STATISTICAL ANALYSIS PLAN

Protocol Number:	SGNTV-003
Protocol Version:	Amendment 4; dated 22-Jan-2024
Protocol Title:	A Randomized, Open-Label, Phase 3 Trial of Tisotumab Vedotin vs Investigator's Choice Chemotherapy in Second- or Third-Line Recurrent or Metastatic Cervical Cancer
Sponsor:	Seagen, Inc.* 21823 30th Drive SE Bothell, WA 98021, USA *As of 14-Dec-2023, Seagen Inc. became part of Pfizer Inc.
Collaborator:	Genmab A/S Kalvebod Brygge 43 DK-1560 Copenhagen V Denmark

APPROVAL SIGNATURES

Product:	Tisotumab Vedotin
Protocol Number/Amendment:	SGNTV-003/ Amendment 4
SAP Version:	Version 3.0
Version Date:	23-APR-2024

The individuals signing below have reviewed and approve this statistical analysis plan.

PPD , Ph.D.
PPD
Seagen Inc.

Date

PPD , Ph.D.
PPD
Seagen Inc.

Date

PPD , Ph.D.
PPD
Seagen Inc.

Date

PPD , M.D..
PPD
Seagen Inc.

Date

PPD , M.D.
PPD
Seagen Inc.

Date

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

List of Abbreviations

2L	second-line
3L	third-line
ADA	anti-drug antibody
ADC	antibody-drug conjugate
ADI	actual dose intensity
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
C1D1	Cycle 1 Day 1
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CR	complete response
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DCO	data cutoff
DOR	duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EORTC	European Organization for Research and Treatment of Cancer
EOS	end-of-study
EOT	end-of-treatment
EQ-5D	standardized instrument developed as a measure of HRQOL
FA	final analysis
FDA	US Food and Drug Administration
FFPE	formalin-fixed paraffin-embedded
G-CSF	granulocyte colony-stimulating factor
GSP	group sequential procedure
HR	hazard ratio
HRQOL	health-related quality of life

2L	second-line
IA	interim analysis
IDI	intended dose intensity
IDMC	independent data monitoring committee
IHC	immunohistochemistry
IPCW	inverse probability of censoring weights
IPD	important protocol deviations
IRB	Institutional Review Board
IRR	infusion-related reaction
ITT	intent-to-treat
IV	intravenous
KM	Kaplan-Meier
LLOQ	lower limit of quantification
LPI	last-patient-in
MedDRA	Medical Dictionary for Regulatory Activities
MMAE	monomethyl auristatin E
NCA	noncompartmental analysis
NCI	National Cancer Institute
NE	not evaluable
ORR	objective response rate
OS	overall survival
PD	progressive disease
PD-1	programmed cell death protein 1
PD-L1	programmed cell death ligand 1
PFS	progression-free survival
PK	pharmacokinetics
PR	partial response
PRO	patient reported outcome
PT	preferred term
Q3W	every 3 weeks
QLQ-C30	Quality of Life of Cancer Patients questionnaire
QLQ-CX24	Quality-of-Life questionnaire cervical cancer module
RECIST	Response Evaluation Criteria in Solid Tumors
RDI	relative dose intensity
r/mCC	recurrent/metastatic cervical cancer
RMST	restricted mean survival time
RNA	ribonucleic acid

2L	second-line
RTSM	randomization and trial supply management
SAE	serious adverse event
SD	stable disease
SOC	system organ class
TAb	total antibody
TEAE	treatment-emergent adverse event
TF	tissue factor
TTR	time to response
TV	tisotumab vedotin
ULN	upper limit of normal
VAS	visual analog scale
WHO	World Health Organization

1. INTRODUCTION

This document outlines the statistical methods to be implemented within the scope of Protocol SGNTV-003, entitled "A randomized, open-label, phase 3 trial of tisotumab vedotin vs investigator's choice chemotherapy in second- or third-line recurrent or metastatic cervical cancer". Results of the planned analyses will become the basis of the clinical study report (CSR) for this protocol.

The purpose of this plan is to provide specific guidelines from which the analysis will proceed. All planned analyses specified in this statistical analysis plan will be performed. Any changes to this plan, in the form of "post hoc" or "data driven" analyses will be identified as such in the final clinical study report. Any changes will either be reflected in amendments to this plan before the database lock or specifically documented in the CSR.

2. STUDY OBJECTIVES

2.1. Primary Objectives

- Demonstrate improvement in clinical efficacy of tisotumab vedotin compared to chemotherapy in subjects with second- or third-line (2L-3L) cervical cancer

2.2. Secondary Objectives

- Assess improvement in clinical efficacy of tisotumab vedotin compared to chemotherapy in subjects with 2L-3L cervical cancer
- Demonstrate improvement in antitumor activity of tisotumab vedotin compared to chemotherapy in subjects with 2L-3L cervical cancer
- Characterize the antitumor response of tisotumab vedotin and chemotherapy in subjects with 2L-3L cervical cancer
- Evaluate the safety and tolerability of tisotumab vedotin
- Assess health-related quality of life (HRQOL)

2.3. Exploratory Objectives

- Investigate the relationship between tumor TF expression and response to tisotumab vedotin
- Assess biomarkers and their association with disease, mechanisms of resistance, and/or response to therapy
- Evaluate the pharmacokinetics (PK) and immunogenicity of tisotumab vedotin

3. STUDY ENDPOINTS

3.1. Primary Endpoint

- Overall survival (OS)

3.2. Secondary Endpoints

Key secondary endpoints are:

- Progression-free survival (PFS) based on Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 as assessed by the investigator
- Confirmed objective response rate (ORR) based on RECIST v1.1 as assessed by the investigator

Other secondary endpoints are:

- Time-to-response (TTR) as assessed by the investigator
- Duration of response (DOR) as assessed by the investigator
- Incidence of adverse events (AEs)
- EQ-5D-5L index
- EQ-5D visual analog scale (VAS)
- EORTC-QLQ-C30
- EORTC-QLQ-CX24

3.3. Exploratory Endpoints

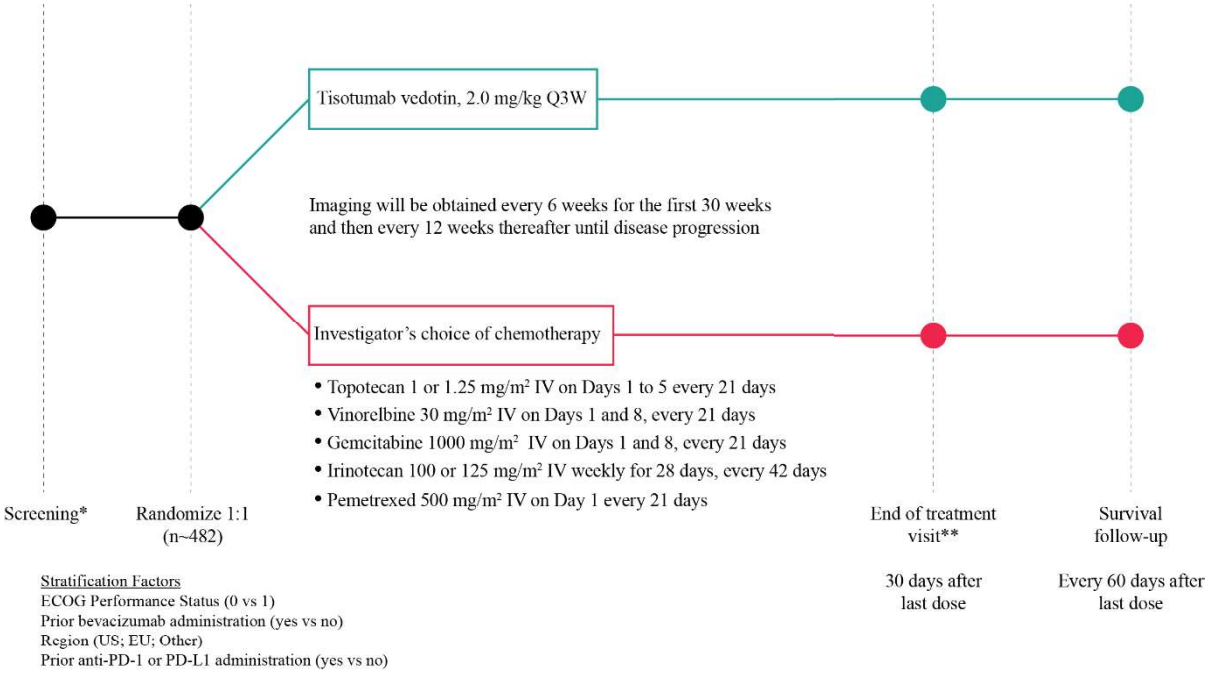
- Tumor TF expression (via immunohistochemistry [IHC] or RNA) in relation to efficacy endpoints
- Baseline characteristics and changes from baseline of biomarkers from peripheral blood and/or formalin-fixed paraffin-embedded (FFPE) tumor tissue in relation to efficacy endpoints
- PK concentrations and anti-drug antibodies (ADAs) associated with tisotumab vedotin

