[redacted]

STATISTICAL ANALYSIS PLAN

Study Protocol Number:	BGB-A317-A1217-301
Study Protocol Title:	A Phase 3, Randomized, Open-Label Study to Compare Ociperlimab (BGB-A1217) Plus Tislelizumab (BGB-A317) Versus Durvalumab in Patients With Locally Advanced, Unresectable, PD-L1-Selected Non-Small Cell Lung Cancer Whose Disease Has Not Progressed After Concurrent Chemoradiotherapy
Date:	07Nov2023
Version:	1.0

[redacted]

SIGNATURE PAGE

Author:	[redacted]

Approval

[redacted]

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

[redacted]

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
3D-CRT	3D conformal radiation
ADAs	antidrug antibodies
AE	adverse event
ALK	anaplastic lymphoma kinase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC	area under the concentration-time curve
BOR	best overall response
BLA	biologics license application
CmDn	cycle m day n
CBR	clinical benefit rate
CI	confidence interval
CK	creatine kinase
CK-MB	creatine kinase cardiac muscle isoenzyme
CR	complete response
CRR	complete response rate
cCRT	concurrent chemoradiotherapy
CT	computed tomography
DCR	disease control rate
DOR	duration of response
DVA	dose volume analysis
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data capture (system)
EGFR	epidermal growth factor receptor
eGFR	estimated glomerular filtration rate
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D-5L	EuroQol 5 dimension 5 level
EOT	End-of-Treatment (Visit)
EVs	extracellular vesicle

[redacted]

GTV	gross tumor volume
HR	hazard ratio
HRQoL	health-related quality of life
IDMC	independent data monitoring committee
imAE	immune-related adverse event
IMRT	intensity modulated radiotherapy
IRC	Independent Review Committee
ITT	Intent-to-Treat (Analysis Set)
ITIM	immunoreceptor tyrosine-based inhibitory motif
ITV	internal target volume
IV	intravenously
LA NSCLC	locally advanced non-small cell lung cancer
MedDRA	Medical Dictionary for Regulatory Activities
NCI-CTCAE	National Cancer Institute-Common Terminology Criteria for Adverse Events
NSCLC	non-small cell lung cancer
BGB-A1217	ociperlimab
OAR	organs at risk
ORR	overall response rate
OS	overall survival
PA	protocol amendment
PD	progressive disease
PD-L1	programmed cell death protein-ligand 1
PGI-S	patient reported global impression of severity
PFS	progression-free survival
PK	pharmacokinetics
PR	partial response
PRTSE	patient reported treatment side-effect burden
PT	preferred term
PTV	planning target volume
PVR	poliovirus receptor
Q1	25th percentile
Q3	75th percentile
QLQ-C30	quality of life questionnaire core 30
QLQ-LC13	quality of life questionnaire-lung cancer 13
QnW	every n weeks

ORR	objective response rate
OS	overall survival
RECIST	Response Evaluation Criteria in Solid Tumors
RT	radiotherapy
SAE	serious adverse event
SAP	statistical analysis plan
SAS	safety analysis set
SC	steering committee
SD	standard deviation
SOC	System Organ Class
T3	Triiodothyronine
T4	Thyroxine
TEAE	treatment-emergent adverse event
TC	tumor cells
TIGIT	T-cell immunoglobulin and ITIM domain
BGB-A317	tislelizumab
TIL	tumor-infiltrating lymphocyte
TMB	tumor mutation burden
TNM	tumor nodes metastasis
TTDM	time to death or distant metastasis
WHO DD	World Health Organization Drug Dictionary
VAS	visual analogue scale

[redacted]

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, under abbreviated clinical study report (CSR) setting, for AdvanTIG-301 protocol amendment (PA) version 2.0: A Phase 3, Randomized, Open-Label Study to Compare Ociperlimab (BGB-A1217) Plus Tislelizumab (BGB-A317) Versus Durvalumab in Patients With Locally Advanced, Unresectable, programmed cell death protein-ligand 1 (PD-L1) Selected Non-Small Cell Lung Cancer (NSCLC) Whose Disease Has Not Progressed After Concurrent Chemoradiotherapy (cCRT).

The study was initially implemented with Protocol Amendment 1.0 (PA 1.0, dated on 16 Apr 2021). In the PA 1.0, patients with newly diagnosed, histologically confirmed, unresectable locally advanced NSCLC and evaluable PD-L1 expression all comers were enrolled; cCRT was given within the study. There were 63 patients who were randomized under the PA 1.0, from29 Jun 2021 to 26 Apr 2022. No patient was randomized under the PA 2.0. The concurrent part (patients randomized under PA 1.0) was integrated into PA 2.0 and managed mainly per Appendix 17 of PA 2.0 after PA 2.0 became effective.

Given the study was early terminated with no patients enrolled under PA 2.0 and only 63 patients randomized under the concurrent part (patients randomized under PA 1.0), the objectives of PA 2.0 are exploratory and have been modified in this SAP by selecting the applicable objectives in PA 2.0 and PA 1.0.

1.1. Study Overview

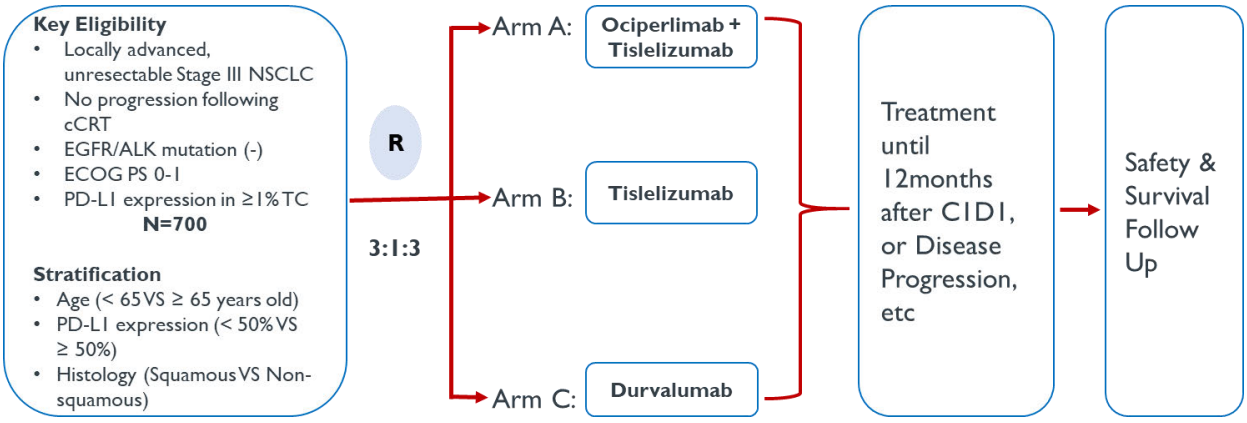
The PA2.0 is an open-label, randomized, multicenter, Phase 3 study to compare the efficacy and safety of anti T-cell immunoglobulin and immunoreceptor tyrosine-based inhibitory motif (ITIM) domains (anti TIGIT) monoclonal antibody ociperlimab plus tislelizumab versus durvalumab in patients with unresectable locally advanced (LA) NSCLC whose disease has not progressed after definitive, platinum based cCRT and with PD L1 expression on ≥ 1% of tumor cells (TC) as assessed by the central lab using the VENTANA PD-L1 (SP263) assay.

Randomization will be stratified by age (< 65 years versus ≥ 65 years), PD-L1 expression on TC (≥ 50% versus < 50%), and histology (squamous versus nonsquamous). The ITT and PD L1 ≥ 50% populations are defined as patients with PD-L1 expression on ≥ 1% and ≥ 50% of TC, respectively.

