STATISTICAL ANALYSIS PLAN

Study Protocol Number:	BGB-HNSCC-201
Study Protocol Title:	A Randomized, Phase 2, Open-Label, Multi-Arm Study of Tislelizumab in Combination With Investigational Agents as First-Line Treatment in Patients With Recurrent or Metastatic Head and Neck Squamous Cell Carcinoma
Regulatory Agency Identification Number:	EU CT Number: 2023-503418-63-00
Date:	12 May 2025
Version:	1.0
NCT:	NCT05909904

SIGNATURE PAGE

Author: [redacted]	[redacted] Signature Date: 12-May-2025 | 15:21:56 PDT

Approval

[redacted]	[redacted] Signature Date: 12-May-2025 | 15:41:15 PDT
[redacted]	[redacted] Signature Date: 18-May-2025 | 19:32:22 PDT
[redacted]	[redacted] Signature Date: 13-May-2025 | 18:20:33 PDT

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS 5
1 INTRODUCTION 7
1.1 Study Design 7
2 GENERAL CONSIDERATIONS 8
2.1 Definitions and Computations 8
2.2 Conventions 9
2.3 Multiplicity Adjustment 9
2.4 Handling of Missing Data 9
3 ANALYSIS SETS 9
4 ANALYSIS SUPPORTING PRIMARY OBJECTIVE(S) 10
4.1 Definition of Endpoint(s) 10
4.2 Statistical Methods 10
4.3 Sensitivity Analysis 11
5 ANALYSES SUPPORTING SECONDARY OBJECTIVE(S) 11
5.1 Key Secondary Endpoint(s) 11
5.1.1 Definition of Endpoint(s) 11
5.1.2 Statistical Methods 11
6 PATIENT CHARACTERISTICS 14
6.1 Patient Disposition 14
6.2 Demographic and Other Baseline Characteristics 14
6.3 Disease History 14
7 SAFETY ANALYSES 16
7.1 Extent of Exposure 16
7.2 Adverse Events 16
7.3 Laboratory Values 19
8 OTHER ANALYSES 20
9 INTERIM ANALYSES 20
9.1 Futility Interim Monitoring 20
7.1 Safety stopping rules 20
10 CHANGES IN THE PLANNED ANALYSIS 21

11 SAMPLE SIZE CONSIDERATIONS 22
12 REFERENCES 23
APPENDIX 1. IMPUTATION OF MISSING OR PARTIALLY MISSING DATES 24
1. Impute partial dates for concomitant medication 24
2. Impute partial dates for adverse events 24
APPENDIX 2. RULES FOR IDENTIFYING MISSING TUMOR ASSESSMENTS 27

LIST OF TABLES

Table 1: Event and Censoring Rules for the Derivation of Progression-free Survival per RECIST Version 1.1 12
Table 2: Event and Censoring Rules for the Analysis of Overall Survival 13
Table 3: Serum Chemistry and Hematology Laboratory Tests 19
Table 4: Safety Stopping Boundaries for Number of Patients With at Least Possibly Treatment-related Deaths 20
Table 5: Statistical Analysis Plan Changes 21
Table 6: Scheduled tumor assessments with time window 27

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Definition
ADI	actual dose intensity
AE	adverse event
BMI	body mass index
BOR	best overall response
BGB-A317	tislelizumab
CBR	clinical benefit rate
CI	confidence interval
CPS	combined positive score
CR	complete response
DCR	disease control rate
dMMR	deficient mismatch repair
DOR	duration of response
ECG	electrocardiogram
eCRF	electronic case report form
ECOG	Eastern Cooperative Oncology Group
EOT	End-of-Treatment
HNSCC	head and neck squamous cell carcinoma
HPV	human papillomavirus
HR	hazard ratio
imAE	immune-mediated adverse event
IRT	Interactive Response Technology
ITT	Intent-to-Treat
LKAD	last known alive date
MedDRA	Medical Dictionary for Regulatory Activities
MMR	mismatch repair
MSI	microsatellite instability
MSI-H	microsatellite instability status-high
MSI-L	microsatellite instability status-low
MSS	microsatellite stability
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events