4. STUDY DESIGN

This is an open-label, randomized (1:1), global, phase 3 trial of tisotumab vedotin versus investigator's choice of chemotherapy in subjects with r/mCC who have received 1 or 2 prior lines of systemic therapy for their recurrent or metastatic disease. Eligible subjects will be randomized to either tisotumab vedotin 2.0 mg/kg every 3 weeks (Q3W) or investigator's choice of chemotherapy (as noted in the "other study treatment" section). Randomization will be stratified by Eastern Cooperative Oncology Group (ECOG) performance status (0 vs. 1), prior bevacizumab administration (yes vs. no), region (US, EU, Other), and prior anti-programmed cell death protein 1 (PD-1) or anti- programmed cell death ligand 1 (PD-L1) administration (yes vs. no). Imaging will be obtained every 6 weeks (42 days ±7 days) for the first 30 weeks and then every 12 weeks (84 days ±7 days) thereafter, calculated from Cycle 1 Day 1 (C1D1) of treatment administration. Imaging must continue until evidence of radiographic disease progression per RECIST v1.1 as assessed by investigator. Survival status will be assessed every 60 days (±7 days) beginning from the day of the last dose of study treatment, or more frequently around the time of a database lock.

An independent data monitoring committee (IDMC), consisting of members who are external and independent of the sponsor study team, will be formed to monitor the safety of subjects in this study on a periodic basis, review safety and efficacy data from this study as per planned analyses, and make recommendations to the sponsor.

Figure 1: Study Design



*The proportion of participants who have not received prior bevacizumab in combination with chemotherapy as 1L treatment may be capped at 50%.

** Some AESIs may be followed longer than 30 days until resolution, improvement, or stabilization.

5. ANALYSIS SETS

This section defines each of the analysis sets that will be utilized. The use of each analysis set will be discussed in Section 7.

5.1. Intent-to-Treat (ITT) Analysis Set

The intent-to-treat (ITT) analysis set will include all subjects who are randomized on or before the date of last-patient-in (LPI) in the global study. Subjects will be included in the treatment group assigned at randomization regardless of any actual treatment received.

5.2. Safety Analysis Set

The safety analysis set will include all subjects who are randomized on or before the date of LPI in the global study and receive any amount of study treatment. Subjects will be analyzed according to the actual treatment received. Subjects who received any dose of tisotumab vedotin will be included in the TV treatment group in the Safety Analysis Set.

5.3. Patient Reported Outcomes Full Analysis Set (PRO FAS)

The PRO FAS will include all subjects who are randomized on or before the date of LPI in the global study and have received any amount of study treatment and have completed baseline and at least one post-baseline PRO assessment. The PRO FAS will be used for PRO analyses.

5.4. Pharmacokinetics Analysis Set

The Pharmacokinetics analysis set includes randomized subjects who are randomized on or before the date of LPI in the global study and have received any amount of TV and have at least one reportable value for ADC, TAb or MMAE concentration. The PK analysis set will be used for PK analyses.

6. STATISTICAL CONSIDERATIONS

6.1. General Principles

In general, descriptive statistics will be presented that include the number of observations, mean, standard deviation, median, minimum and maximum for continuous variables, and the frequencies and percentages for categorical variables.

Unless otherwise specified, confidence intervals (CIs) will be calculated at 2-sided 95% level. The 2-sided 95% exact CI using Clopper-Pearson methodology will be calculated for the response rates where applicable (e.g., ORR) (Clopper 1934). For time-to-event endpoints, the median survival time will be estimated using the Kaplan-Meier (KM) method; the associated 95% CI will be calculated based on the complementary log-log transformation (Collett 1994).

Any analysis that is not described in this plan will be considered exploratory and will be documented in the CSR as a post hoc analysis or a change to the planned analysis.

To comply with regulatory electronic submission guidelines, listings of all clinical data will be submitted as electronic data sets. To facilitate data review for the study report, only pertinent data listings will be created and attached to the appendix of the CSR.

All statistical tables, listings, and figures will be produced using SAS®, version 9.4 or more recent. Sample size calculations were performed using EAST®, version 6.5. Other statistical software, if used, will be described in the CSR.

6.2. Determination of Sample Size

Approximately 482 subjects will be randomized in a 1:1 ratio to one of the 2 study arms.

This sample size is determined to provide sufficient power to demonstrate superiority of the tisotumab vedotin arm against the control arm based on the primary endpoint OS. Specifically, this sample size, with a total number of approximately 336 OS events, provides an overall power of 90% for the hypothesis testing on OS at a 2-sided 5% significance level to detect an HR $\neq$ 1 when the true HR is 0.7, with the following assumptions:

- median OS of 12.9 months in the tisotumab vedotin arm
- median OS of 9 months in the chemotherapy arm

- drop-out rate of 5% per year
- an interim analysis (IA) at approximately 75% (252) of the total number of OS events using the Lan deMets (O'Brien Fleming) boundary (Lan 1983)
- accrual duration of approximately 23 months with at least 12 months of follow-up
- 1:1 randomization ratio

The IA and final analysis (FA) on OS are estimated to occur at approximately 27 and 35 months after the first subject is randomized into the study.

At the OS IA and FA, it is estimated that the number of PFS events will be approximately 413 and 457, respectively, which will provide an overall power of 95% at a 2-sided 5% level of significance assuming an HR of 0.7 with a median PFS of 4.3 months in the tisotumab vedotin arm and a median PFS of 3 months in the control arm.

For ORR, it is estimated that all 482 subjects will be fully randomized into the study, based on the assumed accrual, and have at least 2 post-baseline scans at the OS IA and FA timings. This provides an overall power of approximately 99% at a 2-sided 5% level of significance assuming an ORR of 25% for the tisotumab vedotin arm and a 10% ORR for the control arm.

6.3. Randomization and Blinding

Subjects will be randomized in a 1:1 ratio to the tisotumab vedotin arm or the control arm of investigator's choice of chemotherapy. The randomization will be stratified based on the following stratification factors:

- Region (US, EU, Other)
- ECOG performance status at baseline (0 vs. 1)
- Prior bevacizumab administration (yes vs. no)
- Prior anti-programmed cell death protein 1 (PD-1) or anti-programmed cell death ligand 1 (PD-L1) administration (yes vs. no)

Although this is an open-label study, until the database lock for pre-specified OS primary analysis, analyses or summaries by treatment arm will be limited to the purpose of IDMC safety review and be conducted by an external vendor that is independent from the study team. In addition, retrospectively, the images for response assessment may be centrally reviewed by the independent radiologists who do not have knowledge of treatment assignment.

6.4. Data Transformations and Derivations

Age: Reported age in years will be used.

Baseline: Unless otherwise specified, baseline values used in all analyses will be the most recent non-missing measurement prior to the first dose of study treatment for safety analyses or randomization for efficacy analyses. If there are multiple values at the same date and time that qualify for baseline definition, for continuous values, the average of these values will be

used as baseline value; for categorical values, the value of the assessment indicating better status will be used as baseline to be conservative for the analyses related to change from baseline.

Study Day: For subjects who received treatment, study day will be calculated relative to the first dose as (assessment date – first dose date + 1) for dates on or after the first dose date. The **first dose date** is the earliest date of administration of any study treatments. For dates prior to the first dose date, study day will be calculated as (assessment date – first dose date).

For subjects who did not receive treatment, study day will be calculated relative to the date of randomization as (assessment date – randomization date + 1) for dates on or after randomization date. For dates prior to the randomization date, study day will be calculated as (assessment date – randomization date).

Other time variables based on two dates, e.g., Start Date and End Date, will be calculated as (End Date–Start Date+1) (in days) unless otherwise specified in the planned analysis section.

The following unit conversion will be implemented unless otherwise specified:

Months = Days/30.4375

Years = Days/365.25

EOT: The end-of-treatment (EOT) date will be the date when the EOT visit is performed; if an EOT visit is not performed then the EOT date will be the date of decision making for treatment discontinuation as entered into the EOT summary CRF.