Approximately 700 patients will be randomized in a 3:1:3 ratio to receive the study treatment in the following 3 arm. Arm A: ociperlimab (900 mg intravenously [IV]) combined with tislelizumab (200 mg IV) every 3 weeks (Q3W), Arm B: tislelizumab 200 mg IV Q3W, and Arm C: durvalumab 10 mg/kg IV once every 2 weeks (Q2W) (or 1500 mg every 4 weeks [Q4W] where the dosage has been approved by the local health authority). Study drugs will be given starting from Cycle 1 Day 1 (C1D1) and continued for up to 12 months, or until progressive disease (PD) per Response Evaluation Criteria in Solid Tumors (RECIST) v1.1, unacceptable toxicity, or death, or until another discontinuation criterion is met, whichever occurs first.

The study design schema of PA 2.0 is shown in Figure 1.

Figure 1 Study Schema of PA 2.0



For all patients randomized with PA 1.0 (defined as Concurrent Part of the study):

The study was initially implemented with Protocol Amendment 1.0 (dated on 16 Apr 2021, PA 1.0). In the PA 1.0, patients with newly diagnosed, histologically confirmed, unresectable LA NSCLC and evaluable PD-L1 expression all comers were enrolled; cCRT was given within the study.

Patients were enrolled and randomized in a 1:1:1 ratio to ociperlimab (BGB-A1217) plus tislelizumab plus cCRT followed by ociperlimab plus tislelizumab (Arm A), tislelizumab plus cCRT followed by tislelizumab (Arm B), or cCRT followed by durvalumab (Arm C) in this study. Randomization will be stratified by age (< 65 years versus ≥ 65 years), PD-L1 expression in TC (≥ 1% versus < 1%), and histology (squamous versus non-squamous).

Independent Review Committee (IRC) were removed for concurrent part (patients randomized under PA 1.0) in PA 2.0. Therefore, the endpoints assessed by IRC will not be analyzed and the endpoints by investigator will be analyzed.

2. STUDY OBJECTIVES

2.1. Primary Objective

- Not applicable

2.2. Secondary Objective

- Not applicable

2.3. Exploratory Objective

- To compare progression free survival (PFS) as assessed by the investigator per RECIST v1.1 between Arm A and Arm C, between Arm B and Arm C, and between Arm A and Arm B

- To compare overall survival (OS) between Arm A and Arm C, between Arm B and Arm C, and between Arm A and Arm B
- To compare objective response rate (ORR), duration of response (DOR), and complete response rate (CRR) as assessed by the investigator per RECIST v1.1 between Arm A and Arm C, between Arm B and Arm C, and between Arm A and Arm B
- To compare time to death or distant metastasis (TTDM) between Arm A and Arm C, between Arm B and Arm C, and between Arm A and Arm B
- To compare safety and tolerability by treatment-emergent adverse events (TEAEs) according to NCI CTCAE v5.0
- To characterize the pharmacokinetics (PK) of ociperlimab and tislelizumab
- To assess host immunogenicity to ociperlimab and tislelizumab
- To evaluate health-related quality of life (HRQoL), patient reported global impression of severity (PGI-S) and patient reported treatment side-effect burden (PRTSE)

3. STUDY ENDPOINTS

3.1. Primary Endpoint(s)

- Not applicable

3.2. Secondary Endpoints

- Not applicable

3.3. Exploratory Endpoints

- PFS, ORR, DOR, BOR, DCR, and CBR by investigators per RECIST v1.1
- OS, TTDM, HRQoL, PGI-S and PRTSE
- Safety and tolerability, defined as AEs (using NCI CTCAE v5.0), laboratory tests, vital signs, ECOG Performance Status, physical examinations, concomitant medications, and dose modification
- Serum concentrations of ociperlimab and tislelizumab at specified timepoints
- Immunogenic responses to ociperlimab and tislelizumab evaluated through detection of anti-drug antibodies

4. SAMPLE SIZE CONSIDERATIONS

PA 2.0

The sample size calculation is based on the number of events regarding primary efficacy analyses of PFS between Arm A and Arm C for comparisons in both the PD L1 ≥ 50% and ITT Analysis

Sets. Exponential distribution is assumed for PFS. To demonstrate efficacy with regard to PFS, the estimates of the number of events required are based on the following assumptions:

- The randomization ratio for Arm A versus Arm B versus Arm C is 3:1:3.
- A steady-state enrollment rate of 20 patients per month and an enrollment ramp-up duration of 12 months.
- PFS evaluation dropout rate of 5% per 12 months.
- The prevalence of patients with PD-L1 ≥ 50% of TC in the ITT Analysis Set is approximately 50%. Enrollment of patients with PD-L1 < 50% on TCs might be stopped, if necessary, to ensure that the population reflects the natural PD L1 expression prevalence.
- Sequential testing procedure is implemented to control the overall alpha at a 2.5% one sided level. PFS analysis in Arm A versus Arm C in the PD-L1 ≥ 50% Analysis Set will be performed first, and PFS analysis in Arm A versus Arm C in the ITT Analysis Set will be carried out only if the PFS analysis in the PD-L1 ≥ 50% Analysis Set yields a statistically significant between-arm difference favoring Arm A.
- One interim analysis is planned when approximately 75% of the final PFS events have occurred for each PFS primary endpoint, using Lan-DeMets O'Brien-Fleming approximation spending function.
- Median PFS in Arm C within both the PD-L1 ≥ 50% and ITT Analysis Sets is 24.9 months.

The statistical assumption and power in the sample size calculation in Arm A and Arm C is summarized below; and the total sample size of the study would be 700.

1-sided alpha	Analysis Set	HR (A vs C)	mPFS in A	mPFS in C	Accrual Duration	Sample size in A+C	# Events at FA in A+C	Power
0.025	PD-L1 ≥ 50%	0.6	41.5	24.9	41.3	300	141	85.2%
0.025	ITT	0.7	35.6	24.9	41.3	600	288	85.1%

Abbreviation: A: Arm A; C: Arm C; FA: final analysis, mPFS: median Progression-free Survival.

Sample size and power is calculated using EAST (Version 6.5) and R (Version 4.1.2).

PA 1.0

The sample size calculation is based on the number of events regarding primary efficacy analyses of PFS for both comparisons between Arm A and Arm C and between Arm B and Arm C, and the difference of CRR by IRC assessment per RECIST v1.1 between Arm A and Arm C to demonstrate PFS superiority or CRR superiority. Exponential distribution is assumed for PFS. To demonstrate efficacy with regards to PFS, the estimates of the number of events required are based on the following assumptions:

[redacted]

1. Median PFS of 14.6 months in Arm C.
2. A steady-state enrollment rate of 35 patients per month and an enrollment ramp-up duration of 10 months and the randomization ratio for Arm A versus Arm B versus Arm C is 1:1:1.
3. PFS evaluation dropout rate of 5% per 12 months.
4. One interim analysis is planned when approximately 75% of final PFS events have occurred for each PFS primary endpoint, using Lan-DeMets O'Brien-Fleming approximation spending function.

Nine hundred patients will be enrolled over a 30.6-month period in order to accumulate approximately 336 events at final analysis of PFS in Arm A and Arm C to have a 90% power for detecting a HR of 0.7 using a 1-sided alpha of 0.0249 in the ITT Analysis Set and accumulate approximately 343 events at final analysis of PFS in Arm B and Arm C to have a 75.2% power for detecting a HR of 0.75 using a 1-sided alpha of 0.0249 in the ITT Analysis Set. With the first 411 patients randomized for the 3 arms, approximately 137 patients in each Arm A and Arm C, the study has an 80% power for detecting a 17% difference in CRR (20% versus 3%) in the comparison of Arm A versus Arm C using a 1-sided alpha of 0.0001.

