STATISTICAL ANALYSIS PLAN

Study Protocol Number:	BGB-A317-312
Study Protocol Title:	A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study of Platinum Plus Etoposide With or Without Tislelizumab (BGB-A317) in Patients With Untreated Extensive-Stage Small Cell Lung Cancer
Date:	24 Apr 2022
Version:	1.0

SIGNATURE PAGE

Author:	[redacted] **Senior Statistician,** **Statistics and Data Science**	[redacted]

Approval

[redacted]	[redacted]
[redacted]	
[redacted]	

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS 5
1. ..... INTRODUCTION 7
2. ..... STUDY OVERVIEW 7
2.1. ..... Study Design 7
2.2. ..... Study Assessment 7
3. ..... STUDY OBJECTIVES 8
3.1. ..... Primary Objective 8
3.2. ..... Secondary Objective 8
3.3. ..... Exploratory Objective 9
4. ..... DEFINITION OF ESTIMANDS 9
5. ..... STUDY ENDPOINTS 10
5.1. ..... Primary Endpoint(s) 10
5.2. ..... Secondary Endpoints 10
5.3. ..... Exploratory Endpoints 10
6. ..... SAMPLE SIZE CONSIDERATIONS 11
7. ..... STATISTICAL METHODS 11
7.1. ..... Analysis Sets 11
7.2. ..... Multiplicity Adjustment 12
7.3. ..... Data Analysis General Considerations 12
7.3.1. ..... Definitions and Computations 12
7.3.2. ..... Conventions 12
7.3.3. ..... Handling of Missing Data 13
7.3.4 ..... Adjustments for Covariates/Stratification 13
7.3.5 ..... Data Integrity 13
7.4 ..... Patient Characteristics 13
7.4.1 ..... Patient Disposition 13
7.4.2 ..... Protocol Deviations 13
7.4.3 ..... Demographic and Other Baseline Characteristics 14
7.4.4 ..... Disease History 14
7.4.5 ..... Prior Anticancer Drug Therapies and Surgeries 15
7.4.6 ..... Prior and Concomitant Medications 15
7.4.7 ..... Medical History 15
7.5 ..... Efficacy Analysis 15
7.5.1 ..... Primary Efficacy Endpoint(s) 15

7.5.2 Secondary Efficacy Endpoints 17
Progression Free Survival (PFS) by investigators 17
Objective Response Rate (ORR) by investigators 17
Duration of Response (DOR) 17
Disease control rate (DCR) 18
Health-Related Quality of Life (HRQoL) 18
7.5.3 Subgroup Analyses 19
7.5.4 Exploratory Efficacy Endpoints 19
7.5.5 Post-treatment Anti-Cancer Therapy 19
7.6 Safety Analyses 20
7.6.1 Extent of Exposure 20
7.6.2 Adverse Events 21
Treatment-emergent AE 22
Immune-mediated AE 22
Infusion-related reaction 23
Death 23
7.6.3 Laboratory Values 23
7.6.4 Vital Signs 24
7.6.5 Electrocardiograms (ECG) 24
7.6.6 Eastern Cooperative Oncology Group (ECOG) Performance Status 24
7.7 Pharmacokinetic Analyses 24
7.8 Immunogenicity Analyses 24
8 INTERIM ANALYSES 25
9 CHANGES IN THE PLANNED ANALYSIS 25
10 REFERENCES 26
APPENDIX 1. IMPUTATION OF MISSING OR PARTIALLY MISSING DATES 27
APPENDIX 2 HEALTH RELATED QUALITY OF LIFE 30

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

<<Include a table of abbreviations if they are used in this document. Abbreviations should be sorted alphabetically.>>

Abbreviation	Definition
AE	adverse event
ALT	alanine aminotransferase
AST	aspartate aminotransferase
AUC	area under the plasma or serum concentration-time curve
BGB-A317	tislelizumab
BOR	best overall response
CR	complete response
C_{trough}	lowest concentration before the next dose
DCR	disease control rate
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data capture (system)
EORTC	European Organisation for the Research and Treatment of Cancer
EOT	End-of-Treatment (Visit)
ES-SCLC	extensive-stage small cell lung cancer
HR	hazard ratio
ICF	informed consent form
ICH	International Council for Harmonisation
IDMC	Independent Data Monitoring Committee
ITT	Intent-to-Treat (Analysis Set)
MedDRA	Medical Dictionary for Regulatory Activities
NCI-CTCAE	National Cancer Institute-Common Terminology Criteria for Adverse Events
ORR	overall response rate
OS	overall survival
PD-1	programmed cell death protein-1
PD-L1	programmed cell death protein ligand-1
PFS	progression-free survival
PFS2	PFS after next line of treatment
PK	pharmacokinetic(s)
PR	partial response

QLQ-C30	Quality of Life Questionnaire – Core 30
QLQ-LC13	Quality of Life Questionnaire supplement lung cancer module
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	Statistical Analysis Plan
SCLC	small cell lung cancer
TEAE	treatment-emergent adverse event

1. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the procedures and the statistical methods that will be used to analyze and report results for BGB-A317-312*:* A Phase 3, Randomized, Double-Blind, Placebo-Controlled Study of Platinum Plus Etoposide With or Without Tislelizumab (BGB-A317) in Patients With Untreated Extensive-Stage Small Cell Lung Cancer (ES-SCLC). This SAP is based on BGB-A317-312 Protocol Amendment 1.0 dated Sep 02, 2020. The focus of this SAP is for the planned interim analysis and the final analysis specified in the study protocol. The analysis details for Pharmacogenomics and Biomarker analyses are not described in this SAP.

2. STUDY OVERVIEW

2.1. Study Design

This is a randomized, double-blind, placebo-controlled, multicenter, Phase 3 study to compare the efficacy of tislelizumab + cisplatin or carboplatin + etoposide (Arm A) and placebo + cisplatin or carboplatin + etoposide (Arm B) as first-line treatment in patients who have previously untreated ES-SCLC. The choice of platinum (cisplatin or carboplatin) will be at the investigator's discretion.

The study consists of a Screening phase (up to 28 days before randomization), a Treatment phase (until disease progression as assessed by the investigator per RECIST v1.1, loss of clinical benefit, unacceptable toxicity, or withdrawal of informed consent, whichever occurs first) that includes an Induction treatment period and a Maintenance treatment period, a Safety Follow-up phase (30 days (± 7 days) after the last dose of study drug (including chemotherapy-only), or before the initiation of a new anticancer treatment, whichever occurs first), and a Survival Follow-up phase (every 3 month (± 14 days) after the Safety Follow-up Visit or as directed by the sponsor until death, loss to follow-up, withdrawal of consent, or study completion by the sponsor).

At randomization, eligible patients will be stratified by the following 3 factors:

- ECOG Performance Status (0 versus 1)
- Investigator-chosen chemotherapy (carboplatin versus cisplatin)
- Brain metastasis (yes versus no)

Randomized patients will be in a 1:1 ratio to receive 1 of the following treatment regimens:

- Induction period (administered on a 3-week cycle for 4 cycles):
  - Arm A: tislelizumab + cisplatin or carboplatin + etoposide
  - Arm B: placebo + cisplatin or carboplatin + etoposide
- Maintenance period (administered once every 3 weeks):
  - Arm A: tislelizumab
  - Arm B: placebo

Crossover between treatment arms is not allowed in this study.

2.2. Study Assessment

Tumor imaging will be performed within 28 days prior to randomization. During the study, tumor imaging will be performed approximately every 6 weeks (± 7 days) for the first 48 weeks following Cycle 1 Day 1, and every 9 weeks (± 7 days) thereafter, regardless of treatment dose delay. The investigator

may perform additional scans or more frequent assessments if clinically indicated. Tumor assessment will continue until radiographic disease progression per RECIST v1.1.

