Janssen Research & Development *

Statistical Analysis Plan

A Phase 3, Open-Label, Randomized Study of Amivantamab and Lazertinib in Combination with Platinum-Based Chemotherapy Compared with Platinum-Based Chemotherapy in Patients with EGFR-Mutated Locally Advanced or Metastatic Non-Small Cell Lung Cancer After Osimertinib Failure
MARIPOSA-2

Protocol 61186372NSC3002; Phase 3

Amendment 3

JNJ-61186372 (amivantamab) and JNJ-73841937 (lazertinib)

Status:	Final
Date:	19 May, 2023
Prepared by:	Janssen Research & Development
Document No.:	EDMS-RIM-577583

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

VERSION HISTORY

SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
Original SAP	16 February 2022	Not Applicable	Initial release
Amendment 1	23 June 2022	1. Add hypothesis tests of amivantamab, carboplatin, and pemetrexed (ACP) vs carboplatin and pemetrexed (CP) for PFS, ORR and OS endpoints, and associated analyses. 2. Increase sample size from 500 to 600 while retaining 2:2:1 randomization ratio. 3. Remove hypothesis testing and analysis for biomarker-defined subgroup. 4. Update multiplicity control strategy using a graphical approach. 5. Add additional analyses for patient report outcomes (PROs).	To add hypothesis testing for ACP versus CP as a dual primary and secondary hypotheses with an increased study sample size to provide sufficient power. To remove biomarker-defined subgroup analyses. To control family-wise type I error rate under dual primary hypothesis testing upon the removal of biomarker-defined subgroup testing. To have a comprehensive assessment of PRO measures.
Amendment 2		1. Add analysis by dosing schedule. 2. Remove NSCLC subtype at screening from baseline characteristic. 3. Add analysis for the extension cohort 4. Add venous thromboembolic (VTE) events to analysis of AESI	To add analysis for lazertinib dosing schedule change per IDMC recommendation. NSCLC subtype at screening was not collected in the study. To further describe the safety and efficacy of the ACP-L dosing schedule versus ACP with additional data. In protocol amendment 4, VTE was included into AESI.
Amendment 3		1. Update definition of time to symptomatic progression (TTSP) 2. Add additional intracranial endpoints (ORR, dor, time to intracranial disease progression) 3. Remove improvement rate and time to first improvement from PRO EORTC QLQ C30 analysis 4. Item-wise descriptive summary of PROMIS-PF data for each timepoint is dropped. Add Time to Symptom	Provides additional clarification regarding the definition of time to symptomatic progression to include both symptomatic progression and death as events. To further evaluate the intracranial efficacy. EORTC QLQ C30 is not expected to be improved for majority of patients.

SAP Version	Approval Date	Change	Rationale
		Deterioration (TTSD) for PROMIS-PF	Scores on individual items of the PROMIS PF are not informative. Adding time to deterioration is primarily to be consistent with analyses of other PROs

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for the Phase 3 study of amivantamab and lazertinib in combination with platinum-based chemotherapy (LACP and ACP-L dosing schedules), and amivantamab in combination with platinum-based chemotherapy (ACP) compared with platinum-based chemotherapy (CP) in patients with EGFR-mutated locally advanced or metastatic non-small cell lung cancer after osimertinib failure. The SAP is to be interpreted in conjunction with the protocol. This SAP covers the planned analysis for the clinical study report (CSR).

Additional analyses for patient-reported outcome (PRO) measures will be described in a separate PRO SAP.

1.1. Objectives and Endpoints

Objectives	Endpoints
Dual Primary	
To assess the efficacy of LACP/ACP-L, compared with CP, in participants with locally advanced or metastatic EGFR Exon 19del or Exon 21 L858R substitution NSCLC	• PFS using RECIST v1.1 guidelines, as assessed by BICR
To assess the efficacy of ACP, compared with CP, in participants with locally advanced or metastatic EGFR Exon 19del or Exon 21 L858R substitution NSCLC	• PFS using RECIST v1.1 guidelines, as assessed by BICR
Secondary	
To further assess the clinical benefit of • LACP/ACP-L vs CP • ACP vs CP	• Objective response (by BICR) • Overall survival • Duration of response (by BICR) • Time to subsequent therapy • PFS after first subsequent therapy (PFS2) • Time to symptomatic progression • Intracranial PFS (by BICR)
To assess the safety of • LACP/ACP-L vs CP • ACP vs CP	• Incidence and severity of adverse events and clinical laboratory abnormalities
To assess the relationship between pharmacokinetics or immunogenicity and selected endpoints (including but not limited to efficacy, safety, and/or patient-reported outcomes)	• Serum amivantamab and plasma lazertinib concentrations, and serum anti-amivantamab antibodies
To assess health-related quality of life and disease-related symptoms in participants treated with • LACP/ACP-L vs CP • ACP vs CP	• NSCLC-SAQ • EORTC-QLQ-C30 • PROMIS-PF
To describe the contribution of lazertinib to the efficacy of LACP/ACP-L (vs ACP)	• Intracranial PFS (by BICR) • Objective response (by BICR) • Duration of response (by BICR) • PFS (by BICR)

Objectives	Endpoints
Exploratory	
To further assess the clinical benefit of • LACP/ACP-L vs CP • ACP vs CP	• Disease control rate • Time to treatment discontinuation
To further assess health-related quality of life in participants treated with • LACP/ACP-L vs CP • ACP vs CP	• EQ-5D-5L • PRO CTCAE
To explore biomarkers predictive of improved outcome, and mechanisms of resistance, in participants treated with • LACP/ACP-L vs CP • ACP vs CP	• Characterization of tumor genetics by NGS of ctDNA at baseline, and changes at progression

ACP-L=amivantamab, carboplatin, pemetrexed and lazertinib(lazertinib started after carboplatin treatment is completed); ;BICR=blinded independent central review; ctDNA=circulating tumor deoxyribonucleic acid; EGFR=epidermal growth factor receptor; EORTC-QLQ-C30=European Organisation of Research and Treatment of Cancer Quality of Life Questionnaire Core 30; EQ-5D-5L=EuroQol 5-dimension 5-level descriptive system; LACP=lazertinib, amivantamab, carboplatin, and pemetrexed; NGS=Next-Generation Sequencing; NSCLC=non-small cell lung cancer; NSCLC-SAQ=Non-Small Cell Lung Cancer - Symptom Assessment Questionnaire; PFS=progression-free survival; PRO CTCAE=Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events; PROMIS-PF=Patient-Reported Outcomes Measurement Information System Physical Function; RECIST=Response Evaluation Criteria in Solid Tumors.

The primary objectives of the extension cohort are to further describe the safety and efficacy of the ACP-L dosing schedule versus ACP

1.2. Study Design

This is a randomized, open-label, active-controlled, parallel, multicenter, Phase 3 study of LACP/ACP-L, and ACP compared with CP in participants with EGFR-mutated locally advanced or metastatic NSCLC who have progressed on or after treatment with osimertinib.

The study will include a Screening phase, a Treatment phase, and a Follow-up phase. Participants must complete screening procedures within 28 days before randomization. To be eligible for participation, all participants must have locally advanced or metastatic nonsquamous NSCLC with Exon 19del or Exon 21 L858R substitution EGFR mutations. All participants must have progressed on or after osimertinib monotherapy as the most recent line of treatment.

The Treatment phase for each participant will begin at Cycle 1 Day 1 and continue in 21-day cycles until the End of Treatment visit, approximately 30 days after last dose of study treatment. Study treatment will continue until documented clinical or radiographic (Response Evaluation Criteria in Solid Tumors [RECIST] v1.1) disease progression, or until the participant meets another criterion for discontinuation of study treatment.

Approximately 600 eligible participants will be randomly assigned to study treatment in a 2:2:1 ratio (Arms A:B:C). Randomization will be stratified by osimertinib line of therapy (first-line vs second-line), history of brain metastases (yes vs no), and Asian race (yes vs no).

To further describe the safety and efficacy of the ACP-L dosing schedule versus ACP, a separate open label randomized extension cohort has been added to the study. Enrollment of participants into the extension cohort may begin after enrollment into the main study is complete and when the

Sponsor opens the extension cohort for enrollment. The extension cohort will have the same eligibility criteria, study procedures and operate within the same investigational sites as the main study

In the extension cohort, approximately 90 eligible participants will be randomly assigned to receive ACP-L or ACP in a 2:1 ratio (Arms A2:C2) – the same ratio as in the main study. Randomization will be stratified by osimertinib line of therapy (first-line vs second-line), history of brain metastases (yes vs no), and Asian race (yes vs no), also as in the main study.

STUDY TREATMENT GROUPS AND DURATION

Main Study Arms

Arm A Dosing Schedule 1 (LACP, from study start until 6 November 2022):

- Lazertinib 240 mg orally, once daily
- Amivantamab by intravenous (IV) infusion in 21day cycles:
  - 1,400 mg (1,750 mg if body weight ≥80 kg) on Cycle 1 Days 1/2, 8, and 15, and Cycle 2 Day 1
  - 1,750 mg (2,100 mg if body weight ≥80 kg) on Day 1 of each 21-day cycle, starting with Cycle 3
- Carboplatin and pemetrexed as in Arm B.

Arm A Dosing Schedule 2 (started on 07 November, 2022, ACP-L):

- Lazertinib 240 mg orally, once daily starting Cycle 5 Day 1 or sooner if carboplatin is discontinued earlier
- Amivantamab by IV infusion in 21day cycles:
  - 1,400 mg (1,750 mg if body weight ≥80 kg) on Cycle 1 Days 1/2, 8, and 15, and Cycle 2 Day 1
  - 1,750 mg (2,100 mg if body weight ≥80 kg) on Day 1 of each 21-day cycle, starting with Cycle 3
- Carboplatin and pemetrexed as in Arm B.

Arm B (CP):

- Carboplatin AUC 5 on Day 1 of each 21-day cycle, for up to 4 cycles
- Pemetrexed 500 mg/m^2 on Day 1 of each 21day cycle, with carboplatin for up to 4 cycles, and then as maintenance until disease progression

Arm C (ACP):

- Amivantamab as in Arm A
- Carboplatin and pemetrexed as in Arm B

Extension Cohort Arms

Arm A2 (ACP-L):

- Lazertinib 240 mg orally, once daily starting Cycle 5 Day 1 or sooner if carboplatin is discontinued earlier
- Amivantamab by IV infusion in 21-day cycles:
  - 1,400 mg (1,750 mg if body weight ≥80 kg) on Cycle 1 Days 1/2, 8, and 15, and Cycle 2 Day 1

  - 1,750 mg (2,100 mg if body weight ≥80 kg) on Day 1 of each 21-day cycle, starting with Cycle 3
- Carboplatin AUC 5 on Day 1 of each 21-day cycle, for up to 4 cycles
- Pemetrexed 500 mg/m2 on Day 1 of each 21-day cycle, with carboplatin for up to 4 cycles, and then as maintenance until disease progression

Arm C2 (ACP):

- Amivantamab by IV infusion in 21-day cycles:
  - 1,400 mg (1,750 mg if body weight ≥80 kg) on Cycle 1 Days 1/2, 8, and 15, and Cycle 2 Day 1
  - 1,750 mg (2,100 mg if body weight ≥80 kg) on Day 1 of each 21-day cycle, starting with Cycle 3
- Carboplatin AUC 5 on Day 1 of each 21-day cycle, for up to 4 cycles
- Pemetrexed 500 mg/m2 on Day 1 of each 2

Study treatments may be withheld, and the dosages may be subsequently modified to manage treatment related toxicity.

Study treatment should continue until one of the following criteria applies: documented disease progression using Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 confirmed by blinded independent central review (Exception: Continuation of study treatment after disease progression may be allowed in accordance with local practice, after consultation with the Medical Monitor, if the investigator believes the participant is deriving clinical benefit); withdrawal of consent; the investigator believes that for safety reasons or tolerability reasons (e.g., adverse event) it is in the best interest of the participant to discontinue study treatment; the participant becomes pregnant; or noncompliance with study treatment or procedure requirements.

Tumor response will be assessed by blinded independent central review (BICR) according to RECIST v1.1. Baseline disease assessments, including brain magnetic resonance imaging (MRI) should be performed no more than 28 days prior to randomization. Repeat imaging will occur at regular intervals, as defined in the Schedule of Activities (SoA) (see study protocol Section 1.3), until disease progression.

Safety will be assessed by physical examinations, laboratory tests, vital signs, electrocardiograms, left ventricular ejection fraction (echocardiogram [ECHO] or multigated acquisition [MUGA]), Eastern Cooperative Oncology Group (ECOG) performance status, monitoring of adverse events, and concomitant medication usage.