Abbreviation	Definition
NE	not evaluable
ORR	objective response rate
OS	overall survival
PD	progressive disease
PD-L1	programmed cell death protein ligand-1
PFS	progression-free survival
PK	pharmacokinetic(s)
pMMR	proficient mismatch repair
PR	partial response
PT	preferred term
RDI	relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumors
R/M	Recurrent/Metastatic
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SOC	system organ class
TA	tumor assessment
TE	treatment-emergent
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 the study BGB-HNSCC-201: A Randomized, Phase 2, Open-Label, Multi-Arm Study of Tislelizumab in Combination With Investigational Agents as First-Line Treatment in Patients With Recurrent or Metastatic Head and Neck Squamous Cell Carcinoma. The focus of this SAP is for the planned interim analysis and the final analysis specified in the study protocol. The analysis details for pharmacokinetics (PK) and biomarker analyses are not described within this SAP. Separate analysis plans may be completed for these analyses if appropriate and will be attached in addition to this SAP to the clinical study report.

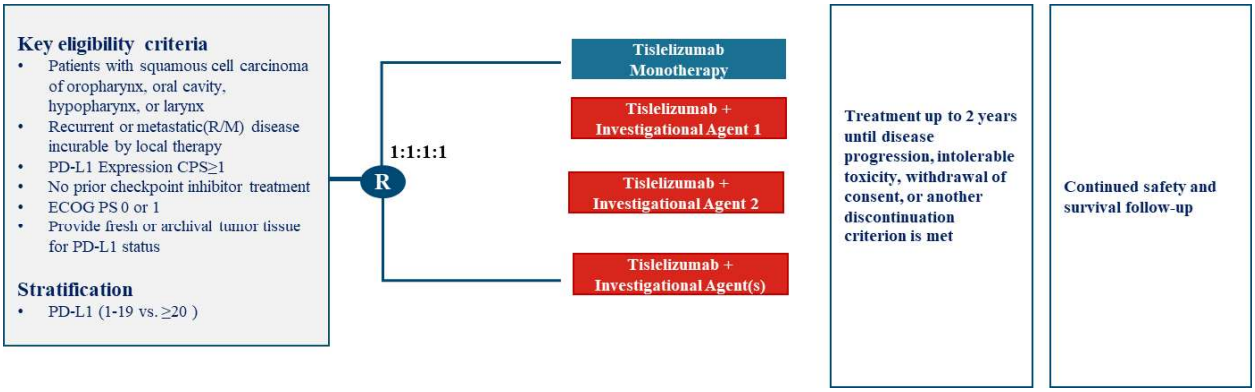
1.1. Study Design

This is a randomized, open-label, multicenter, Phase 2, multi-arm study designed to evaluate the efficacy and safety of tislelizumab and tislelizumab in combination with investigational agent(s) in first-line recurrent or metastatic (R/M) HNSCC with PD-L1 positive expression combined positive score (CPS) ≥ 1. Patients will be randomized to receive either tislelizumab + investigational agent(s) (experimental arm[s]) or tislelizumab monotherapy (reference arm). The initial randomization ratio between each of the experimental arms and the common reference arm will be 1:1.

Randomization in each treatment arm will be stratified based on PD-L1 expression (CPS 1 to 19 versus ≥ 20). An 1:1 randomization ratio is initially applied to each of the experimental arms and the common reference arm. The randomization ratio may be modified to account for emerging safety data, efficacy data, and feasibility considerations. The total number of patients depends on the number of experimental arms and the timing when they are added to the study.

The study design schematic is presented below.

Figure 1: Study Schema



Abbreviations: CPS, combined positive score; ECOG PS, Eastern Cooperative Oncology Group Performance Status; PD-L1, programmed cell death protein ligand-1; R/M, recurrent or metastatic; R, randomization.

Treatments in each arm will be administered as follows:

- Reference arm: tislelizumab
- Experimental arm(s): tislelizumab + investigational agent(s)

The details regarding specific treatment regimens/requirements for investigational agents are provided in each appendix.

Study treatments in all arms will be administered for up to 2 years until disease progression, intolerable toxicity, withdrawal of informed consent, or another discontinuation criterion is met, whichever occurs first.

Tumor response will be assessed by the investigators using standard Response Evaluation Criteria in Solid Tumors Version 1.1 (RECIST v1.1) criteria. Tumor imaging must be performed within 28 days before randomization/enrollment. On-study tumor assessments will occur every 6 weeks (± 7 days) from randomization for the first 52 weeks and then every 12 weeks (± 7 days) thereafter.