Patients who discontinue study treatment early for reasons other than disease progression (eg, toxicity) will continue to undergo tumor assessments following the original plan until the patient begins a subsequent anticancer treatment, experiences disease progression, withdraws consent, is lost to follow-up, death, or until the study terminates, whichever occurs first.

After informed consent has been signed but prior to the administration of the study drug, only SAEs should be reported.

After initiation of study drug, all AEs and SAEs, regardless of relationship to study drug, will be reported until either 30 days after last dose of study drugs including chemotherapy drugs or initiation of new anticancer therapy, whichever occurs first. ImAEs (serious or nonserious) should be reported until 90 days after the last dose of tislelizumab or placebo, regardless of whether or not the patient starts a new anticancer therapy. All SAEs considered related to the study drug(s) that are brought to the attention of the investigator should be reported regardless of time since the last dose of treatment.

In this double-blind, placebo-controlled study, all patients and personnel involved in the conduct and interpretation of the study, including the investigators, BeiGene study team, and site personnel, will be blinded to the treatment assignment. Randomization data will be kept strictly confidential; filed securely by the appropriate groups for BeiGene, the Interactive Response Technology (IRT) and the Independent Data Monitoring Committee (IDMC); and will be accessible only to authorized persons per SOPs until the time of unblinding.

Unblind safety and efficacy monitoring will be performed by an IDMC. The IDMC may recommend modifications to the study, including termination due to safety and/or efficacy concerns. The functions and membership of the IDMC will be described in the IDMC Charter.

3. STUDY OBJECTIVES

3.1. Primary Objective

- To evaluate the efficacy of tislelizumab + cisplatin or carboplatin + etoposide compared with placebo + cisplatin or carboplatin + etoposide in the Intent-to-Treat (ITT) Analysis Set as measured by overall survival (OS)

3.2. Secondary Objective

- To evaluate the efficacy of tislelizumab + cisplatin or carboplatin + etoposide compared with placebo + cisplatin or carboplatin + etoposide in the ITT Analysis Set as measured by investigator-assessed progression-free survival (PFS) according to RECIST v1.1
- To evaluate the efficacy of tislelizumab + cisplatin or carboplatin + etoposide compared with placebo + cisplatin or carboplatin + etoposide in the ITT Analysis Set as measured by investigator assessed overall response rate (ORR), duration of response (DOR), and disease control rate (DCR) according to RECIST v1.1
- To evaluate the effect of tislelizumab + cisplatin or carboplatin + etoposide compared with placebo + cisplatin or carboplatin + etoposide on patients' healthrelated quality of life (HRQoL) according to the EORTC Quality of Life Questionnaire – Core 30 (QLQC30) and the supplemental lung cancer module (QLQLC13)

- To evaluate the safety and tolerability of tislelizumab in combination with cisplatin or carboplatin and etoposide compared with cisplatin or carboplatin and etoposide

3.3. Exploratory Objective

- To assess PFS after next line of treatment (PFS2)
- To explore potential predictive biomarkers in archival and/or fresh tumor tissue and/or blood (or blood derivatives), including but not limited to PDL1 expression by immunohistochemistry (IHC), multiplex IHC (mIHC), gene expression profiling (GEP), tumor mutation burden (TMB), microsatellite instability (MSI), and blood tumor mutation burden (bTMB), to evaluate the association between these biomarkers and response to study treatment or mechanism of resistance
- To characterize PK of tislelizumab in patients with small cell lung cancer (SCLC)
- To evaluate host immunogenicity to tislelizumab by assessing antidrug antibodies (ADAs) against tislelizumab in patients with SCLC
- To examine the patients' quality of life (QoL) as measured by EQ5D-5L

4. DEFINITION OF ESTIMANDS

Primary scientific question of interest is: will the addition of tislelizumab to chemotherapy doublet (platinum and etoposide) prolong survival in first-line patients who have untreated ES-SCLC, regardless of any anticancer therapy received subsequently.

The primary estimand is described by the following attributes:

1. Treatment of interest:
   Experimental treatment constitutes tislelizumab plus chemotherapy (either cisplatin or carboplatin + etoposide). Control treatment constitutes placebo plus chemotherapy (either cisplatin or carboplatin + etoposide).
2. Population:
   Patients with ES-SCLC who had not received first-line treatment.
3. Primary variable:
   Overall survival (see Section 5.1)
4. Handling of intercurrent events:
   - Anticancer therapy subsequent to assigned therapy that is started prior to death: Any incidence will be ignored, i.e., any death or patients' data collected after the new anticancer therapy will be considered for analysis. (treatment policy strategy)
   - Insufficient study treatment or discontinuation of any study drug: Early discontinuation or interruption or any study drug due to any cause (including COVID-19) will be ignored, i.e., any death or patients' data collected after the interruption or discontinuation of treatment will be considered for analysis. (treatment policy strategy)
   - Any other unforeseen intercurrent events: OS analysis will take into account all deaths regardless of any unforeseen intercurrent events
5. Population-level summary:

   Hazard ratio (HR) of OS comparing tislelizumab plus platinum and etoposide versus placebo plus platinum and etoposide, estimated using Cox proportional hazard model stratified by prespecified stratification factors (see Section 7.3.4)

5. STUDY ENDPOINTS

5.1. Primary Endpoint(s)

- OS, defined as the time from randomization to death from any cause

5.2. Secondary Endpoints

- PFS, defined as the time from randomization to the first occurrence of disease progression as determined by investigator using RECIST v1.1 or death from any cause, whichever occurs first
- ORR, defined as the proportion of patients with PR or CR as determined by the investigator using RECIST v1.1
- DOR, defined as the time from the first occurrence of a documented objective response to the time of relapse, as determined by the investigator per RECIST v1.1, or death from any cause, whichever comes first
- DCR, defined as the proportion of patients whose best overall response (BOR) is CR, PR or stable disease per RECIST v1.1
- CBR, defined as the proportion of patients whose best overall response (BOR) is CR, PR or durable stable disease per RECIST v1.1
- Incidence and severity of treatment-emergent adverse events (TEAEs) graded according to National Cancer Institute-Common Terminology Criteria for Adverse Events (NCI-CTCAE) v5.0
- Percentage of patients with clinically meaningful changes post baseline, defined as a 10-point improvement or worsening (detailed in Section 7.5.2) in:
  - Global health status (GHS) and physical function (PF) per the QLQ-C30
  - Dyspnoea, coughing, haemoptysis, dysphagia, chest pain, pain in arms and shoulders, and peripheral neuropathy symptoms of QLQ-LC13
- Time to deterioration, defined as the time from randomization to the first occurrence of worsening scores (10-point change; detailed in Section 7.5.2) confirmed at the following visit or death from any cause, in:
  - GHS and PF per the QLQ-C30
  - Dyspnoea, coughing, haemoptysis, dysphagia, chest pain, pain in arms and shoulders, and peripheral neuropathy symptoms of QLQ-LC13

5.3. Exploratory Endpoints

- PFS2, defined as the time from randomization to the objective disease progression after next line of treatment or death from any cause, whichever occurs first
- Status of predictive and prognostic biomarkers including but not limited to PD-L1 expression, multiplex immunohistochemistry (mIHC), GEP, TMB, and MSI in archival and/or fresh tumor tissue and bTMB in blood before study treatment and/or at disease progression, and the association between the biomarkers and disease status or response to study treatment
- Summary of serum concentrations of tislelizumab
- Assessments of immunogenicity of tislelizumab by determining the incidence of ADAs

- QoL is defined as changes in patients' general well-being measured by the scores of the EQ5D-5L descriptive 5-dimension scores and the visual analog scale (VAS).