Health-related quality of life in participants will be assessed by patient-reported outcomes (PROs) measures including Non-Small Cell Lung Cancer - Symptom Assessment Questionnaire (NSCLC-SAQ), European Organization of Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORTC-QLQ-C30), Patient-Reported Outcomes Measurement Information System –

Physical Function (PROMIS-PF), EuroQol five-dimensional descriptive system (EQ-5D-5L), and Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE).

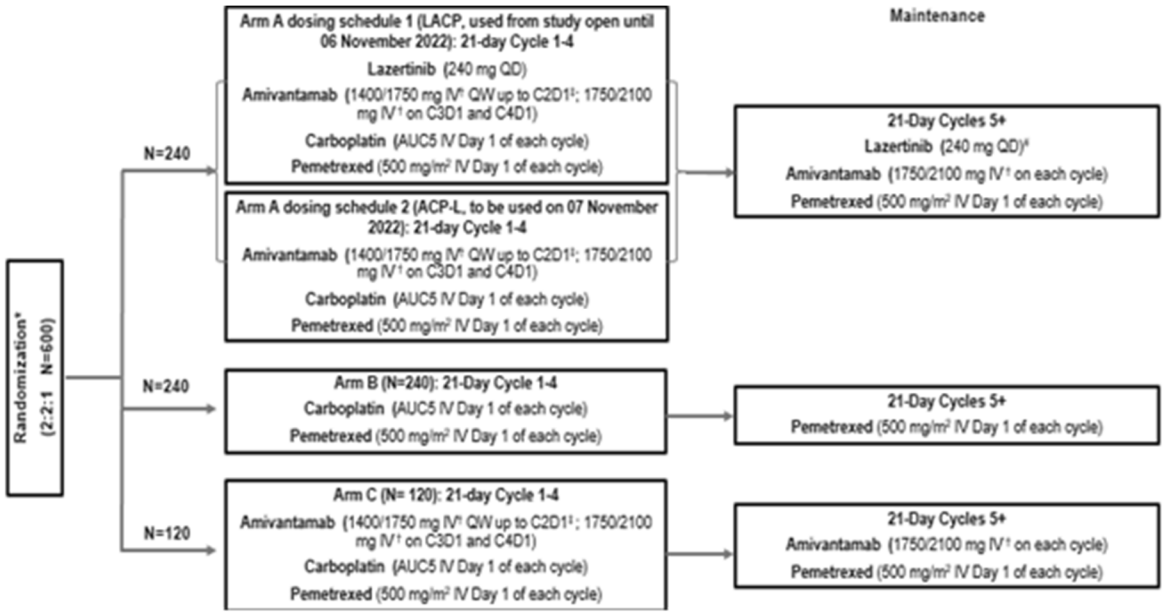
All study evaluations will be conducted according to the SoA in study protocol.

The primary efficacy endpoint is progression-free survival (PFS). Analysis of the primary endpoint will be performed after approximately 350 PFS events from all 3 arms combined have occurred.

An IDMC will be commissioned for this study for periodic review of safety and tolerability data. Details on IDMC are described in Section 5.8.1.

A diagram of the study design is provided in Figure 1, as follows:

Figure 1A: Schematic Overview of the Study



ACP-L=amivantamab, carboplatin, pemetrexed, and lazertinib (lazertinib started after carboplatin treatment is completed); AUC=area under the concentration-time curve; C#D#=Cycle # Day #; IV=intravenously; LACP=lazertinib, amivantamab, carboplatin, and pemetrexed (all drugs started C1D1); QD=once daily; QW=once weekly.

Arm A Dosing schedule 2 (ACP-L) started on 7 November 2022.

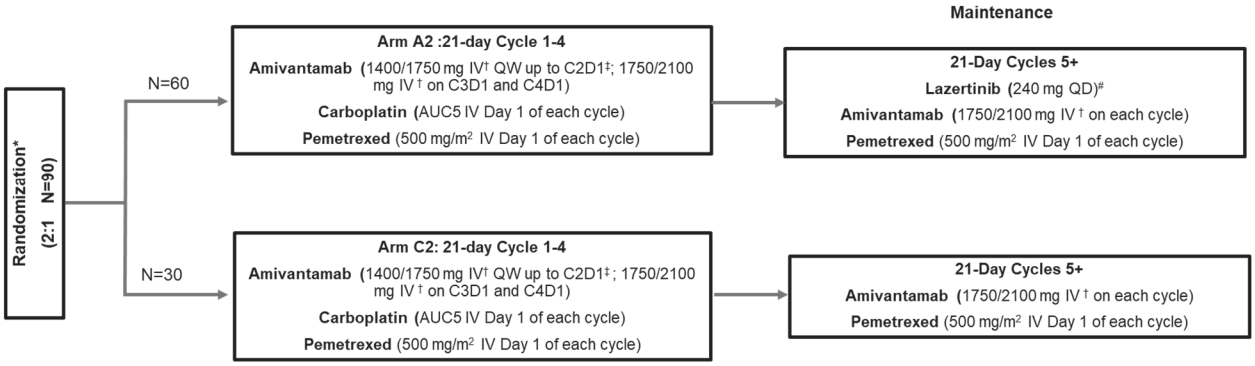
* Stratification factors: osimertinib line of therapy (first-line vs second-line), history of brain metastases (yes vs no), Asian race (yes vs no)

† Doses shown by body weight (<80 kg/≥80 kg)

‡ Cycle 1: Days 1/2 (split dose), 8, and 15; Cycle 2: Day 1

\# Lazertinib for participants receiving dosing schedule 2 in Arm A may start sooner if carboplatin discontinued earlier than Cycle 4.

Figure 1B: Schematic Overview of the Extension Cohort



ACP-L=amivantamab, carboplatin, pemetrexed, and lazertinib (lazertinib started after carboplatin treatment is completed); AUC=area under the concentration-time curve; C#D#=Cycle # Day #; IV=intravenously; QD=once daily; QW=once weekly.

* Stratification factors: osimertinib line of therapy (first-line vs second-line), history of brain metastases (yes vs no), Asian race (yes vs no)

† Doses shown by body weight (<80 kg/≥80 kg)

‡ Cycle 1: Days 1/2 (split dose), 8, and 15; Cycle 2: Day 1

\# Lazertinib for participants ~~receiving dosing schedule 2~~ in Arm A2 may start sooner if carboplatin discontinued earlier than Cycle 4.

Randomization

Central randomization will be implemented in this study. In the main study, participants will be randomly assigned to 1 of 3 study treatment groups based on a computer-generated randomization schedule prepared before the study by or under the supervision of the sponsor. In the extension cohort, participants will be randomly assigned to 1 of 2 study treatment groups based on a computer-generated randomization schedule prepared before the extension cohort by or under the supervision of the sponsor. In both the main study and the extension cohort, the randomization will be balanced by using randomly permuted blocks and will be stratified by osimertinib line of therapy (first-line vs second-line), history of brain metastases (yes vs no), and Asian race (yes vs no). The interactive web response system (IWRS) will assign a unique study treatment code, which will dictate the treatment assignment and matching study treatment kit for the participant. The requestor must use his or his own user identification and personal identification number when contacting the IWRS and will then give the relevant participant details to uniquely identify the participant.

Blinding

This is an open-label study.

2. STATISTICAL HYPOTHESES

This study has dual primary hypotheses to be tested in participants with osimertinib-relapsed EGFR-mutated locally advanced or metastatic NSCLC:

- LACP/ACP-L will demonstrate superior progression-free survival (PFS) compared with CP
- ACP will demonstrate superior progression-free survival (PFS) compared with CP

The secondary hypotheses are that LACP/ACP-L and ACP, compared with CP independently in the aforementioned population, will demonstrate superior ORR and overall survival.

To control the overall Type I error rate at 5% for the hypotheses testing of primary and secondary endpoints, a graphical approach in a group sequential design setting (Maurer and Bretz 2013) as detailed in section 2.1 will be applied.

Two interim analyses and one final analysis are planned for OS (see Section 5.8). The final analysis of OS will be conducted at the end of study. The O'Brien Fleming alpha spending function as implemented by the Lan-DeMets method will be used to determine the statistical boundary for each of the interim analyses of OS.

2.1. Multiplicity

Multiplicity due to multiple hypotheses of interest and multiple analyses will be addressed using a graphical approach applied in a group sequential design setting as described by Maurer and Bretz (2013). A primary analysis for PFS and key secondary endpoints (ie, ORR and OS) is planned in the study. In addition, OS will be analyzed at 2 additional timepoints, at 75% information fraction of OS and final OS analysis.

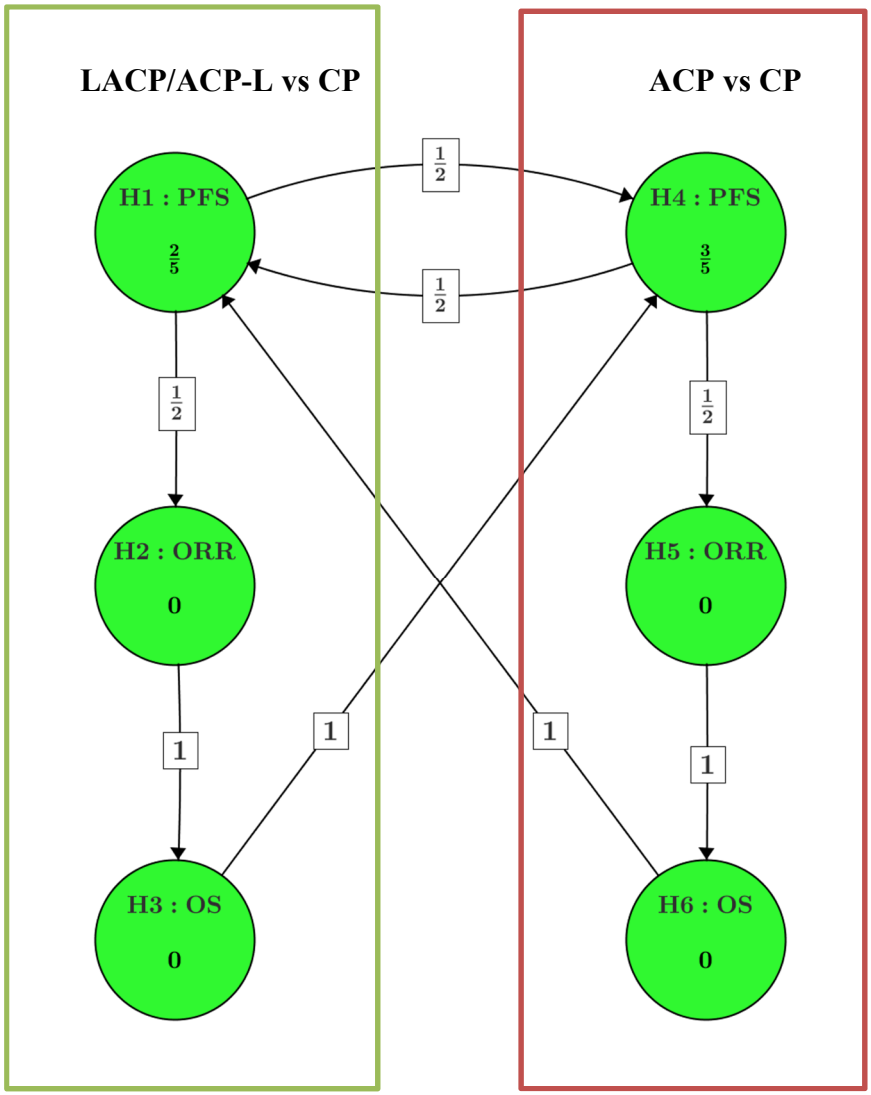
The proposed graphical testing strategy controlling an overall family-wise Type I error rate at the two-sided significance level of 0.05 is presented by Figure 2. The 6 elementary hypotheses are grouped into the 2 families of F1 = {H1: PFS, H2: ORR, H3: OS} and F2 = {H4: PFS, H5: ORR, H6: OS} for the comparison of LACP/ACP-L vs CP and ACP vs CP, respectively. A sequential (hierarchical) testing approach is utilized within each family starting with PFS, followed by ORR, and OS.

Figure 2 schematically shows the distribution of initial amount of alpha (local significance level) allocated to the 6 hypotheses; the local significance levels are represented as numerical values within the circular nodes corresponding to each hypothesis of interest. Particularly H1 and H4 have the initial assigned local alpha of $\frac{2}{5}\alpha$ and $\frac{3}{5}\alpha$, respectively. The local significance levels are re-allocated after each hypothesis is rejected per the algorithm described (Maurer and Bretz 2013) according to the directed edges with the associated weights, and the weights are represented as numerical values within the squares. For example, after the primary endpoint (H1:PFS) rejection, the local significance level associated with H1:PFS is split between the corresponding 2

hypotheses, the half of $\frac{2}{5}\alpha$ is passed to test the next hypothesis of interest in the same family (H2:ORR) and the half of $\frac{2}{5}\alpha$ is allocated to test the primary endpoint in the other family (H4:PFS). If all the hypotheses are rejected in a family, the total level is allocated to the other family.