Response Assessment Date: At each response assessment time point, scans to evaluate target and non-target lesions can be performed on multiple dates. The date of response will be determined as the following:

- If the time point response is complete response (CR) or partial response (PR), then the latest date of all radiologic scans at the given response assessment visit will be the date of response.
- If the time point response is stable disease (SD) or non-CR/non-PD, then the earliest date of all radiologic scans at the given response assessment visit will be the date of response.
- If the time point response is PD, then the earliest date that PD has been documented will be the date of PD, i.e., the earliest of:
    - Date of target lesion assessments when the target lesion response is PD.
    - Date of non-target lesion assessments when the lesion status is unequivocal progression.
    - Date of documenting new lesions.

For subjects whose best overall response is a confirmed CR or PR, the date of objective response will be the date of initial documentation of response (i.e., CR or PR that is subsequently confirmed).

In the cases where an overall response of PD is based on an equivocal new lesion that is later assessed as unequivocal in a subsequent scan, the overall response of PD should be back-

dated to the date when the equivocal new lesion was first identified. If an equivocal new lesion was later absent or confirmed to be a benign lesion, then this new lesion is not considered to define an overall response of PD. In cases where PD occurs on a date after an equivocal new lesion is identified, but the progression is not due to a change of the equivocal new lesion to an unequivocal lesion, but rather from progression of other lesions, the PD date will not be back-dated, but will be the date when definitive PD is recorded.

Adequate Tumor Assessment: An adequate tumor assessment must include a radiologic scan with the overall disease response of CR, PR, SD, or PD. Scans with the overall response not evaluable (NE) will not be considered an adequate response assessment for the purpose of PFS and DOR censoring.

6.5. Handling of Dropouts and Missing Data

Except for the scenarios covered in this section, missing data will not be imputed.

Subjects who do not have at least two (initial response and confirmation scan) post-baseline response assessments will be counted as non-responders for analysis of the confirmed ORR.

Missing AE dates will be imputed for the purpose of calculating duration of events and determining the treatment-emergent status. See Appendix A for the imputation rule and for treatment-emergent classification.

Missing disease diagnosis date will be imputed for the purpose of calculating the time from diagnosis to date of randomization or date of first dose (Appendix C).

Missing subsequent cancer-related therapy start date will be imputed for the purpose of deriving the time-to-event endpoints as applicable (Appendix E).

For prior therapies start dates, if month and year are present and only day is missing, the day will be imputed as the first day of the month. If month or year is missing, no imputation will be performed.

For prior therapies end dates, if month and year are present and only day is missing, the day will be imputed as the last day of the month or 21 days before the first dose of study drug, whichever is earlier. If the imputed end date is earlier than start dates, the end day will be imputed as the day after the start date. If month or year is missing, no imputation will be performed. Censoring for time-to-event endpoints will be described in Section 7 with each planned analysis, as applicable.

Unless otherwise specified, lab values which are recorded or provided as being less than the lower limit of quantification (LLOQ) will be included in figures and analysis as LLOQ/2 and listed as "<LLOQ" in the listings.

Missing data for PROs will be handled according to the user manual for each individual PRO.

6.6. Multicenter Studies

There are multiple centers in this study, however it is not anticipated that any center will accrue enough subjects to warrant an analysis by center.

6.7. Multiple Comparison/Multiplicity

Hypothesis testing will be performed on the primary endpoint OS and two key secondary efficacy endpoints, PFS and ORR, while controlling the overall Type I error rate of 5% (2-sided) using a hierarchical approach with group sequential testing (Maurer 2013).

The endpoints are tested in the hierarchical order as follows: OS first, then PFS, and finally ORR. Hypothesis testing starts with OS and will only proceed to the hypothesis testing on the next endpoint, PFS, when the efficacy boundary is crossed for OS at the first time (IA or FA). Similarly, the hypothesis testing on ORR will only occur after the efficacy boundary is crossed for PFS.

For OS, the Lan DeMets (O'Brien-Fleming) spending function will be used (Lan 1983). With approximately 252 and 336 events estimated to occur at the IA and FA respectively, the OS will be tested at a 2-sided nominal 0.019/0.044 level at the IA/FA corresponding to a HR boundary of 0.75/0.80.

For PFS, the Lan DeMets (Pocock) spending function will be used (Lan 1983). With approximately 413 and 457 events estimated to occur at the OS IA and FA respectively, the PFS will be tested at a 2-sided nominal 0.047/0.025 level at the IA/FA corresponding to an HR boundary of 0.822/0.811.

For ORR, it is estimated that all 482 subjects will be fully randomized into the study and have the opportunity to have at least two post-baseline scans by the OS IA at 27 months. The ORR will be tested at 2-sided 5% level corresponding to an odds ratio boundary of 1.67.

The overall type 1 error for the 3 key endpoints, OS, PFS, and ORR, is controlled at a 2-sided 5% level using the approach of Maurer and Bretz (Maurer 2013). For hypothesis testing on all endpoints, the actual nominal significance level to be applied will be adjusted to reflect the actual number of events realized at the time of the IA/FA using the specified spending function for each endpoint.

6.8. Examination of Subgroups

As exploratory analyses, subgroup analyses may be conducted for selected endpoints (i.e., primary and key secondary efficacy endpoints). Subgroups will be defined using data recorded in the EDC. Subgroups may include but are not limited to the following:

- Age (<65, ≥65 years)
- Region (US, EU, Asian, Other)
- Race (White, Non-White)
- Ethnicity (Hispanic, Non-Hispanic)
- ECOG performance status (PS) at baseline (0, 1)
- Prior bevacizumab administration (yes, no)
- Prior anti-PD-1/PD-L1 therapy administration (yes, no)

- Histology (squamous cell carcinoma and adenosquamous carcinoma, adenocarcinoma)
- Chemotherapy agent
- Number of prior recurrent/metastatic systemic regimens (1, 2)

For the above subgroups, analyses of OS and PFS by treatment arm will be conducted using log-rank tests and Cox proportional hazards model stratified by randomization stratification factors except for region, as specified in Section 6.9. If the subgroup is a stratification factor, then the stratified models will control for the rest of stratification factors. For subgroup variables with two levels, subgroup analysis may not be performed on the variable if the total number of subjects in a subgroup is too small (e.g., < 10% of the total sample size). For subgroup variables with more than two levels, pooling across levels may be considered when the number of subjects within one level is too small.

In addition, estimates of the median OS and PFS time will be provided for the control arm by chemotherapy agent with the two-sided 95% CI using the complementary log-log transformation method (Collett 1994).

Additional subgroup analyses may be performed for specific countries for regulatory submission purposes.

6.9. Covariates

Stratified analyses specified in Section 7 will include adjustment for the stratification factors as recorded at randomization, with the exception that the stratification factor of region will not be included in the stratified analyses to reduce sparse strata.

In the situation where there is insufficient information in a stratum (e.g., if there are very few subjects enrolled in a stratum or there is no informative event in a stratum by combined treatment arms), pooling of this stratum with another stratum may be considered for stratified analyses.

6.10. Timing of Analyses

The OS FA is planned when approximately 336 OS events have occurred in the ITT analysis set. One OS IA is planned when approximately 252 (75% information) of the total expected 336 OS events occur.

If OS is statistically significant at either the IA or FA, a formal statistical test in PFS and the confirmed ORR (by investigator assessment) between the two arms will be performed sequentially.

7. PLANNED ANALYSES

7.1. Disposition

Analysis set: ITT Analysis Set

Subject disposition will be summarized by treatment arm and total for subjects in the ITT analysis set, unless specified otherwise. Subjects who are on study treatment and subjects

who discontinued study treatment will be summarized along with the reasons for discontinuation of treatment. Subjects who are ongoing in the study and subjects who discontinued from the study will be summarized along with the reasons for discontinuation of study. The number and percentage of subjects in long term follow-up will be summarized. The number of subjects who were randomized and the number and percentage of subjects in each analysis set will be summarized.

A by-subject listing of subject disposition data with reasons for study treatment and study discontinuation and a separate one with long-term follow-up assessments will be provided.

The number of subjects enrolled in each country and at each site will be summarized by treatment arm and total.

Summary of screening data will be provided for all screened subjects. The number of subjects who signed the informed consent, the number of screen failures and the percentage relative to the total number of subjects screened, and the reasons for screen failure will be summarized. A listing of subjects who failed screening will be produced with reasons for screen failure and available demographic information.

7.2. Demographic and Baseline Characteristics

Analysis set: ITT Analysis Set

Demographics and baseline characteristics, including age, ethnicity, race and baseline weight will be listed and summarized. Summaries will be presented for each arm and the total.