6. SAMPLE SIZE CONSIDERATIONS

The sample size calculation is based on the number of events required to demonstrate the OS superiority of Arm A to Arm B in the ITT Analysis Set. Driven by external data from studies described in Protocol Section **Error! Reference source not found.** , statistical assumptions were updated in protocol amendment version 1.0 and the number of events required to demonstrate efficacy with regard to OS is estimated based on the following assumptions:

- Median OS of 12 months in Arm B.
- A one-sided α of 0.025 and 80% power for the target OS HR of 0.74 in the ITT Analysis Set.
- One OS interim analysis planned in the ITT Analysis Set when approximately 80% of total required OS events occurred, with Lan-DeMets approximation to O'Brien-Fleming boundary.
- Exponential survival distribution is assumed for both Arm A and Arm B.

Assuming approximate 455 patients are to be enrolled and randomized at 1:1 ratio over a 22-month period at a steady-state enrollment rate of 26 patients per month with enrollment ramp-up duration of 9 months, and assuming a dropout rate of 5% per year for OS endpoints, a total of 353 OS events are required for the ITT Analysis Set for the final OS analysis.

7. STATISTICAL METHODS

7.1. Analysis Sets

The Intent-to-Treat (ITT) analysis set consists of all randomized patients. The ITT analysis set is the primary analysis set and will be used for demography and efficacy analyses.

The safety analysis set includes all patients randomized and received any dose of any study drug, and patients will be grouped by actual treatment received. Patients randomized to arm A but never took any dose of Tislelizumab will be included in arm B. Patients randomized to arm B but took any dose of Tislelizumab will be included in arm A. The safety analysis set is used for all safety analyses.

The per-protocol analysis set consists of all the patients in the ITT analysis set, excluding those with critical protocol deviations. Critical protocol deviations are a subset of important protocol deviations. Criteria for exclusion from the PP will be determined and documented before the database lock for the primary analysis. This will be the secondary analysis set for efficacy analysis when there are over 10% ITT patients with critical protocol deviations.

The HRQoL analysis set includes all randomized patients who received any dose of study drug and completed at least one HRQoL assessment. This will be the analysis set for HRQoL analysis.

PK Analysis Set includes all patients who received ≥ 1 dose of tislelizumab per the protocol for whom any post dose PK data were available.

Antidrug antibody (ADA) Analysis Set includes all patients who received ≥ 1 dose of tislelizumab for whom both baseline ADA and ≥ 1 postbaseline ADA results were available.

7.2. Multiplicity Adjustment

Hypothesis testing of the secondary endpoint PFS in ITT analysis set will be performed at the same time as the OS analysis, either at interim or final analysis. Only when the superiority of OS is demonstrated, full alpha of 1-sided 0.025 will be passed to hypothesis testing of PFS.

Multiplicity adjustment of interim analysis is described in Section 8.

7.3. Data Analysis General Considerations

7.3.1. Definitions and Computations

Study drugs include tislelizumab, cisplatin, carboplatin and etoposide.

Study day will be calculated in reference to the date of the first dose of study drug for safety analyses. For derivation of the duration of any efficacy endpoint, the reference date will be date of randomization. For assessments conducted on or after the date of first dose date, study day will be calculated as (assessment date –first dose date + 1). For assessments conducted before first dose date, study day is calculated as (assessment date – first dose date). If no dose is given, then the date of first randomization will be used. There is no study day 0.

Unless otherwise specified, a baseline value is defined as the last non-missing value collected before or on randomization for ITT analysis set. If not available, last non-missing value collected before or on first dose date will be used. For safety analysis set, a baseline value is defined as the last non-missing value collected before or on first dose date.

Unscheduled measurements will not be included in by-visit table summaries and graphs but will contribute to best/worst case value where required (e.g. shift table). Listings will include scheduled, unscheduled and retest data.

Study follow-up duration is defined as the duration from the randomization date to the study discontinuation date (e.g. death, consent withdrawal, lost to follow-up) or to cutoff date if a patient is still ongoing.

All calculations and analyses will be conducted using SAS® Version 9.4 or higher.

7.3.2. Conventions

Unless otherwise specified, the following conventions will be applied to all analyses:

- 1 year = 365.25 days. Number of years is calculated as (days/365.25) rounded up to 1 significant digit.
- 1 month = 30.4375 days. Number of months is calculated as (days/30.4375) rounded up to 1 significant digit.
- Age will be calculated as the integer part of (date of informed consent – date of birth + 1)/365.25.
- P-values will be rounded to 4 decimal places. P-values that is less than 0.0001 will be presented as '< 0.0001' and p-values that is larger than 0.9999 will be presented as '> 0.9999'.
- Time-to-event or duration of image-based event endpoints will be based on the actual date the event occurs (i.e., radiograph was obtained) rather than the associated visit date.
- For laboratory results collected in numerical range with the symbol ‘<’ or ‘>=’, use the rule: if results >= x then set as x; if <x then set as x/2.

- For by-visit observed data analyses, percentages will be calculated based on the number of patients with non-missing data as the denominator, unless otherwise specified.
- For continuous endpoints, summary statistics will include n, mean, standard deviation, median, Q1, Q3 and range (minimum and maximum).
- For discrete endpoints, summary statistics will include frequencies and percentages.

7.3.3. Handling of Missing Data

Handling of missing data related to primary estimand is elaborated in Section 7.5.1. Missing dates or partially missing dates will be imputed conservatively for adverse events and prior/concomitant medications/procedures. Specific rules for the handling of missing or partially missing dates for adverse events and prior/concomitant medications/procedures are provided in 0. Other missing data will not be imputed unless otherwise specified elsewhere in this SAP.

By-visit endpoints will be analyzed using observed data unless otherwise specified. For observed data analyses, missing data will not be imputed, and only the observed records will be included.

7.3.4 Adjustments for Covariates/Stratification

The value of the stratification factors used at randomization (recorded in electronic data capture (EDC) system), including ECOG performance status (0 vs 1), and investigator-chosen platinum (cisplatin vs carboplatin) will be used in stratified log-rank test for primary endpoint OS, and stratified Cox proportional hazard model for OS and secondary endpoint PFS. In order to avoid over-stratification, brain metastasis (yes vs no) will not be included as number of patients enrolled with brain metastasis is small.

The actual value of the stratification factors (collected in IRT) and other baseline covariates may be used in statistical models as covariates as sensitivity or supplementary analyses for endpoints.

7.3.5 Data Integrity

The data set for analysis should be an accurate and complete representation of the patients' relevant outcomes from the clinical database. All data should be complete and reviewed up to a pre-specified cutoff date as specified in the Data Extract and Snapshot Plan. Consistency checks and appropriate source data verification should be completed as specified in the Site Monitoring Plan.

7.4 Patient Characteristics

7.4.1 Patient Disposition

The number (percentage) of patients randomized, treated, discontinued from the study, reasons for discontinued from the study, and the duration of study follow-up will be summarized in the ITT analysis set. Study follow-up summary will be descriptively. The patients who discontinued treatment and the primary reason for the end of treatment will be summarized among patients who were treated. The reasons for treatment/study discontinuation will be summarized according to the categories in CRF, and those related to COVID-19 will also be summarized, if applicable.

7.4.2 Protocol Deviations

Protocol deviation criteria will be established together with its category/term of important and not important. Patients with important protocol deviations will be identified and documented before the database lock. Important protocol deviations will be summarized for all patients in the ITT analysis set.