The O'Brien Fleming alpha spending function as implemented by the Lan-DeMets method will be used to determine the statistical boundary for each of the interim analyses of OS.

Figure 2: Graphical Testing Strategy Controlling an Overall Family-wise Type I Error Rate



ACP-L=amivantamab, carboplatin, pemetrexed, and Lazertinib; LACP=lazertinib, amivantamab, carboplatin, and pemetrexed combination therapy; ACP= amivantamab, carboplatin, and pemetrexed combination therapy; CP=carboplatin and pemetrexed; PFS=progression-free survival; OS=overall survival; ORR=objective response rate.

The primary and secondary analyses for the hypothesis testing described above will be based on the data from the main study only, and the data from the extension cohort will not be included

3. SAMPLE SIZE DETERMINATION

A total of approximately 600 participants will be randomized in a 2:2:1 ratio (Arm A: Arm B: Arm C) in the main study. The median PFS for CP is estimated to be 5.5 months [Mok 2017, Soria 2015]. Assuming a median PFS of 8.5 months for LACP/ACP-L and ACP, respectively, with an approximate 16-month accrual period and an additional 3-month follow-up, a total of 350 PFS events in all 3 arms combined will provide approximately 93% power for LACP/ACP-L over CP, and 83% power for ACP over CP to detect a 35% reduction in the risk of either disease progression or death (hazard ratio (HR) of 0.65 for LACP/ACP-L vs CP and ACP vs CP, respectively), with a log-rank test, assuming an overall family-wise Type I error rate at two-sided significance level of 5%. The sample size calculation has taken into consideration an annual dropout rate of 5%.

4. POPULATIONS (ANALYSIS SETS) FOR ANALYSIS

Fhe following populations are defined for the main study:

Population	Description
Full Analysis Set	All randomized participants, classified according to their assigned treatment arm regardless of the actual treatment received.
Safety Analysis Set	Randomized participants who receive at least 1 dose of study treatment.
Pharmacokinetics / Immunogenicity Analysis Set	Randomized participants who receive at least 1 dose of study treatment and have at least 1 evaluable postbaseline concentration measurement.[a]
Biomarkers Analysis Set	Randomized participants who receive at least 1 dose of study treatment and have at least 1 biomarker measurement.

a. Participants may be removed from the estimation of certain pharmacokinetic parameters on an individual basis due to, for example, missing pharmacokinetic samples such that the pharmacokinetic parameters cannot be appropriately derived. These participants will be identified at the time of the analyses along with their reason for removal.

The Full Analysis Set will be used to summarize the study population and characteristics, as well as efficacy data; the Safety Population will be used to summarize the safety data, unless otherwise specified.

Participants enrolled in the extension cohort (Arms A2 and C2) are not considered part of the Full Analysis Set or Safety population in the main study.

5. STATISTICAL ANALYSES

5.1. General Considerations

All statistical hypothesis tests and 95% confidence interval presented will be 2-sided.

Hypothesis testing of the primary efficacy endpoint and key secondary efficacy endpoints will be performed for LACP/ACP-L (Arm A) versus CP (Arm B) and ACP (Arm C) versus CP (Arm B), respectively. The hypothesis testing will be based on the data from the main study only, and the

data from the extension cohort will not be included. Comparison of LACP/ACP-L (Arm A) with ACP (Arm C) will be performed to describe the contribution of lazertinib based on intracranial PFS, ORR, DoR, and PFS, using summary statistics and nominal p-values; there will be no formal hypothesis testing for this comparison. Additional analysis based on the open-label randomized extension cohort (see Section 5.10) will be performed to further describe the safety and efficacy of the ACP-L dosing schedule versus ACP, but similarly there is no formal hypothesis testing.

5.1.1. Visit Windows

Participants should start study drug within 72 hours after randomization, and continue in 21-day cycles until the End of Treatment visit. Visit windowing will be based on cycles. Unless otherwise specified, data to be analyzed or presented over time will be presented by cycle, day and time point (as appropriate) that are recorded in CRF.

5.1.2. Study Day/Relative Day

Study day or relative day is defined as:

- Reference date (Day 1) = randomization date for efficacy assessment, or first dose date of study drug for safety assessment.
- Study Day = assessment date – reference date + 1 for assessment performed on or after the reference date; assessment date – reference date for assessment performed before the reference date.

There is no 'Day 0'. First dose date will be the date of starting any of the study drug (amivantamab or lazertinib or carboplatin or pemetrexed).

5.1.3. Treatment Group and Study Drug

In this study, treatment group refers to treatment of LACP/ACP-L, CP or ACP. Study drug refers to each study agent within a treatment group.

5.1.4. Baseline Measurement

Baseline measurement is defined as the closest non-missing measurement taken on or prior to the first study drug administration (including time if time is available). If the first administration date is missing or the administration is not done, then the baseline measurement is the closest non-missing measurement taken on or prior to the corresponding visit date (if visit date is not available, then randomization date should be used).

5.2. Participant Dispositions

Screened participants and reasons for screen failures will be summarized overall.

The number of participants in the following disposition categories will be summarized throughout the study by treatment group and overall:

- Participants randomized
- Participants who received study treatment
- Participants who discontinued study treatment
- Reason for discontinuation of study treatment
- Participants who terminated study prematurely
- Reason for termination of study
- Participants who completed the study

A listing of participants will be provided for the following categories:

- Participants who discontinued study treatment
- Participants who terminated study prematurely

5.3. Primary Endpoint Analysis

5.3.1. Definition of Endpoint

Progression-free Survival (PFS) is defined as the time from randomization until the date of objective disease progression or death, whichever comes first, based on blinded independent central review (BICR) using RECIST v1.1. Participants who have not progressed or have not died at the time of analysis will be censored at their last evaluable RECIST v1.1 assessment date.

Key censoring rules for PFS are summarized below.

Key censoring rules for PFS

Situation	Date of Censoring
No evaluable baseline or postbaseline disease assessment	Censored at the date of randomization
Lost to follow-up or withdraw from study	Censored at the date of last evaluable disease assessment
No documented disease progression or death	Censored at the date of last evaluable disease assessment
Documented disease progression or death after 2 or more consecutive missed/unevaluable disease assessments*	Censored at the date of last evaluable disease assessment before the missed/unevaluable visits

*If no evaluable disease assessment before the consecutive missed/unevaluable visits, participants will be censored at the date of randomization.

PFS is calculated in months as follows:

- PFS = (date of PD/death or censoring – date of randomization + 1) / (365.25/12).

5.3.2. Estimand

Estimand Scientific Question of Interest: What is the relative effect of amivantamab and lazertinib in combination with platinum-based chemotherapy compared with platinum-based chemotherapy in prolonging PFS? And what is the relative effect of amivantamab in combination

with platinum-based chemotherapy compared with platinum-based chemotherapy in prolonging PFS?

Study intervention:

- Experimental: LACP/ACP-L and ACP
- Control: CP

Population: patients with osimertinib-relapsed EGFR-mutated locally advanced or metastatic NSCLC

Variable: PFS

Summary Measure (Population-level summary): hazard ratio (HR) for LACP/ACP-L vs CP and ACP vs CP

Intercurrent events and their corresponding strategies:

Intercurrent Events	**Name of Strategy for Addressing Intercurrent Events and Its Description**
Study treatment discontinuation due to any reason	**Treatment Policy strategy:** use time to disease progression or death, regardless of whether or not study treatment discontinuation had occurred
Study treatment switching to other anticancer therapy	**Treatment Policy strategy:** use time to disease progression or death, regardless of whether or not started subsequent anticancer therapies
Death	**Composite Variable strategy:** death being a component of the variable

5.3.3. Primary Analysis Methods

The primary efficacy endpoint will be first analyzed in the Full Analysis Set using the log-rank test stratified by osimertinib line of therapy (first-line vs second-line), history of brain metastases (yes vs no), and Asian race (yes vs no). The p-values (LACP/ACP-L vs CP, ACP vs CP) generated from the stratified log-rank test will be used for the primary hypotheses testing. Following the graphical testing procedure described in section 2.1, the treatment effect on PFS for LACP/ACP-L vs CP will be first tested at two-sided alpha of 0.02, and that of ACP vs CP will be tested at two-sided alpha of 0.03. If either of the comparisons is not statistically significant, it can be tested again with the initial alpha plus half of the alpha passed from the other statistically significant test on PFS. The HR for PFS will be calculated, along with its 95% CI, from a stratified Cox model with treatment as the sole explanatory variable, using the same stratification factors as for the log-rank test. In addition, the comparison for LACP/ACP-L vs ACP will also be carried out using the same analysis model.

The median PFS with 95% CI will be estimated using Kaplan-Meier method. The Kaplan-Meier PFS curve will also be plotted by treatment group. In addition, PFS rates with 95% CI will be estimated by Kaplan-Meier method at landmarks (e.g. at 6-month and 12-month, etc.) and reported for each treatment group. The number and percentage of participants who had a PFS event or were censored will be reported, and reasons for PFS event and censoring will be summarized.

For assessment of internal consistency and investigation of homogeneity of the treatment effect across subgroups, a subgroup analysis on pre-specified subgroups defined in Section 5.7.6 will be conducted. Forest plots of subgroup analysis will be generated.

5.3.4. Sensitivity Analysis

The following sensitivity analyses will be conducted to evaluate the robustness of the primary analysis of PFS.

5.3.4.1. Unstratified Analysis of PFS

Sensitivity analysis using unstratified log-rank test will also be performed.

5.3.4.2. Assess Hazards Proportional Assumption

The proportional hazards assumption will be examined by plotting log(-log[estimated survival distribution function]) against log(survival time). In addition, a treatment by logarithm-transformed time interaction term will be added into the primary Cox model and tested. A p-value greater than 0.05 for the interaction term will be interpreted as no statistical evidence against the proportional hazard assumption.

5.3.5. Supplementary Analysis

5.3.5.1. Censored for Death/PD after Start of Subsequent Anticancer Therapy

Supplementary analysis will be performed using progression or death prior to the start of the subsequent anticancer therapy as events. Participants who have not progressed or have not died before the initiation of subsequent therapy will be censored at the date of the last evaluable disease assessment prior to the start of subsequent therapy.

5.3.5.2. Not Censored for Missing More Than One Disease Evaluation

Additional supplementary analysis will be performed using all progression or death, whichever occur first, as event regardless missed/unevaluable disease assessment for 2 or more consecutive visits.

5.4. Secondary Endpoints Analysis

5.4.1. Objective Response Rate (ORR)

5.4.1.1. Definition

ORR is defined as the proportion of participants who achieve either a complete response (CR) or partial response (PR), as defined by BICR using RECIST v1.1. Data obtained up until progression or last evaluable disease assessment in the absence of progression will be included in the assessment of ORR. However, any CR or PR, which occurred after a subsequent anticancer therapy was received, will not be included in the numerator for the ORR calculation. Participants who do not have a tumor response assessment for any reason will be considered non-responders and will be included in the denominator when calculating the response rate.

5.4.1.2. Estimand

Estimand Scientific Question of Interest: What is the relative effect of amivantamab and lazertinib in combination with platinum-based chemotherapy compared with platinum-based chemotherapy in ORR? And what is the relative effect of amivantamab in combination with platinum-based chemotherapy compared with platinum-based chemotherapy in ORR?

Study intervention:

- Experimental: LACP/ACP-L and ACP
- Control: CP

Population: patients with osimertinib-relapsed EGFR-mutated locally advanced or metastatic NSCLC

Variable: objective response

Summary Measure (Population-level summary): odds ratio for LACP/ACP-L vs CP and ACP vs CP

Intercurrent events and their corresponding strategies:

Intercurrent Events	Name of Strategy for Addressing Intercurrent Events and Its Description
Study intervention switching to other anticancer therapy	**Hypothetical strategy:** use the best response until subsequent anti-cancer therapy

5.4.1.3. Analysis Methods

Objective response will be analyzed using a logistic regression model stratified by osimertinib line of therapy (first-line vs second-line), history of brain metastases (yes vs no), and Asian race (yes vs no). The results of the analysis will be presented in terms of an odds ratio together with its associated 95% confidence interval and corresponding p-value. Following the graphical testing procedure, if the testing for the primary endpoint of PFS is statistically significant in at least one of the pair of comparisons (LACP/ACP-L vs CP or ACP vs CP), ORR in the same family of comparison will be tested at two-sided alpha reallocated after previous hypotheses rejection per the algorithm defined in the section 2.1. In addition, the comparison for LACP/ACP-L vs ACP will also be performed using the same analysis model.

The same analysis will be carried out for ORR based on confirmed PR or CR from subsequent assessments. The confirmation by subsequent assessments should be performed not less than 4 weeks after the criteria for PR or CR are first met.