Sample size and power is calculated by EAST (version 6.4.1) and R (version 4.0.3).

5. DEFINITION OF PRIMARY ESTIMANDS

Not applicable

6. STATISTICAL METHODS

6.1. Analysis Sets

All the analysis sets are for concurrent part of PA 2.0 (patients randomized under PA 1.0).

The ITT Analysis Set includes all randomized patients.

The Safety Analysis Set (SAS) includes all randomized patients who received any dose of study treatment.

The PK Analysis Set includes all patients who receive any dose of any component of study drugs and for whom any postdose PK data are available.

The Immunogenicity Analysis Set includes all patients who receive any dose of any component of study drugs and for whom both baseline antidrug antibody (ADA) and at least 1 postbaseline ADA result are available.

6.2. Data Analysis General Considerations

All descriptive statistics for continuous variables will be reported using mean, standard deviation (SD), median, 25th percentile (Q1), 75th percentile (Q3), minimum (Min), maximum (Max) and n.

[redacted]

Categorical variables will be summarized as number (percentage) of patients. Time-to-event variable: number of non-missing observations (N), median, minimum and maximum. Kaplan-Meier event rates may also be provided if applicable for specific time-to-event variable.

The study Table Listing Graph shells will be provided in a separate document, which will show the content and format of all tables, listings, and graphs in detail.

6.2.1. Definitions and Computations

Study drugs include ociperlimab, tislelizumab, durvalumab, chemotherapy (cisplatin, etoposide, carboplatin, paclitaxel and pemetrexed) and radiotherapy.

Baseline Measurements:

- For efficacy evaluation: Unless otherwise specified, a baseline value is defined as the last non-missing value collected prior to or at the time of randomization date. This rule also applies to the stratification factors age, PD-L1 expression in TC and histology.
- Safety variables: a baseline value is defined as the one which is the latest available valid measurement taken prior to or on the first study drug administration date. If any randomized patients are not treated, the baseline will be the last value on or prior to the randomization date.

Unscheduled Visits: 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 and unscheduled data.

Study Follow-up Duration: the duration from the randomization date to the study discontinuation date (e.g., death, consent withdrawal, lost to follow-up) or to cutoff date for a patient who is still ongoing in the study.

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

6.2.2. 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.
- P-values will be rounded to 4 decimal places. P-values that round to 0.0000 will be presented as '< 0.0001' and p-values that round to 1.000 will be presented as '> 0.9999'.
- Duration of image-based event endpoints (such as PFS) will be based on the actual date the radiograph was obtained rather than the associated visit date.
- For laboratory results collected as < or >, a numeric value, 0.0000000001 will be subtracted or added, respectively, to the value.
- 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.

[redacted]

6.2.3. Handling of Missing Data

Missing data will not be imputed unless otherwise specified elsewhere in this SAP. 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 Appendix A.

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.

6.2.4. Multiplicity Adjustment

Multiplicity will not be adjusted for the exploratory objectives.

6.2.5. 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.

6.3. Subject Characteristics

6.3.1. Subject Disposition

The number (percentage) of subjects randomized, treated, permanently discontinued from the study, reasons for discontinued from the study, remained on the study and the duration of study follow-up will be summarized in the ITT analysis set. The patients who discontinued treatment and the primary reason for the end of treatment will be summarized among patients who were treated.

6.3.2. Protocol Deviations

Important protocol deviation criteria will be established, and patients with important protocol deviations will be identified and documented before the database lock. Protocol deviations will be listed by each category.

6.3.3. 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
- Race
- Ethnicity
- Geographic Region (Asian versus Non-Asian)
- Weight

[redacted]

- BMI
- ECOG (0 versus 1)
- Smoking status (Never, Current/Former)

6.3.4. Disease History

The number and percentage of patients reporting a history of disease and characteristic will be summarized in the ITT Analysis Set. Categorical disease characteristics variables include disease stage, PD-L1 expression in tumor cells by central lab, histology, epidermal growth factor receptor (EGFR) mutation status, anaplastic lymphoma kinase (ALK) rearrangement status, and Tumor-nodes-metastasis (TNM) stage at study entry. Continuous disease history variables include time from initial diagnosis to study entry (day) and baseline target lesion sum of diameters.

6.3.5. Prior Anticancer Drug Therapies and Surgeries

A summary for prior anticancer drug therapies and surgeries is not planned for the study.

6.3.6. Prior and Concomitant Medications

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

Prior medications are defined as medications that were stopped before the day of the first dose of study drug. Concomitant medications are 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 (as of the Safety Follow-up Visit).

The number and 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.

6.3.7. Medical History

Medical History will be coded using MedDRA of the version currently in effect at BeiGene at the time of database lock. The number and percentage of patients reporting a history of any medical condition, as recorded on the eCRF, will be summarized by System Organ Class (SOC) and preferred term (PT) in the ITT Analysis Set.

A listing of medical history will be provided.

6.4. Efficacy Analysis

Progression Free Survival (PFS)

PFS per investigator is defined as the time from randomization to the first documented disease progression as assessed by the investigator according to RECIST v1.1, or death from any cause, whichever occurs first. Censoring rules for PFS are specified in Table 5.

PFS assessed by the investigator per RECIST v1.1 between Arm A versus Arm C, Arm B versus Arm C and Arm A versus Arm B in the ITT Analysis Set will be analyzed. A log-rank test stratified by the IRT value of stratification factors at randomization (i.e., PD-L1 expression in TC [<1% versus ≥1%], tumor histologic type [squamous versus non-squamous], and age [<65 versus ≥65]) will be used to test the PFS differences between Arm A versus Arm C, Arm B versus Arm C, and Arm A versus Arm B. One-sided nominal p-value using stratified log-rank test will be reported, for descriptive purpose only. The median, Q1 and Q3 of PFS will be assessed at selected timepoints (6, 12, 18 months) if estimable and will be calculated for each treatment arm and presented with a 2-sided 95% confidence intervals (CIs). Kaplan-Meier survival probabilities over time for each arm will be plotted using Kaplan-Meier method.

The treatment effect will be estimated by fitting a Cox regression model with treatment arm as a factor and the IRT value of stratification factors (i.e., PD-L1 expression in TC [<1% versus ≥1%], tumor histologic type [squamous versus non-squamous], and age [<65 versus ≥65]) as strata. From this model, the hazard ratio (HR) of PFS will be estimated and presented with a 2-sided 95% CI. Unstratified Cox regression will also be performed to provide HR and corresponding 95% CI.

Overall Survival (OS)

OS is defined as the time from randomization to death from any cause. OS will be analyzed in the ITT Analysis Set. Data for patients who are not reported as having died at the time of analysis will be censored at the date the patients were last known to be alive. 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 for other alive patients. Data for patients who do not have postbaseline information will be censored at the date of randomization.

Note: 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.

The OS will be analyzed similarly as PFS.

Time to Death or Distant Metastases (TTDM)

Distant metastasis is defined as any new lesion that is metastatic lesion (M1a, M1b, or M1c lesions) per RECIST v1.1 or proven by biopsy.

The TTDM will be analyzed similarly as PFS except that comparisons of TTDM between Arm A versus Arm B will not be summarized.