Safety will be assessed throughout the study by monitoring adverse events (AEs)/serious adverse events (SAEs) toxicity grades assigned per National Cancer Institute Common Terminology Criteria for Adverse Events Version 5.0 and laboratory results. Vital signs, physical examinations, ECOG Performance Status changes, electrocardiogram (ECG) results, and other examinations will also be used for safety assessments.

2. GENERAL CONSIDERATIONS

2.1. Definitions and Computations

Study drugs include tislelizumab (BGB-A317), BGB-A425 and LBL-007.

To derive the study day and the duration of any safety endpoints in safety analysis set, the reference start date is the first treatment (dose) date. To derive the study day and duration of any endpoints in intent to treat (ITT) analysis set, the reference start date is the randomization date. For assessments conducted on or after the date of reference date, the study day will be calculated as (assessment date –reference date + 1). For assessments conducted before the date of reference date, study day is calculated as (assessment date – reference date). There is no study day 0. In the situation where the event date is partial or missing, the date will appear partial or missing in the listings.

Unless otherwise specified, a baseline value for efficacy evaluation is defined as the last non-missing value collected before randomization. A baseline value for safety evaluation is defined as the last non-missing value collected before the first dose of study drug. If no dose was given then the randomization date will be used instead.

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

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.
- 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 are less than 0.0001 will be presented as '< 0.0001' and p-values that are larger than 0.9999 will be presented as '> 0.9999'.
- Duration of disease assessment-based event endpoints (such as progression-free survival (PFS) and disease-free survival) 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 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.

2.3. Multiplicity Adjustment

Since no formal hypothesis is tested in this study, multiplicity adjustment data is not applicable.

2.4. 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 the handling of missing or partially missing dates for adverse events and prior/concomitant medications/procedures are provided in Appendix 1.

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.

Patients without postbaseline tumor assessment will be considered as non-responders for ORR.

3. ANALYSIS SETS

The Intent-to-Treat (ITT) Analysis Set is defined as all patients who undergo randomization in one of the treatment arms. Patients will be analyzed according to the treatment arm/group they are randomized. The ITT Analysis Set will be used as the primary analysis population for all efficacy and demographic analyses unless specified otherwise.

The Safety Analysis Set includes all patients who receive at least 1 dose of study drug(s). Patients will be analyzed according to the treatment they received. All safety and exposure analyses will be conducted using the Safety Analysis Set unless specified otherwise.

The Efficacy Evaluable Analysis Set includes all patients who receive at least 1 dose of study drug(s), have evaluable disease at baseline, and have at least 1 evaluable postbaseline tumor response assessment unless any clinical disease progression or death occurred before the first scheduled postbaseline tumor assessment. The Efficacy Evaluable Analysis Set will be used as the supportive analysis population for efficacy endpoints.

The clinical disease progression is defined by 'Symptomatic Deterioration' as the primary reason for study treatment discontinuation, recorded on the eCRF.

The PK Analysis Set includes all patients who receive at least 1 dose of study drug(s) and have ≥ 1 measurable post-dose concentration(s).

The Biomarker Analysis Set includes all patients who receive at least 1 dose of study drug(s) and have ≥ 1 evaluable biomarker measurement.

4. ANALYSIS SUPPORTING PRIMARY OBJECTIVE(S)

4.1. Definition of Endpoint(s)

The primary efficacy endpoint for this study is the confirmed objective response rate (ORR) as assessed by the investigators per Response Evaluation Criteria in Solid Tumors (RECIST) v1.1.

Confirmed objective response rate (ORR) is defined as the proportion of patients who have a confirmed complete response (CR) or a confirmed partial response (PR), as assessed by the investigators per Response Evaluation Criteria in Solid Tumors (RECIST) v1.1.

4.2. Statistical Methods

The confirmed ORR will be analyzed for the ITT analysis set. The corresponding two-sided 95% confidence interval (CI) calculated from Clopper Pearson exact method will be also presented. The difference in the confirmed ORR between any experimental arm and the reference arm will be evaluated using Mantel-Haenszel odds ratio and/or risk difference, stratified by PD-L1 expression (CPS = 1 to 19 versus ≥ 20). The corresponding 95% CI for the odds ratio, as well as the 95% CI and 80% CI for the risk difference, will be provided. The primary analysis will be performed using the PD-L1 expression per randomization stratification (per IRT).