5.4.2. Overall Survival (OS)

5.4.2.1. Definition

OS is defined as the time from the date of randomization until the date of death due to any cause. Any participant not known to have died at the time of analysis will be censored based on the last recorded date on which the participant was known to be alive.

5.4.2.2. Estimand

Estimand Scientific Question of Interest: What is the relative effect of amivantamab and lazertinib in combination with platinum-based chemotherapy compared with platinum-based chemotherapy in prolonging OS? And what is the relative effect of amivantamab in combination with platinum-based chemotherapy compared with platinum-based chemotherapy in prolonging OS?

Study intervention:

- Experimental: LACP/ACP-L and ACP
- Control: CP

Population: Patients with osimertinib-relapsed EGFR-mutated locally advanced or metastatic NSCLC

Variable: OS

Summary Measure (Population-level summary): HR for LACP/ACP-L vs CP and ACP vs CP

Intercurrent events and their corresponding strategies:

Intercurrent Events	**Name of Strategy for Addressing Intercurrent Events and Its Description**
Study intervention switching to other anticancer therapy	**Treatment Policy strategy:** use time to death, regardless of whether or not started subsequent anticancer therapies

5.4.2.3. Analysis Methods

The comparison between LACP/ACP-L and ACP over CP in OS will be carried out using the similar methodology and model as for the primary analysis of PFS in the Full Analysis Set. Two interim analyses of OS will be performed. The final analysis of OS will be conducted at the end of study. If the testing of both PFS and ORR shows statistical significance in either of the families of comparison (LACP/ACP-L vs CP and ACP vs CP), the analysis of OS will be carried out in the same family using a total 2-sided alpha reallocated after previous hypotheses rejection per the algorithm defined in the section 2.1. The O'Brien Fleming alpha spending function as implemented by the Lan-DeMets method will be used to determine the statistical boundary for each of the interim analyses of OS (see Section 5.8 for more details).

A subgroup analysis on pre-specified subgroups defined in Section 5.7.6 will be conducted. Forest plots of subgroup analysis will be generated.

Sensitivity Analysis

Sensitivity analysis using unstratified log-rank test will be performed.

Supplementary Analysis

Analysis will be carried out using Inverse Probability of Censoring Weighting (IPCW) (Robins and Finkelstein 2000)**Error! Reference source not found.** to adjust for confounding from treatment crossover. The weights to reduce the bias will be estimated from baseline covariates and time-dependent

covariates predictive of treatment crossover such as baseline disease burden, occurrence of serious adverse event before crossover, based on a logistic regression model. Hazard ratio and its 95% confidence interval will be estimated based on a Cox regression analysis with IPCW.

5.4.3. Duration of Response (DoR)

5.4.3.1. Definition

DoR is defined as the time from the date of first documented response (PR or CR) until the date of documented progression or death, whichever comes first, for participant who have CR or PR. The end of response should coincide with the date of progression or death from any cause used for the PFS endpoint. If a participant does not progress following a response, then his/her duration of response will be until the PFS censoring time. Participants who started a subsequent anticancer therapy in the absence of progression will be censored at the last disease assessment before the start of subsequent therapy.

5.4.3.2. Analysis Methods

A Kaplan-Meier plot for duration of response and median duration of response with 95% confidence interval (calculated from the Kaplan-Meier estimate) will be presented by treatment group.

5.4.4. Time to Subsequent Therapy (TTST)

5.4.4.1. Definition

TTST is defined as the time from the date of randomization to the start date of the subsequent anticancer therapy following study treatment discontinuation, or death, whichever comes first. Participants alive and not starting subsequent therapy will be censored on the date on which the participant was last known alive date.

5.4.4.2. Analysis Methods

TTST will be analyzed using the similar method as the primary analysis of PFS.

5.4.5. Progression-free Survival After the First Subsequent Therapy (PFS2)

5.4.5.1. Definition

The PFS2 is defined as the time from randomization until the date of second objective disease progression, after initiation of subsequent anticancer therapy, based on investigator assessment (after that used for PFS) or death, whichever comes first. Any deaths are considered as PFS2 events. Participants alive and for whom a second disease progression has not been observed should be censored at the last time known to be alive and without a second disease progression (i.e., last disease assessment).

Key censoring rules for PFS2

Situation	Date of Censoring
No postbaseline disease assessment	Randomization

Disease progression on study treatment and no subsequent therapy	The date of last disease assessment
Two or more subsequent therapy without a progression	The last disease assessment date prior to the start of 2nd line of subsequent therapy
Treated beyond progression	The last date of disease assessment

5.4.5.2. Analysis Methods

PFS2 will be analyzed using the similar method as the primary analysis of PFS.

5.4.6. Time to Symptomatic Progression (TTSP)

5.4.6.1. Definition

TTSP is defined as the time from randomization to documentation in the eCRF of any of the following (whichever occurs earlier): onset of new symptoms or symptom worsening that is considered by the investigator to be related to lung cancer and requires either a change in anticancer treatment and/or clinical intervention to manage symptoms, or death. The TTSP for a participant who does not experience any of these events will be censored on the date on which the participant was last known to be event free.

5.4.6.2. Analysis Methods

TTSP will be analyzed using the similar method as the primary analysis of PFS.

In addition, similar analyses will be carried out for the following two components of time to symptomatic progression:

- the time from randomization to documentation in the eCRF of any of the following (whichever occurs earlier): onset of new symptoms or symptom worsening that is considered by the investigator to be related to lung cancer and requires a change in anticancer treatment.

- the time from randomization to documentation in the eCRF of any of the following (whichever occurs earlier): onset of new symptoms or symptom worsening that is considered by the investigator to be related to lung cancer and requires clinical intervention to manage symptoms.

5.4.7. Intracranial Progression-free Survival

5.4.7.1. Definition

Intracranial PFS is defined as the time from randomization until the date of objective intracranial disease progression or death, whichever comes first, based on BICR using RECIST v1.1. Specifically, intracranial disease progression is defined as having progression of brain metastasis or occurrence of new brain lesion. Participants who have not progressed intracranially or died will be censored at their last evaluable intracranial disease assessment date.

5.4.7.2. Analysis Methods

Intracranial PFS will be analyzed using the similar method as the primary analysis of PFS. In addition, the comparison for LACP/ACP-L vs ACP will also be performed using the same analysis model.

Similar analysis will be repeated in the subgroup of randomized participants who had history of brain metastasis at screening. The corresponding Cox model will be stratified by osimertinib line of therapy and race.

5.4.8. Intracranial Objective Response Rate (ORR)

5.4.8.1. Definition

Intracranial ORR is defined as the proportion of participants who achieve either an intracranial complete response (CR) or partial response (PR), as defined by BICR using RECIST v1.1. Data obtained up until intracranial progression or last intracranial evaluable disease assessment in the absence of progression will be included in the assessment of intracranial ORR. However, any intracranial CR or PR, which occurred after a subsequent anticancer therapy was received, will not be included in the numerator for the ORR calculation. Participants who do not have a tumor response assessment for any reason will be considered non-responders and will be included in the denominator when calculating the response rate.

5.4.8.2. Analysis Methods

Intracranial ORR will be analyzed using the similar method as the primary analysis of ORR for subjects with baseline intracranial disease by BICR. In addition, the comparison for LACP/ACP-L vs ACP will also be performed using the same analysis model.

5.4.9. Intracranial Duration of Response (DoR)

5.4.9.1. Definition

Intracranial DoR is defined as the time from the date of first documented intracranial response (PR or CR) until the date of documented intracranial progression or death, whichever comes first, for participant who have intracranial CR or PR. The end of response should coincide with the date of intracranial progression or death from any cause used for the intracranial PFS endpoint. If a participant does not progress following an intracranial response, then his/her duration of response will be until the intracranial PFS censoring time. Participants who started a subsequent anticancer therapy in the absence of intracranial progression will be censored at the last intracranial disease assessment before the start of subsequent therapy.

5.4.9.2. Analysis Methods

Intracranial DoR will be analyzed using the similar method as the primary analysis of DoR. In addition, the comparison for LACP/ACP-L vs ACP will also be performed using the same analysis model.

5.4.10. Time to Intracranial Disease Progression

5.4.10.1. Definition

Time to intracranial disease progression is defined as the time from randomization until the date of objective intracranial disease progression, based on BICR using RECIST v1.1. Specifically, intracranial disease progression is defined as having progression of brain metastasis or occurrence of new brain lesion. Participants who have not progressed intracranially will be censored at their last evaluable intracranial disease assessment date.

5.4.10.2. Analysis Methods

Time to intracranial disease progression will be analyzed using the similar method as the primary analysis of PFS in Full Analysis Set.

5.4.11. Patient Reported Outcome - NSCLC-SAQ

5.4.11.1. Definition

NSCLC-SAQ is a 7-item PRO measure designed for use in adults to assess symptoms of advanced NSCLC. The NSCLC-SAQ has a seven-day recall period. It contains five domains and accompanying items that were identified as symptoms of NSCLC: cough (1 item), pain (2 items), dyspnea (1 item), fatigue (2 items), and appetite (1 item). Each item uses a response scale between 0 to 4, with higher scores indicating more severe symptomatology. A single score of pain is created by selecting the higher severity of either item (‘chest pain’ or ‘general pain’). A single score of fatigue is calculated using the mean of 2 items (‘low energy’ and ‘tire easily’). All five of these domains must be non-missing to compute a total score, with a response range from 0 to 20.

5.4.11.2. Analysis Methods

Compliance Rates

Compliance rates for completion of NSCLC-SAQ at each time point will be generated based on the actual number of assessments received over the number of expected.

Change from Baseline

Descriptive statistics will be reported for the five individual scores and total score at baseline and at each visit for absolute value and for change from baseline.

The change of scores from baseline over time will be assessed using mixed-effects model for repeated measures (MMRM) analysis based on restricted maximum likelihood (REML). The model will include participants as a random effect, and baseline value, treatment group, time in week, treatment-by-time interaction, and stratification factors as fixed effects. The treatment comparison will be based on the least squares means and the 2-sided 95% CI will be estimated. Randomized participants who receive at least 1 dose of study treatment and have at least 1 evaluable postbaseline measurement will be included in the analysis. Line plots of mean and the change from baseline with standard error over time by treatment group will be displayed.

Time to Deterioration

Time to deterioration in NSCLC-SAQ measures is defined as the time from randomization until the date of the first clinically meaningful deterioration, or death. A clinically meaningful change in individual score is defined as an increase from baseline of ≥1, and a clinically meaningful change in total score is defined as an increase from baseline of ≥2. Participants who have not shown a deterioration or have not died at the time of analysis will be censored at their last PRO assessment date.

Time to deterioration in NSCLC-SAQ measures will be analyzed using the similar method as the primary analysis of PFS.

Deterioration in each individual score and total score (as defined previously) will also be summarized using count and percentage by treatment and study visit.

Each individual item will also be summarized using count and percentage by treatment and study visit.

5.4.12. Patient Reported Outcome - EORTC-QLQ-C30

5.4.12.1. Definition

EORTC-QLQ-C30 measure cancer patients' functioning for all cancer types. It includes 30 items resulting in 5 functional scales (physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning), 1 global health status/quality of life (GHS/QoL) scale, 3 symptom scales (fatigue, nausea and vomiting, and pain), and 6 single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties). The instrument contains 28 items using a verbal rating scale with 4 response options: "Not at All," "A Little," "Quite a Bit," and "Very Much" (scored 1 to 4). Two additional items use response options (1 to 7): 1 = Very Poor, to 7 = Excellent. All scale and item scores will be linearly transformed to be in the range from 0 to 100 according to the algorithm in EORTC QLQ-C30 scoring manual, version 3.0 (Fayers et al 2001)[4]. A higher score represents a higher ("better") level of functioning, or a higher ("worse") level of symptoms.

5.4.12.2. Analysis Methods

Compliance Rates

Compliance rates for completion of EORTC QLQ-C30 at each time point will be generated based on the actual number of assessments received over the number of expected.

Change from Baseline

For EORTC QLQ-C30 domain scores (GHS/QoL, functional scales, symptom scales), and single symptom items, descriptive statistics will be reported at baseline and at each visit for absolute value and for change from baseline.

The change of scores from baseline over time will be assessed using mixed-effects model for repeated measures (MMRM) analysis based on restricted maximum likelihood (REML). The model will include participants as a random effect, and baseline value, treatment group, time in week, treatment-by-time interaction, and stratification factors as fixed effects. The treatment comparison will be based on the least squares means and the 2-sided 95% CI will be estimated. Randomized participants who receive at least 1 dose of study treatment and have at least 1 evaluable postbaseline measurement will be included in the analysis.

Line plots of mean and the change from baseline with standard error over time by treatment group will be displayed.