Objective Response Rate (ORR)

ORR with confirmation is the proportion of patients who had a confirmed CR or PR as assessed by the investigator per RECIST v1.1 in all randomized patients. Patients without any postbaseline assessment will be considered non-responders. Similar method used to evaluate CRR will be applied to the analysis of ORR. In addition, the number and percentage of patients for each of the Best Overall Response (BOR) categories (e.g., CR, PR, SD, PD, NE) will be presented by

[redacted]

treatment arm. Similar method used to evaluate ORR with confirmation and confirmed BOR will be applied to the analysis of ORR without confirmation and unconfirmed BOR. ORR by the investigator per RECIST v1.1 in the ITT Analysis Set between Arm A and Arm C will be summarized. The 2-sided 95% Cis for the difference in ORR will be calculated, as well as the Clopper-Pearson 95% Cis for the ORR within each arm. The 1-sided p-value from Cochran-Mantel-Haenszel method stratified by stratification factors will be provided for descriptive purpose.

ORR by the investigator per RECIST v1.1 in the ITT Analysis Set between Arm B and Arm C and between Arm A and Arm B will be summarized similarly.

A waterfall plot of (confirmed) best percent change in sum of target lesion diameters from baseline will be provided for each treatment arm. Patients will be ordered by the percentage, from the smallest to the largest in percentages.

Duration of Response (DOR)

DOR is defined as the time from date of the first objective response to the first documented disease progression as assessed by the investigator according to RECIST v1.1, or death from any cause, whichever occurs first. DOR will be analyzed among the responders and ITT Analysis Set based on assessment by the investigator, between Arm A versus Arm C. All the censoring rules for PFS primary analysis would be applied to DOR as well.

The Median and other quantiles of DOR and the cumulative probability of DOR at every 3 months if estimable, will be calculated for each treatment arm and presented with 2 sided 95% Cis.

Disease Control Rate (DCR)

DCR is defined as the proportion of patients who achieve CR, PR, or SD. DCR will be summarized similarly as ORR in the ITT Analysis Set.

Clinical Benefit Rate (CBR)

CBR is defined as the proportion of patients who achieve CR, PR, or durable SD (SD ≥ 24 weeks). CBR will be summarized similarly as ORR in the ITT Analysis Set.

EORTC-QLQ-C30

The European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 (QLQ-C30) consists of 30 questions which can be combined to produce five functional scales (Physical, Role, Cognitive, Emotional, and Social), a global health status/QoL scale, and a number of single items assessing additional symptoms commonly reported by cancer subjects (e.g., dyspnea, loss of appetite, insomnia, constipation, and diarrhea), and financial impact of the disease. The QLQ-C30 will be scored according to the EORTC scoring manual (Fayers et al 2001). Higher scores on the global health status and functioning scales indicate better health status/function but higher scores on symptom scales/items represent greater symptom severity. Please refer to Table 1 for scoring of QLQ-C30 details.

An outcome variable consisting of a score from 0 to 100 will be derived for each of the symptom scales/symptom items, the functional scales and the global health status scale in the QLQ-C30. Each scale of the QLQ-C30 will be transformed so that scale scores will range from 0 to 100. The transformation will proceed in two steps. First, the average of the items contributing to a subscale

[redacted]

will be calculated to compute the raw score of the scale. Next, a linear transformation will be applied to 'standardize' the raw score.

- For all scales, the raw score (RS), is the mean of the component items:

$$\text{RS} = (I_1 + I_2 + \cdots + I_n)/n$$

- For functional scales, the derivation formula is as follows:

$$\text{Score} = \left\{1 - \frac{(\text{RS} - 1)}{\text{range}}\right\} \times 100$$

- For symptom scales/items and global health status / QoL, the derivation formula is as follows:

$$\text{Score} = \left\{\frac{(\text{RS} - 1)}{\text{range}}\right\} \times 100$$

For examples, the raw score and the functional scale of emotional functioning are as follows,

$$\text{RS}_{\text{EF}} = (Q_{21} + Q_{22} + Q_{23} + Q_{24})/4, \text{ EF score} = \left\{1 - \frac{\text{RS}_{\text{EF}} - 1}{3}\right\} \times 100$$

The raw score and the functional scale of fatigue are as follows,

$$\text{RS}_{\text{FA}} = \frac{Q_{10} + Q_{12} + Q_{18}}{3}, \text{ FA score} = \left\{\frac{\text{RS}_{\text{FA}} - 1}{3}\right\} \times 100$$

Table 1 Scoring of QLQ-C30 version 3.0

Scale name	Scale	Number of items	Item range	Items
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

Item range is the difference between the possible maximum and the minimum response to individual items.

[redacted]

EORTC-QLQ-LC13

The European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire-Lung Cancer 13 (QLQ-LC13) is a lung cancer specific module from the EORTC comprising 13 questions to assess lung cancer symptoms (cough, haemoptysis, dyspnoea and site-specific pain), treatment related side-effects (sore mouth, dysphagia, peripheral neuropathy and alopecia) and pain medication.

The dyspnoea scale is only used if all 3 items have been scored, otherwise the items are treated as single-item measures. The developers of EORTC-QLQ-LC13 indicate that it is highly preferred not to use the LC13 alone (without the core module QLQ-C30), since the module has been designed to be used together with the core questionnaire, and the content validity is based upon this combination. The response options and scoring system are the same as for the QLQ-C30, and the administration is similar. The recall period for items is the past 7 days, and response options include either a 4-point Likert scale or yes/no options. The total score for the instrument ranges from 0 to 100. A high score for a symptom scale or item represents a high level of symptomatology or problems (Fayers et al 2001). Please refer to Table 2 for scoring of QLQ-LC13 details.

Table 2 Scoring of QLQ-LC13

Scale name	Scale	Number of items	Item range	Items
Symptom scales/ items				
Dyspnoea	LCDY	3	3	33, 34, 35
Cough	LCCO	1	3	31
Haemoptysis	LCHA	1	3	32
Sore mouth	LCSM	1	3	36
Dysphagia	LCDS	1	3	37
Peripheral neurophathy	LCPN	1	3	38
Alopecia	LCHR	1	3	39
Pain in the chest	LCPC	1	3	40
Pain in arm or shoulder	LCPA	1	3	41
Pain in other parts	LCPO	1	3	42
Did you take any medicine for pain?		1	2	43
If yes, how much did it help?			4	

Item range is the difference between the possible maximum and the minimum response to individual items.
The dyspnoea scale should only be used if all three items have been answered. Some respondents ignore question 35 because they never climb stairs; in this case, the score for the dyspnoea scale would be biased if it were based upon the other two items. Hence if item 35 is missing then items 33 and 34 should be used as single-item measures.

EQ-5D-5L

The EuroQol 5 dimension 5 level (EQ-5D-5L) comprises a descriptive module and a Visual Analogue scale (VAS). The descriptive module comprises of 5 dimensions: mobility, self-care, usual activities, pain/ discomfort, and anxiety/depression. Each dimension has 5 levels: 1= no problems, 2 = slight problems, 3 = moderate problems, 4 = severe problems, and 5 = extreme problems. Higher scores indicate lower quality of life. The EQ VAS measures respondent's self-rated health status on a 0 to 100 scale, with 100 = 'the best health you can imagine' and 0 = 'the worst health you can imagine'. Higher scores on VAS indicate higher health status.

[redacted]

Descriptive modules of EQ-5D-5L will be summarized by visit and dimension in an ordinal scale.

PGI-S

PGI-S is a 1-item questionnaire designed to assess patient's impression of disease severity. The PGI-S item asks the respondent to best describe how severe their symptoms are now ("Check the one number that best describes how your lung cancer symptoms are now") on a 5-point scale scored as: "not at all" (1), "mild" (2), "moderate" (3), "very" (4), "extremely" (5). Higher scores indicate higher severity of symptoms.

Descriptive modules of PGI-S will be summarized by visit in an ordinal scale.