Critical protocol deviation that significantly impacts efficacy or safety evaluation will be reviewed prior to data base lock.

Protocol deviations that are related to COVID-19 will be summarized and all patients affected by COVID-19-related protocol deviations will be listed along with site information and specific deviation description.

7.4.3 Demographic and Other Baseline Characteristics

Demographics and other baseline characteristics will be summarized using descriptive statistics in the ITT analysis set, including the following variables:

- Age (continuously and by categories [<65 or >=65 years])
- Sex
- Ethnicity
- Race
- Region
- Baseline weight (kg)
- Height (m)
- BMI
- Smoking status
- baseline LDH (categorized as <=ULN or >ULN)

Meanwhile, stratification factors per IRT and per EDC will be summarized in ITT analysis set:

- ECOG Performance (0 versus 1)
- Investigator-chosen chemotherapy (carboplatin versus cisplatin)
- Brain metastasis (yes versus no)

7.4.4 Disease History

The number (percentage) of patients reporting a history of disease and characteristics, as recorded on the eCRF, will be summarized in the ITT analysis set.

Disease characteristics include:

- Time from initial diagnosis to study entry
- Time from extensive-stage diagnosis to study entry
- Stage at initial diagnosis
- AJCC staging at study entry
- Metastatic site(s) at study entry
- Number of metastatic sites
- Target lesion sum of diameter

7.4.5 Prior Anticancer Drug Therapies and Surgeries

Prior anti-cancer drug therapies, prior anti-cancer radiotherapy, and prior anti-cancer surgeries including number (%) of patients with any use and treatment intent will be summarized in the ITT analysis set. In addition, number of prior lines and best overall response to last drug therapy will be included in the prior anti-cancer drug therapies summary.

The therapies and surgeries with the same sequence/regimen number are counted as one prior therapy/surgery.

7.4.6 Prior and Concomitant Medications

Prior medications are defined as medications that stopped before the first dose of study drug. Concomitant medications will be defined as medications that (1) started before the first dose of study drug and were continuing at the time of the first dose of study drug, or (2) started on or after the date of the first dose of study drug up to 30 days after the patient's last dose or the initiation of a new anti-cancer therapy.

Prior and concomitant medications will be coded using the version of World Health Organization Drug Dictionary (WHO DD) drug codes effective for the study at the time of database lock. They will be further classified to the appropriate Anatomical Therapeutic Chemical (ATC) code.

The number (percentage) of patients reporting prior and concomitant medications will be summarized by ATC medication class and WHO DD preferred name in the safety analysis set. A listing of prior and concomitant medications will be provided.

7.4.7 Medical History

Medical History will be coded using Medical Dictionary for Regulatory Activities (MedDRA) effective for the study at the time of database lock. The number (percentage) of patients reporting a history of any medical condition, as recorded on the CRF, will be summarized by system organ class and preferred term in the ITT analysis set. A listing of medical history will be provided.

7.5 Efficacy Analysis

7.5.1 Primary Efficacy Endpoint(s)

Primary estimand is defined in Section 4. Details of the statistical methods used in OS derivation and analysis are provided in this section including pre-defined sensitive analyses of the primary estimand and supplementary analyses of OS.

Overall survival is defined as time from randomization date to the documented death date for patients who died prior to or on the clinical cutoff date. For patients who are alive by the clinical cutoff date, OS will be censored at the last known alive date (LKADT). The last known alive date will be defined as either the clinical data cutoff date for patients who are still on treatment, or last available date showing patients alive or cut-off date whichever comes first.

Every effort should be made to ensure complete death dates. In the rare case, if day of death date is missing, death date is imputed as the max (last available date showing patients alive + 1, first day of year/month of death date). The patient with imputed death date will be considered as an event for OS analysis. Death with missing month and/or year will not be imputed for OS analysis and will not be considered as OS event and will be censored at LKADT.

The null and alternative hypotheses to be evaluated for OS are:

$$H_0 : OS_A \leq OS_B,$$

$$H_a : OS_A > OS_B,$$

where OS_A and OS_B represent the OS of arm A (the experimental treatment) and arm B (the control group), respectively. The primary efficacy analyses of OS will be tested using stratified log-rank test with ECOG performance status and investigator-chosen platinum as recorded in EDC as strata. In order to avoid over-stratification, brain metastasis (yes vs no) will not be included as number of patients enrolled with brain metastasis is small.

The survival distribution of OS will be estimated using Kaplan-Meier method. The results will be plotted graphically (Kaplan-Meier curves) by treatment. The plots will display the number of patients at risk at equidistant time point. The median OS for each treatment group will be provided along with the approximate 95% confidence intervals using the method of Brookmeyer and Crowley (Brookmeyer & Crowley, 1982). Additionally, 25% and 75% percentiles will also be provided. The cumulative probability of OS at every 3 months if estimable, will be calculated for each treatment arm and presented with 2-sided 95% CIs. The CIs will be estimated using Greenwood's formula (Greenwood, 1926).

The HR and its 2-sided 95% CI will be estimated from a stratified Cox regression model with Efron's method of tie handling, and with the same stratification factors above. Unstratified analysis will also be presented.

Handling of intercurrent events is described in Section 5.1.

Sensitivity analyses

Sensitivity analysis 1: OS will be analyzed using an unstratified Cox model, and the treatment effect will be summarized by the hazard ratio with its 95% confidence interval.

Sensitivity analysis 2: OS will be analyzed with stratification factors as recorded from IRT

Supplementary analyses

Supplementary analysis 1 "OS analysis based on Restricted Mean Survival Time method": This analysis targets an estimand which has the same attributes as the primary estimand except the population level summary will be different in restricted mean survival time (RMST) between two treatment groups. In order to account for the possible non-proportional hazard effect, the restricted mean survival time (RMST) (RMST, Uno H, Claggett B, Tian L, Inoue E, et al. 2014) will be computed for OS separately using the area under the curve from baseline to the minimum of the largest observed time on each of the two treatment groups. RMST will be computed for each treatment arm and the difference with its 95% CI will be displayed.

Supplementary analysis 2 "OS analysis adjusted for baseline covariates": OS will be also analyzed by adjusting multivariate covariates at baseline. The analysis addresses a different scientific question. i.e., will the addition of tislelizumab to chemotherapy doublet prolong survival, adjusting for covariates liver metastasis, LDH, and smoking status. Covariate adjusted multivariate Cox regression provides a conditional treatment effect and a stratified Cox regression model will be performed adjusted for the above mentioned key baseline prognostic factors.

Supplementary analyses exploring the impact on OS from other confounding factors, including but not limited to between-arms imbalance of clinically meaningful baseline characteristics, post-treatment anti-cancer therapy, or COVID-19, might be performed as appropriate.

7.5.2 Secondary Efficacy Endpoints

Progression Free Survival (PFS) by investigators

The methods used to analyze OS will be applied to the analysis for PFS. Table 1 shows the derivation rule for PFS.

Table 1: Censoring Rules for Progression-free Survival Per RECIST Version 1.1

	Derivation rules	**Outcome**
No progression at the time of data cut-off or withdrawal from study or lost to follow up	Date of last adequate radiologic assessment prior to or on date of data cut-off or withdrawal from study or lost to follow up	Censored
New anticancer therapy started prior to disease progression or death	Last adequate radiological assessment before the new anticancer therapy (hypothetical strategy)	Censored
No baseline or post-baseline tumor assessments without death within 13 weeks after randomization	Date of randomization	Censored
No baseline or post-baseline tumor assessments with death within 13 weeks after randomization	Date of death	Event
Death or progression after more than one consecutive missed visit	Date of last adequate radiologic assessment before missed tumor assessments	Censored

PFS will be analyzed in similar methodology as OS.