Time to Symptom Deterioration (TTSD)

Time to symptom deterioration in EORTC QLQ-C30 measures is defined as the time from randomization until the date of the first clinically meaningful deterioration, or death. A clinically meaningful change is defined as a decrease for GHS/QoL and functional scales or an increase for symptom scales/items in the score from baseline of ≥10 (Osoba et al 1998)[6]. Participants who have not shown a deterioration or have not died at the time of analysis will be censored at their last PRO assessment date.

Time to symptom deterioration in EORTC QLQ-C30 measures will be analyzed using the similar method as the primary analysis of PFS.

Deterioration in each individual measures (as defined previously) will also be summarized using count and percentage by treatment and study visit.

5.4.13. Patient Reported Outcome – PROMIS-PF

5.4.13.1. Definition

Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function (PF) (short-form) is used to characterize and better understand overall health, level of physical disability, and general well-being. Physical function is a foundation for commonly used general and cancer-specific PRO measures.

5.4.13.2. Analysis Methods

Compliance Rates

Compliance rates for completion of PROMIS-PF at each time point will be generated based on the actual number of assessments received over the number of expected.

Following analyses will be conducted based on Total score (T-score).

Change from Baseline

The change of scores from baseline over time for PROMIS-PF T-score will be assessed using mixed-effects model for repeated measures (MMRM) analysis based on restricted maximum likelihood (REML). The model will include participants as a random effect, and baseline value, treatment group, time in week, treatment-by-time interaction, and stratification factors as fixed effects. The treatment comparison will be based on the least squares means and the 2-sided 95% CI will be estimated. Randomized participants who receive at least 1 dose of study treatment and have at least 1 evaluable postbaseline measurement will be included in the analysis.

Time to Symptom Deterioration (TTSD)

Time to symptom deterioration in PROMIS-PF T-score measures is defined as the time from randomization until the date of the first clinically meaningful deterioration, or death. A clinically meaningful change is defined as an decrease for T-score from baseline of ≥4. Participants who have not shown a deterioration or have not died at the time of analysis will be censored at their last PRO assessment date.

Time to symptom deterioration in PROMIS-PF T-score measures will be analyzed using the similar method as the primary analysis of PFS.

5.5. Tertiary/Exploratory Endpoint(s) Analysis

5.5.1. PFS Based on Investigator Assessment

PFS based on investigator review using RECIST v1.1 will be analyzed in a similar manner as described in Section 5.3.3 for PFS based on BICR.

5.5.2. Disease Control Rate (DCR)

5.5.2.1. Definition

DCR is defined as the percentage of participants achieving complete or partial response, or stable disease as defined by BICR using RECIST v1.1.

5.5.2.2. Analysis Methods

DCR will be analyzed using the similar method as the analysis of ORR.

5.5.3. Time to Treatment Discontinuation (TTD)

5.5.3.1. Definition

TTD is defined as the time from randomization to discontinuation of all study treatments for any reason, including disease progression, treatment toxicity, death, and will be utilized to capture clinical benefit for participants continuing treatment beyond RECIST v1.1 defined disease progression.

5.5.3.2. Analysis Methods

TTD will be analyzed using the similar method as the primary analysis of PFS.

5.5.4. Time to Response

5.5.4.1. Definition

Time to response (i.e. time to first response) is defined as the time from the date of randomization to the date of first documentation of a response (PR or CR) prior to any disease progression and subsequent anticancer therapy, as defined by BICR using RECIST v1.1., for participants who have PR or CR as their best response.

5.5.4.2. Analysis Methods

Descriptive statistics (mean, standard deviation, median, and range) will be provided to summarize time to response.

5.5.5. Patient Reported Outcome - EQ-5D-5L

5.5.5.1. Definition

The EQ-5D-5L is a validated tool to measure health status and health utility. It is a 5-item questionnaire that assesses 5 domains including mobility, self-care, usual activities, pain/discomfort and anxiety/depression plus a visual analog scale (VAS) rating “health today” with anchors ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). The scores for the 5 separate questions are categorical and cannot be analyzed as cardinal numbers. However, the scores for the 5 dimensions are used to compute a single utility score ranging from zero (0.0) to 1 (1.0) representing the general health status of the individual (but allows for values less than 0 by UK scoring algorithm).

5.5.5.2. Analysis Methods

Compliance Rates

Compliance rates for completion of EQ-5D-5L at each time point will be generated based on the actual number of assessments received over the number of expected.

Change from Baseline

Descriptive statistics will be reported for the VAS and utility score at baseline and at each visit for absolute value and for change from baseline.

The change in the VAS and utility score from baseline over time will be analyzed using MMRM in a similar manner to EORTC QLQ-C30 measures. Randomized participants who receive at least 1 dose of study treatment and have at least 1 evaluable postbaseline measurement will be included in the analysis.

Line plots of mean and the change from baseline with standard error over time by treatment group will be displayed.

5.5.6. Patient Reported Outcome – PRO-CTCAE

5.5.6.1. Definition

The National Cancer Institute (NCI) Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) is a validated tool to assess the symptomatic toxicities self-reported by participants in clinical trials**.**

5.5.6.2. Analysis Methods

Compliance rates for completion of PRO-CTCAE at each time point will be generated based on the actual number of assessments received over the number of expected.

PRO-CTCAE data will be summarized descriptively by treatment group using count and percentage. Shift table from baseline to worst postbaseline value will be provided.

5.6. Safety Analyses

All safety analyses will be based on the safety analysis set and based on actual treatment received, unless otherwise specified.

For all continuous safety variables, descriptive statistics by treatment group will include the N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized by treatment group using frequency counts and percentages.

5.6.1. Extent of Exposure

All the exposure information will be summarized based on safety analysis set by treatment group, and for each study drug within a treatment group.

Study treatment duration is defined as (date of last dose of study treatment – date of first dose of study treatment) +1. Descriptive statistics for duration of study treatment will be presented in months by treatment group.

The total number of administration cycles for each participant will be summarized by treatment group using descriptive statistics. Cumulative duration will be provided by cycle (≥1 cycle, ≥2

cycles, …). Total number of infusion for each participant will be summarized by descriptive statistics.

Total dose days of study drug, defined as the total number of days that study drug was administered to the participant (excluding days off study drug), will be summarized for lazertinib. Total dose administrated for lazertinib will be summarized by descriptive statistics.

The relative dose intensity (%) defined as the ratio of total received dose versus total prescribed dose will be summarized by descriptive statistics.

The number of interruptions during the infusion due to AE will be summarized.

The number (%) of participants with a dose reduction/dose not administrated will be summarized. Reasons for dose reduction/dose not administrated will also be summarized.

The number (%) of participants with cycle delay will be summarized by treatment group. The reason for cycle delay will also be summarized.

5.6.2. Compliance of Disease Evaluation

Tumor assessment will occur at regular intervals, as defined per SoA in study protocol. Descriptive statistics will be provided for imaging assessments separately, by treatment and overall for:

- Number of participants missed at least 1 scheduled disease evaluation
- Number of participants missed 2 or more consecutive scheduled disease evaluation
- Number of missed scheduled disease evaluation per participant

5.6.3. Adverse Events

The verbatim terms used in the CRF by investigators to identify adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). Any AE occurring at or after the initial administration of study treatment through the day of last dose plus 30 days, or until the start of subsequent anticancer therapy (if earlier), is considered to be treatment emergent. If the event occurs on the day of the initial administration of study treatment, and either event time or time of administration are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered to be treatment emergent unless it is known to be prior to the first administration of study treatment based on partial onset date or resolution date. All reported treatment-emergent adverse events (TEAEs) will be included in the analysis. For each adverse event, the number and percentage of participants who experience at least 1 occurrence of the given event will be summarized by treatment group.

The incidence (%) of TEAEs will be summarized overall, by MedDRA system organ class (SOC) and preferred term (PT), by toxicity grade, and by relationship to study drug administration.

5.6.3.1. Treatment Emergent Adverse Events

An overview of TEAEs reported through the study will be provided. The overview will include summaries of participants with TEAEs, with TEAEs related to each study drug, with TEAEs of maximum toxicity grade of 1 to 5, Serious TEAEs, TEAEs leading to discontinuation of any study drug, and deaths due to TEAE.

5.6.3.1.1. All TEAEs

- Incidence (%) of TEAEs by SOC and PT

5.6.3.1.2. Toxicity Grade 3 or higher TEAEs

- Incidence (%) of toxicity grade 3 or higher TEAEs by SOC and PT

5.6.3.1.3. Study Drug-Related TEAEs

- Incidence (%) of TEAEs by relationship to each treatment/study drug, and by SOC and PT
- Incidence (%) of TEAEs with toxicity grade 3 or higher by relationship to each treatment/study drug, and by SOC and PT
- Incidence (%) of TEAEs leading to study drug interruption/dose reduction by relationship to each treatment/study drug, and by SOC and PT
- Incidence (%) of TEAEs leading to study drug discontinuation by relationship to each treatment/study drug, and by SOC and PT

5.6.3.1.4. Serious TEAEs

- Incidence (%) of serious TEAEs by SOC and PT
- Incidence (%) of serious TEAEs by toxicity grade, and by SOC and PT
- Incidence (%) of serious TEAEs by relationship to each treatment/study drug, and by SOC and PT
- Listing of participants with any serious TEAEs

5.6.3.1.5. TEAEs Leading to Study Drug Interruption/Dose Reduction

Incidence (%) of TEAEs leading to each study drug/dose reduction will be summarized respectively by SOC and PT. The summaries will be presented for all toxicity grades and for toxicity grade 3 or higher.

5.6.3.1.6. TEAEs Leading to Discontinuation of Study drug

Incidence (%) of TEAEs leading to each study drug discontinuation will be summarized by SOC and PT. The summaries will be presented by all toxicity grades and toxicity grade 3 or higher. The AEs leading to discontinuation of any study drug are based on AEs recorded in the AE CRF page with an action taken of drug withdrawal for any study drug.

5.6.3.2. Adverse Events of Special Interest

Adverse events of special interest are pneumonitis/interstitial lung disease (ILD), rash, infusion-related-reaction (IRR), and venous thromboembolic (VTE) events. The MedDRA preferred terms associated with each of these categories are identified in Appendix 6 Adverse Events of Special Interest. Additional information will be collected for pneumonitis/ILD and IRR events.

Treatment-emergent adverse events of special interest will be included for analysis. Incidence (%) for the following AEs will be provided for each AE of special interest as appropriate:

- TEAEs by PT
- TEAEs by toxicity grade
- TEAEs of toxicity grade 3 or higher by PT
- Serious TEAEs by PT
- TEAEs by relationship to study drug
- Serious TEAEs by PT
- Serious TEAEs by relationship to study drug
- TEAEs leading to study drug discontinuation by PT
- TEAEs leading to study drug discontinuation by relationship to study drug
- TEAEs leading to death by PT

Additional analyses will be provided based on information collected in CRF.

Pneumonitis/ILD

Relative onset day (since day 1) of pneumonitis/ILD will be summarized by descriptive statistics (N, mean, standard deviation, median, and range).

All information related to pneumonitis/ILD collected in CRF page will be presented in listing.

Rash

Relative onset day (since day 1) and duration will be summarized for rash by descriptive statistics (N, mean, standard deviation, median, and range) in days.

Number of participants with rash will be summarized by cycle and by toxicity grade.

IRR

Incidence (%) of IRR leading to infusion modification (infusion interrupted, infusion rate decreased, and infusion aborted) will be presented.

Relative onset day (since day 1), and duration will be summarized for IRR by descriptive statistics (N, mean, standard deviation, median, and range) in days.

Number of participants with IRR will be summarized by cycle and by toxicity grade.

VTE

Relative onset day (since day 1) of VTE will be summarized by descriptive statistics (N, mean, SD, median, and range). Duration of VTE will be summarized by descriptive statistics (N, Median with 95% CI, and range).

The number (%) of patients with VTE that require hospitalization, hospitalization for management of symptoms will be presented. The number (%) of patients with VTE used prophylactic anticoagulation therapy will be presented. Incidence (%) of treatment-emergent symptoms of VTE will be presented by SOC and PT. The number (%) of patients with VTE compared with use of anticoagulants will be presented.

5.6.3.3. Deaths

5.6.3.3.1. Death Due to TEAEs

The number of participants who died due to TEAEs will be summarized by preferred term and relationship to study drug. The TEAEs included in this table are AEs with outcome of death or toxicity grade of 5 recorded in the AE CRF page within 30 days of the last dose or until the start of subsequent anticancer therapy (if earlier), or considered as related to study drug by investigator.

A listing of participants who died due to TEAE will be provided.

5.6.3.3.2. All Deaths

A summary of all death and cause of death will be tabulated. Specifically, the number of participants who died during the study will be summarized. The primary cause of death collected on the death information CRF page will be reported.

