STATISTICAL ANALYSIS PLAN

Study Title:	Open-Label, Global, Multicenter, Randomized, Phase 3 Study of Sacituzumab Govitecan Versus Docetaxel in Participants with Advanced or Metastatic Non-Small Cell Lung Cancer (NSCLC) With Progression on or After Platinum-Based Chemotherapy and Anti-PD-1/PD-L1 Immunotherapy
Name of Test Drug:	Sacituzumab Govitecan
Study Number:	GS-US-577-6153
Protocol Version:	Amendment 2
Protocol Date:	22 December 2022
Analysis Type:	Interim and Final Analyses
Analysis Plan Version:	1.0
Analysis Plan Date:	12 July 2023
Analysis Plan Author:	PPD

TABLE OF CONTENTS

TABLE OF CONTENTS 2
LIST OF IN TEXT TABLES 4
LIST OF IN-TEXT FIGURES 4
LIST OF ABBREVIATIONS 5
1. INTRODUCTION 6
1.1. Study Objectives and Endpoints 6
1.2. Study Design 7
1.3. Sample Size and Power 8
2. TYPE OF PLANNED ANALYSIS 9
2.1. Interim Analyses 9
2.1.1. DMC 9
2.2. Final Analysis 9
3. GENERAL CONSIDERATIONS FOR DATA ANALYSES 11
3.1. Analysis Sets 11
3.1.1. Intent-to-Treat Analysis Set 11
3.1.2. Safety Analysis Set 11
3.1.3. Pharmacokinetic Analysis Set 11
3.1.4. Immunogenicity Analysis Set 11
3.1.5. Quality of Life Analysis Set 12
3.2. Participant Grouping 12
3.3. Strata and Covariates 12
3.4. Multiple Comparisons 12
3.5. Data Handling Conventions and Transformations 13
3.6. Assessment of COVID-19 Impact 15
4. PROTOCOL DEVIATIONS 16
5. PARTICIPANT INFORMATION 17
5.1. Participant Enrollment and Disposition 17
5.2. Extent of Study Drug Exposure 17
5.2.1. Exposure to Study Drug 18
5.2.2. Relative Dose Intensity 18
5.3. Demographics and Baseline Characteristics 19
5.3.1. Demographics 19
5.3.2. Baseline Disease Characteristics 20
5.4. Medical History 20
5.4.1. Prior Anti-Cancer Therapy 20
5.4.2. Prior Radiation Therapy 20
5.4.3. Medical History Excluding Cancer 20
5.5. Prior and Concomitant Medications 20
6. EFFICACY ANALYSES 22
6.1. Primary Efficacy Endpoint 22
6.1.1. Definition of the Primary Efficacy Endpoint 22
6.1.2. Statistical Hypothesis for the Primary Efficacy Endpoint 22
6.1.3. Analysis of the Primary Efficacy Endpoint 22
6.2. Secondary Efficacy Endpoints 22
6.2.1. Definition of Secondary Efficacy Endpoints 23

6.2.2. Analysis Methods for Secondary Efficacy Endpoints 24
6.3. Subgroup Analyses 27
6.4. Other Secondary Endpoints 28
6.4.1. Definition of Other Secondary Endpoints 28
6.4.2. Analysis Methods for Other Secondary Endpoints 29
CCI [redacted]

7. SAFETY ANALYSES 32

7.1. Adverse Events and Deaths 32
7.1.1. Adverse Event Dictionary 32
7.1.2. Adverse Event Severity 32
7.1.3. Relationship of Adverse Events to Study Drug 32
7.1.4. Serious Adverse Events 32
7.1.5. Treatment-Emergent Adverse Events 32
7.1.6. Summaries of Adverse Events and Deaths 33
7.1.7. Additional Analysis of Adverse Events 34
7.2. Laboratory Evaluations 36
7.2.1. Summaries of Numeric Laboratory Results 37
7.2.2. Graded Laboratory Values 37
7.2.3. Liver-related Laboratory Evaluations 38
7.2.4. Shifts Relative to the Normal Range 38
7.3. Body Weight and Vital Signs 38
7.4. Other Safety Measures 38
7.4.1. Pregnancy 38
7.4.2. Post Study Anticancer Therapies 38

8. PHARMACOKINETIC AND IMMUNOGENICITY ANALYSES 39

8.1. Pharmacokinetic Analyses 39
8.2. Immunogenicity Analyses 39

9. BIOMARKER ANALYSES 40

10. METHODS FOR HANDLING MISSING DATA 41

11. REFERENCES 43

12. SOFTWARE 44

13. SAP REVISION 45

14. APPENDICES 46

Appendix 1. Study Schema 47
Appendix 2. Study Procedures Table 48
Appendix 3. NSCLC-SAQ 49
CCI [redacted]

LIST OF IN TEXT TABLES

Table 1. Summary of Interim and Final Analyses Strategy....10
Table 2. Censoring rules for PFS....25
Table 3. Censoring Rules for Time to First Deterioration in Shortness of Breath Domain as Measured by NSCLC-SAQ/ in NSCLC-SAQ Total Score....27
Table 4. Censoring Rules for DOR....28
CCI [redacted]
Table 6. Definitions of Adverse Events of Special Interest....35

LIST OF IN-TEXT FIGURES

Figure 1. Hierarchical Testing Procedures....13

LIST OF ABBREVIATIONS

AE	Adverse event
BPM	Beats per minute
CRF	Case report form
CRO	Contract research organization
FDA	Food and Drug Administration
HLT	High level term
HLGT	High level group term
HR	Hazard Ratio
ICH	International Conference on Harmonization
LLT	Lower level term
MedDRA	Medical dictionary for regulatory activities
PK	Pharmacokinetics
PT	Preferred term
Q1	First quartile
Q3	Third quartile
SAP	Statistical analysis plan
SOC	System organ class
TFLs	Tables, figures, and listings
WHO	World Health Organization

1. INTRODUCTION

This statistical analysis plan (SAP) describes the statistical analysis methods and data presentations to be used in tables, figures, and listings (TFLs) in the clinical study report (CSR) for Study GS-US-577-6153. This SAP is based on protocol amendment 2 dated 22 December 2022. The SAP will be finalized before interim analysis. Any changes made to the SAP after the final data base lock will be documented in the CSR.

1.1. Study Objectives and Endpoints

Primary Objective	Primary Endpoint
• To compare the overall survival (OS) of sacituzumab govitecan (SG) versus docetaxel.	• OS is defined as the time from the date of randomization until death due to any cause in the Intent-to-Treat (ITT) Analysis Set.
Secondary Objectives	**Secondary Endpoints**
• To compare the effect of SG versus docetaxel on the following:	
• Progression-free survival (PFS) as assessed by the investigator per Response Evaluation Criteria in Solid Tumors (RECIST) Version 1.1.	• PFS is defined as the time from the date of randomization until the date of objective disease progression or death (whichever comes first) as assessed by the investigator per RECIST Version 1.1.
• Objective response rate (ORR) as assessed by the investigator per RECIST Version 1.1.	• ORR is defined as the proportion of participants who achieve a complete response (CR) or PR that is confirmed at least 4 weeks later as assessed by the investigator per RECIST Version 1.1.
• Duration of response (DOR) as assessed by the investigator per RECIST Version 1.1.	• DOR is defined as the time from the first documentation of CR or PR to the earlier of the first documentation of PD or death from any cause (whichever comes first) as assessed by the investigator per RECIST Version 1.1.
• Disease control rate (DCR) as assessed by the investigator per RECIST Version 1.1.	• DCR is defined as the proportion of participants who achieve a CR, PR, or stable disease (SD) as assessed by the investigator per RECIST Version 1.1.
• Safety and tolerability.	• Incidence of treatment-emergent adverse events (TEAEs) and clinical laboratory abnormalities.
• Quality of life (QOL) using NSCLC Symptom Assessment Questionnaire (NSCLC-SAQ).	• Time to first deterioration in shortness of breath domain as measured by NSCLC-SAQ • Time to first deterioration in NSCLC-SAQ total score

1.2. Study Design

Study GS-US-577-6153 is an open-label, global, multicenter, randomized, Phase 3 study to compare the efficacy and safety of SG versus docetaxel in participants with advanced or metastatic NSCLC with progression on or after platinum-based chemotherapy and anti-PD-1/PD-L1 immunotherapy received either in combination or sequentially. Participants with actionable genomic alterations will also be included if they have received prior treatment with an appropriate tyrosine kinase inhibitor.