Additional analyses exploring the impact on PFS from other confounding factors, including but not limited to COVID-19, might be performed per needs.

Objective Response Rate (ORR) by investigators

Best overall response (BOR), defined as the best response recorded from randomization until data cut or the start of new anticancer treatment. Patients with no post-baseline response assessment (due to any reason) will be considered non-responders for BOR. ORR is defined as the number of patients whose BOR is CR or PR divided by the number of randomized patients in each arm. Patients with no baseline response assessment (for any reason) will be considered as non-responders. Confirmation on response is required as primary analysis, and analysis based on unconfirmed response might be presented as well. The 2-sided 95% CIs for the odds ratio in ORR will be calculated, as well as Clopper-Pearson 95% CIs for ORR and its corresponding Clopper-Pearson 95% CI for each of the response categories (CR, PR, SD, and PD) will be presented by treatment arm.

Duration of Response (DOR)

Duration of Response (DOR) is defined as progression/death event free time counted from the first objective response date to the first documented radiological PD date/or death date, whichever occurs first. Confirmation on response is required as primary analysis, and analysis based on unconfirmed response

might be presented as well. All the censoring rules for PFS should be applied to DOR. DOR will be analyzed in the responders only and in similar methodology as PFS except that no hypothesis testing will be performed and no HR will be reported.

Disease control rate (DCR)

Disease control rate (DCR) defined as the proportion of patients whose best overall response (BOR) is CR, PR, or SD including non-CR/non-PD ≥ 6 weeks. Confirmation on response is required as primary analysis, and analysis based on unconfirmed response might be presented as well. DCR by investigators will be analyzed similarly to ORR.

Clinical benefit rate (CBR)

Clinical benefit rate (CBR) defined as the proportion of patients who have CR, PR, or SD including non-CR/non-PD of ≥ 24 weeks in duration. Confirmation on response is required as primary analysis, and analysis based on unconfirmed response might be presented as well. CBR assessed by investigators will be analyzed similarly to ORR.

Health-Related Quality of Life (HRQoL)

All HRQoL analyses will be based on HRQoL analysis set, unless otherwise specified.

Compliance

Compliance for EORTC QLQ-C30, QLQ-LC13 modules, defined as the proportion of questionnaires actual received out of the expected number, in the HRQoL analysis set will be summarized for each scheduled assessment and treatment arm.

Change from baseline by visit

Descriptive analysis will be performed for all the scales and single items of the PRO questionnaires in all cycles. Details of HRQoL scoring are specified in Appendix 2 according to the algorithm described in the EORTC QLQ-C30 and EORTC QLQ-LC13 scoring manual (Fayers 2001).

A mixed effect model analysis for measuring clinically meaningful changes will be performed on selected PRO endpoints (GHS and PF in QLQ-C30, coughing, dysphagia, and chest pain in QLQ-LC13). Differences in change from baseline between arms at the key clinical cycles of 4 and 6, if data allows, will be assessed in the mixed models which include baseline score, stratification factors, treatment arm, visit, and treatment arm by visit interaction as fixed effects and visit as a repeated measure with an unstructured covariance structure.

Time to deterioration

Time to deterioration (TTD) is defined as the time from randomization to first occurrence of a worsening score confirmed at the subsequent visit or death from any cause. The minimum important clinically meaningful difference change (e.g. worsening) in symptoms of QLQ-C30 and QLQ-LC13 is defined as ≥10 points increase from baseline (Osoba et al 1998; King, 1996; Maringwa et al 2011). The clinically meaningful deterioration in function and global health status/quality of life is defined as ≥10 points decrease from baseline. The median TTD of selected PRO endpoints (PF in QLQ-C30, coughing and chest pain in QLQ-LC13) will be calculated using Kaplan-Meier estimates, and presented with 2-sided 95% CIs.

7.5.3 Subgroup Analyses

To determine if the treatment effect is consistent across various subgroups, the median OS in each subgroup along with unstratified OS hazard ratio and its 95% CI will be estimated and plotted within each category of the following variables:

- age (≤ 65 vs > 65 years)
- gender (Female vs Male)
- ECOG PS (0 vs 1)
- brain metastasis (yes vs no)
- choice of platinum (cisplatin vs carboplatin)
- liver metastasis (yes vs no)
- baseline LDH (<=ULN vs >ULN)
- Smoking status (never vs smoker) where smoker includes both current and former smoker
- AJCC staging at study entry (III vs IV)

For patients with evaluable PD-L1 expression, subgroups with PD-L1 score <1% vs >=1% will be analyzed in a similar way but in a separate output and not included in forest plot, and other cutoffs and algorithms might be explored as well. PD-L1 expression is determined by PD-L1 score assessed by tumor area positive score (TAP), which is defined as the total percentage of the tumor area covered by tumor cells with any membrane staining above background and tumor-associated immune cells with any staining above background using Ventana PD-L1 (SP263) assay.

If a subgroup includes only few patients, then the subgroup variable might be ignored.

7.5.4 Exploratory Efficacy Endpoints

Progression-free survival after next-line of treatment (PFS2) is defined as time from the randomization date to the first documented disease progression on next-line therapy or death from any cause, whichever occurs first. Next-line therapy is defined as the first subsequent treatment initiated after the discontinuation of the study treatment . The first documented progression on next-line treatment will be recorded by investigator (i.e. captured on the post treatment discontinuation anti-cancer systemic therapy CRF page). Patients alive and for whom a disease progression on next-line therapy has not been observed will be censored at the last known alive date. If a second new anti-cancer therapy is introduced, then PFS2 will be censored at the end date of prior line of therapy. Any death prior to initiation of next-line therapy will be considered as an event for PFS2.

Kaplan-Meier (KM) method as described in the PFS and OS analyses will be used in the analysis of PFS2. The median PFS2 and the cumulative probability of PFS2 estimated at every 3 months will be calculated using Kaplan-Meier estimates for each treatment arm and presented with 2-sided 95% CIs computed by Brookmeyer and Crowley method. The HR of PFS2 from stratified Cox model will be estimated and presented with a 2-sided 95% CI.

7.5.5 Post-treatment Anti-Cancer Therapy

Post treatment anti-cancer therapy is defined as the anti-cancer therapy started after the last dose of study drug(s). A summary of number and percentage of patients who received subsequent systematic anticancer therapy, chemotherapy, I/O, target therapy(refer to coding), others will be provided by arm in ITT analysis set.

Separate flags of start date of new anti-cancer therapy for efficacy and safety analyses are derived individually.

- As for efficacy analysis, start date of new anti-cancer therapy could be the earliest of date of prohibited anti-cancer therapy taken during treatment, date of the post-treatment systemic anti-cancer therapy or include other anti-cancer therapy such as post-treatment surgery and radiotherapy as deem appropriate.

- The start date of new anti-cancer therapy in defining TEAE for safety is always the first date of new systemic anti-cancer therapy taken after the last study treatment.

Since the Tumor response per RECIST or event driven endpoints have not been commonly used for the efficacy evaluation of Tradition Chinese Medicine, The ORR, PFS or OS benefit of Chinese herbal medicines and Chinese patent medicines has not yet been established. Therefore, they will not be considered as new anti-cancer therapy in the efficacy and safety analyses.

Patient data listings of post-treatment anti-cancer therapy, radiotherapy, and procedure or surgery will be provided.