The similar summaries will be presented for participants who died within 30 days of last study drug dose.

5.6.4. Additional Safety Assessments

5.6.4.1. Clinical Laboratory Tests

Clinical laboratory tests will be displayed for the participants included in the safety analysis set.

Descriptive statistics will be presented for all chemistry and hematology laboratory tests at scheduled time points. Change from baseline over time will be summarized and displayed. Plots for selected laboratory tests change over time may be provided.

NCI-CTCAE version 5.0 will be used to derive toxicity grades for clinical laboratory tests when applicable. Shift tables from baseline to worst value on treatment (from treatment start to 30 days

after last dose date or until the start of subsequent anti-cancer therapy, whichever is later) will be provided. The worst toxicity grade during the treatment will be tabulated.

An eDISH plot of peak ALT/ AST versus peak BILI will be provided along with a listing of participants who had ALT/ AST values > 3xULN or BILI values > 2xULN.

Laboratory criteria for potential Hy's Law cases are defined as:

- Peak aminotransaminases (AT, either ALT or AST) of >3xULN (Upper Limit of Normal);
- Total bilirubin ≥2xULN;
- Alkaline phosphatase (ALP) <2xULN prior to or on the same date of the first occurrence of total bilirubin ≥2x ULN;

Note: data from all the on-treatment (postbaseline) visits are combined to check the above laboratory criteria.

- All potential Hy's Law cases based on the laboratory criteria will be presented.

5.6.4.2. Vital Signs and Physical Examination Findings

Continuous vital sign parameters including temperature, respiratory rate, oxygen saturation, pulse/heart rate, blood pressure (systolic and diastolic), as well as weight from physical examination will be summarized at each scheduled timepoint. Change from baseline will be summarized over time. Descriptive statistics (mean, standard deviation, median, minimum and maximum) will be presented.

Post baseline physical examination findings were collected as AEs, and therefore will not be summarized.

5.6.4.3. Electrocardiogram

Electrocardiograms (ECG) will be performed at Screening (in triplicate), C1D1 (before dose and in single), and as clinically indicated post-treatment. If ECG is performed within 72 hours before the first dose of study treatment, the assessment does not need to be repeated at C1D1.

A listing of clinically relevant ECG abnormalities will be provided.

5.6.5. Left Ventricular Ejection Fraction (LVEF)

During the Screening Phase each participant will undergo a baseline LVEF assessment performed locally by ECHO or MUGA scan to demonstrate eligibility (ie, a LVEF within the normal range). Participant may receive ECHO or MUGA scan as clinical indicated post-treatment.

A listing of clinically relevant LVEF abnormalities will be provided.

5.6.6. ECOG Performance Status

Frequencies of ECOG performance over time will be summarized. In addition, shift from baseline to worst score during treatment will be provided.

5.7. Other Analyses

5.7.1. Pharmacokinetics

Serum samples will be collected from participants in Arm A and Arm C for PK and immunogenicity assessments of amivantamab at the time points outlined in Table 2 of study protocol. Plasma samples will be collected from participants in Arm A for the evaluation of PK of lazertinib.

PK analyses will be performed on the PK analysis set, defined as randomized participants who received at least 1 dose of a corresponding study drug and have at least 1 evaluable postbaseline concentration measurement.

Concentrations below the lowest quantifiable concentration or missing data will be labeled as such in the concentration data presentations. Concentrations below the lower quantifiable concentration will be treated as zero in the summary statistics.

Descriptive statistics (N, mean, SD, median, range, CV (%) and IQ range) will be used to summarize amivantamab concentrations at each sampling time point and for each PK parameter of amivantamab and lazertinib. PK data may be displayed graphically, such as mean +/- SD PK concentrations over time by study intervention.

Amivantamab concentrations will be presented based on the following baseline body weight categories at each time point:

- <80 kg
- ≥ 80 kg

All participants and samples excluded from the analysis will be clearly documented.

The pharmacokinetic serum/plasma concentration-time data collected from this study will be combined with similar data from other studies to perform population PK and assess the relationship between PK or immunogenicity and selected safety and efficacy endpoints. Details will be provided in a population PK and exposure-response analysis plan and results of the analysis will be presented in a separate report.

5.7.2. Immunogenicity

The incidence (%) of antibodies to Amivantamab will be summarized based on Immunogenicity Analysis Set, defined as all participants who receive at least 1 dose of Amivantamab and have appropriate samples for detection of antibodies to Amivantamab (i.e., participants with at least 1 sample obtained after their first dose of Amivantamab).

A listing of participants who are positive for antibodies to Amivantamab will be provided. The maximum titers of antibodies to Amivantamab will be summarized for participants who are positive for antibodies to Amivantamab.

Other immunogenicity analyses may be performed to further characterize the immune responses that are generated.

5.7.3. Pharmacodynamics (PD)/Biomarkers

Analyses are planned to explore PD and other biomarkers that may be indicative of the mechanisms of action of the study intervention or predictive of efficacy. Correlation of baseline expression levels or changes in expression levels with response or time-to-event endpoints could identify responsive (or resistant) subgroups. Any PD or other biomarker measures will be listed, tabulated, and plotted, as appropriate.

Assessment of additional genes or biomarkers (DNA, RNA, or protein) relevant to lung or other cancers or the mechanism of action of study interventions, may also be performed in blood samples collected during study to better understand mechanisms of response or resistance to study interventions.

Alterations in blood may be evaluated for correlation with response to study interventions, tumor burden, and disease progression as data warrant.

Plasma mutation data derived from ctDNA and tumor tissue NGS analyses collected from this study will be used to perform mutational analysis and assess the relationship of individual mutations, and classes of mutations, to efficacy endpoints. IHC analyses on tissue specimens collected from this study will be used to assess the relationship of exploratory endpoints to efficacy endpoints. Additional exploratory endpoints may be explored from serum samples collected from this study and may be used to understand the relationship of these endpoints to efficacy endpoints. Results of these analyses will be presented in a separate report.

Analysis

Biomarker analyses are dependent upon the availability of appropriate biomarker assays and clinical response rates. Biomarker analysis may be deferred or not performed, if during or at the end of the study, it becomes clear that the analysis will not have sufficient scientific value for biomarker evaluation, or if there are not enough samples or responders to allow for adequate biomarker evaluation. In the event the study is terminated early or shows poor clinical efficacy, completion of biomarker assessments is based on justification and intended utility of the data.

5.7.4. Pharmacokinetic/Pharmacodynamic Relationships

The relationship between PK and PD measures may be evaluated by PK/PD modeling. Participants may be grouped by dose schedule or clinical response. Results of PD and exploratory biomarker analyses will be presented in separate reports.

5.7.5. Health Economics

Not Applicable

5.7.6. Definition of Subgroups

The following pre-specified subgroup analyses are to be performed for the efficacy and/or safety endpoints. Additional subgroup analyses may be planned if deemed necessary.

Definition of Subgroups

Subgroup	Definition
Age Group	<65 years, ≥65 years; <75 years, ≥75 years
Sex	Male, Female
Race	Asian, Non-Asian
Weight	<80 kg, ≥80 kg
History of brain metastasis	Yes, No
Osimertinib line of therapy	First-line, Second-line
ECOG performance status score	0, 1
History of smoking	Yes, No

5.8. Interim Analyses

There is no interim analysis planned for PFS.

For both LACP/ACP-L vs CP and ACP vs CP, two interim analyses are planned for OS. The first interim analysis for OS will be performed at the time of the analysis for PFS, when approximately 170 deaths (all 3 Arms combined, approximately 43% of the total planned OS events) are anticipated. The second interim analysis for OS will be performed approximately 32 months after the first participant is randomized, when approximately 300 deaths (all 3 arms combined, approximately 75% of the total planned OS events). Final analysis for OS will be conducted at approximately 48 months after the first participant is randomized, when 400 deaths (all 3 arms combined) are anticipated.

Following the testing algorithm for multiple hypotheses using the graphical procedure as described in Section 2.1, OS will be tested in the FAS at the two-sided alpha re-allocated from previous rejected hypotheses to the respective comparisons of LACP/ACP-L vs CP and ACP vs CP. The significance level at the interim analyses for OS will be determined based on the O'Brien Fleming alpha spending approach as implemented by the Lan-DeMets method.

5.8.1. Independent Data Monitoring Committee

An IDMC consisting of two medical experts in the relevant therapeutic area and one statistician not otherwise participating in the study, has been established to review safety results.

The IDMC will review safety results for LACP/ACP-L in the ongoing Phase 1 study prior to enrollment of participants in this Phase 3 study. An IDMC meeting will occur after first 50 participants have been randomized and treated for at least 2 cycles. Regular safety review meetings will occur approximately every 4 months thereafter.

The frequency of meetings may change at any moment as per IDMC request. During safety analysis for IDMC review, enrollment will continue.

Other meetings or data reviews may be scheduled at the discretion of the IDMC or upon request by the Sponsor Committee.

5.9. Analysis by Arm A Dosing Schedule

Following the 04 November 2022 safety data review for MARIPOSA-2, the IDMC recommended modification of Arm A to withhold lazertinib during administration of carboplatin due to an apparent imbalance in AEs affecting Arm A. The lazertinib dosing schedule for participants randomized to Arm A was thus modified in Amendment 5 of the protocol, such that new participants randomized to Arm A began treatment with amivantamab, carboplatin, and pemetrexed, and only started lazertinib after treatment with carboplatin treatment was complete.

As outlined in previous sections, the primary analysis will be based on safety and full analysis set irrespective of Arm A dosing schedule change. However, to show that the two dosing schedules have no clinical relevant impact on key efficacy endpoints, supplemental descriptive analyses will be performed for subgroup of subjects randomized prior to 07 November 2022 (S1) and subgroup of subjects randomized from 07 November 2022 onward (S2). Subjects in arm A from S1 (A-S1) are intended to receive LACP, and subjects in arm A from S2 are intended to receive ACP-L. Subjects from arm B from both subgroups (B-S1 and B-S2) will receive CP as treatment, and subjects from arm C from both subgroups (C-S1 and C-S2) will receive ACP as treatment.

5.9.1. Subject Distribution by Dosing Schedule

Following by-treatment and overall summaries will be generated for subgroups S1 and S2, respectively:

- Number of subjects randomized, treated, discontinued and ongoing
- Primary reasons for study treatment discontinuation

5.9.2. Demographics and Baseline characteristics by Dosing Schedule

The demographics and baseline characteristics outlined in Appendix 2 will be summarized by by-treatment and overall for subgroups S1 and S2, respectively.

5.9.3. Study Drug Administration by Dosing Schedule

Descriptive statistics will be provided on treatment duration and dose modifications by treatment for treated subjects from subgroups S1 and S2, respectively.

5.9.4. Safety Analysis by Dosing Schedule

The safety analysis outlined in Section 5.6 will be performed among treated subjects for A-S1vs B, A-S2 vs B, A-S1 vs B-S1 and A-S2 vs B-S2.

5.9.5. Efficacy Analysis by Dosing Schedule

The primary PFS analysis and statistical inference will be based on the full analysis set regardless of the Arm A dosing schedule. In order to assess the consistency of treatment benefit, supplemental analysis by dosing schedule subgroup will be performed for A-S1vs B, A-S2 vs B, A-S1 vs B-S1 and A-S2 vs B-S2 on the following key efficacy endpoints: PFS (by BICR)

- Objective Response (by BICR)
- Overall Survival
- Duration of Response (by BICR)
- Time to Subsequent Therapy (TTST)
- Progression-free Survival After the First Subsequent Therapy (PFS2)
- Time to Symptomatic Progression (TTSP)
- Intracranial PFS (by BICR)

The key efficacy endpoints for the subgroups will be analyzed using similar methods as the analysis for full analysis set. There will be no formal hypothesis testing for these comparisons.

To describe the contribution of lazertinib to LACP or ACP-L, same efficacy analyses will be performed for A-S1vs C, A-S2 vs C, A-S1 vs C-S1 and A-S2 vs C-S2 using summary statistics; there will be no formal hypothesis testing for these comparisons.

5.10. Statistical Analysis of Extension Cohort

The primary objectives of the extension cohort are to further describe the safety and efficacy for the ACP-L dosing schedule versus ACP with additional data.
To further describe the contribution of lazertinib to ACP-L, similar efficacy analyses as for main study will be performed for pooled data from all subjects randomized to receive ACP-L in main study and extension cohort (A-S2 and A2) vs pooled data from all subjects randomized to receive ACP in main study and extension cohort (C and C2) using summary statistics; and there will be no formal hypothesis testing for these comparisons. Similar safety analyses as for main study will be performed for the pooled data among treated subjects.

5.10.1. Participant Disposition

Screened participants and reasons for screen failures will be summarized overall.