This study will be conducted in approximately 250 centers globally. Participation will include screening, randomization, treatment, and follow-up. Screening will last no longer than 28 days to confirm eligibility and establish disease characteristics prior to randomization and treatment. The study schema and study procedures table are presented in Appendix 1 and Appendix 2, respectively.

Approximately 580 eligible participants will be randomly assigned in a 1:1 ratio to receive either SG (Investigational Arm A) or docetaxel (Control Arm B). Randomization will be stratified based on histology (squamous vs nonsquamous), response to last prior immune therapy received (best response PD/SD vs CR/PR on immune therapy), and if they have received prior therapy for actionable genomic alteration (yes vs no).

SG will be administered at 10 mg/kg via IV infusion on Days 1 and 8 of a 21-day cycle. Docetaxel will be administered at 75 mg/m^2 via IV infusion on Day 1 of a 21-day cycle. Participants will receive study drug until PD, death, unacceptable toxicity, or another treatment discontinuation criterion is met. Follow-up will begin at the time of the completion of the end of treatment (EOT) visit, which will occur 30 days (± 7) after the last dose of study drug. All participants will be followed for survival until 1 of the early discontinuation criteria from the study is met or completion of survival follow-up, whichever occurs first.

An independent data monitoring committee (DMC) will be convened at regular intervals to assess the progress of this study, review safety data, and conduct the interim efficacy analysis.

1.3. Sample Size and Power

Approximately 580 eligible participants will be randomly assigned in a 1:1 ratio to receive either SG or docetaxel over a planned nonuniform accrual period of 19 months.

The sample size is estimated based on the primary endpoint of OS. The assumed OS treatment effect under the alternative hypothesis is an average HR of 0.70 for SG versus docetaxel, which is characterized by a median OS of 11 months on docetaxel treatment and a median OS of 15.7 months on SG treatment assuming OS is exponentially distributed. A total of 336 death events are required to detect a statistically significant difference at overall 1-sided alpha of 2.5% with 90% power based on a log-rank test. Assuming 2% annual drop off rate for OS, it is estimated that approximately 580 participants will provide 336 death events after approximately 29 months of survival follow-up after the first participant is randomized.

EAST® 6.5 was used to calculate the sample size.

2. TYPE OF PLANNED ANALYSIS

The study has 1 planned interim analysis and 1 final analysis. Table 1 below summarize the strategy for interim and final analyses. Note the efficacy boundary will be updated based on the actual number of events occurred at IA.

2.1. Interim Analyses

There will be 1 planned interim superiority analysis and 1 final analysis based on the primary endpoint of OS for this study.

The superiority interim efficacy analysis will be performed when approximately 242 death events (72% of the final target OS events) have occurred across the SG and docetaxel treatment arms. This is predicted to occur approximately 23 months after the first participant is randomized.

The Lan-DeMets spending function that approximates an O'Brien Fleming approach will be used to account for multiplicity introduced by including an interim analysis for superiority. The alpha level applied at the interim depends upon the proportion of information available. If 242 death events (72% of the target 336 events with 42% maturity) have occurred at the time of the analysis, the 1-sided significance level to be applied for the OS at the interim and final analysis would be 0.82% and 2.25%.

2.1.1. DMC

To ensure the integrity of the study and to oversee the planned interim analyses, an independent DMC, consisting of therapeutic area experts and biostatisticians who are not employed by Gilead, are not investigators for this study, and do not have any major conflict of interest, will be convened at regular intervals to assess the progress of this study, review safety data, and conduct the interim efficacy analysis.

The planned analyses for DMC meetings and communication process will be provided in the DMC charter.

2.2. Final Analysis

The final analysis of the primary endpoint of OS will be conducted in the Intent-to-Treat (ITT) Analysis Set when approximately 336 death events have occurred across the treatment arms (336/580 events with 58% maturity), which is projected to be approximately 29 months of survival follow-up after the first participant is randomized.

Table 1. Summary of Interim and Final Analyses Strategy

Analysis	Key Endpoint	Expected # of Events (Information fraction)	Estimated Timing after First Participant Randomized	Primary Purpose of Analysis	Stopping Boundary (1-sided alpha)
IA	OS	~ 242 death events (72% of the final target OS events)	~23 months	Superiority	HR $\leq$ 0.734 ($p \leq 0.0082$)
FA	OS	~ 336 death events	~29 months	Superiority	HR $\leq$ 0.803 ($p \leq 0.0225$)

3. GENERAL CONSIDERATIONS FOR DATA ANALYSES

Analysis results will be presented using descriptive statistics. For categorical variables, the number and percentage of participants in each category will be presented; for continuous variables, the number of participants (n), mean, standard deviation (StD) or standard error (SE), median, first quartile (Q1), third quartile (Q3), minimum, and maximum will be presented.

By-participant listings will be presented for all participants in the Intent-to-Treat (ITT) Analysis Set and sorted by participant ID number, visit date, and time (if applicable). Data collected on log forms, such as AEs, will be presented in chronological order within the participant. The treatment arm to which participants were randomized will be used in the listings. Age, sex at birth, race, and ethnicity will be included in the listings, as space permits.

3.1. Analysis Sets

Analysis sets define the participants to be included in an analysis. Analysis sets and their definitions are provided in this section. The analysis set will be identified and included as a subtitle of each table, figure, and listing.

For each analysis set, the number and percentage of participants eligible for inclusion, as well as the number and percentage of participants who were excluded and the reasons for their exclusion, will be summarized by treatment arm.

A listing of reasons for exclusion from analysis sets will be provided by participant.

3.1.1. Intent-to-Treat Analysis Set

The ITT Analysis Set will include all randomized participants according to the treatment arm to which the participant is randomized. To be randomized, subjects need to be enrolled first, which includes all participants who signed informed consent form and received a study participant identification number in the study after screening.

This is the primary analysis set for all efficacy analyses.

3.1.2. Safety Analysis Set

The Safety Analysis Set will include all participants who received at least 1 dose of any study drug, with treatment assignments designated according to the actual treatment received. This is the primary analysis set for all safety analyses.

3.1.3. Pharmacokinetic Analysis Set

The PK analysis will be conducted on the Pharmacokinetic Analysis Set, defined as all randomized participants who received at least 1 dose of SG per the protocol and have at least 1 evaluable posttreatment serum concentration of SG, total SN-38, free SN-38, or total antibody (hRS7 IgG).

3.1.4. Immunogenicity Analysis Set

The immunogenicity analysis will be conducted on the Immunogenicity Analysis Set, defined as all randomized participants who received at least 1 dose of SG and have at least 1 evaluable posttreatment anti-SG antibody test result.

3.1.5. Quality of Life Analysis Set

Quality of life assessments will be analyzed for participants in the ITT Analysis Set.

3.2. Participant Grouping

For analyses based on the ITT Analysis Set, participants will be grouped according to the treatment to which they were randomized.

For analyses based on the Safety Analysis Set, participants will be grouped according to the actual treatment received.

For the PK Analysis Set, participants will be grouped according to the actual treatment they received.

3.3. Strata and Covariates

Participants will be randomly assigned to treatment arms via the interactive web response system (IWRS) in a 1:1 ratio using a stratified randomization schedule. Stratification will be based on the following variables:

- Histology (squamous vs non-squamous)
- Response to last prior immune therapy received (best response PD/SD vs CR/PR on immune therapy)
- Received prior therapy for actionable genomic alteration (yes vs no)

If there are discrepancies in stratification factor values between the IWRS and the clinical database, the values recorded in the clinical database will be used for analyses.

