. For all three arms, the estimated median ***g*** with its 95% Wilcoxon CI will be reported.

A sensitivity analysis to only fit the ***gd*** model [**Equation (1)**] will be conducted.

4.2.4.3. Analysis of DepOR

DepOR will be analyzed among participants with measurable lesions at baseline. Absolute and percentage change from baseline in target lesion tumor size (the sum of the longest diameters of

target lesions) will be summarized using descriptive statistics at each visit for all 3 treatment arms. The best DepOR values will be calculated from assessments before disease progression or before starting subsequent anticancer treatment. Waterfall plots will be provided for each arm.

4.2.5. Analysis of Exploratory Endpoints

4.2.5.1. Analysis of PFS2, TFST, and TSST

Time-to-event endpoints of PFS2, TFST, and TSST will be analyzed on all 3 treatment arms using the stratified log-rank test and the stratified Cox proportional hazard model as described above for the primary endpoint analysis. The median event time will be evaluated with a 95% CI using the Kaplan-Meier method for each arm.

4.2.5.2. Correlation between TGR and ctDNA clearance with PFS and OS in the study treatment arms

PFS and OS KM curves and median survival times by the quartiles of tumor growth rates of each treatment arms will be generated to characterize the correlation between tumor growth rates and PFS and OS.

Correlation between ctDNA clearance with PFS and OS will be analyzed for available data.

4.2.5.3. Description of CNS Efficacy

CNS efficacy will be assessed on the cFAS which includes all participants at Screening/baseline who have been identified as having at least 1 measurable brain lesion at baseline based on Investigator assessment.

CNS efficacy endpoints will include PFS, OS, DoR, ORR, DCR, and DepOR based on Investigator assessment on the CNS scans per RANO criteria.

For time-to-event CNS endpoints, including PFS, OS, and DoR, the median event time will be evaluated with a 95% CI for each treatment arm using the Kaplan-Meier method.

For binary CNS endpoints, including ORR and DCR, the point estimate of the rate with its exact 95% CI will be calculated for each arm.

DepOR will be analyzed using the same method as for the secondary endpoint of DepOR.

4.3. Safety Analyses

Safety analyses will be performed on the SAS and will be presented by the actual treatment arm as defined in Section 3.

4.3.1. Extent of Exposure

Extent of exposure to study drug parameters listed below will be summarized using descriptive statistics.

Cumulative total actual dose (mg) = sum of the actual daily dose subject administrated.

Total exposure time (days) = (date of last dose of study drug - date of first dose of study drug) + (1 for oral drugs or dose interval for intravenous drugs).

Actual exposure time (days) = total exposure time – total duration of days not receiving study drug.

Actual dose intensity (mg/day) = cumulative total actual dose (mg) / total exposure time (days).

Relative dose intensity (%) = cumulative total actual dose (mg) / total planned dose (mg) *100.

Medication compliance for oral drugs (%) = actual exposure time (days) / total exposure time (days) × 100%.

The number of participants with any dose adjustment will be presented for entire treatment period. The number of participants with dose reductions, dose interruption will also be summarized.

4.3.2. Adverse Events

Adverse events coded by MedDRA and graded per the current version of the NCI CTCAE version 5.0 will be presented in tables and listings using MedDRA system organ class and preferred term.

Treatment-Emergent Adverse Events (TEAEs) are defined as AEs that start on or after the first dose of any study intervention or a medical history condition worsened in severity after the first dose, through 28 days after the last dose of any study intervention. Any events that start after 28 days after the last dose of any study intervention, or after a participant receives a subsequent anticancer therapy, will be included in the data listing but will not be considered as a TEAE.

Statistical analysis of treatment related AEs will utilize the Investigator-assessed relationship between study intervention and the occurrence of each AE, which include “Definitely Related”, “Probably Related”, and “Possibly Related” as detailed in Section 10.2.3. (Assessment of Causality) of the Protocol V4. If the relationship of AE is missing, then AE will be summarized as related to study treatment.

In the event that a participant has repeat episodes of the same AE, then the event with the highest severity and/or strongest causal relationship to treatment will be used for the purpose of statistical tabulations.