Disease specific characteristics as recorded in the EDC will be listed and summarized by treatment arm and total for the following:

- ECOG performance status
- Prior bevacizumab administration
- Prior anti-PD-L1/PD-1 administration
- Histology
- FIGO stage at initial diagnosis
- Disease status at study entry (local recurrence (pelvis) vs. metastatic disease)
- Time (in months) from the earliest recurrent/metastatic diagnosis to randomization
- Tumor TF expression

All prior anti-cancer therapies will be listed. Prior anti-cancer systemic therapies will be summarized by treatment arm and total.

7.3. Protocol Deviations

Analysis set: ITT Analysis Set

Important protocol deviations (IPD) (defined as protocol violations by Seagen) are those that represent a divergence from the protocol that could have a significant effect on the integrity

of the study data, or on the subject's rights, safety, or welfare. IPDs also include exemptions to the study inclusion/exclusion criteria and will be summarized by category. All IPDs will be listed.

7.4. Treatment Administration

Analysis set: Safety Analysis Set

Treatment administration will be summarized using counts and percentages for categorical variables and summary statistics (mean, median, standard deviation, and range) for continuous variables. The following information will be summarized for each study treatment:

- Total number of treatment cycles per subject
- Duration of treatment (months)
- Total cumulative dose (mg)
- Dose modification by modification type (dose reduction, dose delay, dose elimination, dose interruption, and dose discontinuation) and overall (i.e. overall drug modifications for a subject), as applicable
- Absolute dose intensity (ADI) and R intensity (RDI)

Duration of treatment (except when calculating dose intensity) is defined as the time from first dose date to the earliest of the following dates:

- Last dose date:
    - For TV, pemetrexed and irinotecan, the last infusion date + 20
    - For topotecan, the last infusion date + 16
    - For vinorelbine and gemcitabine, the last infusion date + 13
- Date of death.
- Start of subsequent cancer-related therapy.
- Analysis data cutoff (DCO) date for an analysis
- End of treatment date.

Intended Dose Intensity (IDI) is defined as the intended dose of drug per unit of time (week). In this study, the IDI for each study treatment is as the following:

- **Tisotumab Vedotin (TV):** The IDI is 0.67 mg/kg per week (i.e., 2.0 mg/kg every 21 days).
- **Topotecan**: The planned dose is 1 or 1.25 mg/m^2 on Days 1 to 5, every 21 days. Thus, the IDI is 1.67 or 2.08 mg/m^2 per week (i.e., 5 or 6.25 mg/m^2 every 21 days).
- **Vinorelbine**: The planned dose is 30 mg/m^2 on Days 1 and 8, every 21 days. Thus, the IDI is 20 mg/m^2 per week (i.e., 60 mg/m^2 every 21 days).

- **Gemcitabine:** The planned dose is 1000 mg/m^2 on Days 1 and 8, every 21 days. Thus, the IDI is 666.67 mg/m^2 per week (i.e., 2000 mg/m^2 every 21 days).
- **Irinotecan**: The planned dose is 100 or 125 mg/m^2 weekly for 28 days, every 42 days. Thus, the IDI is 66.67 or 83.33 mg/m^2 per week (i.e., 400 or 500 mg/m^2 every 42 days).
- **Pemetrexed**: The IDI is 166.67 mg/m^2 per week (i.e., 500mg/m^2 every 21 days).

Cumulative Dose is the sum of the actual dose amount per unit that a subject received across all cycles. For TV, the unit is mg/kg; for topotecan, vinorelbine, gemcitabine, irinotecan and pemetrexed, the unit is mg/ m^2.

Actual Dose Intensity (ADI) = cumulative actual dose / duration of treatment (week), where duration of treatment (week) is computed as [(last dose date – first dose date) + 1] / 7, where the last dose date is defined as above.

Relative Dose Intensity (RDI) = ADI/IDI × 100%.

A by-subject listing of study drug administration will be provided.

7.5. Efficacy Analyses

The analysis for efficacy endpoints, including the primary endpoint of OS, the key secondary endpoints of PFS and ORR, and other secondary and exploratory endpoints will be conducted using the ITT analysis set, unless otherwise specified. Subjects will be analyzed based on their randomized treatment arm.

For efficacy endpoints, the randomization stratification factors per Randomization and Trial Supply Management (RTSM) system will be used as strata in stratified analysis, except for region. In addition, analysis of OS, PFS and ORR may be repeated for each of the subgroups specified in Section 6.8.

7.5.1. Primary Endpoint

The primary endpoint of this study is OS. The study will be considered positive if the OS comparison is statistically significant between treatment arms. Please refer to Section 6.7 for multiplicity adjustment for the primary hypothesis.

No formal multiplicity adjustment will be used for sensitivity analyses or subgroup analyses.

Overall Survival (OS)

Analysis set: ITT Analysis Set

OS is defined as the time from the date of randomization to the date of death from any cause. Specifically,

OS = Date of death or Censoring date – Date of randomization +1.

In the absence of death before the analysis cutoff date, OS will be censored at the date the subject is last known to be alive.

The null hypothesis for the primary endpoint is that the OS of the TV arm has no difference from that of the control arm. To test this hypothesis, a stratified log-rank test comparing two arms in the ITT analysis set will be conducted controlling for the randomization stratification factors [i.e., ECOG status (0 vs. 1), prior bevacizumab administration (yes vs. no) and prior anti-PD-L1/PD-1 administration (yes vs. no)].

Kaplan-Meier curves and estimates of the median OS time will be provided for each arm with the two-sided 95% CI using the complementary log-log transformation method (Collett 1994). KM estimates of the 25th and 75th percentiles of the OS time and the observed minimum and maximum OS time will also be reported. In addition, the estimated probability of OS at landmark time points (for example, every 3 months) and their 95% CIs will also be reported. The hazard ratio and 95% CI will be estimated using a stratified Cox proportional hazards regression model controlling for the randomization stratification factors, except for region.

The testing strategy for the efficacy endpoint OS is described in Section 6.7. The Lan-DeMets O'Brien-Fleming approximation spending function will be used to obtain the efficacy boundaries at the IA and final OS analyses.

A sample SAS code for the stratified log-rank test is provided below.

```
** ostime = OS time;
** censor = Censor Indicator (1 = censored)
** Trt = Treatment Arm (TV vs. Control)
** SF1 = Stratification Factor 1
** SF2 = Stratification Factor 2
** SF3 = Stratification Factor 2
ODS OUTPUT HomTests=chisq;

PROC LIFETEST DATA = osdata;
        TIME ostime*censor(1);
        STRATA SF1 SF2 SF3 / GROUP = Trt TEST=logrank;
RUN;
```

For the purpose of describing the treatment effect, a hazard ratio between the TV arm and the control arm and its 95% CI will be estimated using a stratified Cox proportional hazards regression model controlling for the stratification factors. A sample SAS code is provided below.

```
PROC PHREG DATA = osdata;
        CLASS Trt(ref='Control');
                STRATA SF1 SF2 SF3;
                MODEL ostime*censor(1) = Trt / TIES=Efron RL;
                HAZARDRATIO Trt;
RUN;
```

Proportional hazards (PH) assumption for treatment arms will be examined with the following approaches:

1. Plot of 'log-negative-log' of the KM estimator of the survival function versus log(time) by treatment arm (SAS PROC LIFETEST plot(lls)). Parallel curves with constant distance over time indicate the PH assumption is met.
2. Plot of Schoenfeld residuals for the treatment variable versus time. Schoenfeld residuals including LOESS curve will be plot in SAS using the PHREG procedure via the OUTPUT statement (keyword=RESSCH). With proportional hazards, the LOESS curve should be parallel to the x-axis. Residuals that do not show any trend indicate that the PH assumption is met.

Sensitivity Analyses for OS

1. **Unstratified Analysis:** OS analysis by unstratified log-rank test to compare the treatment arms. Additionally, an unstratified Cox regression model will be used to estimate the hazard ratio and the corresponding 95% CI for the treatment effect. The same censoring method as for the primary analysis of OS will be used.
2. **Mis-stratification:** In the case of > 5% inconsistency in stratification factors between RTSM and eCRF, the stratification factors collected at eCRF will be used as strata in both stratified log-rank test and stratified Cox proportional hazards regression model.
3. **Initiation of Subsequent Cancer-related Therapy:** Start subsequent cancer-related therapy is likely to bias the analyses of OS, more specifically, may underestimate OS comparison when the distribution of subsequent therapies is imbalanced between 2 arms. To reduce such bias, the following OS sensitivity analysis may be planned:

   Inverse probability of censoring weights (IPCW) method [Robins and Finkelstein, 2000]. Subjects who took subsequent therapy will be censored at the time of sequent cancer-related therapy, but subjects are weighted according to their probability to take subsequent therapy.
4. **Non-proportional hazard:** In the case that the PH assumption is violated, OS may be analyzed based on a restricted mean survival time (RMST) up to τ (Royston 2011; Uno 2014), and τ will be the minimum of (largest observed OS event time for the TV arm, largest observed OS event time for the control arm). Comparison of RMST between treatment arms will be performed using the KM curve-based test. RMST for each treatment arm and difference in RMST between the treatment arms will be provided with the corresponding 95% CIs.