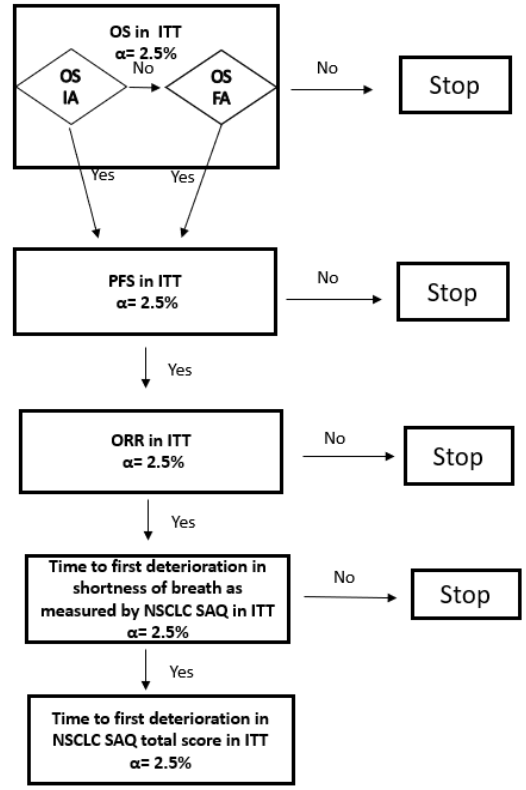
Based on a blinded review of OS event counts by stratum prior to the first efficacy interim analysis, if there are <10 event counts in one or more strata, stratification factors will be combined for analysis to ensure sufficient number of events in each stratum. Specifically, prior therapy for actionable genomic alteration (yes vs no) is collapsed, leaving stratification by histology and response to last prior immune therapy received.

3.4. Multiple Comparisons

The overall Type I error rate for this study is strictly controlled at a 1-sided alpha of 2.5%. A Lan DeMets alpha spending function that approximates an O'Brien Fleming approach will be used to account for multiplicity introduced by including an interim analysis for superiority. The OS efficacy interim analysis will be tested at the 1-sided significance level of 0. 82% if 72% of the death events (242/336) is available at the time of the analysis. If the OS interim analysis is not positive, the OS final analysis will be tested at the 1-sided significance level of 2.25%. Note that alpha levels for the OS interim and final analysis are based on the actual observed events and will be adjusted accordingly.

If the primary OS analysis is positive at either interim or final, the final analysis based on the full dataset of the key secondary efficacy endpoints of PFS, ORR, time to first deterioration in shortness of breath as measured by NSCLC SAQ and time to first deterioration in NSCLC SAQ total score will be formally tested sequentially at the 1-sided alpha of 2.5% when the above hypotheses in the hierarchy were also statistically significant (Figure 1). All of the secondary efficacy endpoints will be mature at the time of the OS interim analysis.

Figure 1. Hierarchical Testing Procedures



FA = final analysis; IA = interim analysis; ITT = intent to treat; NSCLC-SAQ = non-small cell lung cancer Symptom Assessment Questionnaire; ORR = objective response rate; OS = overall survival; PFS = progression-free survival

3.5. Data Handling Conventions and Transformations

Baseline value is defined as the last measurement that was observed on or prior to the date of randomization, unless otherwise specified.

The following conventions will be used for the imputation of date of birth when it is partially missing or not collected:

- If only month and year of birth is collected, then "15" will be imputed as the day of birth
- If only year of birth is collected, then "01 July" will be imputed as the day and month of birth
- If year of birth is missing, then date of birth will not be imputed

In general, age collected at Day 1 (the randomization date) (in years) will be used for analyses and presented in listings. If age at Day 1 is not available for a participant, then age derived based on date of birth and the Day 1 visit date will be used instead. If an enrolled participant was not dosed with any study drug, the randomization date will be used instead of the Day 1 visit date. For screen failures, the date the first informed consent was signed will be used for the age derivation.

Non-PK data that are continuous in nature but are less than the lower limit of quantitation (LOQ) or above the upper LOQ will be imputed as follows:

- A value that is 1 unit less than the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "< x" (where x is considered the LOQ). For example, if the values are reported as < 50 and < 5.0, values of 49 and 4.9, respectively, will be used to calculate summary statistics. An exception to this rule is any value reported as < 1 or < 0.1, etc. For values reported as < 1 or < 0.1, a value of 0.9 or 0.09, respectively, will be used to calculate summary statistics.

- A value that is 1 unit above the LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "> x" (where x is considered the LOQ). Values with decimal points will follow the same logic as above.

- The LOQ will be used to calculate descriptive statistics if the datum is reported in the form of "≤ x" or "≥ x" (where x is considered the LOQ).

Sparse PK concentration values that are below the limit of quantitation (BLQ) will be presented as "BLQ" in the data listing.

Natural logarithm transformation will be used for analyzing concentrations and PK parameters in intensive PK samples. Concentration values that are BLQ will be presented as "BLQ" in the concentration data listing. Values that are BLQ will be treated as 0 at predose time points, and one-half the value of the LOQ at postdose time points for summary purposes. The following conventions will be used for the presentation of summary and order statistics for intensive PK concentrations:

- If at least 1 participant has a concentration value of BLQ for the time point, the minimum value will be displayed as "BLQ."

- If more than 25% of the participants have a concentration data value of BLQ for a given time point, the minimum and Q1 values will be displayed as "BLQ."

- If more than 50% of the participants have a concentration data value of BLQ for a given time point, the minimum, Q1, and median values will be displayed as "BLQ."

- If more than 75% of the participants have a concentration data value of BLQ for a given time point, the minimum, Q1, median, and Q3 values will be displayed as "BLQ."

- If all participants have concentration data values of BLQ for a given time point, all order statistics (minimum, Q1, median, Q3, and maximum) will be displayed as "BLQ." PK parameters that are BLQ will be imputed as one-half LOQ before log transformation or statistical model fitting.

3.6. Assessment of COVID-19 Impact

This study was ongoing during the novel coronavirus (COVID-19) pandemic which has an impact on the study conduct. Some participants were unable to attend onsite visits due to shelter in place guidelines, site closures, or other reasons. The following special situations due to COVID-19 will be handled in the analysis:

- Study treatment or study discontinuation due to COVID-19
- Protocol deviations due to COVID-19
- Adverse events (AEs) due to COVID-19

The counts and percentages will be presented by treatment arm.

4. PROTOCOL DEVIATIONS

Participants who did not meet the eligibility criteria for study entry but enrolled in the study, will be summarized regardless of whether they were exempted by the sponsor or not. The summary will present the number and percentage of participants who did not meet at least 1 eligibility criterion and the number of participants who did not meet specific criteria by treatment arm based on the ITT Analysis Set. A by-participant listing will be provided for those participants who did not meet at least 1 eligibility (inclusion or exclusion) criterion. The listing will present the eligibility criterion (or criteria if more than 1 deviation) that participants did not meet and related comments, if collected.

Protocol deviations occurring after participants entered the study are documented during routine monitoring. The number and percentage of participants with important protocol deviations by deviation reason (e.g., nonadherence to study drug, violation of select inclusion/exclusion criteria) will be summarized by treatment arm for the ITT Analysis Set. A by-participant listing will be provided for those participants with important protocol deviation.

5. PARTICIPANT INFORMATION

5.1. Participant Enrollment and Disposition

A summary of participant enrollment will be provided by treatment arm for each country, investigator within a country, and overall. The summary will present the number and percentage of participants enrolled. For each column, the denominator for the percentage calculation will be the total number of participants analyzed for that column.

A similar enrollment table will be provided by randomization stratum. The denominator for the percentage of participants in the stratum will be the total number of enrolled participants. If there are discrepancies in the value used for stratification assignment between the IWRS and the clinical database, the value collected in the clinical database will be used for the summary. A listing of participants with discrepancies in the value used for stratification assignment between the IWRS and the clinical database at the time of data finalization will be provided.

The randomization schedule used for the study will be provided as an appendix to the CSR.