PRTSE

PRTSE is a 1-item questionnaire designed to assess patient reported treatment related side effect burden. The PRTSE item asks the respondent to describe how much the patient is bothered by the side effect of the administered treatment on a 5-point scale scored as: "not at all bothered" (1), "mildly bothered" (2), "moderately bothered" (3), "very bothered" (4), or "extremely bothered" (5).

Descriptive modules of PRTSE will be summarized by visit in an ordinal scale.

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. No imputation will be performed for missing scale score. Completion and compliance rates will be summarized at each timepoint by treatment arm for QLQ-C30, QLQ-LC13, EQ VAS, PGI-S, and PRTSE.

6.5. Safety Analyses

All safety analyses will be performed by treatment arms based on the safety analysis set.

6.5.1. Extent of Exposure

The following measures of the extent of exposure will be summarized:

- Duration of exposure for ociperlimab, tislelizumab and durvalumab:
  - Duration of tislelizumab and ociperlimab (months) is defined as,

    (min (date of cutoff, death date, last dose date + 20) – first dose date + 1)/30.4375
  - Duration of durvalumab Q2W (months) is defined as,

    (min (date of cutoff, death date, last dose date + 13) – first dose date + 1)/30.4375
  - Duration of durvalumab Q4W (months) is defined as,

    (min (date of cutoff, death date, last dose + 27) – first dose date + 1)/30.4375
- Number of treatment cycles/weeks received is defined as
  - The total number of treatment cycles for immunotherapy in which at least one dose of the ociperlimab, tislelizumab or durvalumab is administered.

[redacted]

- The total number of treatment cycles for chemotherapy with non-missing doses for cisplatin plus etoposide, cisplatin plus pemetrexed, and carboplatin plus pemetrexed are administered.

  Note that the frequency of administration and duration of treatment for carboplatin plus paclitaxel are not cycle based. Therefore, the number of weeks received will be summarized instead.

- The number of cycles/weeks received will also be summarized in category of 1 cycle/week, 2 cycles/weeks, …

Note that the average cycle length will not be calculated for carboplatin plus paclitaxel.

- Number of fractions received: the number of fractions taken in radiotherapy will be calculated as the sum of numbers of completed fractions. Number of fractions received will also be summarized in category of 1-5, 6-10, 11-15, 16-20, 21-25, and 26-30.
- Cumulative dose received: defined as
  - IO and chemotherapy (mg): the sum of all actual doses given from first to last administration. The summary will include study drug administration in all scheduled and unscheduled visits prior to the cutoff date.
  - Radiotherapy (Gy): Following the above summary for IO. In addition, the cumulative dose received will be categorized as: <56 Gy, 56 to <57 Gy, 57 to <=60 Gy, and >60 Gy.
- Actual dose intensity: defined as the cumulative dose received by a patient divided by the duration of exposure.
  - Ociperlimab, tislelizumab (mg/cycle): 21* total cumulative dose (mg) / (last dose up to cutoff date +21 – first dose date) [note: 21 days each cycle]
  - Durvalumab Q2W (mg/kg/cycle): [14*total cumulative dose (mg/kg) / (last dose up to cutoff date + 14 – first dose date)]/ weight at each visit [note: 14 days each cycle for durvalumab Q2W; planned dose: 10 mg/kg]
  - Durvalumab Q4W (mg/cycle): 28*total cumulative dose (mg) / (last dose up to cutoff date + 28 – first dose date) [note: 28 days each cycle for durvalumab Q4W; planned dose 1500 mg]
  - The actual dose intensity of chemotherapy is listed in the Table 3. And the estimation of GFR (eGFR) can be done by using the Calvert formula to measure creatinine clearance.
    - Creatinine clearance for male (mL /min): $\frac{(140\text{-age})(\text{weight})}{(72)(SC_r)}$
    - Creatinine clearance for female (mL /min): $\frac{(0.85)(140\text{-age})(\text{weight})}{(72)(SC_r)}$
- Relative dose intensity (%): defined as

$$\frac{\text{ADI}}{\text{Planned Dose Intensity}} \times 100$$

where planned dose intensity is defined as the planned one cycle dose or planned one week dose with detail referring to Table 3.

Table 3 Example Formula to Calculate ADI and Planned Dose Intensity of Chemotherapy by Cycle

	Individual Chemotherapy	ADI	Planned dose per cycle or per week
Cisplatin + Etoposide	Cisplatin	$\frac{\sum_1^{\#\text{ of cycles}} \frac{\text{actual dose}}{\text{BSA}}}{\max\left(\frac{\text{last dose up to cutoff date} + 21 - \text{first dose date}}{28}, \text{number of cycles of corresponding non-missing drug administration}\right)}$	2 x 50 mg/m^2/cycle
	Etoposide	$\frac{\sum_1^{\#\text{ of cycles}} \frac{\text{actual dose}}{\text{BSA}}}{\max\left(\frac{\text{last dose up to cutoff date} + 24 - \text{first dose date}}{28}, \text{number of cycles of corresponding non-missing drug administration}\right)}$	5 x 50 mg/m^2/cycle
Carboplatin + Paclitaxel	Carboplatin	$\frac{\sum_1^{\#\text{ of weeks}} \frac{\text{actual dose}}{\text{GFR}+25}}{\max\left(\frac{\text{last dose up to cutoff date} + 7 - \text{first dose date}}{7}, \text{number of weeks of corresponding non} - \text{missing drug administration}\right)}$	AUC 2 per week
	Paclitaxel	$\frac{\sum_1^{\#\text{ of weeks}} \frac{\text{actual dose}}{\text{BSA}}}{\max\left(\frac{\text{last dose up to cutoff date} + 7 - \text{first dose date}}{7}, \text{number of weeks of corresponding non-missing drug administration}\right)}$	40-50 mg/m^2/ per week
Cisplatin + Pemetrexed	Cisplatin	$\frac{\sum_1^{\#\text{ of cycles}} \frac{\text{actual dose}}{\text{BSA}}}{\max\left(\frac{\text{last dose up to cutoff date} + 21 - \text{first dose date}}{21}, \text{number of cycles of corresponding non} - \text{missing drug administration}\right)}$	75 mg/m^2/cycle

	Pemetrexed	$\frac{\sum_{1}^{\#\text{ of cycles}} \frac{\text{actual dose}}{\text{BSA}}}{\max\left(\frac{\text{last dose up to cutoff date} + 21 - \text{first dose date}}{21}, \text{number of cycles of corresponding non} - \text{missing drug administration}\right)}$	500 mg/m^2/cycle
Carboplatin + Pemetrexed	Carboplatin	$\frac{\sum_{1}^{\#\text{ of cycles}} \frac{\text{actual dose}}{\text{GFR}+25}}{\max\left(\frac{\text{last dose up to cutoff date} + 21 - \text{first dose date}}{21}, \text{number of cycles of corresponding non} - \text{missing drug administration}\right)}$	AUC 5 per cycle
	Pemetrexed	$\frac{\sum_{1}^{\#\text{ of cycles}} \frac{\text{actual dose}}{\text{BSA}}}{\max\left(\frac{\text{last dose up to cutoff date} + 21 - \text{first dose date}}{21}, \text{number of cycles of corresponding non} - \text{missing drug administration}\right)}$	500 mg/m^2/cycle

Number of cycles received by patient as a quantitative variable and by category (i.e., number (%) of patient receiving at least 1 cycle, at least 2 cycles etc.), duration of exposure, cumulative dose, ADI and RDI will be summarized by descriptive statistics.