The best overall response (BOR) is defined as the best response recorded from the date of randomization until progressive disease (PD) or up to start of new anticancer therapy, whichever comes first. If the first tumor assessment occurs after the new anticancer therapy, the BOR is considered as not evaluable (NE).

- Confirmed CR is defined as at least 4 weeks apart (in-between) 2 CRs. Single "NE" between two CRs, CR NE CR is considered as confirmed CR. More than one NE between CR is considered as unconfirmed. Any assessments, except NE, between two CRs are considered as data issue and need to be queried.

- Confirmed PR is defined as at least 4 weeks apart (in-between) the first PR and the last PR/CR. Single "NE" or "Stable Disease (SD)" between two PRs, (PR NE PR, or PR SD PR) is considered as confirmed. More than one NE/SD after PR is considered as unconfirmed. If it is found that there is a PD in between 2 PRs, query of the data.
- Stable Disease (SD) is defined as the assessments of SD which is at least 5 weeks after the reference start date.

4.3. Sensitivity Analysis

A sensitivity analysis for ORR stratified by PD-L1 expression per actual central/local lab results (CPS <1 vs 1-19 vs >= 20) will be conducted similarly for the primary analysis for ORR stratified by PD-L1 expression per IRT. For patients with PD-L1 results from both central and local labs, the central lab results will be used for this analysis.

5. ANALYSES SUPPORTING SECONDARY OBJECTIVE(S)

5.1. Key Secondary Endpoint(s)

5.1.1. Definition of Endpoint(s)

Secondary efficacy endpoints will include the following endpoints:

- **Duration of Response (DOR)** is defined as the time from the first determination of a confirmed response per RECIST v1.1 until the first documentation of progression or death, whichever occurs first. DOR is only summarized for patients with a confirmed response.
- **Clinical Benefit Rate (CBR)** is defined as the proportion of patients with a best overall response of a confirmed CR, a confirmed PR, or a durable stable disease (SD) (SD duration ≥ 24 weeks).
- **Disease Control Rate (DCR)** is defined as the proportion of patients with a best overall response of a confirmed CR, a confirmed PR, or an SD.
- **Progression-free Survival (PFS)** is defined as the time from the date of randomization to the date of the first documentation of PD assessed by the investigators per RECIST v1.1 or death, whichever occurs first.
- **Overall Survival (OS)** is defined as the time from the date of randomization to the date of death due to any cause.

5.1.2. Statistical Methods

The secondary endpoints will be summarized for each treatment arm. The difference between any experimental arm and the reference arm will be reported.

CBR and DCR

The CBR and DCR will be analyzed for the ITT analysis set. The corresponding two-sided 95% confidence interval (CI) calculated from Clopper-Pearson exact method will be also presented.

PFS

The distribution of PFS, including median, Q1 and Q3, and event-free rates at 3, 6, 9 and 12 months, will be estimated using the Kaplan-Meier method for each treatment group. Ninety-five percent CIs for median and Q1 and Q3 of PFS will be estimated using the method of Brookmeyer and Crowley (Brookmeyer and Crowley, 1982), and 95% CIs for progression-free survival rates will be estimated using Greenwood's formula (Greenwood, 1926).

PFS will be censored at the last adequate tumor assessment if one of the following occurs: absence of event; the event occurred after a new anticancer therapy is given; the event occurred after two or more consecutive missing tumor assessments. For the cases of missing baseline tumor assessment, an early death occurring within 13 weeks (i.e., 12 weeks for two tumor assessments + 1 week) from randomization will be considered a PFS event. Clinical or symptomatic progressions without supporting radiologic data will not be considered as PFS events. The event and censoring rules for derivation of PFS are presented in Table 1: Event and Censoring Rules for the Derivation of Progression-free Survival per RECIST Version 1.1Table 1.

The hazard ratio (HR) and its 2-sided 95% CI will be explored from a Cox regression model with treatment group as a factor stratified by the PD-L1 (CPS = 1-19 vs ≥ 20) per IRT.