A summary of participant disposition will be provided by treatment arm. This summary will present the number of participants screened, the number of participants randomized, the number of participants randomized but not treated, and the number of participants in each of the categories listed below:

- ITT Analysis Set
- Safety Analysis Set
- Continuing study drug
- Discontinued study drug with reasons
- Ongoing in study (including long term OS follow-up)
- Discontinued study with reasons

The following by-participant listings will be provided by participant identification (ID) number in ascending order to support the above summary tables:

- Reasons for study drug discontinuation, or study discontinuation
- Reasons for screen failure (will be provided by screening ID number in ascending order)
- Lot number and kit ID

5.2. Extent of Study Drug Exposure

The extent of exposure to study drug will be examined by assessing the total duration of exposure to study drug, the total number of doses received, the total number of treatment cycles received for each study drug, and total cumulative dose administered (mg/kg for SG and mg/m^2 for docetaxel), relative dose intensity (%), and number (%) of participants with dose modifications (i.e., dose delay, dose reduction, or dose interruption) due to adverse events or other reasons.

5.2.1. Exposure to Study Drug

Total duration of exposure to study drug will be defined as last dosing date minus first dosing date plus 1, regardless of any temporary interruptions in study drug administration, and will be expressed in months using up to 1 decimal place (e.g., 4.5 months).

The total duration of exposure to study drug will be summarized using descriptive statistics and using the number and percentage of participants exposed through the following time periods: 3 months, 6 months, 12 months, and 24 months etc., if applicable.

The number of doses participants were exposed to study drug as well as the number of cycles participants were exposed to study drug will be summarized using descriptive statistics.

Number and percentage of participants who have dose reductions and the reasons for dose reduction, number and percentage of participants with total, 0, 1, or 2 dose reductions, time to the first dose reduction, number and percentage of participants who had dose interruptions will be summarized by treatment arm.

A listing of study drug administration will be provided by participant ID number (in ascending order) and visit (in chronological order), including dosing date/time, planned dosage, actual dosage administered, infusion outcome and reason for dose delay or dose interruption (if applicable).

5.2.2. Relative Dose Intensity

Relative dose intensity (RDI) is the percentage of the total amount of study drug administered relative to the total amount of study drug expected to be administered during a participant's actual on-treatment period based on the study drug regimen.

RDI for Docetaxel = (Total Amount of Study Drug Received (in mg/m2) / Total Amount of Study Drug Planned by Protocol (in mg/m2)) x 100

RDI for SG = (Total Amount of Study Drug Received (in mg/kg) / Total Amount of Study Drug Planned by Protocol (in mg/kg)) x 100

Specifically for SG as an example,

Delivered dose (mg) for each infusion is as recorded in CRF Study Drug Administration page.

Delivered dosage (mg/kg) of each infusion in a cycle is calculated by dividing the delivered dose (in mg) by body weight (in kg) at the beginning of the cycle (the body weight according to which the prescribed dose is calculated and prepared per the Protocol).

Total amount of study drug received (mg/kg) for each subject is defined as the sum of all delivered dosages (mg/kg) of all infusions the subject received in the study.

Total amount of study drug planned by protocol (mg/kg) for each subject is defined as the product of the assigned dose of SG (10 mg/kg) and number of doses the subject was scheduled to receive during the subject's treatment period.

Example:

- Docetaxel 75 mg/m2 via IV infusion on Day 1 of a 21-day cycle (i.e., once every 3 weeks)

	C1 D1	C2 D1	C3 D1	C4 D1
BSA calculated based on height and weight at beginning of each cycle	2.4 m2	2.4 m2	2.1 m2	2.1 m2
Received	180 mg	0 mg (missed)	180 mg	150 mg

RDI = (180/2.4 + 0/2.4 + 180/2.1 + 150/2.1) / (75 * 4) = 77.4%

Where the body surface area (BSA) is based on the BSA calculated at the beginning of each cycle, and the denominator is based on the protocol originally planned dose.

- SG 10 mg/kg via IV infusion on Days 1 and 8 of a 21-day cycle (i.e., 2 weekly doses plus 1 week without treatment)

	C1 D1	C1 D8	C2 D1	C2 D8
Weight	70 kg		65 kg	
Received	700 mg	700 mg	0 mg (missed)	465 mg

RDI = (700/70 + 700/70 + 0/65 + 465/65) / (10 * 4) = 67.9%

Where weight is measured at the beginning of each cycle, and the denominator is based on the protocol originally planned dose.

Descriptive statistics for relative dose intensity with the number and percentage of participants belonging to categories (e.g., < 70%, ≥ 70 to < 90%, ≥ 90% to < 110%, ≥ 110%) will be provided by treatment arm using the Safety Analysis Set.

5.3. Demographics and Baseline Characteristics

5.3.1. Demographics

Participant demographic variables (i.e., age, sex, race, ethnicity, and region) and baseline characteristics (ECOG, stratification factors collected in clinical database) will be summarized by treatment arm and overall using descriptive statistics for continuous variables and using number and percentage of participants for categorical variables. The summary of demographic data will be provided for the ITT Analysis Set.

A by-participant demographic listing, including the informed consent date, will be provided by participant ID number in ascending order.

5.3.2. Baseline Disease Characteristics

Summaries of baseline disease characteristics will include, but not be limited to, the following variables: disease stage at screening, histology, time from initial diagnosis to date of randomization, time from end of last therapy to date of randomization, time from metastatic disease diagnosis to date of randomization, metastasis location/differentiation, PDL-1 score at baseline, UGT1A1 status are relevant to SG arm and will be summarized as well. The summary of these baseline characteristics will be provided for the ITT Analysis Set.

A by-participant listing of other baseline characteristics will be provided by participant ID number in ascending order.

5.4. Medical History

5.4.1. Prior Anti-Cancer Therapy

Summary of prior anticancer therapy will include the following: number of prior lines of therapies, treatment setting (neoadjuvant, adjuvant, advanced/metastatic), best response for the last therapy before entering study.

Prior anticancer therapy will be summarized by treatment setting, line of therapy and regimen in the ITT population.

By-participant listing will be generated according to information as collected from CRFs.

5.4.2. Prior Radiation Therapy

Prior radiation therapy will include anatomical site, duration of therapy, time from last therapy to randomization date. The last therapy refers to the therapy whose end date is the last before randomization date.

By-participant listing will be generated according to information as collected from CRFs.

5.4.3. Medical History Excluding Cancer

Medical history excluding cancer will be summarized for the ITT population and listed by participant for each treatment arm. Medical history excluding cancer will be summarized by system organ class (SOC) and preferred term (PT) and sorted by frequency in SOC and by decreasing frequency in PT.

5.5. Prior and Concomitant Medications

Prior and concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary.

Prior medications include medication with a start and end date prior to the first dose of study drug.

Concomitant medications include medications that were taken at any time while on study treatment, including medications that were started before the first administration of study drug but ongoing at the time of first study drug, or that were initiated on or after first administration but prior to 30 days after last administration of study drug. If an end date is missing or the medication is ongoing during study treatment, the medication will be included as concomitant medication.

Prior medications and concomitant medications will be summarized according to WHO Drug ATC Classification.

Prior and concomitant medications listings will be provided using the ITT Analysis Set.

6. EFFICACY ANALYSES

6.1. Primary Efficacy Endpoint

6.1.1. Definition of the Primary Efficacy Endpoint

The primary efficacy endpoint is overall survival, defined as the interval from the date of randomization to death from any cause. For participants who were not known to have died at the time of the analysis, OS data will be censored at the last date that they were known to be alive. In the event that the participant has withdrawn consent, the vital status of the participant can be obtained by site personnel from publicly available resources under applicable local laws.

6.1.2. Statistical Hypothesis for the Primary Efficacy Endpoint

The null hypothesis to be tested is that SG (treatment arm) is not superior to docetaxel (control arm) in OS. Using $S_T(t)$ and $S_C(t)$ to denote the OS survival distribution functions of SG and docetaxel, respectively, the statistical hypotheses to be tested for OS will be:

H_0: $S_T(t) \leq S_C(t)$ at all time points t

H_1: $S_T(t) > S_C(t)$ (SG is superior to docetaxel in terms of OS) for some $t > 0$

6.1.3. Analysis of the Primary Efficacy Endpoint

The primary analysis of OS will be performed in the ITT population using the stratified log-rank test, stratified by the stratification factors collected in clinical database. Medians, Q1, Q3 and the proportion of participants who are alive at 6 and 12 months from randomization will be provided based on Kaplan-Meier method. The 95% CI of median will be calculated based on the Brookmeyer and Crowley method. Kaplan-Meier curves will be provided by treatment arm.