The number of participants in the following disposition categories will be summarized for subjects in extension cohort (A2 and C2), as well as for all subjects to treat with ACP-L (A-S2 and A2) and all subjects to treat with ACP (C and C2), by treatment group and overall:

- Participants randomized
- Participants who received study treatment
- Participants who discontinued study treatment
- Reason for discontinuation of study treatment
- Participants who terminated study prematurely
- Reason for termination of study
- Participants who completed the study

A listing of participants will be provided for the following categories:

- Participants who discontinued study treatment
- Participants who terminated study prematurely

5.10.2. Demographics and Baseline characteristics

The demographics and baseline characteristics outlined in Appendix 2 will be summarized by treatment and overall for subjects in the extension cohort (A2 and C2), as well as all subjects randomized to ACP-L (A-S2 and A2) and all subjects randomized to ACP (C and C2), respectively.

To assess the consistency of the populations enrolled prior to Arm A dose schedule change implementation and the populations enrolled after the change, demographic and baseline characteristics including age, BMI, baseline ECOG, histology grade at screening, cancer stage at screening and initial diagnosis NSCLC subtype will be summarized for subjects enrolled in arm C before arm A dosing schedule change implementation (C-S1) and all subjects in Arm A and C enrolled after Arm A dose schedule change implementation (A-S2+A2+ C-S2+C2). P-values from t-test of selected key continuous/numerical variables or chi-squared test of selected key categorical variables may be used to assist the assessment between the two groups.

5.10.3. Study Drug Administration

Descriptive statistics will be provided on treatment duration and dose modifications by treatment for treated subjects in the extension cohort (A2 and C2), as well as all subjects treated with ACP-L (A-S2 and A2) and all subjects treated with ACP (C and C2), respectively.

5.10.4. Safety Analyses

The safety analysis outlined in Section 5.6 will be performed among all subjects treated with ACP-L (A-S2 and A2) and all subjects treated with ACP (C and C2), respectively.

For subjects enrolled after dosing schedule change implementation in main study and extension cohort, overall Summary of adverse events, summary of serious adverse events, and adverse events with toxicity grade 3 or higher will be provided for subjects treated with ACP-L (A-S2 and A2) vs subjects treated with ACP (C-S2 and C2). Similar analyses will be performed as well for treated subjects in the extension cohort (A2 vs C2).

5.10.5. Efficacy Analyses

To describe the contribution of lazertinib to ACP-L with additional data, key efficacy analyses will be performed for all subjects to treat with ACP-L (A-S2 and A2) vs all subjects to treat with ACP (C and C2) using summary statistics, similar as the analyses performed for full analysis set in main study; there will be no formal hypothesis testing for these comparisons.

The key efficacy endpoints to be analyzed will include:

- PFS (by BICR)
- Objective Response (by BICR)
- Duration of Response (by BICR)
- Intracranial PFS (by BICR)

Additional supplementary analyses will also be performed to assess any potential bias introduced because of nonconcurrent randomization, as subjects in C-S1 enrolled prior to Arm A dose schedule change implementation but all subjects in ACP-L enrolled afterwards. The efficacy analyses described above will also be conducted for

- ACP-L vs ACP for subjects enrolled post Arm A dosing schedule change only (A-S2 and A2 vs C-S2 and C2).
- ACP-L vs ACP based on all subjects to treat with ACP-L (A-S2 and A2) and all subjects to treat with ACP (C and C2), but subjects in C-S1 (subjects enrolled before Arm A dosing schedule change) will be given partial weight (i.e. 50% weight).

SUPPORTING DOCUMENTATION

Appendix 1 List of Abbreviations

ACP	amivantamab, carboplatin, and pemetrexed
ACP-L	amivantamab, carboplatin, pemetrexed and lazertinib
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	anatomic and therapeutic class
BICR	blinded independent central review
BILI	bilirubin
BMI	body mass index
CI	confidence interval
CP	carboplatin and pemetrexed
CR	complete response
CRF	case report form
CTCAE	common terminology criteria for adverse events
CV	coefficient of variation
DCR	disease control rate
DoR	duration of response
ECG	electrocardiogram
ECHO	echocardiogram
ECOG	Eastern Cooperative Oncology Group
EGFR	epidermal growth factor receptor
EQ-5D-5L	EuroQol five-dimensional descriptive system (5-level version)
FAS	full analysis set
ILD	interstitial lung disease
IDMC	independent data monitoring committee
IPCW	inverse probability of censoring weighting
IQ	interquartile
IRR	infusion-related reaction
LACP	lazertinib, amivantamab, carboplatin, and pemetrexed
LLN	lower limit of normal
LVEF	left ventricular ejection fraction
IWRS	interactive web response system
MedDRA	medical dictionary for regulatory activities
MET	mesenchymal-epithelial transition
MUGA	multigated acquisition
NCI CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NSCLC	non-small cell lung cancer
NSCLC-SAQ	Non-Small Cell Lung Cancer - Symptom Assessment Questionnaire
ORR	objective response rate
OS	overall survival
PD	pharmacodynamic(s)
PFS	progression free survival
PK	pharmacokinetic(s)
PR	partial response
PRO	patient reported outcomes
QT	uncorrected QT interval
QTc	corrected QT
QTcF	corrected QT interval by Fridericia
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	statistical analysis plan

SD	standard deviation
SoA	schedule of activities
TEAE	treatment-emergent adverse event
TKI	tyrosine kinase inhibitor
TTSP	time to symptomatic progression
ULN	upper limit of normal
US	United States
VTE	Venous thromboembolic
WHO-DD	world health organization drug dictionary

Appendix 2 Demographics and Baseline Characteristics

The number of participants in each analysis set will be summarized and listed by treatment group and overall. In addition, the distribution of participants by country and site ID will be presented unless otherwise noted.

Table below presents a list of the demographic variables that will be summarized by treatment group and overall.

Continuous Variables	**Summary Type**
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum].
Weight (kg)	
Height (cm)	
Body Mass Index (BMI) (kg/m^2)	
Categorical Variables	
Age (<65 years, ≥65 years; <75 years, ≥75 years)	Frequency distribution with the number and percentage of participants in each category.
Sex (male, female, undifferentiated)	
Weight (<80 kg, ≥80 kg)	
Race[a] (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Multiple)	
Race (Asian, non-Asian)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	
Baseline ECOG performance status (0, 1)	
Osimertinib line of therapy (First-line, Second-line)	
History of Smoking (Yes, No)	

[a]If multiple race categories are indicated, the Race is recorded as 'Multiple'

The following table presents a list of the baseline characteristics variables that will be summarized by treatment group and overall.

Continuous Variables	**Summary Type**
Time since initial lung cancer diagnosis (months)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum].
Time since metastatic disease diagnosis (months)	
Number of prior lines of systemic therapy taken from early stage	
Categorical Variables	
NSCLC subtype at initial diagnosis (adenocarcinoma, large cell carcinoma, squamous cell carcinoma, other)	Frequency distribution with the number and percentage of participants in each category.
Histology grade at initial diagnosis (moderately differentiated, poorly differentiated, well differentiated, other)	
Cancer stage at initial diagnosis (IA, IB, IIA, IIB, IIIA, IIIB, IIIC, IVA, IVB)	
Histology grade at screening (moderately differentiated, poorly differentiated, well differentiated, other)	
Cancer stage at screening (IIIB, IIIC, IVA, IVB)	
Location of metastasis at screening (bone, liver, brain, lymph node, adrenal gland, lung, other)	
Prior systemic therapy (adjuvant, neo-adjuvant, curative/palliative/any other intent, concurrent chemoradiation) taken from early stage	

A summary of stratification factors (osimertinib line of therapy, history of brain metastasis and Asian race) used in the randomization based on IWRS will be provided to evaluate whether or not randomization process was appropriately executed in the study.

Appendix 3 Protocol Deviations

In general, the following list of major protocol deviations may have the potential to impact participants' rights, safety or well-being, or the integrity and/or result of the clinical study. Participants with major protocol deviations will be identified prior to database lock and the participants with major protocol deviations will be summarized by category.

- Developed withdrawal criteria but not withdrawn
- Entered but did not satisfy criteria
- Received a disallowed concomitant treatment
- Received wrong treatment or incorrect dose
- Any major PD due to COVID
- Other

A listing of all major protocol deviations including participant ID, type of deviation, and reason will be provided.

Appendix 4 Prior and Concomitant Medications

Prior and Concomitant medications collected in the CRF page will be coded using the World Health Organization Drug Dictionary (WHO-DD) and summarized for each treatment group and overall.

Prior medications will be summarized by ATC level/preferred terms and treatment. The number and percentage of participants who received prior systemic therapy will be summarized.

Summaries of concomitant medications will be presented by ATC level/preferred terms. The proportion of participants who receive each concomitant medication will be summarized as well as the proportion of participants who receive at least 1 concomitant medication.

The incidence (%) of pre-infusion and post-infusion medication will be presented by ATC level/preferred terms.

Appendix 5 Medical History

Medical history collected at baseline or screening visit will be summarized by system-organ class and preferred term for each treatment group and overall.

Appendix 6 Adverse Events of Special Interest

Adverse events of special interest are defined as follows:

AE of Special Interest Category	Preferred Term
Infusion Related Reaction	INFUSION RELATED REACTION
Rash	ACNE ACNE CONGLOBATA ACNE CYSTIC ACNE FULMINANS ACNE PUSTULAR ACNE VARIOLIFORMIS ACUTE GENERALISED EXANTHEMATOUS PUSTULOSIS DERMATITIS DERMATITIS ACNEIFORM DERMATITIS EXFOLIATIVE DERMATITIS INFECTED DRUG ERUPTION EPIDERMOLYSIS ERYTHEMA ERYTHEMA MULTIFORME EXFOLIATIVE RASH FOLLICULITIS HERPES GESTATIONIS IMPETIGO HERPETIFORMIS MACULE MUCOCUTANEOUS RASH NODULAR RASH PALMAR ERYTHEMA PAPULE PERINEAL RASH PRIDE SYNDROME PUSTULE RASH RASH ERYTHEMATOUS RASH FOLLICULAR RASH MACULAR RASH MACULO-PAPULAR RASH MACULOVESICULAR RASH MORBILLIFORM RASH PAPULAR RASH PRURITIC RASH PUSTULAR RASH VESICULAR SJS-TEN OVERLAP SKIN EXFOLIATION SKIN LESION STEVENS-JOHNSON SYNDROME TOXIC EPIDERMAL NECROLYSIS TOXIC SKIN ERUPTION
Pneumonitis/Interstitial Lung Disease	ACUTE INTERSTITIAL PNEUMONITIS INTERSTITIAL LUNG DISEASE PNEUMONITIS
Venous Thromboembolic (VTE)	THROMBOSIS

AE of Special Interest Category	Preferred Term
	EMBOLISM Preferred terms from SMQ= Embolic and thrombotic events, venous

Appendix 7 Laboratory Toxicity Grading

The grading scale use for lab assessments is based on 'Common Terminology Criteria for Adverse Events (CTCAE) v5.0'.

Pre-baseline measurements will use the same grading ranges as applied to baseline measurements. In case a test has two sets of ranges – one for baseline normal and one for baseline abnormal, the one for baseline normal will be applied for all measurements taken pre-baseline and on baseline.