7.5.2. Key Secondary Efficacy Endpoints

Progression Free Survival Assessed per Investigator

Analysis set: ITT Analysis Set

PFS is defined as the time from the date of randomization to the first documented disease progression (as assessed by investigator per RECIST v1.1) or death from any cause, whichever occurs first. Specifically, PFS is derived as

PFS Event Date (Date of first documented PD or death) or Censoring date–Date of randomization + 1

Censoring scheme for the primary analysis of PFS are described below in Table 1 .

Table 1: Censoring scheme for primary analysis of PFS per RECIST v1.1 per investigator

Scenario	Progression/Censor Date	Outcome
No post-baseline tumor assessments[a]	Date of randomization	Censored
No documented disease progression or death	Date of last tumor assessment of CR, PR, SD, or non-CR/non-PD	Censored
New anti-cancer treatment (systemic, radiation, or surgery) started before PD or death observed	Date of last CR, PR, SD, or non-CR/non-PD on or prior to date of new anti-cancer treatment	Censored
Progressive disease (PD)	Date of PD	Event
Death before first PD assessment[b]	Date of death	Event
Death or progression right after two or more consecutive missed tumor assessments	Date of last tumor assessment of CR, PR, SD, or non-CR/non-PD prior to the missed visits	Censored

[a] Censoring date will be date of randomization when 1) no post-baseline tumor assessment and no death, or 2) no post-baseline tumor assessment and with death after a prolonged duration (eg., after the planned 2nd response assessment date);
[b] Patient with no post-baseline tumor assessment and with death within a prolonged duration (eg., before the planned 2nd response assessment date) will be counted as a death event.

The two arms will be compared for PFS using a stratified log-rank test controlling for the stratification factors. Kaplan-Meier curves and estimates of the median PFS time will be provided for each arm with the two-sided 95% CI using the complementary log-log transformation method (Collett 1994). KM estimates of the 25th and 75th percentiles of the PFS time and the observed minimum and maximum PFS time will also be reported. In addition, the estimated probability of PFS at landmark time points (for example, every 3 months) will be reported. The hazard ratio and 95% CI will be estimated using a stratified Cox proportional hazards regression model controlling for the randomization stratification factors.

The testing strategy for the efficacy endpoints OS, PFS and ORR is described in Section 6.7. Hypothesis testing starts with OS and will only proceed to the hypothesis testing on the next endpoint, PFS, when the efficacy boundary is crossed for OS at the first time (IA or FA).

PH assumption will be examined using the same method as described in Section 7.5.1.

Sensitivity Analyses for PFS

1. **New Subsequent cancer-related therapy before PD/death:** For subjects who received subsequent cancer-related therapy before PD or death, two sensitivity analyses will be performed:

    - Not to consider any subsequent cancer-related therapies (whether systemic, radiation, or other) as a censoring reason.

- Consider any subsequent cancer-related therapies (whether systemic, radiation, or other) as an event.

2. **Missing Tumor Assessments:** To explore the potential impact of missing tumor assessments on PFS, subjects who missed two or more consecutive scheduled assessments before death or PD are considered to have had an event, with the earlier date of death or progression as event date .
3. **Mis-stratification:** In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the stratification factors collected at eCRF will be used as strata in both stratified log-rank test and stratified Cox proportional hazards regression model.

Confirmed Objective Response Rate as Assessed by Investigator

Analysis set: ITT Analysis Set

Confirmed ORR per investigator is defined as the proportion of subjects with a confirmed CR or PR per RECIST v.1.1 per assessment by investigator. Only response assessments before the first documented PD or new anti-cancer therapies will be considered. For a response to be considered as confirmed, the subsequent response needs to be at least 4 weeks after the initial response. The minimum criteria for SD duration are defined as ≥ 5 weeks after the date of randomization. A subject will have a best overall response of SD if there is at least one SD assessment that meet the minimum criteria for SD duration and the subject does not qualify for CR or PR. Subjects with unconfirmed CR or PR will be considered to have SD as their best overall response on the date of the corresponding unconfirmed CR and PR if the minimum criteria for SD duration is met.

The primary analysis of confirmed ORR will be conducted for subjects in the ITT analysis set with measurable disease at baseline. The ORR and its 2-sided 95% exact CI will be calculated for each treatment arm using the Clopper-Pearson method (Clopper 1934).

Comparison in confirmed ORR between treatment arms will be performed using a Cochran-Mantel-Haenszel (CMH) chi-squared test stratifying for the stratification factors. The common odds ratio across strata will be estimated with its 95% CI and 2-sided p-value for testing of superiority of TV over chemotherapy.

The testing strategy for the efficacy endpoints OS, PFS and confirmed ORR is described in Section 6.7. The hypothesis testing on confirmed ORR will only occur after the efficacy boundary is crossed for PFS.

Sensitivity Analyses for Confirmed ORR

Mis-stratification: In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the stratification factors collected at eCRF will be used as strata in the CMH test.

7.5.3. Other Secondary Efficacy Endpoints

Duration of Response (DOR)

Analysis set: ITT Analysis Set (Responders)

DOR is defined as the time from the date of the first confirmed objective response (CR or PR that is subsequently confirmed) to the date of the first documented PD per RECIST v1.1 or death from any cause, whichever occurs first. The same derivation of PD date and censoring rules as described in Table 1 for primary PFS analysis will apply for DOR.

Analysis of DOR will include the subgroup of subjects in the ITT analysis set who achieve a confirmed CR or confirmed PR based on investigator assessment. KM curves and the estimates of median DOR with corresponding 95% CIs for each treatment arm will be provided with the two-sided 95% CI using the complementary log-log transformation method (Collett 1994).

Time to Response (TTR)

Analysis set: ITT Analysis Set (Responders)

TTR is defined as the time from the randomization date to the date of the first confirmed objective response (CR or PR that is subsequently confirmed).

Analysis of TTR will include the subgroup of subjects in the ITT analysis set who achieve a confirmed CR or confirmed PR based on investigator assessment. TTR will be summarized using the descriptive statistics.

7.6. Safety Analyses

The safety Analysis Set will be used to summarize all safety endpoints. If notable imbalance in one or more baseline characteristics are observed between treatment arms, subgroup analyses may be performed to evaluate their potential impact on safety.

Adverse events will be coded using Medical Dictionary for Regulatory Activities (MedDRA, version 24.0 or higher).

Laboratory values will be graded using the National Cancer Institute's Common Toxicity Criteria for Adverse Events (CTCAE, version 5.0 or higher).

Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary (version 2019Mar B3 or higher).

7.6.1. Adverse Events

Analysis set: Safety Analysis Set

AEs will be summarized by MedDRA preferred term (PT) in descending frequency of occurrence unless otherwise specified. For incidence reporting, if a subject reports more than one AE that was coded to the same system organ class (SOC) or PT, the subject will be counted only once for that specific SOC or PT. For summaries by severity, only the worst grade for an AE will be counted for a particular subject.

A treatment-emergent AE (TEAE) is defined as a newly occurring or worsening AE after the first dose of study treatment and with onset date on or before 30 days after the last dose of study treatment. See Appendix B for details regarding treatment-emergent classification.

Overall summaries of TEAEs by treatment arm will be presented. In addition, the following summaries of AEs will be provided by PT, unless otherwise specified:

- All TEAEs
- TEAEs by PT
- TEAEs by SOC and PT
- TEAEs by SOC, PT, and maximum severity. At each SOC or PT, multiple occurrences of events within a patient are counted only once at the highest severity.
- Grade 3 to 5 TEAEs
- Grade 3 or 4 TEAEs
- Treatment-emergent Serious Adverse Events (SAEs)
- Treatment-emergent SAEs by SOC and PT
- TEAEs leading to dose interruption/ dose reduction/dose delay/study treatment discontinuation
- Grade 5 TEAEs
- Treatment-related TEAEs
- Treatment-related Grade 3 to 5 TEAEs
- Treatment-related Grade 3 or 4 TEAEs
- Treatment-related treatment-emergent SAEs
- Treatment-related TEAEs leading to study treatment discontinuation
- Treatment-related Grade 5 TEAEs

In addition, exposure-adjusted analyses may be performed to account for imbalanced treatment exposure between the two treatment arms if appropriate.