Overall Summary of TEAEs will summarize the incidence rate of TEAE, grade ≥ 3 TEAE, SAE, AE leading to dose modification (reduction and interruption), permanent treatment discontinuation, study discontinuation and death on each treatment arm. All the above-mentioned AE parameters that is related study treatment per investigator assessment will also be summarized in the AE overview table.

The following summaries will be produced by PT within SOC: TEAEs, grade ≥ 3 TEAEs, SAEs AE leading to dose modification (reduction and interruption), permanent treatment discontinuation, study discontinuation and death on each treatment arm.

All the summary tables mentioned in the last paragraph will be repeated on the treatment related AE parameters.

TEAEs and related TEAEs will be summarized by System Organ Class, Preferred Term and CTCAE Maximum Grade.

4.3.3. Adverse Events of Clinical Interest

The following events/groups of events were considered AECIs:

- Diarrhea
- Hepatotoxicity
- Elevated blood CPK
- Interstitial lung disease
- Cardiac failure
- QT prolongation
- Ocular toxicity

Where available and appropriate, these AECIs will be analyzed using SMQs. However, diarrhea has been evaluated based on the single PT, "diarrhea," and ocular toxicity has been evaluated using a company-specific basket of MedDRA PTs. The search strategy of PTs for each AECI is summarized in Appendix 1, which may be updated based on the newer MedDRA version.

A summary of treatment emergent AECI by PT will be performed for each prespecified AECI group including:

- TEAEs
- Grade > 3 TEAEs

Additional summaries of time to onset of first AE for each AECI category and each preferred term within it; time to onset of first CTCAE grade three or higher will be produced.

In addition to the descriptively summarizing AEs incidence number and percentage by treatment arms, 95% Clopper-Pearson confidence interval will be calculated for selected AECI categories or preferred term within it including diarrhea, rash and CPK elevation to facilitate informal comparisons between two arms, especially aumolertinib arm versus Osimertinib arm. In addition, to the investigator reported toxicity, these same events will be analyzed using the patient reported outcomes using NCI PRO-CTCAE (see Section 4.4.1).

4.3.4. Deaths

All deaths will be listed along with the reason for death, if known. For those deaths attributed to an AE, the listing will include the PT of the AE. A summary of deaths including reasons for death will be produced.

4.3.5. Clinical Laboratory Tests

Selected laboratory tests that are performed at scheduled time points per Protocol will be summarized. All lab results will be listed.

For each treatment arm, the actual value and the change from baseline will be summarized for selected clinical hematology and chemistry laboratory tests.

The severity of selected hematology and chemistry laboratory tests will be evaluated by the NCI CTCAE version 5.0 criteria if the tests have a corresponding CTCAE classification. These calculated grades will be summarized by cycle and maximum post-baseline grade over the entire study.

Additionally, the following summary will be produced:

- Summary of Subjects with Clinically Significant Abnormal Clinical Hematology Laboratory Tests by Visit
- Summary of Subjects with Clinically Significant Abnormal Clinical Biochemistry Laboratory Tests by Visit
- Summary of Subjects with Clinically Significant Abnormal Clinical Urinalysis Laboratory Tests by Visit
- Hematology Results Shift from Baseline according to Clinically Significant by Visit
- Clinical Biochemistry Results Shift from Baseline according to Clinically Significant by Visit
- Urinalysis Results Shift from Baseline according to Clinically Significant by Visit
- Summary of Normal or Abnormal Not Clinically Significant Hematology Laboratory Tests at Baseline that Turned to Abnormal Clinically Significant after Start of Treatment
- Summary of Normal or Abnormal Not Clinically Significant Clinical Biochemistry Laboratory Tests at Baseline that Turned to Abnormal Clinically Significant after Start of Treatment
- Summary of Normal or Abnormal Not Clinically Significant Urinalysis Laboratory Tests at Baseline that Turned to Abnormal Clinically Significant after Start of Treatment
- Scatter plots of maximum AST/ALT versus maximum total bilirubin during treatment with Hy's Law Ranges

4.3.6. Vital Signs

For each treatment arm, the actual value and the change from baseline of at each study evaluation will be summarized for vital signs including body temperature, systolic and diastolic blood pressures, pulse rate (after 5 minutes resting) and respiration rate, as well as body weight.