7.6 Safety Analyses

All safety analyses will be performed by treatment arms based on the safety analysis set. Safety and tolerability will be assessed, where applicable, by incidence, severity, and change from baseline values for all relevant parameters including AEs, laboratory values, vital signs and ECG findings. Descriptive statistics (e.g., n, mean, standard deviation, median, Q1, Q3, minimum, maximum for continuous variables; n [%] for categorical variables) will be used to analyze all safety data.

7.6.1 Extent of Exposure

The following measures of the extent of exposure will be summarized with descriptive statistics for each study drug separately. One cycle is defined as 21 days of treatment.

- Duration of exposure: defined as the duration from the first dose date of study drug to the last dose date of the study drug, and calculated as (last date of exposure – date of first dose +1)
  - If patients discontinued treatment (with non-missing EOT date), using min (cutoff date, death date, last dose date + 20) as the “last date of exposure” for tislelizumab, cisplatin and carboplatin; min (cutoff date, death date, last dose date + 18) as the “last date of exposure” for etoposide
  - otherwise, if patient has treatment ongoing, using cutoff date as the “last date of exposure”
- Number of treatment cycles received: defined as the total number of treatment cycles in which at least one dose of the study drug is administered.
- Total dose received per patient: defined as the cumulative dose received of the study drug during the treatment period of the study, and calculated as sum (all actual dose per administration at all visits prior to or on the cutoff date)
- Actual dose intensity (ADI): defined as the total dose received by a patient divided by the duration of exposure, and calculated as
  - For tislelizumab, ADI (mg/cycle) = 21*total cumulative dose (mg) / (last dose date prior to cut off date+ 21 - first dose date)
  - For chemotherapy, derivations are shown in Table 2.

- Relative dose intensity (RDI): defined as the ratio of the actual dose intensity and the planned dose intensity. Planned dose intensity is defined as the planned dose on study day 1 by a patient divided by the duration of exposure. RDI is calculated as
  - For tislelizumab, RDI (%) = ADI/200 *100%; Planned dose intensity (mg/cycle) = 200mg/cycle
  - For chemotherapy, derivations are shown in Table 2.

Table 2. ADI, Planned dose and RDI for Chemotherapy

	ADI(mg/m²/cycle for Cisplatin and Etposide; AUC/cycle for Carboplatin)	Planned dose per cycle	RDI
Cisplatin	$\frac{\sum_1^{\#of\ cycles} \frac{actual\ dose}{BSA} \times 21}{date\ of\ last\ dose\ up\ to\ cutoff + 21 - first\ dose\ date}$	75 mg/m²	$\frac{ADI}{75}$
Carboplatin	$\frac{\sum_1^{\#of\ cycles} \frac{actual\ dose}{(GFR + 25)} \times 21}{date\ of\ last\ dose\ up\ to\ cutoff + 21 - first\ dose\ date}$	AUC 5	$\frac{ADI}{5}$
Etoposide	$\frac{\sum_1^{\#of\ cycles} \frac{actual\ dose}{BSA}}{\max(\frac{date\ of\ last\ dose\ up\ to\ cutoff + 19 - first\ dose\ date}{21}, number\ of\ cycles\ in\ last\ dosing\ CRF\ page)}$	100 *3 mg/m²	$\frac{ADI}{300}$

Patient data listings will be provided for all dosing records.

7.6.2 Adverse Events

AEs will be graded by the investigators using CTCAE version 5.0. The AE verbatim descriptions (investigator reported terms from the eCRF) will be classified into standardized medical terminology using MedDRA. Adverse events will be coded to the MedDRA (version for the study at the time of database lock) lowest level term closest to the verbatim term, along with the linked MedDRA preferred term (PT) and primary system organ class (SOC).

A treatment-emergent adverse event (TEAE) is defined as an AE that had onset or increase in severity level date on or after the date of the first dose of study drug through 30 days following study drug discontinuation or initiation of new anticancer therapy, whichever is earlier. Summary tables will only include TEAEs. All AEs, treatment-emergent or otherwise, will be presented in patient data listings. COVID-19 related adverse events will be summarized separately.

Treatment-related AEs include those events considered by the investigator to be "related" to study drug or with a missing assessment of the causal relationship.

For AE grading, if the grade is missing for one of the treatment-emergent occurrences of an adverse event, the maximal grade on the remaining occurrences with the same preferred term of the same patient will be used. If the patient has no other TEAE with the same preferred term, then impute as the maximum grade on all TEAE with the same preferred term; If the grade is missing for all the occurrences, do not impute, a “missing” category will be added in the summary table.

For incidences, a patient will be counted only once by the highest severity grade within an SOC and PT, even if the patient experienced more than 1 TEAE within a specific SOC and PT.

Incidences of TEAEs by induction and maintenance phase will also be provided, and definitions of the phases are as below:

- Induction phase: defined as time from first dose date to 1 day before the first dose in maintenance phase
- Maintenance phase: defined as time from the first monotherapy dose date (all chemotherapies are completed) to end of TEAE reporting period

Treatment-emergent AE

An AE overview table, including the number and percentage of patients with

- TEAEs
- TEAEs with Grade 3 or above
- treatment-emergent serious adverse events (SAEs)
- treatment-related TEAEs
- TEAEs leading to treatment discontinuation
- TEAEs leading to dose modification
- infusion-related reactions
- TEAEs leading to death

The incidence of following TEAEs will be reported as the number (percentage) of patients by SOC and PTs:

- treatment-related TEAEs
- TEAEs with grade 3 or above
- serious TEAEs
- TEAEs leading to treatment discontinuation
- TEAEs leading to treatment modification
- TEAEs leading to death

Immune-mediated AE

Immune-mediated AEs will be identified from all AEs that had an onset date or a worsening in severity from baseline (pretreatment) on or after the first dose of study drug (tislelizumab/placebo) and up to 90 days from the last dose of study drug (tislelizumab/placebo), regardless of whether the patient starts a new anticancer therapy. The identification of immune-mediated adverse events is described in immune-mediated adverse event charter. Immune-mediated adverse events are of special interest and summarized by category.

Summaries of the following incidence of immune-mediated adverse events will be provided:

- Immune-mediated adverse events by category and maximum severity
- Immune-mediated adverse events with NCI-CTCAE grade ≥3 by category
- Immune-mediated adverse events leading to treatment discontinuation by category
- Immune-mediated adverse events leading to death by category

- Immune-mediated adverse events leading to dose modification by category
- Immune-mediated adverse events treated with systematic corticosteroid by category
- Time to onset of Immune-mediated adverse events by category
- Time to resolution of Immune-mediated adverse events by category
- Duration of resolved Immune-mediated adverse events by category

Infusion-related reaction

For infusion related reaction (IRR), a summary of incidence by SOC, PT and maximum severity will be provided,. Summaries of IRRs, IRRs with NCI-CTCAE grade ≥3 and IRRs leading to treatment discontinuation will also be provided.

Death

All deaths and causes of death will be summarized by treatment arms, including those occurred during the study treatment period and those reported during the survival follow-up period after treatment completion/discontinuation.

7.6.3 Laboratory Values

Laboratory safety tests will be evaluated for selected parameters described in **Error! Reference source not found.**3.

Descriptive summary statistics (n, mean, standard deviation, median, minimum, maximum for continuous variables; n [%] for categorical variables) for laboratory parameters and their changes from baseline will be summarized by visit. Laboratory values will be summarized by visit and by worst postbaseline visit. The summary tables will report lab assessments up to 30 days after the last dose date.