All AEs, SAEs, AEs leading to treatment discontinuation, and AEs leading to death will be listed.

Adverse Events of Special Interest

Treatment-emergent adverse events of ocular AEs, peripheral neuropathy (PN), and bleeding will be considered adverse events of special interest (AESIs). The search criteria for AESIs will be maintained in a separate document and will be finalized prior to database lock.

Incidence of AESIs will be summarized by treatment arm. AESIs will be summarized by PT and maximum severity. In addition, serious AESIs, AESIs that are related to study drug, lead to dose modification will be summarized.

For AESIs, time to onset, improvement, or resolution will be summarized by descriptive statistics (mean, standard deviation, median, minimum, and maximum). Time to onset will be summarized at the subject level. Time to resolution or improvement will be summarized at the event level.

- Time to onset of a specific AESI is defined as time (days) from the date of the first dose to the start date of the first treatment emergent event (start date of first event – date of first dose + 1).
- Resolution is defined as an event status of 'recovered/resolved' or 'recovered/resolved with sequelae'. Time to resolution is defined as time (days) from the start date of the event to the end date of the same event (end date of the event – start date of the event + 1).
- For events that are not resolved, improvement is defined as a decrease by at least one grade from the worst grade as of the latest assessment. The date of improvement is defined as follows: for events that did not resolve and decrease by one grade or more from the worst post-baseline severity as of the last recorded severity (i.e., severity did not subsequently worsen), the date of improvement is the date of severity improvement. Time (days) to improvement is defined as time from the worst grade of the event to the date that the event is improved.
- Time to resolution is computed from the start date of the first treatment emergent episode of the event or start date of newly onset event after the first dose of treatment drug to the date of resolution. Time to improvement is computed from the start date of the worst grade of the event.

Infusion-related Reactions

Infusion-related reactions (IRR) are defined as any event indicated as IRR by the investigator. The incidence of IRR will be summarized by preferred term and severity. The incidence of IRR leading to dose modification will be summarized.

Ophthalmological Exam and Eye Examination Data

The data for ophthalmological exam and eye examination at baseline and other applicable visits will be listed. Summary statistics of ophthalmological exam and eye examination results will be tabulated by treatment arm and visit where appropriate. Shift tables comparing the post-baseline to baseline results will be presented by treatment arm. Summaries may include but are not limited to visual acuity, Schirmer's test, slit lamp test, and inspections of the conjunctivas and corneas including staining, intraocular pressure, and fundoscopy.

7.6.2. Clinical Laboratory Parameters

Analysis set: Safety Analysis Set

Clinical laboratory results will be presented in standardized units. Analysis of clinical laboratory data will include data up to 30 days after the last dose of study treatment or the EOT visit, whichever is later, or all available data for subjects who were still on treatment at the time of analysis.

For selected laboratory tests (hematology, serum chemistry, and coagulation), both observed values and changes from baseline will be summarized with descriptive statistics by treatment arm at each scheduled visit.

Shift from baseline to the maximum post-baseline NCI CTCAE grade will be summarized for each laboratory test by treatment arm with number and percentage of subjects with Grade 1, 2, 3, 4. Shift tables will be based on the subjects in safety analysis set with the baseline and at least one post-baseline laboratory value.

Treatment-emergent laboratory abnormalities are defined as values that increase at least 1 toxicity grade from baseline at any post-baseline time point. If the relevant baseline laboratory value is missing, then any abnormality of at least Grade 1 observed within the time frame specified above will be considered treatment emergent. Treatment-emergent laboratory abnormalities will be summarized by treatment arm.

Laboratory results and NCI CTCAE grades for hematology, serum chemistry, and coagulation will be presented by subject in data listings. Normal ranges will be documented and out-of-range values will be flagged. A separate listing of laboratory results with CTCAE grade 3 or higher will be presented.

Liver Safety Assessment

The liver safety assessment will be summarized based on the measurements of alkaline phosphatase (ALP), alanine transaminase (ALT), total bilirubin, aspartate transaminase (AST), and their combination as defined below. The number and percentage of subjects meeting the following criteria post-baseline will be summarized by treatment arm.

- ALT: >3x upper limit of normal (ULN), >5xULN, >10xULN, or >20xULN
- AST: >3xULN, >5xULN, >10xULN, or >20xULN
- ALT or AST: >3xULN, >5xULN, >10xULN, or >20xULN
- ALP: >1.5xULN
- Total bilirubin: >2xULN
- (ALT or AST >3xULN) and Total bilirubin >2xULN*
- (ALT or AST >3xULN) and Total bilirubin >2xULN and ALP <2xULN*

*Combination of values measured on the same day or within 1 day apart.

In addition, subjects with the post-baseline liver function test results that are consistent with the Hy's law criteria will be listed.

7.6.3. ECOG Performance Status

Analysis set: Safety Analysis Set

Shifts from baseline to the best and worst post-baseline score will be tabulated by treatment arm.

A by-subject listing of ECOG performance status will be provided.

7.6.4. Vital Signs

Analysis set: Safety Analysis Set

For vital signs (including temperature, blood pressure, respiratory rate, and heart rate), the baseline, on-treatment and changes from baseline will be summarized with descriptive statistics by treatment arm at each scheduled visit.

A by-subject listing of vital signs will be provided.

7.6.5. Electrocardiogram

12-lead Electrocardiograms (ECGs) will be collected at screening, at the end-of-treatment and as clinically indicated during the treatment period. A by-subject listing of ECG parameters will be provided.

7.6.6. Concomitant Medications

Analysis set: Safety Analysis Set

Summaries of concomitant medications will include the number and percentage of subjects by WHO Drug Anatomical Therapeutic Chemical (ATC) classification level 2 (therapeutic subgroup) and level 4 (chemical subgroup) and preferred term by treatment arm. Multiple occurrences of the same medication within a subject will be summarized only once. If any concomitant medication is classified into multiple ATC classes, the medication will be summarized separately under each of these ATC classes. The summary tables will be sorted on ascending alphabetical order of drug class and then decreasing frequency of drug name in a given drug class. Listing of concomitant medications by subject will also be provided.

In addition, concomitant medications for treatment of AESIs will be presented in separate listings and summarized in tables.

7.6.7. Pregnancy

Positive pregnancy test will be listed by subject.

7.7. Pharmacokinetics and Immunogenicity Endpoints

7.7.1. Pharmacokinetics

Analysis set: PK Analysis Set

Tisotumab vedotin antibody-drug conjugate (ADC), total antibody (TAb), and MMAE

concentrations will be summarized with descriptive statistics at each PK sampling time point.

7.7.2. Antidrug Antibody (ADA) Incidence Rate

Analysis set: Safety Analysis Set

The ADA incidence rate is defined as the proportion of subjects that develop ADA as defined confirmed positive at any time during the study. Titers will be determined as surrogate ADA concentration.

ADA incidence will be summarized for all treated subjects who received TV.

A by-subject listing for ADA status at each time point and the titer for subjects with positive

ADA status will be provided.

7.7.3. Pharmacodynamic and Mechanism of Action Biomarkers

The analyses for pharmacodynamic biomarkers and for biomarkers related to drug mechanism(s) of action will be described in a separate Biomarker Analysis Plan.

7.8. Patient Reported Outcomes

Electronic PRO (ePRO) assessments will include the EORTC Quality of Life Core 30 (QLQ-C30), EORTC-QLQ-CX24, and EQ-5D (EQ-5D-5L and EQ VAS) questionnaires.

7.8.1. EORTC QLQ-C30

Analysis set: PRO FAS

The EORTC-QLQ-C30 questionnaire is composed of 30 questions for which the answers ranges either from 1 (Not at all) to 4 (Very much) for items 1 to 28, or from 1 (Very poor) to 7 (Excellent) for items 29 to 30.

The EORTC QLQ-C30 scale scores will be calculated using the EORTC QLQ-C30 Scoring Manual (Fayers 2001). These include 5 functional scales, 3 symptom scales, a global health status/QoL scale, and 6 single items. Each of the multi-item scales includes a different set of items (i.e., no item occurs in more than one scale).

All of the scales and single-item measures range in score from 0 to 100, with a high scale score representing a higher response level (e.g., a high level of functioning, a high QoL, or a high level of symptomatology/problems).

The 15 scores computed from the EORTC-QLQ-C30 questionnaire will be summarized by treatment arm using descriptive statistics at baseline, and at each scheduled visit for actual values and changes from baseline.