The number and percentage of patients with dose modification to ociperlimab/tislelizumab/ durvalumab, to cisplatin plus etoposide, to carboplatin plus paclitaxel, to cisplatin plus pemetrexed, to carboplatin and pemetrexed, and to radiotherapy will be summarized, respectively. Reasons for dose modification to ociperlimab, tislelizumab, and durvalumab include dose delay and infusion interruption. Reasons for dose modification to chemotherapy include infusion interruption, dose delay and dose reduction. Reasons for dose modification of radiotherapy only includes dose delay.

6.5.2. Adverse Events

Aes will be graded by the investigators using CTCAE version 5.0. The AE verbatim descriptions (investigator terms from the eCRF) will be classified into standardized medical terminology using MedDRA. Adverse events will be coded to the MedDRA (Version 25.0 or higher) lowest level term closest to the verbatim term, along with the linked MedDRA PT and primary system organ class (SOC).

Treatment-emergent Adverse Events

A treatment-emergent adverse event (TEAE) is defined as an AE that had onset or a worsening in severity from baseline (pretreatment) on or after the first dose of study drug and up to 30 days after study drug discontinuation or initiation of new anti-cancer therapy, whichever occurs first. Only those AEs that were treatment-emergent will be included in the summary tables of TEAE. All AEs, treatment-emergent or otherwise, will be presented in patient data listings.

An overall summary of TEAEs will summarize the number (%) of patients with:

- At least one TEAE
- At least one TEAE with grade 3 or above
- At least one serious TEAE
- At least one TEAE leading to death
- At least one TEAE leading to treatment discontinuation
    - At least one TEAE leading to discontinuation of ociperlimab
    - At least one TEAE leading to discontinuation of tislelizumab/durvalumab
    - At least one TEAE leading to discontinuation of chemotherapy
    - At least one TEAE leading to discontinuation of radiotherapy
- At least one TEAE leading to treatment modification of any component of study drug
    - At least one TEAE leading to treatment modification of ociperlimab
        - Leading to infusion interruption
        - Leading to dose delay
        - Leading to infusion rate decrease
    - At least one TEAE leading to treatment modification of tislelizumab/durvalumab
        - Leading to infusion interruption
        - Leading to dose delay
        - Leading to infusion rate decrease
    - At least one TEAE leading to treatment modification of chemotherapy
    - At least one TEAE leading to treatment modification of radiotherapy
- At least one TEAE related to any component of study drug
    - At least one TEAE related to ociperlimab
    - At least one TEAE related to tislelizumab/durvalumab
    - At least one TEAE related to chemotherapy
    - At least one TEAE related to radiotherapy
    - At least one TEAE with grade ≥3 and related to any component of study drug
        - At least one TEAE with grade ≥3 and related to ociperlimab
        - At least one TEAE with grade ≥3 and related to tislelizumab/durvalumab
        - At least one TEAE with grade ≥3 and related to chemotherapy
        - At least one TEAE with grade ≥3 and related to radiotherapy
    - At least one serious TEAE related to any component of study drug
        - At least one serious TEAE related to ociperlimab

[redacted]

        - At least one serious TEAE related to tislelizumab/durvalumab
        - At least one serious TEAE related to chemotherapy
        - At least one serious TEAE related to radiotherapy
    - At least one TEAE leading to death related to any component of study drug
        - At least one TEAE leading to death related to ociperlimab
        - At least one TEAE leading to death related to tislelizumab/durvalumab
        - At least one TEAE leading to death related to chemotherapy
        - At least one TEAE leading to death related to radiotherapy
    - At least one TEAE leading to treatment discontinuation and related to any component of study drug
    - At least one TEAE leading to treatment modification and related to any component of study drug

The incidence of TEAEs will be reported as the number (percentage) of patients with

- TEAEs by SOC and PT (any grade and grade ≥3)
- Treatment-related TEAE by SOC and PT (any grade and grade ≥3)
- Treatment-related TEAE of ociperlimab by SOC and PT (any grade and grade ≥3)
- Treatment-related TEAE of tislelizumab/durvalumab by SOC and PT (any grade and grade ≥3)
- Treatment-related TEAE of chemotherapy by SOC and PT (any grade and grade ≥3)
- Treatment-related TEAE of radiotherapy by SOC and PT (any grade and grade ≥3)
- Serious TEAE by SOC and PT
- Serious treatment related TEAE by SOC and PT
- Serious TEAE related to ociperlimab by SOC and PT
- Serious TEAE related to tislelizumab/durvalumab by SOC and PT
- Serious TEAE related to chemotherapy by SOC and PT
- Serious TEAE related to radiotherapy by SOC and PT
- TEAE leading to death by SOC and PT
- Treatment-related TEAE leading to death by SOC and PT
- TEAE leading to treatment discontinuation by SOC and PT
- TEAE leading to ociperlimab discontinuation by SOC and PT
- TEAE leading to tislelizumab/ durvalumab discontinuation by SOC and PT
- TEAE leading to chemotherapy discontinuation by SOC and PT

[redacted]

- TEAE leading to radiotherapy discontinuation by SOC and PT
- TEAE leading to dose modification by SOC and PT
- TEAE leading to dose modification of ociperlimab by SOC and PT
- TEAE leading to dose modification of tislelizumab/durvalumab by SOC and PT
- TEAE leading to dose modification of chemotherapy by SOC and PT
- TEAE leading to dose modification of radiotherapy by SOC and PT
- Treatment-related TEAE leading to treatment discontinuation by SOC and PT
- Treatment-related TEAE leading to dose modification by SOC and PT

Treatment-related AEs include those events considered by the investigator to be related to study treatment or with missing assessment of the causal relationship. For patients with multiple occurrences of the same event will be counted only once, and the maximum grade per CTCAE v5.0 will be used if CTCAE grade is needed.

Immune-mediated Adverse Events

Immune-mediated AEs (imAE) 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 and up to 90 days from the last dose of study drug, regardless of whether the patient starts a new anti-cancer therapy.

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

- Overview of immune-mediated adverse events
- Immune-mediated adverse events by category and PT (any grade and grade ≥3)
- Immune-mediated adverse events by category and worst grade
- Serious immune-mediated adverse events by category and PT
- Immune-mediated adverse events leading to treatment discontinuation of any study drug by category and PT
- Immune-mediated adverse events leading to treatment modification of any study drug by category and PT
- Summary of immune-mediated adverse events treated with systemic corticosteroids by category

Deaths

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

Number and causes of deaths, classified by deaths within 30 days of last dose of study drug and deaths more than 30 days after the last dose, will be summarized based on Safety Analysis Set.

[redacted]

6.5.3. Radiotherapy Quality Assurance

A formal radiotherapy quality assurance (QA) program will be a mandatory component of this study. Prior to the randomization of the first subject, sites will need to have undergone a "benchmarking" process where a 3-dimensional CT scan of an anonymized clinical case with stage III NSCLC will be electronically transmitted to participating sites by the QA program, and the local radiation oncologists from the sites will be required to contour target volumes and organs-at-risk volumes, in accordance with the quality assurance criteria specified for thoracic radiation therapy as defined in the protocol and the approved RT QA program. All subjects treated by radiation in the study will submit their data to the RT QA team to perform a retrospective radiation prescription. This will be a confirmation of the treatment delivered, and an overall assessment of compliance with protocol stipulations regarding radiation treatment. In case there is any deviation from the protocol, this information will be conveyed to the institution.