Laboratory parameters (Table 3) that are graded in NCI CTCAE Version 5.0 will be summarized by shifts from baseline CTCAE grades to maximum post-baseline grades. In the summary of laboratory parameters by CTCAE grade, parameters with CTCAE grading in both high and low directions will be summarized separately.

Table 3: Serum Chemistry and Hematology Laboratory Tests

Serum Chemistry	**Hematology**	**Thyroid Function**
Alanine aminotransferase (ALT)	Hemoglobin	Free Triiodothyronine (FT3)
Aspartate aminotransferase (AST)	Platelet counts	Free Thyroxine (FT4)
Creatinine	White blood cell (WBC) count	Thyroid Stimulating Hormone (TSH)
Potassium	Neutrophil (Absolute)	
Sodium		
Creatine kinase (CK)		
Creatine kinase-cardiac muscle isoenzyme (CK-MB)		
Total Bilirubin		
LDH		

Serum Chemistry	Hematology	Thyroid Function
Calcium		

7.6.4 Vital Signs

Descriptive statistics for vital sign parameters (systolic and diastolic blood pressure, pulse rate, temperature, weight) and changes from baseline will be presented by visit.

7.6.5 Electrocardiograms (ECG)

The number and percentage of patients satisfying the following QTcF conditions at any time post-baseline will be summarized:

- >450, > 480, or > 500 msec
- ≤ 30 msec increase from baseline, > 30 and ≤ 60 msec increase from baseline, or > 60 msec increase from baseline

7.6.6 Eastern Cooperative Oncology Group (ECOG) Performance Status

A shift table from baseline to worst post-baseline in ECOG performance status will be summarized. ECOG status will be summarized by visit and treatment arm.

7.7 Pharmacokinetic Analyses

Tislelizumab post-dose and C_{trough} (pre-dose) will be tabulated and summarized for each cycle at which these concentrations are collected. Descriptive statistics will include means, medians, ranges, standard deviations, coefficient of variation (CV%), geometric mean and geometric CV%, as appropriate.

Additional PK analyses, including population PK analyses and exposure-response (efficacy, safety endpoints) analyses may be conducted as appropriate and the results from these analyses will be reported separately from the CSR.

7.8 Immunogenicity Analyses

Human anti-drug antibodies (ADA) to tislelizumab will be assessed during the study as defined in the protocol.

ADA attributes:

- Treatment-boosted ADA is defined as ADA positive at baseline that was boosted to a 4-fold or higher level following drug administration.
- Treatment-induced ADA is defined as ADA negative at baseline and ADA positive post-baseline.
- Persistent ADA response is defined as Treatment-induced ADA detected at 2 or more time points during treatment or follow-up, where the first and last ADA positive samples are separated by 16 weeks or longer; or detected in the last time point.
- Transient ADA response is defined as Treatment-induced ADA detected only at 1 time point during treatment or follow-up, excluding last time point; or detected at 2 or more timepoints during treatment or follow-up, where the first and last positive samples (irrespective of any negative samples in between) are separated by less than 16 weeks and the last time point is negative. Transient ADA is a treatment-induced response that is not considered persistent.
- Neutralizing ADA is defined as ADA that inhibits or reduces the pharmacological activity.

ADA response endpoints:

- ADA incidence is defined as sum of treatment-emergent ADA, which include both treatment-induced and treatment-boosted ADA-positive patients, as a proportion of the ADA evaluable population.
- ADA prevalence is defined as proportion of all patients that are ADA positive, including pre-existing ADA, at any time point.

The immunogenicity results will be summarized using descriptive statistics by the number and percentage of patients who develop detectable ADA. The incidence of positive ADA and neutralizing ADA will be reported for evaluable patients. The effect of immunogenicity on PK, efficacy, and safety may be evaluated if data allow, and reported separately from the CSR.

8 INTERIM ANALYSES

There will be one interim efficacy analysis of OS performed in the ITT Analysis Set. The interim efficacy analysis of OS will be performed with approximately 282 OS events observed in the ITT Analysis Set. The analysis timing is projected at approximately 36 months after the first patient randomization. The final OS analysis will be conducted with approximately 353 OS events projected to occur roughly 51 months after the first patient randomization.

The interim boundary for OS is based on the Lan-DeMets approximation to O'Brien-Fleming boundary. The interim and final analysis timing, and stopping boundaries are summarized in Table . The stopping boundaries in Table may be updated based on the actual OS events observed in the ITT Analysis Set at the time of interim analysis.

Table 4. Analysis Timing and Stopping Boundaries for OS in the ITT Analysis Set

Type of Analysis	Time (Months)	# Events	One-Sided Testing Boundary	
			p-Value Boundary	Approx. HR Threshold
Interim Analysis	36	282	0.0122	0.765
Final analysis	51	353	0.0214	0.806

Abbreviations: HR, hazard ratio; ITT, Intent-to-Treat; OS, overall survival

An interim analysis for OS will be performed by an independent statistician external to the Sponsor. The independent statistician will work with the blinded study statistician to provide statistical outputs to the IDMC as described in the IDMC charter and perform any ad-hoc analyses requested by the IDMC.

9 CHANGES IN THE PLANNED ANALYSIS

If the SAP needs to be revised, the sponsor will determine how the revision impacts the study and how the revision should be implemented. The details of the revision will be documented and described in the clinical study report.

10 REFERENCES

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics. 1982;38:29-41.

Greenwood M. The Natural Duration of Cancer. Reports on Public health and medical subjects. 1926;33:1-26.

ICH.Addendum on Estimands and Sensitivity Analyses in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials: E9(R1). 2019.

Mantel N, Haenszel W. Statistical Aspects of Analysis of Data from Retrospective Studies of Disease. Journal of the National Cancer Institute. 1959;22:719–748.

Sato T. On the Variance Estimator of the Mantel-Haenszel Risk Difference. Biometrics. 1989; 45:1323–1324.

APPENDIX 1. IMPUTATION OF MISSING OR PARTIALLY MISSING DATES

In general, missing or partial dates will not be imputed at the data level. The following rules will apply for the specific analysis and summary purposes mentioned below only.

1. Prior/Concomitant Medications

When the start date or end date of a medication/therapy/procedure is partially missing, the date will be imputed to determine whether the medication/therapy/procedure is prior or concomitant. The following rules will be applied to impute partial dates for medications.

If start date of a medication/therapy/procedure is partially missing, impute as follows:

- If both month and day are missing, then set to January 01
- If only day is missing, then set to the first of the month
- If the imputed start date > death date, then set to death date

If end date of a medication/therapy/procedure is partially missing, impute as follows:

- If both month and day are missing, then set to December 31
- If only day is missing, then set to last day of the month
- If the imputed end date > death date, then set to death date

If the year of start date or year of end date of a medication/therapy/procedure is missing, or the start date or end date is completely missing, do not impute.

2. Adverse Events

If year of the start date is missing or start date is completely missing, do not impute. Impute AE end date first if both AE start date and end date are partially missing.

If end date of an adverse event is partially missing, impute as follows:

- If both month and day are missing, then set to December 31
- If only day is missing, then set to last day of the month
- If the imputed end date > death date, then set to death date

If year of the end date is missing or end date is completely missing, only if subject is died and AE is ongoing, then impute end date to death date, otherwise do not impute.

If start date of an adverse event is partially missing, impute as follows:

- If both month and day are missing and year = year of treatment start date, then set to treatment start date
- If both month and day are missing and year ≠ year of treatment start date, then set to January 01
- If day is missing and month and year = month and year of treatment start date, the set to treatment start date
- If day is missing and month and year ≠ month and year of treatment start date, the set to first of the month
- If the imputed AE start date is after AE end date (maybe imputed), then update AE start date with AE end date as final imputed AE start date. If the imputed end date > death date, then set to death date.