In addition, the treatment effect will be assessed by HR along with its 95% CI using the Cox proportional hazards regression model with a single treatment covariate stratified by the stratification factors, with tie handled using Efron's method.

The follow-up time for OS, defined as the interval from date of randomization to the last known alive date for participants who do not have a death record, or death date for those who are dead, as of data cutoff date will be summarized by treatment arms using descriptive statistics.

6.2. Secondary Efficacy Endpoints

There are 4 key secondary efficacy endpoints in the hierarchical testing structure as specified in Section 3.4, in addition to the primary efficacy endpoint. The 4 key secondary efficacy endpoints are PFS, ORR, time to first deterioration in shortness of breath domain as measured by NSCLC-SAQ and time to first deterioration in NSCLC-SAQ total score.

6.2.1. Definition of Secondary Efficacy Endpoints

Secondary efficacy endpoints include:

6.2.1.1. PFS

PFS is defined as the time interval from date of randomization to date of the first documentation of definitive disease progression or date of death from any cause, whichever occurs first. Definitive disease progression is determined based on Response Evaluation Criteria in Solid Tumors (RECIST) Version 1.1, as defined in Protocol Amendment 2 Appendix 6. The progression should be identified by relevant radiographic, imaging assessments by the investigator. Participants without progression or death will be censored at the last adequate post-baseline tumor assessment date.

If participants received a new anticancer therapy or have ≥ 2 consecutive missing postbaseline tumor assessments immediately before documented progression or death, they will be censored at the last adequate postbaseline tumor assessment time before starting the new anticancer therapy or before the ≥ 2 consecutive missing postbaseline tumor assessments, whichever is earlier. If participants do not have an adequate baseline tumor assessment or adequate postbaseline tumor assessments, they will be censored on the date of randomization unless participants died before or on the second scheduled postbaseline tumor assessment without receiving anticancer therapy.

The adequate tumor assessments for determining censoring dates include complete response (CR), partial response (PR), stable disease (SD) and progression (PD). Non-evaluable (NE) is considered as not adequate. The censoring rules for primary analyses are summarized in Table 2 in Section 6.2.2.

When the date of initiation of anticancer therapy other than the study drug is incomplete or missing, the following algorithm will be followed:

- If the day is missing but the month and year are available, then the imputed day will be the first day of the month.
- If day and month are missing but year is available, then the imputed day and month will be 01 Jan or the last day of the month for the last adequate disease assessment if they have the same year, whichever is later.

Every attempt will be made to ensure that complete death dates are recorded.

6.2.1.2. ORR

ORR is defined as the proportion of ITT participants who achieve complete response (CR) or partial response (PR) that is confirmed at least 4 weeks after initial documentation of response as assessed by the investigator per RECIST Version 1.1. Tumor response assessments after the date of participants receiving new anticancer therapy will be excluded from the analysis. Participants, who do not have sufficient baseline or on-study tumor status information is not adequate for response status assessment or received new anticancer therapy prior to achieving CR or PR, will be considered as non-responders.

6.2.1.3. Time to Response

Time to response (TTR) will be calculated as the time from date of randomization to date of first documented response (CR or PR). Time to response is calculated for participants who achieved a CR or PR.

6.2.1.4. Time to First Deterioration (TTD) in Shortness of Breath Domain as Measured by NSCLC-SAQ and Time to First Deterioration in NSCLC-SAQ Total Score

Participant-reported disease-related symptoms will be assessed by the NSCLC-SAQ, a 7-item PRO instrument covering 5 symptom concepts of NSCLC: cough, pain, dyspnea, fatigue, and appetite. Each item uses a 5-point verbal rating scale from either "No <symptom> at All" to "Very severe <symptom>" or from "Never to Always," depending on the item's question structure relative to either intensity or frequency. These correspond to scores of 0, 1, 2, 3, and 4, respectively. The scoring method for each domain score and total score is described below.

- Cough Domain Score: score of the cough item, or missing if skipped
- Fatigue Domain Score: if both items are present, compute the mean; use the score from 1 item if the other is missing; or set to missing if both are skipped
- Pain Domain Score: if both items are present, use the most severe of both; use the score from 1 item if the other is missing; or set to missing if both are skipped
- Dyspnea/Shortness of Breath Domain Score: score of the shortness of breath item, or missing if skipped
- Appetite Domain Score: score of the poor appetite item, or missing if skipped
- NSCLC-SAQ Total Score: a total score that sums all five domain scores at each visit. If any domain score is missing, the score is not computed. This total score ranges between 0 and 20, with higher scores indicating more severe symptoms.

The NSCLC-SAQ will be assessed on Day 1 prior to dosing of every cycle for the duration that participants receive SG or docetaxel treatment until and at the end of treatment (EOT) visit.

6.2.2. Analysis Methods for Secondary Efficacy Endpoints

6.2.2.1. PFS

The censoring rules of the primary and sensitivity analysis for PFS are summarized in Table 2 below.

6.2.2.1.1. Primary Analysis for PFS

The primary analysis of PFS will compare the PFS distributions of 2 treatment arms using the stratified log-rank test, stratified by the stratification factors at randomization using the ITT Analysis Set. Median PFS will be derived by Kaplan-Meier estimates and 95% confidence interval will be calculated based on Brookmeyer and Crowley method. The proportion of participants who are progression-free at 3 and 6 months from randomization will be derived from the KM curve.

In addition, the HR between the 2 treatment arms and its 95% CI will be estimated using the Cox proportional hazards regression model with a single treatment covariate, stratified by the stratification factors at randomization. The analysis strategy for the situation where there is insufficient information in a stratum is detailed in Section 3.3.

6.2.2.1.2. Sensitivity Analysis for PFS

To evaluate the robustness of the primary PFS analysis, sensitivity Analysis of PFS will use the same censoring rule of the primary PFS definition except that objective documented progression or death are counted as events regardless of missed study visits or initiation of new anticancer therapy.

Table 2. Censoring rules for PFS

Case	Outcome	Date of Event/Censoring	
		Primary Analysis (PA)	**Sensitivity Analysis (SA)**
No adequate response assessment after randomization			
No baseline tumor assessment or alive	Censored	Randomization Date	Randomization Date
Died prior to second scheduled assessment without initiation of other anti-cancer therapy	Event	Death Date	Death Date
Died immediately after ≥ 2 consecutive missing or new anti-cancer therapy, if any	PA: Censored SA: Event	Randomization Date	Death date
Continued scheduled response assessments until objective PD or death			
PD or death documented after ≤1 missed disease assessments, and before other anti-cancer therapy, if any	Event	PD or death date, whichever occurs first	PD or death date, whichever occurs first
PD or death immediately after ≥ 2 consecutive missed disease assessments, or after other anti-cancer therapy, if any	PA: Censored SA: Event	Date of last adequate tumor assessment prior to the earlier date of ≥ 2 consecutive missed disease assessments and other anti-cancer therapy, if any	Death date or PD date, whichever occurs first
Continued scheduled response assessments without objective PD or death			
Other anticancer treatment is initiated	Censored	Date of last adequate tumor assessment before starting other anticancer treatment	Date of last adequate tumor assessment
Other anticancer treatment is not initiated	Censored	Date of last adequate tumor assessment	Date of last adequate tumor assessment

6.2.2.2. ORR

The ORR as assessed by the investigator will be compared between the 2 treatment arms using the Cochran Mantel-Haenszel test adjusted by the stratification factors using the ITT analysis set. The ORR and the corresponding 95% CI based on Clopper-Pearson method of each treatment arm will be presented. The difference in ORR between SG and docetaxel and its 95% CI will also be reported.

6.2.2.3. Time to Response

TTR will be summarized with descriptive statistics only for confirmed responders (PR or better) based on disease assessments conducted prior to or at the initiation of the next line of anticancer therapy.