In addition, compliance and completeness to the questionnaire, as defined below, will also be summarized by treatment arm. Compliance will be presented overall and by visit, while completeness will only be presented by visit.

Compliance (percentage) will be computed on a subject level as the number of visits for which the subject filled in the questionnaire (at least one item answered) divided by the number of visits scheduled for PRO data taking into account which visit the subject reached. Of note, all screening visits have been pooled as one visit.

Compliance will also be computed on a visit level as how many subjects filled in the questionnaire (at least one item answered) divided by the total number of subjects who reached that visit.

Completeness will be computed on a visit level as how many subjects filled in the questionnaire (at least one item answered) divided by the total number of subjects in the FAS population, regardless of subject drop-outs.

For a particular scale, if at least half of the items in a scale have been answered for a timepoint, then the score will be calculated using the average of all items that were completed; otherwise, the scale score will be set to missing. For single-item measures, if the item is missing, the scale score will be set to missing (Fayers 2001).

A by-subject listing of answers to the EORTC QLQ-C30 questionnaire and scales scores will be provided.

7.8.2. EORTC-QLQ-CX24

Analysis set: PRO FAS

The EORTC QLQ-CX24 questionnaire is meant for use among cervical cancer subjects varying in disease stage and treatment modality. It should always be complemented by the QLQ-C30.

The EORTC-QLQ-CX24 questionnaire is composed of 24 questions for which the answers range from 1 (Not at all) to 4 (Very much).

Four functional scales and 5 symptom scales will be calculated using the EORTC QLQ-CX24 Scoring Manual (Fayers 2001). The 9 scores computed from the EORTC-QLQ-CX24 questionnaire will be summarized by treatment arm using descriptive statistics at baseline, and at each scheduled visit for actual values and changes from baseline. In addition, compliance and completeness to the questionnaire will also be summarized as described in the previous Section 7.8.1.

Details of handling missing values in calculating scale scores are described in the previous Section 7.8.1.

A by-subject listing of answers to the EORTC QLQ-CX24 questionnaire and scales scores will be provided.

7.8.3. EuroQol-5D-5L

Analysis set: PRO FAS

The EQ-5D-5L questionnaire is a 5-item self-reported measure of functioning and well-being, which assesses 5 dimensions of health, including mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension comprises 5 levels (no problems, slight problems, moderate problems, severe problems and extreme problems). Responses to the 5 items are then converted to a weighted health state index (utility score) based on values derived from general population samples (Herdman 2011). This health utility score is between 0 and 1, where 0 is death and 1 is perfect health. In addition to the utility score, this

questionnaire also records the respondent's self-rated health status on a vertical graduated (0 to 100) VAS.

The health utility score, VAS and dimension scores will be summarized by treatment arm using descriptive statistics at baseline, and at each scheduled visit for actual values and changes from baseline. In addition, compliance and completeness to the questionnaire will also be summarized as described in the previous Section 7.8.1.

Details of handling missing values in calculating scale scores are described in the previous Section 7.8.1.

A by-subject listing of the EQ-5D-5L questionnaire and scales scores will be provided.

7.9. Additional Analyses

7.9.1. Deaths

Analysis set: ITT Analysis Set

The number of total deaths, deaths that occur within 30 days of last study treatment, and deaths that occur more than 30 days after last study treatment will be summarized by treatment arm and total. In addition, the primary cause of death will be summarized by treatment arm and total.

A by-subject listing of death information will be provided.

7.9.2. Subsequent Cancer-Related Therapy

Analysis set: ITT Analysis Set

The number and percentage of subjects who receive subsequent cancer-related therapies, including palliative radiotherapy and systemic therapy for progressive disease will be summarized by treatment arm.

In addition, time from last dose of study treatment to the first subsequent cancer-related therapy and time from last dose of study treatment to the first subsequent systemic cancer-related therapy for progressive disease will be summarized by treatment arm.

A by-subject listing of subsequent cancer-related therapy will be provided.

8. INTERIM ANALYSIS

One interim efficacy analysis for the primary endpoint OS is planned when approximately 252 (75%) of the total expected 336 OS events occur. The IDMC will evaluate the overall safety and efficacy data at the IA to make recommendations regarding whether to terminate or continue the trial. Overall type I error for OS will be controlled using the O'Brien-Fleming spending function (Lan 1983).

9. CHANGES FROM PLANNED ANALYSES

9.1. Changes from the Original SAP

9.1.1 Version 2 Changes

Section# and Name	Description of Change
Section 5 Analysis set	Clarified the definition of the analysis sets
Section 6.4 Data transformation and derivation	Clarified study day definition and clarified PD backdating rule for new lesion as specified in RECIST 1.1
Section 6.5 Handling of dropouts and missing data	Removed sentence regarding PK measurements
Section 6.8 Examination of subgroups	Updated list of subgroups
Section 7.2 Demographic and baseline characteristics	Updated list of disease specific characteristics
Section 7.4 Treatment administration	Clarified duration of treatment and IDI
Section 7.5.1 Primary endpoint	Clarified on the OS censoring
Section 7.5.2 Key secondary endpoints	Clarified a minimum required length for SD as a best overall response
Section 7.6.1 Adverse events	Clarified on algorithm to compute date of improvement Updated summary of TEAE presentations Clarified dose modification definition Removed a sentence that is not applicable to AESI
Section 7.7.1 Pharmacokinetics	Updated the analysis summary
Section 9 Changes from planned analyses	Replaced original Section 9.1 Changes from the original protocol with new Section 9.1 Changes from the original SAP
Appendix F Statistical analysis plan for the China population	Added
Appendix G Imputation of death date	Added

9.1.2 Version 3 Changes

Section# and Name	Description of Change
Cover page	Updated protocol version and date Added a note that as of 14-Dec-2023, Seagen Inc. became a part of Pfizer Inc
Signature page	Updated SAP version and date Updated list of individuals to review and approve the SAP
Footer	Updated SAP version and date
Appendix F Statistical analysis plan for the China population	Added the analysis timing for the China population

10 REFERENCES

Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365-76.

Collett D (1991). Statistical inference for binary data. Modelling Binary Data. London, Chapman and Hall: 23-26.

Collett D (1994). Interval-censored survival data. Modelling survival data in medical research. Boca Raton, Fla., Chapman & Hall/CRC: 237-251.

Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.

Fayers PM. Interpreting quality of life data: population-based reference data for the EORTC QLQ-C30. Eur J Cancer. 2001;37(11):1331-4.

Greimel ER, Kuljanic Vlasic K, Waldenstrom AC, et al. The European Organization for Research and Treatment of Cancer (EORTC) Quality-of-Life questionnaire cervical cancer module: EORTC QLQ-CX24. Cancer. 2006;107(8):1812-22.

Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.

Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika. 1983;70(3):659-63.

Maurer W, Bretz F. Multiple Testing in Group Sequential Trials Using Graphical Approaches. Statistics in Biopharmaceutical Research. 2013;5(4):311-20.

Robins JM and Finkelstein DM (2000). Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics 56: 779-88.

Robins JM and Tsiatis AA (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Commun. Statist. - Theory Meth 20: 2609-2631.

Royston P and Parmar MK (2011). The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. Stat Med 30: 2409-21.

Uno H, Claggett B, Tian L, Inoue E, Gallo P, Miyata T, Schrag D, Takeuchi M, Uyama Y, Zhao L, Skali H, Solomon S, Jacobus S, Hughes M, Packer M and Wei LJ (2014). Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol 32: 2380-5.

Westfall PH and Krishen A (2001). Optimally weighted, fixed sequence and gatekeeper multiple testing procedures. Journal of Statistical Planning and Inference 99: 25-40.

APPENDIX A. IMPUTATION OF PARTIALLY UNKNOWN ADVERSE EVENT DATES

For an adverse event (AE) with a partial start or end date, if it can be determined that the event occurred prior to the date of first dose of study treatment, the partial date will not be imputed; Otherwise, the partial date will be imputed using the rules described below. AE start dates should be imputed before imputation of AE condition end date in all cases.

Incomplete AE Start Date:

AE day only is missing

- If the month/year is the same as the month/year of first dose of any study treatment, then AE start date will be imputed as the first dose date of any study treatment.
- If the month/year is after the month/year of first dose of any study treatment, then AE start date will be imputed as the first day of the month.

AE day and month are missing

- If the year is the same as the year of first dose of any study treatment, then AE start date will be imputed as the first dose date of any study treatment.
- If the year is after the year of first dose of any study treatment, then AE start date will be imputed as January 1st.

AE day, month and year are missing

- AE start date will be imputed as the first dose date of any study treatment.

