STATISTICAL ANALYSIS PLAN

Protocol Number:	SGN22E-003
Protocol Version:	Amendment 8; 15-Feb-2023
Protocol Title:	An open-label, randomized, controlled phase 3 study of enfortumab vedotin in combination with pembrolizumab versus chemotherapy alone in previously untreated locally advanced or metastatic urothelial cancer
Study Name	EV-302
Investigational Product	Enfortumab vedotin
IND Number:	116360
Sponsor:	Seagen Inc. 21823 30th Drive SE Bothell, WA 98021, USA
Collaborators:	Astellas Pharma Global Development, Inc. (APGD) 1 Astellas Way Northbrook, IL 60062, USA Merck Sharp & Dohme LLC 126 East Lincoln Ave., P.O. Box 2000, Rahway, New Jersey 07065 USA

APPROVAL SIGNATURES

Investigational Product:	Enfortumab vedotin
Protocol Number/Version:	SGN22E-003 / Amendment 8
Study Name	EV-302
SAP Version:	4; 22-Jun-2023

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

[redacted]
Seagen Inc.

Date

[redacted]
Seagen Inc.

Date

[redacted]
Seagen Inc.

Date

[redacted]
Seagen Inc.

Date

[redacted]
Seagen Inc.

Date

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

ADC	antibody-drug conjugate
ADI	absolute dose intensity
AE	adverse event
AESI	adverse event of special interest
ALT	alanine aminotransferase
AST	aspartate aminotransferase
ATA	antitherapeutic antibodies
AUC	area under the curve
BICR	blinded independent central review
BPI-SF	Brief Pain Inventory – Short Form
CI	confidence interval
CPS	combined positive score
CR	complete response
CRF	case report form
DCR	disease control rate
DOR	duration of response
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EDC	electronic data capture
EORTC	European Organisation for Research and Treatment of Cancer
EOT	end of treatment
ePRO	electronic patient reported outcome
ER	emergency room
EV	enfortumab vedotin
FA	final analysis
HR	hazard ratio
HRU	health resource utilization
IA	interim analysis
iCPD	confirmed immune progressive disease
IDI	intended dose intensity
IDMC	independent Data Monitor Committee
IPCW	inverse probability of censoring weights
iRECIST	modified RECIST v1.1 for immune-based therapeutics
IRR	infusion related reaction
ITT	intent-to-treat
iUPD	unconfirmed immune progressive disease
IV	intravenous
MedDRA	Medical Dictionary for Regulatory Activities
MMAE	monomethyl auristatin E
MMRM	mixed model for repeated measures
mOS	median overall survival
mPFS	median progression-free survival
mUC	metastatic urothelial cancer
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NE	not evaluable
ORR	objective response rate

OS	overall survival
PD	progressive disease
PD-L1	programmed cell death ligand-1
PFS	progression-free survival
PK	pharmacokinetic
PN	peripheral neuropathy
PR	partial response
PRO	patient reported outcome
PT	preferred term
QLQ-C30	Quality of Life Core 30
QOL	quality of life
RDI	relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumors
RMST	restricted mean survival time
RTSM	Randomization and Trial Supply Management
SAE	serious adverse event
SAP	statistical analysis plan
SD	stable disease
SMQ	standardized MedDRA query
SOC	system organ class
SSQ	sponsor specified query
TEAE	treatment-emergent adverse events
TTPP	time to pain progression
ULN	upper limit of normal
VAS	visual analogue scale
WHO	World Health Organization

1 INTRODUCTION

This document outlines the statistical methods to be implemented within the scope of Protocol SGN22E-003, entitled "An open-label, randomized, controlled phase 3 study of enfortumab vedotin in combination with pembrolizumab versus chemotherapy in previously untreated locally advanced or metastatic urothelial cancer". Results of the proposed analyses will become the basis of the clinical study report (CSR) for this protocol.

The purpose of this statistical analysis plan (SAP) is to provide specific guidelines from which the analysis will proceed. All planned analyses specified in this document 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.

This SAP describes the analysis for the global portion of the study and Japan-specific safety run-in. A separate China SAP describes the analysis for the China population.