TTR (months) = (Date of response - date of randomization + 1)/30.4375

6.2.2.4. Time to First Deterioration (TTD) in Shortness of Breath Domain as Measured by NSCLC-SAQ and Time to First Deterioration in NSCLC-SAQ Total Score

For shortness of breath domain, a 1-point or greater worsening from baseline represents a clinically meaningful deterioration. For NSCLC-SAQ total score, a 2-point or greater worsening from baseline represents a clinically meaningful deterioration.

6.2.2.4.1. Primary Analysis for TTD

The time to first deterioration will be calculated as the time from date of randomization to the time a participant experienced change from baseline ≥ the pre-specified threshold value for deterioration. Participants whose baseline scores are so poor that attaining a score change exceeding or equaling the pre-specified threshold for deterioration is impossible will be censored on the date of randomization. For shortness of breath domain, baseline score of 4 is considered as a poor baseline. For the NSCLC-SAQ total score, baseline total score of 19 or 20 is considered a poor baseline. Participants who never experience a meaningful deterioration will be censored at the time of the last non-missing assessment, and death will be considered as an event (Table 3 for censoring rules for the time to first deterioration in shortness of breath domain and time to first deterioration in total score).

Table 3. Censoring Rules for Time to First Deterioration in Shortness of Breath Domain as Measured by NSCLC-SAQ/ in NSCLC-SAQ Total Score

Situation	Date of Deterioration or Censoring	Situation Outcome
Change from baseline ≥ the pre-specified threshold for worsening on at least one assessment visit before two missing consecutive assessment visits	Date of first assessment visit where change from baseline >= pre-specified threshold value for worsening	Event
No worsening from baseline ≥ the pre-specified threshold for worsening at any assessment visit	Date of last assessment with a non-missing score	Censored
Intermittent missing scores for ≥ two consecutive assessment visits	Date of last assessment with a non-missing score prior to the >= 2 consecutive missing assessment visits	Censored
Death*	Date of death	Event
No baseline score and/or post-baseline score	Date of randomization	Censored
Baseline score that was too poor to have an event of interest	Date of randomization	Censored

Note: one intermittent missing visit will be assumed to have no worsening.
*Death after two consecutive assessment visits will not be considered as an event and will be censored at the date of last assessment with a non-missing score prior to the >= 2 consecutive missing assessment visits.

The Kaplan-Meier method will be used to estimate the TTD curve for each treatment arm. The median TTD will be calculated based on Kaplan-Meier estimate and its 95% CI will be obtained based on Brookmeyer and Crowley method. The treatment difference in TTD will be assessed by the stratified log-rank test. A stratified Cox proportional hazard model with Efron's method of tie handling and with a single treatment covariate will be used to assess the magnitude of the treatment difference (i.e., HR). The HR and its 95% CI will be reported. The same stratification factors used at randomization will be used as the stratification factors in both the stratified log-rank test and the stratified Cox model.

6.3. Subgroup Analyses

To evaluate whether the treatment effect is consistent across various subgroup populations, the between-group treatment effect for OS and PFS will be estimated with a 95% CI, and plotted graphically for the following subgroups, including but not limited to:

- Stratification factors
  - Histology (squamous vs non-squamous)
  - Response to last prior immune therapy received (best response PD/SD vs CR/PR on immune therapy)
  - Received no prior therapy for actionable genomic alteration

- Geographic region (US vs. Canada/Western EU/Australia vs. rest of world)
- Age group (<65 versus ≥65)
- Race (White vs Non-White)
- Baseline ECOG status (0 vs 1)
- Gender (Male vs Female)

If the number of participants in a category of a subgroup variable is less than 10% of the ITT population, the subgroup analysis will not be performed for this category of the subgroup variable, and this subgroup variable will not be displayed in the forest plot. The subgroup analyses for PFS and OS will be conducted using an unstratified Cox model.

6.4. Other Secondary Endpoints

6.4.1. Definition of Other Secondary Endpoints

6.4.1.1. Duration of Response

Duration of response (DOR) is defined as the time interval from the date of first confirmed response (CR or PR) to the earlier date of the first documentation of definitive disease progression or death from any cause as assessed by the investigator per RECIST Version 1.1. The analyses of DOR will be based on the participants in the ITT analysis set who achieve a CR or PR. The censoring rules for DOR are similar to the censoring rules for the primary PFS analysis and detailed in Table 4.

DOR in months will be derived as (date of event/censoring – date of first response [CR or PR] + 1) / 30.4375.

Table 4. Censoring Rules for DOR

Case	Outcome	Date of Event/Censoring
Continued scheduled response assessments until objective PD or death		
PD or death documented after ≤1 missed disease assessments, and before other anti-cancer therapy, if any	Event	PD or death date, whichever occurs first
PD or death immediately after ≥ 2 consecutive missed assessments, or after other anti-cancer therapy, if any	Censored	Date of last adequate tumor assessment prior to the earlier date of >= 2 consecutive missed disease assessments and other anti-cancer therapy, if any
Continued scheduled response assessments without objective PD or death		
Other anticancer treatment is initiated	Censored	Date of last adequate tumor assessment before starting other anticancer treatment
Other anticancer treatment is not initiated	Censored	Date of last adequate tumor assessment

6.4.1.2. Disease Control Rate

Disease control rate (DCR) is defined as the proportion of participants who achieve a CR, PR, or stable disease (SD) as assessed by the investigator per RECIST Version 1.1.

6.4.2. Analysis Methods for Other Secondary Endpoints

6.4.2.1. Duration of Response

A Kaplan-Meier analysis will be performed for DOR, median DOR will be derived by Kaplan-Meier estimates and 95% CIs will be calculated based on Brookmeyer and Crowley method {Brookmeyer 1982}. The milestone DOR rate at 3 months, 6 months will be derived from the KM curve.

6.4.2.2. Disease Control Rate

The counts and percentages, along with the 2-sided 95% CIs using Clopper-Pearson exact method, will be presented by treatment arm.

CCI [redacted]

[redacted]

[redacted]

[redacted]

CCI [redacted]

CCI

CCI

7. SAFETY ANALYSES

7.1. Adverse Events and Deaths

7.1.1. Adverse Event Dictionary

Clinical and laboratory adverse events (AEs) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). System organ class (SOC), high level group term (HLGT), high level term (HLT), preferred term (PT), and lower-level term (LLT) will be provided in the AE dataset.

7.1.2. Adverse Event Severity

Adverse events are graded by the investigator as Grade 1, 2, 3, 4, or 5 according to NCI CTCAE Version 5.0. The severity grade of events for which the investigator did not record severity will be categorized as "missing" for tabular summaries and data listings. The missing category will be listed last in summary presentation.

7.1.3. Relationship of Adverse Events to Study Drug

Related AEs are those for which the investigator selected "Related" on the AE CRF to the question of "Related to Study Treatment." Relatedness will always default to the investigator's choice, not that of the medical monitor. Events for which the investigator did not record relationship to study drug will be considered related to study drug for summary purposes. However, by-participant data listings will show the relationship as missing.

7.1.4. Serious Adverse Events

Serious adverse events (SAEs) will be identified and captured as SAEs if the AEs met the definitions of SAEs that were specified in the study protocol. SAEs captured and stored in the clinical database will be reconciled with the SAE database from the Gilead Global Participant Safety Department before data finalization.

7.1.5. Treatment-Emergent Adverse Events

7.1.5.1. Definition of Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as any AEs with an onset date on or after the study drug start date and no later than 30 days after last dose of study drug.

7.1.5.2. Incomplete Dates

If the onset date of the AE is incomplete and the AE stop date is not prior to the first dosing date of study drug, then the month and year (or year alone if month is not recorded) of onset determine whether an AE is treatment emergent. The event is considered treatment emergent if both of the following 2 criteria are met:

- The AE onset is the same as or after the month and year (or year) of the first dosing date of study drug, and
- The AE onset date is the same as or before the month and year (or year) of the date corresponding to 30 days after the date of the last dose of study drug

An AE with completely missing onset and stop dates, or with the onset date missing and a stop date later than the first dosing date of study drug, will be considered to be treatment emergent. In addition, an AE with the onset date missing and incomplete stop date with the same or later month and year (or year alone if month is not recorded) as the first dosing date of study drug will be considered treatment emergent.