Table 1: Event and Censoring Rules for the Derivation of Progression-free Survival per RECIST Version 1.1

No.	Situation	Date of Progression or Censoring	Outcome
1	No baseline and/or any post-baseline tumor assessments and without death within 13 weeks	Date of Randomization	Censored
2	No baseline and/or any post-baseline tumor assessments, but died within the two scheduled disease assessment windows from randomization date	Date of death	Event
3	Progression documented between/at scheduled visits	Date of first radiologic PD assessment	Event
4	No progression or death by the time of data cut-off or withdrawal from study or lost to follow up or other EOS reasons	Date of last adequate radiologic assessment prior to or on date of data cut-off or withdrawal from study or lost to follow up or other EOS reasons*	Censored
5	New anticancer treatment started prior to PD/death	Date of last adequate radiologic assessment prior to or on date of new anticancer treatment*	Censored

6	Death before first documented PD assessment	Date of death	Event
7	Progression or death after two or more consecutive missing visits**	Date of last adequate radiologic assessment before the missing tumor assessments or reference start date	Censored

*Adequate tumor assessment is a radiologic assessment of CR, PR, SD, non-CR/non-PD or PD as determined per RECIST 1.1.

** The algorithm to identify missing tumor assessments (TAs) are presented in Appendix 2.

DOR

The confirmed DOR will be analyzed in the confirmed responders in the ITT analysis set only. All the censoring rules for PFS in Table 1 will be applied to DOR. The distribution of DOR, including median, Q1, Q3, and DOR rates at 3, 6, 9 and 12 months, will be estimated using the Kaplan-Meier method for each treatment group. 95% CIs for median, Q1, and Q3 of DOR will be estimated using the method of Brookmeyer and Crowley (Brookmeyer and Crowley, 1982). And 95% CIs for DOR rates will be estimated using Greenwood's formula (Greenwood, 1926).

OS

The OS will be summarized in a similar way as the PFS. The event and censoring rules for the analysis of OS are presented in Table 2.

Table 2: Event and Censoring Rules for the Analysis of Overall Survival

Sequence	Situation	Date of Death Event or Censoring	Outcome
1	Death prior to/on the clinical cutoff date	Death date	Event
2	Alive by data cutoff	Last known alive date*	Censored
3	Lost to follow-up/withdrawal consent/discontinued from study due to other reasons	Last known alive date	Censored

*The last known alive date (LKAD) for patients who are still alive is defined as

- For patients who are still on treatment, LKAD is the clinical cutoff date
- For other patients, LKAD is the min (last available date showing patients alive, cutoff date).

6. PATIENT CHARACTERISTICS

6.1. Patient Disposition

The number (percentage) of patients enrolled, 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. The patients who discontinued treatment and the primary reason for the end of treatment will be summarized among patients who were treated.

A listing of patient disposition will be provided.

6.2. Demographic and Other Baseline Characteristics

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

- Age
- Age Group (<65 vs ≥65 years)
- Sex Race (Male vs Female)
- Ethnicity
- Country
- ECOG Performance Status at Baseline (0 vs 1)
- Height (cm)
- Weight (kg)
- BMI (kg/m^2)
- PD-L1 expression (CPS) per IRT (1 - 19 vs >= 20)
- Tabacco use (Current vs Former vs Never)

6.3. 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 the following variables:

- Primary location of the disease (Hypopharynx, Larynx, Oral cavity, Oropharynx, Other)
- Disease status at study entry (Recurrent, Metastatic, Recurrent and metastatic)
- AJCC stage at initial diagnosis
- AJCC stage at study entry
- MSI or MMR status (MSI-H/dMMR, MSI-L/MSS/pMMR, Unknown)

- HPV (Positive, Negative, Unknown)
- Prior Platinum (Yes, No)
- Time from initial diagnosis to randomization (months)
- Time from initial diagnosis of recurrent disease to randomization (months)
- Time from initial diagnosis of metastatic disease to randomization (months)
- Baseline sum of diameters of target lesions (mm)

7. 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 and laboratory values.