Text in gray italic in the table is present in the grading scale, but is not applied by Janssen when grading lab data.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Blood and lymphatic system disorders					
Anemia	Hemoglobin (Hgb) <LLN - 10.0 g/dL; <LLN - 6.2 mmol/L; <LLN - 100 g/L	Hemoglobin (Hgb) <10.0 - 8.0 g/dL; <6.2 - 4.9 mmol/L; <100 - 80g/L	Hemoglobin (Hgb) <8.0 g/dL; <4.9 mmol/L; <80 g/L; *transfusion indicated*	*Life-threatening consequences; urgent treatment indicated*	Clinical signs and symptoms are not taken into consideration for grading.
Leukocytosis	-	-	>100,000/mm3; >100 x 10e9 /L	*Clinical manifestations of leucostasis; urgent treatment indicated*	Clinical signs and symptoms are not taken into consideration for grading; Added ranges in SI unit (x 10e9 /L)
Investigations					
Activated partial thromboplastin time prolonged	>ULN - 1.5 x ULN	>1.5 - 2.5 x ULN	>2.5 x ULN; *bleeding*	-	Clinical signs and symptoms are not taken into consideration for grading.
Alanine aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Alkaline phosphatase increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Aspartate aminotransferase increased	>ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Blood bilirubin increased	>ULN - 1.5 x ULN if baseline was normal;	>1.5 - 3.0 x ULN if baseline was normal;	>3.0 - 10.0 x ULN if baseline was normal;	>10.0 x ULN if baseline was normal;	Ranges defined for "abnormal baseline" are

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
	> 1.0 - 1.5 x baseline if baseline was abnormal	>1.5 - 3.0 x baseline if baseline was abnormal	>3.0 - 10.0 x baseline if baseline was abnormal	>10.0 x baseline if baseline was abnormal	applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
CD4 lymphocytes decreased	<LLN - 500/mm3; <LLN - 0.5 x 10e9 /L	<500 - 200/mm3; <0.5 - 0.2 x 10e9 /L	<200 - 50/mm3; <0.2 x 0.05 - 10e9 /L	<50/mm3; <0.05 x 10e9 /L	
Cholesterol high	>ULN - 300 mg/dL; >ULN - 7.75 mmol/L	>300 - 400 mg/dL; >7.75 - 10.34 mmol/L	>400 - 500 mg/dL; >10.34 - 12.92 mmol/L	>500 mg/dL; >12.92 mmol/L	
CPK increased	>ULN - 2.5 x ULN	>2.5 x ULN - 5 x ULN	>5 x ULN - 10 x ULN	>10 x ULN	
Creatinine increased	Creatine Kinase >ULN - 1.5 x ULN	Creatine Kinase >1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	Creatine Kinase >3.0 x baseline; >3.0 - 6.0 x ULN	Creatine Kinase >6.0 x ULN	
Fibrinogen decreased	<1.0 - 0.75 x LLN; if abnormal, <25% decrease from baseline	<0.75 - 0.5 x LLN; if abnormal, 25 - <50% decrease from baseline	<0.5 - 0.25 x LLN; if abnormal, 50 - <75% decrease from baseline	<0.25 x LLN; if abnormal, 75% decrease from baseline; absolute value <50 mg/dL	Ranges defined for "abnormal" are applied only on values < LLN. Grade 0 will be assigned to values > ULN.
GGT increased	>ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	>2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	>5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	>20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal	Ranges defined for "abnormal baseline" are applied only if baseline > ULN. If baseline < LLN, then ranges for "normal baseline" are applied.
Haptoglobin decreased	<LLN	-	-	-	
Hemoglobin increased	Increase in >0 - 2 g/dL; Increase in >0 - 20 g/L	Increase in >2 - 4 g/dL; Increase in >20 - 40 g/L	Increase in >4 g/dL; Increase in >40 g/L	-	The increase indicates the level of increase above normal (above ULN). Applied as, e.g. grade 1 (g/dL): >ULN – ULN+2 g/dL; Added ranges in SI unit (g/L).
INR increased	>1.2 - 1.5; *>1 - 1.5 x baseline if on anticoagulation; monitoring only indicated*	>1.5 - 2.5; *>1.5 - 2.5 x baseline if on anticoagulation; dose adjustment indicated*	>2.5; *>2.5 x baseline if on anticoagulation; bleeding*	-	Concomitant therapy or clinical signs and symptoms are not taken into consideration for grading.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Lipase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; *>2.0 - 5.0 x ULN and asymptomatic*	>2.0 - 5.0 x ULN with signs or symptoms; *>5.0 x ULN and asymptomatic*	>5.0 x ULN and with signs or symptoms	"Asymptomatic" ranges are not taken into consideration for grading, i.e. worst case grading is applied.
Lymphocyte count decreased	<LLN - 800/mm3; <LLN - 0.8 x 10e9/L	<800 - 500/mm3; <0.8 - 0.5 x 10e9 /L	<500 - 200/mm3; <0.5 - 0.2 x 10e9 /L	<200/mm3; <0.2 x 10e9 /L	
Lymphocyte count increased	-	>4000/mm3 - 20,000/mm3; >4 - 20 x 10e9 /L	>20,000/mm3; >20 x 10e9 /L	-	Added ranges in SI unit (x 10e9 /L).
Neutrophil count decreased	<LLN - 1500/mm3; <LLN - 1.5 x 10e9 /L	<1500 - 1000/mm3; <1.5 - 1.0 x 10e9 /L	<1000 - 500/mm3; <1.0 - 0.5 x 10e9 /L	<500/mm3; <0.5 x 10e9 /L	Both Neutrophils and segmented neutrophils are graded using these criteria.
Platelet count decreased	<LLN - 75,000/mm3; <LLN - 75.0 x 10e9 /L	<75,000 - 50,000/mm3; <75.0 - 50.0 x 10e9 /L	<50,000 - 25,000/mm3; <50.0 - 25.0 x 10e9 /L	<25,000/mm3; <25.0 x 10e9 /L	
Serum amylase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN; *>2.0 - 5.0 x ULN and asymptomatic*	>2.0 - 5.0 x ULN *with signs or symptoms; >5.0 x ULN and asymptomatic*	>5.0 x ULN *and with signs or symptoms*	"Asymptomatic" ranges are not taken into consideration for grading, i.e. worst case grading is applied.
White blood cell decreased	<LLN - 3000/mm3; <LLN - 3.0 x 10e9 /L	<3000 - 2000/mm3; <3.0 - 2.0 x 10e9 /L	<2000 - 1000/mm3; <2.0 - 1.0 x 10e9 /L	<1000/mm3; <1.0 x 10e9 /L	
Metabolism and nutrition disorders					
Acidosis	pH <normal, but ≥7.3	-	pH <7.3	*Life-threatening consequences*	pH <normal is implemented as pH <LLN. Clinical signs and symptoms are not taken into consideration for grading.
Alkalosis	pH >normal, but <=7.5	-	pH >7.5	*Life-threatening consequences*	pH >normal is implemented as pH >ULN. Clinical signs and symptoms are not taken

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
					into consideration for grading.
Hypercalcemia	Corrected serum calcium of >ULN - 11.5 mg/dL; >ULN - 2.9 mmol/L; Ionized calcium >ULN - 1.5 mmol/L	Corrected serum calcium of >11.5 - 12.5 mg/dL; >2.9 - 3.1 mmol/L; Ionized calcium >1.5 - 1.6 mmol/L; *symptomatic*	Corrected serum calcium of >12.5 - 13.5 mg/dL; >3.1 - 3.4 mmol/L; Ionized calcium >1.6 - 1.8 mmol/L; *hospitalization indicated*	Corrected serum calcium of >13.5 mg/dL; >3.4 mmol/L; Ionized calcium >1.8 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hyperkalemia	Potassium >ULN - 5.5 mmol/L	Potassium >5.5 - 6.0 mmol/L; *treatment initiated*	Potassium >6.0 - 7.0 mmol/L; *hospitalization indicated*	Potassium >7.0 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypermagnesemia	Magnesium >ULN - 3.0 mg/dL; >ULN - 1.23 mmol/L	-	Magnesium >3.0 - 8.0 mg/dL; >1.23 - 3.30 mmol/L	Magnesium >8.0 mg/dL; >3.30 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypernatremia	Sodium >ULN - 150 mmol/L	Sodium >150 - 155 mmol/L; *treatment initiated*	Sodium >155 - 160 mmol/L; *hospitalization indicated*	Sodium >160 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypertriglyceridemia	Triglycerides 150 mg/dL - 300 mg/dL; 1.71 mmol/L - 3.42 mmol/L	Triglycerides >300 mg/dL - 500 mg/dL; >3.42 mmol/L - 5.7 mmol/L	Triglycerides >500 mg/dL - 1000 mg/dL; >5.7 mmol/L - 11.4 mmol/L	Triglycerides >1000 mg/dL; >11.4 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hypoalbuminemia	Albumin <LLN - 3 g/dL; <LLN - 30 g/L	Albumin <3 - 2 g/dL; <30 - 20 g/L	Albumin <2 g/dL; <20 g/L	*Life-threatening consequences; urgent treatment indicated*	Clinical signs and symptoms are not taken into consideration for grading.
Hypocalcemia	Corrected serum calcium of <LLN - 8.0 mg/dL; <LLN - 2.0 mmol/L;	Corrected serum calcium of <8.0 - 7.0 mg/dL; <2.0 - 1.75 mmol/L;	Corrected serum calcium of <7.0 - 6.0 mg/dL; <1.75 - 1.5 mmol/L;	Corrected serum calcium of <6.0 mg/dL; <1.5 mmol/L;	Clinical signs and symptoms are not taken into consideration for grading.

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
	Ionized calcium <LLN - 1.0 mmol/L	Ionized calcium <1.0 - 0.9 mmol/L; *symptomatic*	Ionized calcium <0.9 - 0.8 mmol/L; *hospitalization indicated*	Ionized calcium <0.8 mmol/L; *life-threatening consequences*	
Hypoglycemia	Glucose <LLN - 55 mg/dL; <LLN - 3.0 mmol/L	Glucose <55 - 40 mg/dL; <3.0 - 2.2 mmol/L	Glucose <40 - 30 mg/dL; <2.2 - 1.7 mmol/L	Glucose <30 mg/dL; <1.7 mmol/L; *life-threatening consequences; seizures*	Clinical signs and symptoms are not taken into consideration for grading. Urine glucose is not graded.
Hypokalemia	*Potassium <LLN - 3.0 mmol/L*	*Symptomatic with* Potassium <LLN - 3.0 mmol/L; treatment indicated	Potassium <3.0 - 2.5 mmol/L; *hospitalization indicated*	Potassium <2.5 mmol/L; *life-threatening consequences*	"Symptomatic" ranges are applied for grade 2, grade 1 not assigned, i.e. worst case applied. Clinical signs and symptoms are not taken into consideration for grading of grade 3 and 4.
Hypomagnesemia	Magnesium <LLN - 1.2 mg/dL; <LLN - 0.5 mmol/L	Magnesium <1.2 - 0.9 mg/dL; <0.5 - 0.4 mmol/L	Magnesium <0.9 - 0.7 mg/dL; <0.4 - 0.3 mmol/L	Magnesium <0.7 mg/dL; <0.3 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading.
Hyponatremia	Sodium <LLN - 130 mmol/L	*Sodium 125-129 mmol/L and asymptomatic*	*Sodium 125-129 mmol/L symptomatic;* *120-124 mmol/L regardless of symptoms* Sodium <130-120 mmol/L	Sodium <120 mmol/L; *life-threatening consequences*	Clinical signs and symptoms are not taken into consideration for grading. Worst case ("<130-120 mmol/L" for grade 3 added by Janssen) is applied across grade 2/3 ranges: 120-129 mol/L assigned to grade 3, grade 2 not used.
Renal and urinary disorders					

CTCAE Term	Grade 1	Grade 2	Grade 3	Grade 4	Janssen implementation notes
Proteinuria	1+ proteinuria; urinary protein ≥ULN - <1.0 g/24 hrs; urinary protein ≥ULN - <1000 mg/day	**Adult:** 2+ and 3+ proteinuria; urinary protein 1.0 - <3.5 g/24 hrs; urinary protein 1000 - <3500 mg/day **Pediatric:** Urine P/C (Protein/Creatinine) ratio 0.5 - 1.9; Urine P/C (Protein/Creatinine) 56.5 – 214.7 g/mol	**Adult:** 4+ proteinuria; urinary protein ≥3.5 g/24 hrs; urinary protein ≥3500 mg/day; **Pediatric:** Urine P/C (Protein/Creatinine) ratio >1.9; Urine P/C (Protein/Creatinine) >214.7 g/mol	-	In case both 24-h urine collection and dipstick are collected, then worst case is taken, as opposed to having 24-h urine collection take precedence over dipstick. Added ranges in SI unit for urinary protein (mg/day) and for urine P/C (g/mol). Pediatric grading is applied to age range 0- <18. Adult grading is applied for ages ≥18.

* Grade 0 is assigned to a lab assessment when the lab test is described in the table, but the lab value is not assigned a grade 1 or higher.

REFERENCES

1. Spiessens B (2010), Debois M. Adjusted significance levels for subgroup analyses in clinical trials. Contemp Clin Trials. 2010;31(6):647-656.
2. Mok TS (2017), Wu Y-L, Ahn M-J, et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. N Engl J Med. 2017;376(7):629-640.
3. Soria JC (2015), Wu YL, Nakagawa K, et al. Gefitinib plus chemotherapy versus placebo plus chemotherapy in EGFR-mutation-positive non-small-cell lung cancer after progression on first-line gefitinib (IMPRESS): a phase 3 randomised trial. Lancet Oncol. 2015;16:990-998.
4. Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log rank test, Biometrics, 2000; 56, 779-788.
5. Fayers, P., Aaronson, N. K., Bjordal, K., Groenvold, M., Curran, D., & Bottomley, A. (2001). EORTC QLQ-C30 Scoring Manual (3rd edition). (3rd ed.) Brussels: European Organisation for Research and Treatment of Cancer.
6. Osoba D., Rodrigues G., Myles J. Interpreting the significance of health-related quality-of-life scores. J Clin Oncol. 1998; 16: 139-144.
7. Maurer W, Bretz F. (2013) Multiple testing in group sequential trials using graphical approaches. Statistics in Biopharmaceutical Research. 2013;5(4):311-320.