If AE condition end date is not missing, and the imputed start date is after the end date, the start date will be set to the AE condition end date.

Incomplete AE End Date:

- If AE outcome is “not recovered/resolved”, “unknown”, or blank, then AE condition end date will not be imputed.

- If AE outcome is “recovering/resolving”, “recovered/resolved”, “recovered/resolved with sequelae”, or “fatal”, then apply the following:

AE day only is missing

- AE condition end date will be imputed as the minimum of (death date, DCO date or data extraction date, last day of the end date month/year, end-of-study (EOS) date)

AE day and month are missing

- If the year is equal to the year of the last dose date, then AE condition end date will be imputed as the minimum of (last dose date + 30, death date, DCO date or data extraction date, December 31st of the end date year, EOS date).

- If the year is not equal to the year of the last dose date, then AE condition end date will be imputed as the minimum of (death date, DCO date or data extraction date, December 31st of the end date year, EOS date).

AE day, month and year are missing

- AE condition end date will not be imputed.

Within a single record, if the imputed stop date is before the start date, then the imputed stop date will be equal to the start date.

APPENDIX B. DEFINITION OF THE TERM "TREATMENT-EMERGENT" WITH RESPECT TO AE CLASSIFICATION

The algorithm below should be used to determine whether an adverse event (AE) is classified as a treatment-emergent adverse event (TEAE). A TEAE is defined as any AE which is newly occurring or worsening in severity, where newly occurring means that the AE was not present at baseline.

With the CRF data variable, TEAE Flag = Yes if records from the AE CRF page meet all three conditions listed below:

1. Onset Period = Started after first dose of any study treatment
2. AE Start Date on or after first dose of study treatment
3. For subjects who have discontinued the study treatment, AE Start Date $\leq$ 30 days after last dose of study treatment

If subjects are still on treatment (no EOT data), then count all AEs satisfying conditions 1 and 2 as TEAE.

In case the starting date of AE is missing, count the event as a TEAE if Onset Period = Started after first dose of any study treatment.

APPENDIX C. IMPUTATION OF PARTIAL MISSING DATES FOR DISEASE DIAGNOSIS

Missing disease diagnosis dates will be imputed as the first day of the month if both month and year are present and only day is missing.

APPENDIX D. RECIST VERSION 1.1 SUMMARY

Response Evaluation Criteria in Solid Tumors

Term	Definition
Complete response (CR)	Disappearance of all target lesions. Any pathological lymph nodes must have reduction in short axis to <10 mm.
Partial response (PR)	A ≥30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.
Progressive disease (PD)	At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 0.5 cm. The appearance of one or more new lesions is also considered progression.
Stable disease (SD)	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.
Measurable lesion	Must be accurately measured in at least one dimension (longest diameter in the plane of measurement is to be recorded) with a minimum size of 10 mm by CT scan (CT slice thickness no greater than 5 mm).

From RECIST Version 1.1 (Eisenhauer 2009)

APPENDIX E. IMPUTATION OF PARTIAL MISSING DATES FOR SUBSEQUENT CANCER-RELATED THERAPY

Subsequent anticancer therapy start date will be imputed if both month and year are present and only day is missing.

If the year of the subsequent anticancer therapy start date is the same as the year of the EOT date,

- If the month of subsequent anticancer therapy start date is the same as the month of the EOT date, then its start date will be imputed as the EOT date.
- If the month of the subsequent anticancer therapy start date is later than the month of the EOT date, then its start date will be imputed as the first day of the month.
- If the month of the subsequent anticancer therapy start date is earlier than the month of the EOT date, then the subsequent anticancer therapy start date will be imputed as the last day of the month.

If the year of the subsequent anticancer therapy start date is later than the year of the EOT date, then the subsequent anticancer therapy start date will be imputed as the first day of the month.

If the EOT date is missing, then the EOT date will be EOS date or 30 days after the last dose of any study drug, whichever is earlier.

APPENDIX F. STATISTICAL ANALYSIS PLAN FOR THE CHINA POPULATION

This section outlines the statistical analysis strategy and procedures for analyses of the China population, which is required by local regulations to support potential marketing authorization in China.

After trial enrollment of the primary study population (global portion) is complete, subjects from China will continue to be randomized at a 1:1 ratio into either the experimental arm or the control arm in China extension until the sample size for the China population reaches approximately 15% of global population (~74 subjects). The China portion will refer to the China extension after the completion of global portion. The China portion is intended to be identical to the global portion of the trial (e.g., inclusion and exclusion criteria, primary and secondary endpoints, study procedures, stratification factors), with the analyses planned for the China population outlined in this Appendix.

China population includes the Chinese subjects in the global portion and the subjects in the China portion. The purpose of this China analyses is to summarize the efficacy and safety in the China population.

All subjects enrolled in China, including subjects enrolled in the global portion and the China portion, will continue their randomized treatment per protocol and continue to be followed up for the purpose of registration in China. The global study is considered completed when the last survival follow-up visit is completed or 5 years after randomization of the last participant in the study, whichever occurs first. The China portion of the study will be considered completed when the above criteria are met for the global study, or approximately 12 months after the last Chinese participant is enrolled, whichever is later. The completion time for the China portion of the study aligns with the global study assumptions.

The analysis for the China population will be conducted when approximately 38 OS events occur in China population, or at approximately 12 months after the last subject is enrolled in the China portion, whichever occurs earlier.

F.1. Hypotheses/Estimation

No hypothesis testing is planned for the China portion of the study.

F.2 The study Endpoints

F.2.1 Primary Endpoints

Primary endpoints are the same as described in Section 3.1.

F.2.2 Secondary Endpoints

Secondary endpoints are the same as described in Section 3.2.

F.2.3 Exploratory Endpoints

Exploratory endpoints are the same as described in Section 3.3.

F.3 Analysis sets

F.3.1 Intent-to-Treatment (ITT) Analysis Set

This analysis set will include all participants who are randomized from China (both the global portion and the China portion of the study). Participants will be included in the treatment group assigned at randomization regardless of the actual treatment received. This patient population will be for the China-specific efficacy analyses.

F.3.2 Safety Analysis Set

This analysis set will include all participants who are randomized from China (both the global portion and the China portion) and who received any amount of the study treatment.

F.4 Planned analysis

Descriptive analyses are planned for the China Population. No formal hypothesis testing is planned.

F.4.1 Efficacy Analyses

Overall Survival

Analyses of OS for China population are the same to that for the global population as described in Section 7.5.1 except for the hypothesis testing components.

Progression Free Survival (PFS) as Assessed by Investigator

Analyses of PFS for China population are the same to that for the global population as described in Section 7.5.2 except for the hypothesis testing components.

Confirmed Objective Response Rate as Assessed by Investigator

Analyses of ORR for China population are the same to that for the global population as described in Section 7.5.2 except for the hypothesis testing components.

Other Secondary Efficacy Endpoints

Analyses of other secondary endpoints for China population are the same to that for the global population as described in Section 7.5.3.

F.4.2 Safety analysis

Safety analyses for the China population are the same to that for the global population as described in Section 7.6 if applicable.

F.4.3 Patient Reported Outcomes

Patient reported outcome analyses are the same for China population to that for the global population as described in Section 7.8.

F.4.4 Pharmacokinetic Analyses

Pharmacokinetic analyses are the same for China population to that for the global population as described in Section 7.7 except for Subsection 7.7.3.

F.5 Interim Analysis

No interim analysis is planned for China population.

F.6 Multiplicity

No multiplicity adjustment will be applied to the analyses for China population because no formal hypothesis testing will be conducted.

F.7 Sample Size for the China portion of study

Approximately 482 participants will be randomized in the global study. Approximately 15% of the planned global study population (n=482) will be enrolled from China. In the event the number of participants enrolled from China is less than 15% of the 482 participants upon the global enrollment completion, enrollment will continue only in China until approximately 15% (~74 Chinese participants) are enrolled.

APPENDIX G. IMPUTATION OF DEATH DATES

Death dates are imputed based on partial dates with only day missing and the alive status of subject is No.

The death date is imputed as the latter of (1st day of the month and year of the partial death date, last-known-alive date).

Signature Page for SGN-CLIN-047520 v1.0

eSignature Approval	PPD 23-Apr-2024 17:32:03 GMT+0000

eSignature Approval	PPD 23-Apr-2024 18:23:19 GMT+0000

eSignature Approval	PPD 23-Apr-2024 19:43:37 GMT+0000

eSignature Approval	PPD 23-Apr-2024 22:45:19 GMT+0000

eSignature Approval	PPD 22-May-2024 17:25:14 GMT+0000

Signature Page for SGN-CLIN-047520 v1.0