7.1. Extent of Exposure

The following measures of the extent of exposure will be summarized for each study drug:

- **Duration of exposure (months):** defined as (last date of exposure – first dose date + 1)/30.4375
    a) If patients discontinued treatment (with non-missing end-of-treatment (EOT) date), "last date of exposure" will be defined as min (cutoff date, study discontinuation date, death date, last dose date + 20).
    b) Otherwise for treatment ongoing patient, using cutoff date as the "last date of exposure"
- **Number of cycles received:** defined as total number of cycles is defined as the total number of cycles with non-missing doses
- **Cumulative dose administered (mg):** defined as the sum of all actual dosages per administration at all visits up to the cutoff date.
- **Actual dose intensity (ADI) (mg/cycle):** defined as 21* cumulative dose administered (mg) / (last dose date up to cutoff date+ 21 - first dose date)
- **Planed dose intensity:**
    - **Planned dose intensity (mg/cycle) for tislelizumab: 200 mg/cycle**
    - **Planned dose intensity (mg/cycle) for BGB-A425: 600 mg/cycle**
    - **Planned dose intensity (mg/cycle) for LBL-007: 600 mg/cycle**
- **Relative dose intensity (RDI) (%):** defined as the ratio of the actual dose intensity and the planned dose intensity*100 (%)

Patient level listings will be provided for all dosing records and for calculated summary statistics.

7.2. 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 27.0 or higher) 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 an onset date or a worsening in severity from baseline (pretreatment) on or after the first dose of study drug(s) and up to 30 days after the last dose of study drug(s) or initiation of new anticancer therapy, whichever occurs first. Summary tables will generally focus on those AEs that were treatment-

emergent (TE). All AEs, treatment-emergent or otherwise, will be presented in patient data listings.

An AE overview table will be provided, including the number and percentage of patients by treatment arm and in total with

- TEAEs
    - Treatment-related
        - Related to tislelizumab
        - Related to BGB-A425
        - Related to LBL-007
- TEAEs with Grade 3 or above
    - Treatment-related
        - Related to tislelizumab
        - Related to BGB-A425
        - Related to LBL-007
- Treatment-emergent serious adverse events (SAEs)
    - Treatment-related
        - Related to tislelizumab
        - Related to BGB-A425
        - Related to LBL-007
- TEAEs leading to death
    - Treatment-related
        - Related to tislelizumab
        - Related to BGB-A425
        - Related to LBL-007
- TEAEs that led to treatment discontinuation
    - Leading to tislelizumab discontinuation
    - Leading to BGB-A425 discontinuation
    - Leading to LBL-007 discontinuation
- TEAEs that lead to treatment modification: treatment modification for tislelizumab, BGB-A425 and LBL-007 includes dose interruption, dose delay and dose rate reduced.
    - Leading to tislelizumab modification
        - Dose interruption

        - Dose delay
        - Dose rate reduced
    - Leading to BGB-A425 modification
        - Dose interruption
        - Dose delay
        - Dose rate reduced
    - Leading to LBL-007 modification
        - Dose interruption
        - Dose delay
        - Dose rate reduced
    - Infusion-related reaction (IRR): IRR is defined as any TEAE checked by the investigator as an IRR in the eCRF IRR checkbox.
        - Grade 3 or above

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.

The incidence of TEAEs will be reported as the number (percentage) of patients with TEAEs by SOC, PT and the worst grade by treatment arms and in total. A patient will be counted only once by the highest severity grade within an SOC and/or PT, even if the patient experienced more than 1 TEAE within a specific SOC and/or PT. The number (percentage) of patients with the following TEAEs will be summarized by SOC and/or PT.

- All TEAEs
    - All TEAEs by SOC and PT
    - Grade 3 or above TEAEs by SOC and PT
- SAEs
    - All SAEs by SOC and PT
- TEAEs leading to death
    - TEAEs leading to death by PT
- TEAEs leading to treatment discontinuation
    - TEAEs leading to treatment discontinuation of any study drug by SOC and PT

An immune-mediated adverse event (imAE) overview table, including the number and percentage of patients with imAE, imAE with Grade 3 or above, imAE that led to death, serious imAE, imAE that led to treatment discontinuation, imAE that led to treatment modification, treated with systemic corticosteroids, treated with other immunosuppressant, and treated with hormone therapy will be provided.

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

7.3. Laboratory Values

Laboratory safety tests will be evaluated for selected parameters described in Table 3.