2 STUDY OBJECTIVES

2.1 Primary Objectives

- To compare progression-free survival (PFS) between the experimental arm (enfortumab vedotin (EV) + pembrolizumab [Arm A]) and the control arm (gemcitabine + cisplatin or carboplatin [Arm B]) by blinded independent central review (BICR)
- To compare overall survival (OS) between Arm A and Arm B

2.2 Secondary Objectives

- To compare objective response rate (ORR) between Arm A and Arm B by BICR
- To compare time to pain progression (TTPP) from the subject perspective between Arm A and Arm B
- To compare average change in pain from the subject perspective between Arm A and Arm B
- To evaluate PFS between Arm A and Arm B by investigator assessment
- To evaluate ORR between Arm A and Arm B by investigator assessment
- To evaluate duration of response (DOR) between Arm A and Arm B
- To evaluate disease control rate (DCR) between Arm A and Arm B
- To evaluate the impact of study treatment on quality of life (QOL), functioning, and symptoms from the subject perspective
- To evaluate the safety profile of each treatment regimen

2.3 Exploratory Objectives

- To assess PFS, ORR, and DOR per the modified RECIST v1.1 for immune-based therapeutics (iRECIST) in Arm A
- To assess subject reported health resource utilization (HRU)
- To assess the pharmacokinetics (PK) of enfortumab vedotin, monomethyl auristatin E (MMAE)
- To assess the development of anti-therapeutic antibody (ATA) to enfortumab vedotin
- To assess biomarkers of biological activity, resistance, and predictive biomarkers of response

3 STUDY ENDPOINTS

3.1 Primary Endpoints

- PFS per RECIST v1.1 by BICR
- OS

3.2 Secondary Endpoints

- ORR per RECIST v1.1 by BICR
- TTPP
- Mean change from baseline in worst pain at Week 26
- PFS per RECIST v1.1 by investigator assessment
- ORR per RECIST v1.1 by investigator assessment
- DOR per RECIST v1.1 by BICR and by investigator assessment
- DCR per RECIST v1.1 by BICR and by investigator assessment
- Mean scores and change from baseline of the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Core 30 (QLQ-C30)
- Mean scores and change from baseline of the EuroQOL 5-dimensions (EQ-5D-5L) VAS and utility scores
- Type, incidence, relatedness, severity and seriousness of adverse events (AEs)
- Type, incidence and severity of laboratory abnormalities
- Treatment discontinuation rate due to AEs

3.3 Exploratory Endpoints

- PFS per iRECIST by investigator assessment (Arm A only)

- ORR per iRECIST by investigator assessment (Arm A only)
- DOR per iRECIST by investigator assessment (Arm A only)
- Cumulative incidence of HRU as reported by subject
- Plasma or serum concentrations of EV and MMAE
- Incidence of ATA to EV
- Exploratory biomarkers of clinical activity

4 STUDY DESIGN

This is a phase 3 open-label, 2-arm randomized, controlled multicenter study to evaluate the combination of EV + pembrolizumab (Arm A) versus gemcitabine + cisplatin or carboplatin (Arm B), in subjects with previously untreated locally advanced or metastatic urothelial cancer. The study is designed to assess the dual primary endpoints of PFS and OS of Arm A compared to Arm B.

Approximately 860 subjects will be randomized in a 1:1 ratio to Arm A or B with the following stratification factors: cisplatin eligibility (eligible or ineligible), PD-L1 expression (high or low), and liver metastases (present or absent).

Subjects in Arm A will receive EV at 1.25 mg/kg, administered as an IV infusion on Days 1 and 8 of every 3-week cycle, and pembrolizumab 200 mg as an IV infusion on Day 1 of every 3-week cycle, after completion of the EV infusion.