3. Subsequent Anti-cancer Therapies/Surgery/Procedure

When the start date of subsequent anti-cancer therapy is partially missing, the following rules will be applied to impute partial dates.

If start date of is partially missing, impute as follows:

- If both month and day are missing, then set to December 31
- If only day is missing, then set to last day of the month
- If the imputed start date > min (death date, study discontinuation date, data cutoff date, start date of the next subsequent therapy), then set to min (death date, study discontinuation date, data cutoff date, start date of the next subsequent therapy)
- The imputed start date must be before or equal to the end date

If stop date of is partially missing, impute as follows:

- If both month and day are missing, then set to December 31
- If only day is missing, then set to last day of the month
- If the imputed stop date > min (death date, study discontinuation date, data cutoff date, start date of the next subsequent therapy), then set to min (death date, study discontinuation date, data cutoff date, start date of the next subsequent therapy)
- The imputed stop date must be after or equal to the end date

If year of the start date/stop date is missing, do not impute.

4. Prior anti-cancer therapy (drug, surgery/procedure, radiotherapy)

Impute end date first. If end date is partially missing, impute as follows:

- If both month and day are missing, then set to December 31
- If only day is missing, then set to the last day of the month
- For start date of prior systemic therapy for cancer, if imputed end date > randomization date – 6 months, then set to randomization date – 6 months
- For start date of prior radiotherapy/surgery, if imputed end date > randomization date, then set to randomization date - 1

If start date is partially missing, impute as follows:

- If both month and day are missing, then set to January 01
- If only day is missing, then set to the first of the month
- If the imputed start date > end date, then set to the end date

If the year of start date or year of end date of a medication/therapy/procedure is missing, or the start date or end date is completely missing, do not impute.

5. Diagnosis

If a diagnosis date is partially missing, impute as follows:

- If both month and day are missing, then set to January 01
- If only day is missing, then set to the first of the month

If a diagnosis date is completely missing, do not impute.

APPENDIX 2 HEALTH RELATED QUALITY OF LIFE

EORTC QLQ-C30 measures HRQoL of general cancer and includes two items that measure Global health status and quality of life (Global health status/QoL). The instrument is also consisted of two Functional Scales and Symptom Scales/ Items. Functional scales are Physical (2 items), Role (2 items), Emotional (4 items), Cognitive (2 items) and Social (2 items) functioning. Symptom scales are Fatigue (3 items), Nausea (2 items), Pain (2 items) symptoms. The single-items measuring symptoms include Dyspnoea, Insomnia, Appetite loss, Constipation, Diarrhoea and Financial Difficulties.

QLQ-LC13, which is the lung cancer module of the QLQ-C30, measure symptoms specific to lung cancer and its treatment. It is comprised of 13 items including Dyspnoea scale (2 items) and single items measuring Coughing, Haemoptysis, Sore month, Dysphagia, Perpherial neuropathy, Alopecia and Pain in chest, Pain arm/shoulder and Pain other parts.

EQ5D-5L measures general HRQoL consists for 5 scales (descriptive dimension) and a Visual Analogue Scale (VAS). The descriptive Dimensions scale includes Mobility, Self-Care, Usual Activities, Pain/Discomfort and Anxiety/Depression.

QLQ-C30 and QLQ-LC13 scale scores as well as EQ-5D-5L will be calculated as described below.

Scoring Process

QLQ-C30 and QLQ-LC13: The principle for scoring applies to all scales/scores: Raw scores are calculated as the average of the items that contribute to the scale.

A linear transformation to standardize the raw scores is utilized, so that the scores are ranged from 0 to 100. Increases in scores for functional domains (e.g., physical, role, social, emotional, etc.) are improvements while increases in scores for symptoms (e.g., fatigue, vomiting and nausea, diarrhea, pain, etc.) are deteriorations.

Missing Items

If at least half of the items for a scale are answered, then all the completed items are used to calculate the score. Otherwise, the scale score is set to missing.

In practical terms, if items I_1, I_2, ... I_n are included in a scale, the procedure is as follows:

Raw Score

For all scores, the raw score (RS), is the mean of the component items

RS = $(I_1+I_2+\ldots+I_n)/n$

Derived Scale

The derived scales are obtained from the raw scores as defined in the EORTC manual. The derived scales have a more intuitive interpretation: larger function scale or global health status / QoL are improvements while larger symptom scales (e.g., pain, nausea, etc.) are deteriorations.

The derivation formulas are as follows.

Linear transformation

Apply the linear transformation to 0-100 to obtain the score S,

Functional scales: $S = \left\{1 - \frac{(RS-1)}{range}\right\} \times 100$

Symptom scales / items: $S = \{(RS-1)/range\} \times 100$

Global health status / QoL: $S = \{(RS-1)/range\} \times 100$

The Index scores

To calculate QlQ-C30 index-score, individual functioning scale are subtracted by 100 to convert them into having the same meaning as symptom/problem scales. These 6 subtracted scores are subsequently summed with the 9 symptom/problem scales, and then divided by 15 (the total number of QLQ-C30 scales). A higher C30 index-score reflects a worse overall HRQOL. This is the mathematical formula:

QlQ- C30 index score = ∑[(100 - Physical functioning score), (100 - Role functioning score), (100 - Emotional functioning score), (100 - Cognitive functioning score), (100 - Social functioning score) , (100 - global QOL score), Fatigue score, Nausea/vomiting score , Pain score, Dyspnoea score, Insomnia score, Appetite loss score, Constipation score, Diarrhea score, Financial Difficulty score] ÷ 15

LC13 index-score is defined as the sum of all 10 QLQ-LC13 symptom/problem scales divided by 10 (the total number of QLQ-LC13 scales). A higher LC13 index-score reflects a worse overall HRQOL. This is the mathematical formula:

LC13 index score = Σ (scores of Dyspnoea, Coughing, Haemoptysis, Sore mouth, Dysphagia, Peripheral neuropathy, Alopecia, Pain in chest, Pain in arm or shoulder, Pain in other parts) ÷ 10

Table 5 Scoring of QLQ-C30

	Scale	Number of items	Item range	Item Numbers
Global health status/ QoL Global health status/QOL	QL2	2	6	29,30
Functional Scales				
Physical functioning	PF2	5	3	1, 2, 3, 4, 5
Role functioning	RF2	2	3	6, 7
Emotional functioning	EF	4	3	21, 22, 23, 24
Cognitive functioning	CF	2	3	20, 25
Social functioning	SF	2	3	26, 27

Symptom Scales/ items				
Fatigue	FA	3	3	10, 12, 18
Nausea and vomiting	NV	2	3	14, 15
Pain	PA	2	3	9, 19
Dyspnoea	DY	1	3	8
Insomnia	SL	1	3	11
Appetite loss	AP	1	3	13
Constipation	CO	1	3	16
Diarrhoea	DI	1	3	17
Financial Difficulties	FI	1	3	28

Table 6 Scoring of QLQ-LC13

	Scale	Number of items	Item range	Item Numbers
Symptom scales/items				
Dyspnoea	LCDY	3	3	3,4,5
Coughing	LCCO	1	3	1
Haemoptysis	LCHA	1	3	2
Sore mouth	LCSM	1	3	6
Dysphagia	LCDS	1	3	7
Peripheral neuropathy	LCPN	1	3	8
Alopecia	LCHR	1	3	9
Pain in chest	LCPC	1	3	10
Pain in arm or shoulder	LCPA	1	3	11
Pain in other parts	LCPO	1	3	12