The total dose of radiotherapy will be 60 Gy, administered in 30 once-daily fractions of 2 Gy and 5 fractions per week for 6 weeks. Weekly clinical assessments are required during the cCRT phase. While 60 Gy in 30 fractions is the target dose of radiation, concern about patient tolerance discovered during treatment planning or delivery, for example, a V20 exceeding that recommended in the protocol, may dictate that a lower dose be administered. Definitive RT will be considered as receiving a minimum of 56 Gy to the planning target volume (PTV). Normalization of the treatment plan will cover 95% of the PTV with the prescription dose. The minimum PTV dose should ideally not fall below 93% of the prescription dose. No field reductions will be permitted and the entire PTV must be treated daily. The gross tumor volume (GTV), Internal Target Volume (ITV) and PTV are defined on all appropriate slices. Organs at risk (OAR) includes spinal cord, lung, heart, esophagus, and brachial plexus. The followings will be summarized to evaluate the

- Radiotherapy (RT) completion summary:
  - Completed as planned (complete at least 56 Gy or 28 fractions), incomplete radiotherapy (complete less than 56 Gy or 28 fractions) and no radiotherapy was given.
  - The number (percentage) of patients whose radiotherapy plan was adapted during the RT.
- Radiotherapy modality (IMRT, 3D-CRT, Other)
- Radiotherapy location
- Radiotherapy overall score rated by quality assurance vendor (Imaging and Radiation Oncology Core) by the following: no deviation (1), minor deviation (2), major deviation (3).
- Radiation contour score will be summarized by the following categories: no deviation (1), minor deviation (2), major deviation (3).
  - Overall RT contour
  - Overall target volume (TV) contour
  - GTV contour

[redacted]

  - Clinical target volume (CTV) contour
  - ITV contour
  - PTV contour
  - Overall OAR contour
  - Spinal cord contour
  - Lung contour
  - Heart contour
  - Esophagus contour
  - Brachial plexus contour
- Radiation dose volume analysis:
  - Overall RT dose volume analysis (DVA) by category
  - Overall TV DVA by category
  - PTV dose 95% (D95) (Gy) (continuously and by category)
  - PTV volume 55.8% (V93) (%) (continuously and by category)
  - PTV D 1cc (inside PTV) Dmax (Gy) (continuously and by category)
  - External PTV D1cc (outside PTV) Dmax (continuously and by category)
  - Overall OAR DVA by category
  - Spinal cord Dmax (Gy) (continuously and category)
  - Lung V20 (%) (continuously and by category)
  - Lung V5 (%) (continuously and by category)
  - Lung Dmean (Gy) (continuously and by category)
  - Heart V50 (%) (continuously and by category)
  - Heart Dmean (Gy) (continuously and by category)
  - Esophagus Dmean (Gy) (continuously and by category)
  - Brachial plexus Dmax (Gy) (continuously and by category)
- RT dose specification in patients who experienced >=grade 3 pneumonitis or radiation pneumonitis
  - Lung V20 (%) (continuously and by category)
  - Lung V5 (%) (continuously and by category)
  - Lung Dmean (Gy) (continuously and by category)

[redacted]

6.5.4. Laboratory Values

Hematology, serum chemistry, and thyroid function results will be summarized/listed for selected parameters described in Table 4.

Abnormal laboratory values will be flagged and identified as those outside (above or below) the normal range. Reference (normal) ranges for laboratory parameters will be included. Plots of laboratory values/change from baseline over time will be provided for selected lab parameters.

Laboratory parameters that are graded in NCI-CTCAE v.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. The lab parameters with grades increased in at least 2 from baseline to worst post-baseline will also be summarized.

Potential Hy's Law for liver injury will also be summarized.

Table 4 Clinical Laboratory Assessments

Serum Chemistry	Hematology	Coagulation	Urinalysis	Thyroid Function
Potassium Sodium Chloride Creatinine Blood Urea Nitrogen Urea Albumin Total Protein Total Bilirubin Direct Bilirubin Alanine aminotransferase (ALT) Aspartate aminotransferase (AST) Alkaline phosphatase (ALP) Calcium Magnesium Phosphate Glucose Lactate dehydrogenase CK CK-MB (Activity) CK-MB (Mass) Troponin T Troponin I	Hemoglobin Hematocrit White Blood Cell Count Platelet Count Neutrophils Absolute Lymphocytes Absolute	Prothrombin time (PT) PTT aPTT International Normalized Ratio (INR)	Glucose Protein Ketones Blood WBC pH	Free Triiodothyronine (FT3) Free Thyroxine (FT4) Thyroid Stimulating Hormone (TSH)

6.5.5. Vital Signs

Patient data listing of vital signs will be provided.

[redacted]

6.5.6. Physical Examination

Analyses of physical examinations are presented in either the adverse event form or medical history section, as clinical significant abnormalities in the physical examination were categorized adverse events, while pre-existing conditions before the study treatment were documented in the medical history Electrocardiograms (ECG).

ECG will be performed at the baseline and multiple time points after the start of treatment.

Abnormal post-baseline QTc results will be summarized with the following categories:

- Patients with increase of >30 msec, increase of >60 msec from baseline
- Patients with post-baseline value of >450 msec, value of >480 msec, value of >500 msec

Patient listing of all ECG recordings will be provided.

6.5.7. Eastern Cooperative Oncology Group (ECOG)

A shift table from baseline to worst post-baseline in ECOG performance score will be summarized.

6.6. Pharmacokinetic Analyses

Pharmacokinetic samples will be collected in this study for patients randomized to Arm A (ociperlimab and tislelizumab) and Arm B (tislelizumab) as outlined in Appendix 1 of Protocol.

The following analysis plan provides the framework for the summarization of the PK data from study AdvanTIG-301. The objective is to summarize available ociperlimab and tislelizumab PK concentrations following IV administration. PK parameters will not be characterized as only sparse samples were collected.

6.6.1. Reporting of Pharmacokinetic Concentrations for Descriptive Statistics

Ociperlimab (Arm A only) and tislelizumab (Arms A and B) serum concentration data will be listed and tabulated by visit/cycle at which these concentrations are collected per the study design. Descriptive statistics will include means, medians, ranges, standard deviations, coefficient of variation (CV%), geometric means, and geometric CV%, as appropriate, and will be based on the PK Analysis Set.

7. IMMUNOGENICITY ANALYSES

Anti-drug antibodies (ADAs) samples will be collected in this study for patients randomized to Arm A (ociperlimab and tislelizumab) and Arm B (tislelizumab) as outlined in Appendix 1 of Protocol.

The scope of ADAs calculations used for characterizing clinical immunogenicity depends on the incidence and kinetics of detected ADA. Therefore, not all parameters described below will be derived or additional parameters may be added based on the results. The effect of immunogenicity on PK, efficacy, and safety may be evaluated if data allows and will be reported separately from the CSR.

The immunogenicity results will be summarized using descriptive statistics by the number and percentage of subjects who develop detectable ADAs for ociperlimab and tislelizumab separately based on Immunogenicity Analysis Set. The incidence of positive and neutralizing ADAs (as applicable) will be reported for ADA-evaluable subjects according to the following definitions:

- **ADA-evaluable patient:** Number of patients with reportable non-missing baseline result and at least one reportable sample taken after drug administration during the treatment or follow-up observation period with reportable result (used for computing treatment-induced ADA incidence).
- **Treatment-emergent ADA:** Sum of treatment-boosted and treatment-induced ADA patients as a proportion of the ADA-evaluable patient population. Synonymous with "ADA Incidence".
- **Treatment-induced ADA:** ADA-evaluable patients that were ADA-negative at baseline and ADA-positive after drug administration during the treatment or follow-up observation period.
- **Treatment-boosted ADA**: Baseline-positive ADA-evaluable patients with significant increases (4-fold or higher) in ADA titer after drug administration during the treatment or follow-up observation period. Baseline-positive ADA-evaluable patient is an ADA-evaluable patient with positive ADA result at baseline.
- **Persistent ADA:** Treatment-induced ADA detected at two or more sampling time points during the treatment (including follow-up period if any), where the first and last ADA-positive samples (irrespective of any negative samples in between) are separated by a period of 16 weeks or longer, or, treatment-induced ADA incidence only in the last sampling time point of the treatment study period or at a sampling time point with less than 16 weeks before an ADA-negative last sample.
- **Transient ADA:** Treatment-induced ADA detected only at one sampling time point during the treatment or follow-up observation period, or two or more time points during the treatment, where the first and last ADA-positive samples are separated by a period of less than 16 weeks, and the subject last sampling time point is ADA-negative.
- **Neutralizing ADA:** ADA-evaluable patients with positive NAb.