4.3.7. ECOG PS

ECOG PS at each study assessment will be listed and summarized.

4.3.8. Physical Examination Findings

For physical examinations, findings considered clinically significant by the Investigator that were present before the participant provided signed informed consent will be reported as medical history; findings considered clinically significant by the Investigator that occurred after the

provision of signed informed consent, and which meet the definition of an AE (or SAE, as applicable) as detailed in protocol Section 10.2, will be reported as AEs/SAEs.

4.3.9. Electrocardiograms

ECG results will be summarized descriptively. Actual values and changes from baseline will be reported for each collection visit/timepoint. The summary of ECG overall interpretation (normal/abnormal NCS/abnormal CS) at each post-baseline assessment, as well as shift from baseline overall interpretation will be produced. QTcF absolute value categorized by ≤450, >450 – 480, >480 – 500, >500, and change from baseline at categories of ≤30, >30 – 60, >60 will be summarized at each assessment, and the maximum post-baseline value and change from baseline.

4.3.10. ECHO or MUGA (LVEF)

At each post-baseline assessment, and among all the post-baseline assessments, subjects with "LVEF value <50% and decrease from baseline of >=10%" and with "LVEF value >=50% and decrease from baseline of >=15%" will be summarized. LVEF at each assessment will be listed.

4.3.11. Additional Safety Assessments

Serum Pregnancy Test and SARS-CoV-2 Test are listed separately if data is available.

4.4. Other Analyses

4.4.1. Patient-reported Outcomes and Quality of Life Analysis

PROs and QoL will be assessed via NCI PRO-CTCAE questionnaire. A descriptive summary over time will be provided.

4.4.2. Pharmacokinetic Analyses

Concentrations of aumolertinib and its metabolite HAS-719 in plasma will be determined using validated bioanalytical methods. A population PK model will be developed to characterize the PK of aumolertinib and its metabolite (as applicable). Data from this study may be combined with data from other studies in a meta-population analysis using nonlinear mixed-effects modeling techniques. Exposure-response analysis to evaluate the relationship between aumolertinib exposure and efficacy/safety endpoints in this study may be performed if the data allows for this analysis to be completed. Details of the population pharmacokinetic and exposure-response analyses will be described in a separate pharmacometrics analysis plan.

4.4.3. Biomarker Analysis

Exploratory biomarker analysis of plasma samples banked for the purpose of characterizing resistance mechanisms in participants treated with aumolertinib alone, versus aumolertinib and chemotherapy, versus Osimertinib will be dependent on the data availability. Any biomarker data generated may be reported separately.

4.4.4. Subgroup Analyses

- EGFR mutation (ex19del versus L858R)
- ECOG PS (0versus 1/2)
- Brain metastasis (yes versus no)

4.4.5. Protocol Deviation

Protocol deviations will be listed by participant. The major protocol deviations will be summarized by the treatment arm.

4.5. Interim Analyses

Two interim analyses (IAs) are planned, as follows:

- The first interim analysis will be performed approximately 12 weeks after enrollment is completed in all 3 treatment arms.
- The second interim analysis will be performed upon occurrence of 70% of the projected PFS events.

The interim analyses will include PFS and the other endpoints listed below:

- ORR and DCR
- ctDNA clearance
- TGR
- AEs

The Lan DeMets approach, which approximates the O'Brien-Fleming spending function, will be used to maintain an overall 2-sided 0.05 Type I error for the primary PFS comparison.

The key secondary endpoint of OS (aumolertinib with chemotherapy vs osimertinib monotherapy) would be tested sequentially at a 2-sided 0.05 Type I error if the result of the primary endpoint analysis (PFS) is statistically significant. Interim OS analyses will be conducted at the time of the primary PFS analysis if the PFS analysis is positive (at one of the IAs or at the final PFS analysis) and all subsequent planned analysis points. Type I error for OS will be controlled using Lan-DeMets method with O'Brien-Fleming boundary.

All the secondary endpoints for descriptive comparison between aumolertib monotherapy and Osimertinib monotherapy will be performed at each IA.

Efficacy endpoints will be analyzed as described in Section 4.2. Safety endpoints will be analyzed as described in Section 4.3.

4.6. Changes to Protocol-planned Analyses

There are no changes to Protocol (V4.0) planned analyses.