Laboratory parameters that are graded by NCI-CTCAE v5.0 will be summarized by shifts from baseline CTCAE grades to maximum post-baseline grades. The parameters with NCI-CTCAE grading in both high and low directions (e.g, glucose, potassium, sodium) will be summarized separately. The lab summaries will report lab assessment up to 30 days following study drug discontinuation or initiation of the first new systemic anticancer therapy, whichever occurs first.

Table 3: Serum Chemistry and Hematology Laboratory Tests

Serum Chemistry	**Hematology**
Alkaline phosphatase	Hemoglobin
Alanine aminotransferase	Platelet counts
Aspartate aminotransferase	White blood cell count
Albumin	Neutrophil
Total bilirubin	Lymphocyte
Calcium	
Creatine kinase	
Creatinine	
Glucose	
Potassium	
Sodium	
Troponin I	
Troponin T	

8. OTHER ANALYSES

Other analyses such as Pharmacokinetics/Pharmacodynamics and biomarkers analysis, may be conducted if appropriate and will be described in a separate analysis plan.

9. INTERIM ANALYSES

9.1. Futility Interim Monitoring

A futility monitoring process will be conducted when any of the experimental arms has randomized approximately 20 patients, and when these patients have been followed up for at least 2 post-baseline tumor assessments. Bayesian posterior probability (to be calculated with a noninformative prior such as Beta (0.5, 0.5) for ORR) will be used to evaluate clinical antitumor activity, and the decision criteria may be based on the difference in ORR between an experimental arm and the reference arm (ie, ΔORR). An experimental arm will be considered futile if the ORR in this group is unlikely to be better than the reference arm during the interim monitoring process (ie, Posterior Probability [ΔORR < 0] ≥ 90%). Enrollment in this experimental arm may be paused if the criterion for futility is met.

7.1 Safety stopping rules

To control the number of patients put at risk in the event of a major safety concern, a pause in enrollment will be triggered by unacceptable conditions such as any death due to reasons other than disease progression and at least possibly related to the study treatment.

The safety stopping boundaries for at least possibly treatment-related deaths is listed in Table 4**Error! Reference source not found.**, which is derived based on the 90% Bayesian posterior probability that the incidence rate of deaths at least possibly related to the study treatment is > 5%, calculated with a weak prior distribution Beta (0.005, 0.095).

Following the stopping boundaries, if the accumulating data in an arm suggest that the incidence rate of at least possibly treatment-related deaths is likely to exceed 5% with 90% probability, the enrollment to that specific arm will be paused, and a comprehensive evaluation of the safety data will be performed by the sponsor. Based on the totality of the data, recommendations, and benefit-risk evaluation, the sponsor will take appropriate safety actions to continue or terminate enrollment in the relevant treatment arm(s).

Table 4: Safety Stopping Boundaries for Number of Patients With at Least Possibly Treatment-related Deaths

Number of Patients in an Arm	Number of Patients With at Least Possibly Treatment-Related Deaths That Will Trigger Pause in Enrollment
1-2	≥ 1
3-11	≥ 2
12-23	≥ 3

24-36	≥ 4
37-40	≥ 5

10. CHANGES IN THE PLANNED ANALYSIS

Table 5 summarizes the major changes in the planned analyses from the statistical section of the study protocol, including the timing, rational and descriptions of the changes. The changes are all made before database lock and not based on any comparative data.

Table 5: Statistical Analysis Plan Changes

SAP version	Approval date	Change made from	Rationale of the change	Description of the change
1.0	This version	Protocol V2.0	This study is documented in the synoptic clinical study report	1. Remove the analysis for the prior and concomitant medications 2. Remove the section on statistical analysis for the exploratory endpoints/analyses described in the protocol. 3. Remove the analysis for vital signs and the descriptive summary for clinical laboratory and their changes from baselines.

11. SAMPLE SIZE CONSIDERATIONS

This study is designed to obtain efficacy, safety, PK, and immunogenicity data on tislelizumab in combination with novel agents compared to tislelizumab alone. No formal hypothesis testing will be performed in the efficacy evaluation as the study is not designed to make explicit power and Type I error considerations for a hypothesis test.