8. CHANGES IN THE PLANNED ANALYSIS

Given the study was early terminated with no patients enrolled under PA 2.0 and only 63 patients randomized under PA 1.0 (the concurrent part of PA 2.0), the objectives of PA 2.0 are exploratory and have been modified in this SAP by selecting the applicable objectives in PA 2.0 and PA 1.0.

[redacted]

9. REFERENCES

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

Clopper, C. and Pearson, ES. The use of confidence or fiducial limits illustrated in the case of the binomial, Biometrika 1934; 26: 404-413.

Eisenhauer EA, Therasse P, Bogaerts J Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer. 2009; 45:228-47.

Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A on behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 Scoring Manual (3rd edition). Brussels: EORTC, 2001.Lan K and DeMets D. Discrete Sequential Boundaries for Clinical Trials. Biometrika. 1983;70(3):659-663.

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

Kaplan E, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association. 1958; 53:457-81.

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.

[redacted]

APPENDIX A. IMPUTATION OF MISSING OR PARTIALLY MISSING DATES

Missing data will not be imputed unless otherwise specified. Missing data will not be imputed in the listings.

Concomitant Medication/Procedure

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:

If start date of a medication/therapy/procedure is partially missing, the imputation rules are as follows.

- If both month and day are missing, then set to January 01
- If only day is missing, then set to the first day 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, the imputation rules are 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 end date of a medication/therapy/procedure is missing, or the start date or end date is completely missing, do not impute.

Adverse Events

When the start date or end date of an adverse event is partially missing, the date will be imputed to determine whether the adverse event is treatment-emergent. The following rules will be applied to impute partial dates for adverse events:

If start date of an adverse event is partially missing, the imputation rules are 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, then set to treatment start date
- If day is missing and month and year ≠ month and year of treatment start date, then set to first day 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

[redacted]

- If the imputed end date > min (death date, end of study date), then set to min (death date, end of study date)

If end date of an adverse event is partially missing, the imputation rules are 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 > min (death date, end of study date), then set to min (death date, end of study date)

If year of the start date or start date is completely missing, do not impute; if year of the end date or end date is completely missing, do not impute.

Disease History and Prior Therapy (drug, surgery/procedure, radiotherapy)

The following rules will be applied to impute partial dates such as initial diagnosis date, initial BCLC staging date, relapse date, therapy date (start/end date), or surgery date.

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 prior systemic therapy for cancer, if imputed end date > randomization date – 6 months, then set to randomization date – 6 months
- For prior radiotherapy/locoregional therapy, 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.

Subsequent anti-cancer therapy collected in the post-treatment page

If start date of subsequent anti-cancer therapy 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 subsequent anti-cancer therapy), then set to min(death date, study discontinuation date, data cutoff date, start date of the next subsequent therapy)

If stop date of subsequent anti-cancer therapy is partially missing, impute as follows:

- If both month and day are missing, then set to December 31

[redacted]

- If only day is missing, then set to last day of the month
- If the imputed stop > min (death date, study discontinuation date, data cutoff date, start date of the next subsequent anti-cancer 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 (imputed) start date. If year of the start date/stop date is missing, do not impute.

[redacted]

APPENDIX B. CENSORING RULES FOR PRIMARY ANALYSIS OF PFS PER RECIST VERSION 1.1

Table 5 Censoring Rules for Primary and Sensitivity Analysis of PFS

No.	Situation	Date of Progression or Censoring	Primary Analysis
1	No baseline or any post-baseline tumor assessments and without death within 19 weeks from reference start date*	Reference start date	Censored
2	Progression documented between scheduled visits	Date of first radiologic PD assessment	Progressed
3	No progression at the time of data cut-off or withdrawal from study	Date of last adequate radiologic assessment prior to or on date of data cut-off or withdrawal from study	Censored
4	New anticancer treatment started	Date of last adequate radiologic assessment prior to or on date of new anticancer treatment	Censored
5	Death before first PD assessment	Date of death	Progressed
6	Death between adequate assessment visits**	Date of death	Progressed
7	Death or progression after more than one missed visit***	Date of last adequate radiologic assessment before missed tumor assessments	Censored
8	No baseline or any post-baseline tumor assessments and died within 19 weeks from reference start date*	Date of death	Progressed

The reference start date is the randomization date.
*19 weeks is defined as two tumor assessments duration plus the protocol allowed window 7 days.
**Adequate tumor assessment is a radiologic assessment of CR, PR, SD, non-CR/non-PD or PD as determined by the investigators.
***More than one missed visit is defined if the duration between the last tumor assessment and death or PD is longer than D2. The D2 is defined as two times protocol specified interval between TAs plus the protocol allowed window around the assessments. Since tumor assessment is scheduled approximately every 9 weeks (±7 days) from randomization, for the first 54 weeks and every 12 weeks (±7 days) thereafter based on RECIST v1.1.

[redacted]

APPENDIX C. RULES FOR IDENTIFYING MISSING TUMOR ASSESSMENTS

Rules for identifying at least two missing tumor assessments

1) Input scheduled TA list:
    a. TA was performed approximately every 9 weeks (±7 days) from randomization, for the first 54 weeks, and every 12 weeks (±7 days) thereafter based on RECIST v1.1: week 9, week 18, week 24, …, week 54, week 66, week 78, week 90,…
2) Identify last evaluable TA before PD or death (LPTADT) and map it to the closest scheduled visit (LPTADT_WK):
    a. In the event of unscheduled TA, choose the closest scheduled visit number as LPTADT_WK. It can be achieved programmatically by following the classification rule (e.g., defining thresholds) described in the below table.
    b. Otherwise (scheduled visit), assign the scheduled visit number (assuming it is coded correctly) to LPTADT_WK.
3) Find the 2nd TA visit LPTADT_WK according to the list in Step 1 (LPTADT_WK_2):
    a. If LPTADT_WK_2 + 1 week < earliest date of PD/death, then PFS is censored at the LPTADT

Table 6 shows how to assign unscheduled TA to a scheduled visit. The Threshold column is defined as the mid-point between current and next visit (except for baseline); it is the upper limit for LPTADT to be mapped to the prior scheduled assessment (step 2a above). For example, if LPTADT is Week 57 for an unscheduled visit, it will be mapped to Week 54 TA since it is within the Threshold for Week 54.

Table 6 An example of scheduled tumor assessments with time window

Weeks	Scheduled week -1	**Scheduled week**	Scheduled week+1	Threshold
Baseline		Baseline		
Every 9 weeks for the first 54 weeks	Week 8	Week 9	Week 10	Week 13
	Week 17	Week 18	Week 19	Week 22
	Week 26	Week 27	Week 28	Week 31
	Week 35	Week 36	Week 37	Week 40
	Week 44	Week 45	Week 46	Week 49
	Week 53	Week 54	Week 55	Week 60
Every 12 weeks afterwards	Week 65	Week 66	Week 67	Week 72
	Week 77	Week 78	Week 79	Week 84
	Week 89	Week 90	Week 91	Week 96
	…	…	…	…