5. Sample Size Determination

Assuming a randomization ratio of 1:1 for aumolertinib with chemotherapy: Osimertinib monotherapy, a true HR of 0.7 (median PFS improves from 19 months to 27.1 months), and an exponential distribution for PFS, 251 PFS events will provide 80% power at a 2-sided Type I error of 0.05. If the result of the primary endpoint analysis is statistically significant, a sequential testing method will be used to test the key secondary endpoint of OS to control the study overall type I error at 2-sided 0.05. Assuming a true HR of 0.70 (ie, median OS improves from 38.6 months to 55.1 months), and an exponential distribution for OS, 200 OS events will provide 71.1% power at a 2-sided Type I error of 0.05.

Approximately 400 participants in total will be enrolled in the aumolertinib with chemotherapy and osimertinib treatment arms (200 per arm) to reach 251 PFS events. The sample size of 100 participants in the aumolertinib monotherapy arm will allow the precision of estimating an incidence rate for toxicity to be within ± 10.3% (eg, incidence rate, 50%; Clopper-Pearson 95% CI, 39.8%, 60.2%). Therefore, approximately 500 total participants are planned to be randomized at a 2:2:1 ratio to receive aumolertinib in combination with chemotherapy, osimertinib monotherapy, and aumolertinib monotherapy, respectively.

Cytel East® version 6.5 was used for the sample size calculations.

6. Supporting Documentation

AECI	MedDRA v24.0 SMQ /PT*
Diarrhea	PT-Diarrhea
Hepatotoxicity	Drug related hepatic disorders – comprehensive search (narrow SMQ) [20000006]
Elevated blood creatine phosphokinase	Rhabdomyolysis/myopathy (SMQ) [20000002]
Interstitial lung disease (IDL)	Interstitial lung disease (narrow-SMQ) [20000042]
Cardiac failure	Cardiac failure (narrow-SMQ) [20000004]
QTc prolongation	Torsade de pointes/QT prolongation (SMQ) [20000001]
Ocular toxicity	Includes the following PTs (selected based on known class effects of EGFR TKIs): • Dry eye • Lacrimation increased • Conjunctival disorder • Conjunctival haemorrhage • Conjunctivitis • Ocular congestion • Eye haemorrhage • Eye discharge • Eye swelling • Ocular discomfort • Vision blurred • Visual impairment • Photophobia • Eye pain

* The search strategy of preferred terms for AECIs will be updated at the analysis based on a newer MedDRA version per the advice from Pharmacovigilance.

7. References

Bruno, R., Mercier, F., & Claret, L. (2014). Evaluation of tumor size response metrics to predict survival in oncology clinical trials. *Clinical pharmacology and therapeutics*, 95(4):386-93.

Chukwueke, U. N., & Wen, P. Y. (2019). Use of the Response Assessment in Neuro-Oncology (RANO) criteria in clinical trials and clinical practice. *CNS oncology*, 8(1):CNS28.

Claret, L., Jin, J. Y., Ferté, C., Winter, H., Girish, S., Stroh, M., . . . Bruno, R. (2018). A Model of Overall Survival Predicts Treatment Outcomes with Atezolizumab versus Chemotherapy in Non-Small Cell Lung Cancer Based on Early Tumor Kinetics. *Clinical cancer research*, 24(14):3292-3298.

Gong, Y., Mason, J., Shen, Y.-L., Chang, E., Kazandjian, D., Blumenthal, G. M., . . . Beaver, J. A. (2020). An FDA analysis of the association of tumor growth rate and overall and progression-free survival in metastatic non-small cell lung cancer (NSCLC) patients. *Journal of Clinical Oncology*, 38(15_suppl):9541-9541.

Stein, W. D., Gulley, J. L., Schlom, J., Madan, R. A., Dahut, W., Figg, W. D., . . . Fojo, T. (2011). Tumor regression and growth rates determined in five intramural NCI prostate cancer trials: the growth rate constant as an indicator of therapeutic efficacy. *Clinical cancer research*, 17(4):907–917.

Wilkerson, J. (2016). *wilkersj/tumgr: Tumor Growth Rate Analysis*. Retrieved from GitHub: https://github.com/wilkersj/tumgr