The initial randomization will use a 1:1 ratio between each of the experimental arms and the common reference arm and will be stratified by the PD-L1 expression (CPS score = 1 to 19 versus ≥ 20). A total of up to 40 patients per experimental arm may be enrolled in this phase of the study. The total sample size of the reference arm depends on the number of experimental arms and the timing of their enrollment. Randomization ratios may be modified to account for emerging safety data, efficacy data, and feasibility considerations.

The study is an activity-estimating study and not designed to make explicit power and Type I error considerations. Bayesian posterior probability may be used to facilitate the evaluation of whether an experimental arm would have a relatively sufficient antitumor activity.

12. 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.

Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JCJM, Kaasa S, Klee MC, Osoba D, Razavi D, Rofe PB, Schraub S, Sneeuw KCA, Sullivan M, Takeda F. The European Organisation for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. Journal of the National Cancer Institute 1993; 85: 365-376.

Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by European Organisation for Research and Treatment of Cancer, Brussels 2001.

APPENDIX 1. IMPUTATION OF MISSING OR PARTIALLY MISSING DATES

Please note:
The last known alive date is only based on complete dates without imputation. In general, all the imputed start date should be prior to/or by last known alive date.

1. Impute partial dates for concomitant medication

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 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 > min(death date, study discontinuation date), then set to min (death date, study discontinuation date)

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 > min (death date, cutoff date, concomitant medication end date), then set to min (death date, cutoff date, concomitant medication 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.

2. Impute partial dates for 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 > min(death date, study discontinuation date), then set to min (death date, study discontinuation date)

If year of the end date is missing or end date is completely missing, 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

3. Impute partial dates related to 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 etc.

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 prior systemic therapy for cancer, if imputed end date > randomization date – 14 days, then set to randomization date – 14 days.
- For prior radiotherapy/locoregional therapy, if imputed end date > randomization date, then set to randomization date – 1 day.

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.

4. Impute partial dates for subsequent anti-cancer therapy as collected in the post-treatment page (same rule applies to safety and efficacy flag)

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

cutoff date, end date of subsequent anti-cancer therapy, start/end date of the next subsequent therapy)

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

Note: if the imputed subsequent anti-cancer therapy date collected from CRF “post-treatment discontinuation anti-cancer systemic therapy” or “post-treatment discontinuation anti-cancer procedure” page is before the last dosing date, send data query.

APPENDIX 2. RULES FOR IDENTIFYING MISSING TUMOR ASSESSMENTS

Two or more consecutive missing visits will be identified as below.

1) Input scheduled TA visit list for each study
2) Identify last evaluable tumor assessment (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 (e.g., following the classification rule per the thresholds as depicted in Table 3).
    b. Otherwise, assign the scheduled visit number to --LPTADT_WK
3) Find the 2nd TA visit after LPTADT_WK (--LPTADT_WK_2)
    If LPTADT_WK_2+1wk < earliest of PD/death date, then censor PFS at the –LPTADT

Table 6 shows how to assign unscheduled TA to a schedule 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 44 for an unscheduled visit, it will be mapped to Week 42 TA since it is within the Threshold for Week 42. Assuming it is SD and the subsequent TA of the patient is PD after Week 58 (ie, 2 consecutive missing assessments), PFS will be censored at LPTADT (Week 44); had the PD occurred prior to/at Week 58, it would be counted as an PFS event.

Table 6: Scheduled tumor assessments with time window

Weeks	Scheduled week -1week	**Scheduled week (LPTADT_WK)**	Scheduled week+1week	Threshold
Every 6 weeks for the first year	Week 5	Week6	Week 7	Week 9
	Week 11	Week 12	Week 13	Week 15
	Week 17	Week 18	Week 19	Week 21
	Week 23	Week 24	Week 25	Week 27
	Week 29	Week 30	Week 31	Week 33
	Week 35	Week 36	Week 37	Week 39
	Week 41	Week 42	Week 43	Week 45
	Week 47	Week 48	Week 49	Week 54
Every 12 weeks afterwards	Week 59	Week 60	Week 61	Week 66
	Week 71	Week 72	Week 73	Week 78
	Week 83	Week 84	Week 85	Week 90
	Week 95	Week 96	Week 97	Week 102
	Week 107	Week 108	Week 109	Week 114
	Week 119	Week 120	Week 121	Week 126
	Week 131	Week 132	Week 133	Week 138
	Week 143	Week 144	Week 145	Week 150