7.1.6. Summaries of Adverse Events and Deaths

Treatment-emergent AEs will be summarized based on the Safety Analysis Set.

7.1.6.1. Summaries of AE Incidence

A brief, high-level summary of the number and percentage of participants who experienced at least 1 TEAE in the categories described below will be provided by treatment arm.

- TEAEs
- TE SAEs
- TE treatment-related AEs
- TE treatment related SAEs
- TEAEs leading to study drug interruption
- TEAEs leading to dose reduction of study drug
- TEAEs leading to study drug discontinuation
- TE treatment-related AEs leading to study drug discontinuation
- TEAEs leading to death
- TE treatment-related AEs leading to death
- TEAEs with CTCAE Grade 3 or higher
- TE treatment-related AEs with CTCAE Grade 3 or higher

In addition, the number and percentage of participants who experienced at least 1 TEAE in above categories will be summarized by SOC, PT, and treatment arm.

Multiple events will be counted only once per participant in each summary. Adverse events will be summarized and listed first in alphabetic order of SOC and then by PT in descending order of frequency of the SG arm within each SOC. For summaries by severity, the most severe severity will be used for those AEs that occurred more than once for a given participant during the study.

In addition, data listings will be provided for the following:

- TE SAEs
- Deaths
- TEAEs with CTCAE Grade 3 or higher
- TEAEs leading to discontinuation of study drug
- TEAEs leading to dose interruption of study drug
- TEAEs leading to dose reduction of study drug
- TEAEs leading to death
- AEs due to COVID-19

7.1.6.2. Summary of Deaths

A summary (number and percentage of participants) of deaths will be provided by treatment arm. Summary will include the following categories:

- All deaths
- Deaths within 30 days of the last dosing of study drug
- Deaths after 30 days of the last dosing of study drug

7.1.7. Additional Analysis of Adverse Events

7.1.7.1. Subgroup Analysis of TEAE

The following subgroup analyses for the TEAEs will be performed by SOC, PT in the Safety Analysis Set.

- Race (White vs. Non-White)
- Region (US vs. Canada/Western EU/Australia vs. rest of world)
- Age (< 65, ≥ 65 years)
- UGT1A1 Genotype in Participants Treated with SG (see Section 7.1.7.3)

7.1.7.2. AE of Special Interest (AESI)

The overall Summary of TEAE of special interest by treatment arm will be presented. In addition, the following summaries will be provided for TEAEs of special interest by Category and PT:

- TEAE of special interest
- TE treatment-related AEs of special interest
- TE treatment related SAEs of special interest
- Serious TEAE of special interest
- TEAEs of special interest leading to study drug discontinuation
- TEAEs of special interest leading to dose interruption of study drug
- TEAEs of special interest leading to dose reduction of study drug
- TE treatment related CTCAE Grade 3 or higher TEAEs of special interest
- CTCAE Grade 3 or higher TEAEs of special interest
- TEAEs of special interest leading to death

A data listing of TEAEs of interest will be provided.

Definitions of AESI, as currently defined, are provided in Table 6, including but not be limited to those listed.

Table 6. Definitions of Adverse Events of Special Interest

AESI	Definition
Serious infections secondary to neutropenia	SOC: Infections and Infestations (serious infections occurring after any grade neutropenia AE)[a]
Severe diarrhea	PT: diarrhea (Grade 3 or higher)
Hypersensitivity[b]	Hypersensitivity SMQ (broad and narrow) and Anaphylactic Reactions SMQ (broad and narrow)
Neutropenia[c]	PTs: neutropenia, neutrophil count decreased, febrile neutropenia

AE = adverse event; AESI = adverse event of special interest; MedDRA = Medical Dictionary for Regulatory Activities; PT = preferred term; SMQ = Standardised MedDRA Query; SOC = system organ class

a SAEs from the Infections and Infestations SOC that occurred within 11 days of neutropenia, febrile neutropenia, or neutrophil count decreased
b For the category of hypersensitivity, only events with onset dates on the day of or 1 day after an infusion are included.
c Grouped AE terms
d All definitions are based on MedDRA Version 26.0 or higher.

7.1.7.3. AE in Participants Treated with SG by UGT1A1 Genotype

For participants treated with SG, the following summaries will be provided by treatment arm and UGT1A1 genotype

- Overall Summary of TEAE in participants treated with SG by UGT1A1 genotype
- Summary of TEAE by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of treatment related TEAE by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of TE SAE by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of treatment related SAE by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of TEAE leading to study drug dose reduction by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of TEAE leading to study drug dose interruption by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of TEAE leading to study drug discontinuation by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of TEAE leading to death by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of treatment related TEAE leading to death by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of Grade 3 or higher TEAE by SOC and PT in participants treated with SG by UGT1A1 genotype
- Summary of Treatment Related Grade 3 or higher AE by SOC and PT in participants treated with SG by UGT1A1 genotype

7.2. Laboratory Evaluations

Laboratory data collected during the study will be analyzed and summarized using both quantitative and qualitative methods. Summaries of laboratory data will be provided for the Safety Analysis Set. The analysis will be based on values reported in conventional units. When values are below the LOQ, they will be listed as such, and the closest imputed value will be used for the purpose of calculating summary statistics as specified in Section 3.6. A by- participant listing for laboratory test results will be provided by participant ID number and time point in chronological order for hematology, serum chemistry, and urinalysis separately. Values falling outside of the relevant reference range and/or having a severity grade of 1 or higher on the CTCAE severity grade will be flagged in the data listings, as appropriate.

7.2.1. Summaries of Numeric Laboratory Results

Descriptive statistics will be provided by treatment arm for each laboratory test specified in the study protocol as follows:

- Baseline values
- Values at each postbaseline time point
- Change from baseline at each postbaseline time point

The mean, median, Q1, Q3, minimum, and maximum values will be displayed to the reported number of digits; StD values will be displayed to the reported number of digits plus 1.

7.2.2. Graded Laboratory Values

CTCAE Version 5.0 will be used to assign toxicity grades (0 to 4) to laboratory results for analysis. Grade 0 includes all values that do not meet the criteria for an abnormality of at least Grade 1. For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (i.e., increased, decreased) will be presented separately. Local labs will be graded based on the central lab normal ranges with in-house macro.

7.2.2.1. Treatment-Emergent Laboratory Abnormalities

Treatment-emergent (TE) laboratory abnormalities are defined as values that increase at least 1 toxicity grade from baseline at any postbaseline time point, up to and including the date of last dose of study drug plus 30 days. If the relevant baseline laboratory value is missing, any abnormality of at least Grade 1 observed within the time frame specified above will be considered treatment emergent.

7.2.2.2. Summaries of Laboratory Abnormalities

The following summaries (number and percentage of participants) for treatment-emergent laboratory abnormalities will be provided by lab test and treatment arm; participants will be categorized according to the most severe postbaseline abnormality grade for a given lab test:

- TE laboratory abnormalities
- TE Grade 3 or 4 laboratory abnormalities

For all summaries of laboratory abnormalities, the denominator is the number of participants with non-missing postbaseline values up to 30 days after the last dosing date.

A by-participant listing of treatment-emergent Grade 3 or 4 laboratory abnormalities will be provided by participant ID number and time point in chronological order. This listing will include all test results that were collected throughout the study for the lab test of interest, with all applicable severity grades displayed.

7.2.3. Liver-related Laboratory Evaluations

Liver function abnormalities after initial study drug administration will be examined and summarized for selected chemistry and hematology parameters.

7.2.4. Shifts Relative to the Normal Range

Shift tables will be presented by showing change in severity grade from baseline to the worst grade postbaseline for selected chemistry parameters (Creatinine/Glucose) and hematology parameters (Hemoglobin/Platelet Count/WBC count/Absolute neutrophil count).

7.3. Body Weight and Vital Signs

Descriptive statistics will be provided by treatment arm for body weight and vital signs as follows:

- Baseline value
- Values at each postbaseline time point
- Change from baseline at each postbaseline time point

Body weight and vital signs measured at unscheduled visits will be included for the baseline value selection. In the case of multiple values in an analysis window, the record with date and time closest to the scheduled visit date and time will be used.

A by-participant listing of vital signs will be provided by participant ID number and time point in chronological order. Body weight will be included in the vital signs listing, if space permits. If not, they will be provided separately.

7.4. Other Safety Measures

7.4.1. Pregnancy

A data listing will be provided for participants experiencing pregnancy during the study.

7.4.2. Post Study Anticancer Therapies

Post study anticancer therapy will be summarized by regimen in ITT population.

By-participant listing will be generated according to information as collected from CRFs.

8. PHARMACOKINETIC AND IMMUNOGENICITY ANALYSES

8.1. Pharmacokinetic Analyses

Participants in the PK Analysis Set will be included for analyses. Descriptive summary of PK concentration (SG, total SN-38, free SN-38 and total antibody -hRS7 IgG) and PK parameters including Cmax (maximum drug concentration) and Ctrough (predose drug concentration) will be summarized by timepoints for each cohort. A listing of PK concentrations will also be generated.

Additional analyses, including population pharmacokinetics (pop-PK), will be described in separate analyses plans and are not in scope for this SAP. Data from this study may be combined with data from other studies with SG for population PK and exposure-response analyses. If applicable, results from such analyses may be summarized in a separate report, rather than in a clinical study report.

8.2. Immunogenicity Analyses

Data on anti-drug antibody (ADA), if applicable, will be listed and may be summarized by number and frequency of positive results for each cohort. Participants in the Immunogenicity Analysis Set will be included for analyses.

9. BIOMARKER ANALYSES

All Trop-2 biomarker analyses will be described in a separate analyses plan (Biomarker Analysis Plan).

10. METHODS FOR HANDLING MISSING DATA

Missing normal ranges for laboratory parameters

When either the lower limit of normal, the upper limit of normal or both are missing or are not machine readable, a standardized reference range will be used.

Missing Data Imputation for Missing Adverse Event/Concomitant Medication Start Date

1) Missing day only

   - If the month and year of the AE/the concomitant medication are the same as the month and year of the first dose date, the first dose date day will be used.
   - If the month and year are before the month and year of the first dose date, the first day of the month will be assigned to the missing day.
   - If the month and year are after the month and year of the first dose date, the first day of the month will be assigned to the missing day.

2) Missing day and month

   - If the year is the same as the year of the first dose date, the first dose date day and month will be used.
   - If the year is prior to the year of the first dose date, December 31 will be assigned to the missing fields.
   - If the year is after the year of the first dose date, January 1 will be assigned to the missing fields.

3) Missing day, month, and year

   - The first dose date will be used.

The imputed start date should be prior or equal to the end date of the AE or medication. If only month is missing, then it will be imputed in a similar way as scenario 2) above. If both year and month missing and day not missing, then it will be imputed in a similar way as scenario 3) above.

Missing Data Imputation for Missing Adverse Event/Concomitant Medication Stop Date

4) Missing day only

- The month and year are the same as the month and year of the first dose date: use the last date of the month.
- The month and year are before the month and year of the first dose date: use the last date of the month.
- The month and year are after the month and year of the first dose date: use the last date of the month.

5) Missing day and month

- The year is the same as the year of the first dose date: use December 31.

6) Missing year

- Uncertain: unable to impute.

7) Missing month

- The year is the same as the year of the first dose date: use December.
- The year is before the year of the first dose date: use December.
- The year is after the year of the first dose date: use December.

If the death date is available and the imputed end date is after the death date, the death date will be used.

If the imputed end date is before the start date of the AE or medication, then make end date equals to start date.

Missing Data Imputation for Disease Diagnosis Date

1) Missing day only

- The first day of the month will be assigned to the missing day

2) Missing day and month

- January 1 will be assigned to the missing fields

3) Missing year

- Uncertain: unable to impute

11. REFERENCES

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

12. SOFTWARE

SAS® Software Version 9.4. SAS Institute Inc., Cary, NC, USA.

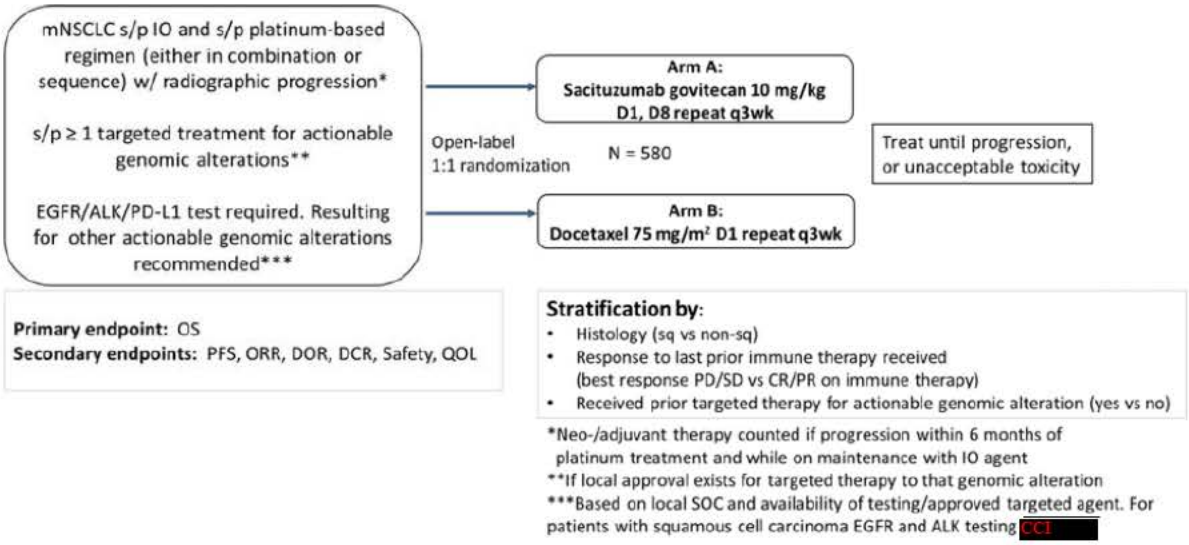
13. SAP REVISION

Revision Date (dd month, yyyy)	Section	Summary of Revision	Reason for Revision

14. APPENDICES

Appendix 1. Study Schema
Appendix 2. Study Procedures Table
Appendix 3. NSCLC-SAQ
Appendix 4. Scoring the QLQ-C30

Appendix 1. Study Schema



ALK = anaplastic lymphoma kinase; CR = complete response; D = day; DOR = duration of response; DCR = disease control rate; EGFR = epidermal growth factor receptor; IO = immuno-oncology; mNSCLC = metastatic non-small cell lung cancer; non-sq = non-squamous; ORR = objective response rate; OS = overall survival; PD = progressive disease; PFS = progression-free survival; PR = partial response; PRO = patient-reported outcomes; q3wk = every 3 weeks; SD = stable disease; SOC = standard of care; s/p = status post; sq = squamous; vs = versus

Appendix 2. Study Procedures Table

See Protocol Amendment 2 Table 1: GS-US-577-6153 Study Procedure Table

Appendix 3. NSCLC-SAQ

Domain	Item		Response
Cough	1. How would you rate your coughing at its worst…?		0, 1, 2, 3, 4
Pain	2. How would you rate the worst pain in your chest…?	Create a single score by selecting the highest severity (i.e., value) on either item	0, 1, 2, 3, 4
	3. How would you rate the worst pain in areas other than your chest…?		
Dyspnea	4. How often did you feel short of breath during usual activities…?		0, 1, 2, 3, 4
Fatigue	5. How often did you have low energy…?	Create a single score by computing the mean of severity (i.e., value) on both items	0, 1, 2, 3, 4
	6. How often did you tire easily…?		0, 1, 2, 3, 4
Appetite	7. How often did you have a poor appetite over the last 7 days?		0, 1, 2, 3, 4
NSCLC-SAQ Total Score (Sum the 5 domains)			Range 0 to 20


CCI

GS-US-577-6153 SAP v1.0

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy hh:mm:ss)
PPD	Biostatistics eSigned	13-Jul-2023 00:41:57
PPD	Clinical Development eSigned	13-Jul-2023 11:06:26