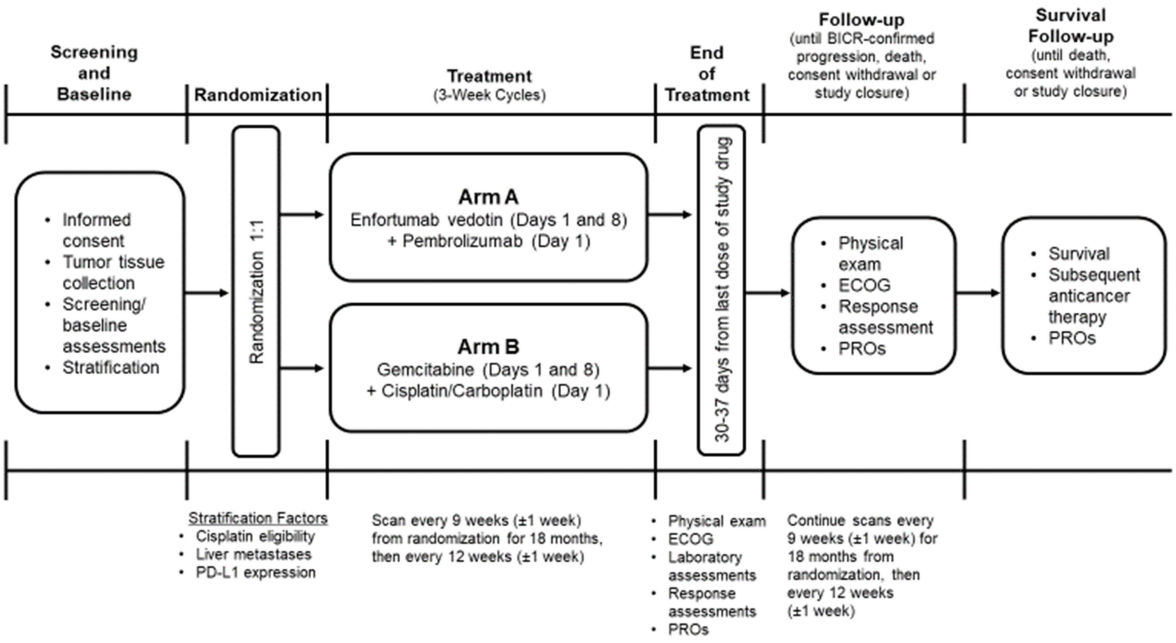
Subjects in Arm B will receive gemcitabine at 1000 mg/m^2 as an IV infusion on Days 1 and 8 of every 3-week cycle, and either cisplatin (70 mg/m^2) or carboplatin (area under the curve [AUC] 4.5 or AUC 5 according to local guidelines) on Day 1 of every 3-week cycle, with adequate pre- and post-hydration, by IV infusion per institutional standards. Crossover to Arm A will not be permitted for subjects in Arm B.

Antitumor activity will be assessed by radiographic imaging at protocol-specified timepoints. Tumor response will be assessed by investigators and BICR according to RECIST v1.1 (Eisenhauer 2009). Subjects may continue on study treatment until progressive disease (PD) (BICR-confirmed or clinical progression), an unacceptable AE, pregnancy, investigator decision, start of a subsequent anticancer therapy, subject decision (non-AE), study termination by the sponsor, completion of the maximum number of drug cycles allowed, or other reason unrelated to an AE. Treatment beyond PD per RECIST v1.1 may be considered for subjects in Arm A who are deriving clinical benefit as defined in the protocol. Subjects treated beyond PD per RECIST v1.1 may continue to be treated until confirmed PD per iRECIST as assessed by the investigator (Seymour 2017).

Safety assessments will include surveillance and recording of AEs, including serious adverse events (SAEs), concomitant medications, physical examination findings, vital signs, protocol-specified laboratory tests, ECOG status, electrocardiograms (ECGs), ophthalmologic examination findings and pregnancy testing.

On a periodic basis, safety will be monitored over the course of the study by an Independent Data Monitoring Committee (IDMC). The IDMC will be responsible for evaluating the results of safety analyses and will make recommendations to the sponsor.

Figure 1: Study design



A Japan-specific safety run-in will assess the safety and PK of the EV + pembrolizumab combination in Japanese subjects with previously untreated locally advanced or metastatic urothelial cancer prior to enrollment of Japanese subjects in the global randomized study. Dose-limiting toxicities (DLTs) will be assessed during Cycle 1 of study treatment.

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.

Arm C (EV + pembrolizumab + cisplatin or carboplatin) was dropped from the study under protocol Amendment 2. Subjects who were randomized to Arm C under the original protocol or protocol Amendment 1 will be included in the analysis sets. Due to small sample size, only limited summary will be presented for Arm C as specified in Section 7. Arm C data will be provided in listings.

5.1 All Enrolled Subjects

All enrolled subjects will include all subjects who signed the inform consent form, meet eligibility criteria, and are randomized in the study or enrolled for Japan-specific safety run-in.

5.2 Intent-to-Treat (ITT) Analysis Set

The ITT analysis set will include all randomized subjects. Subjects will be analyzed according to the treatment arm assigned at randomization regardless of the actual treatment received.

5.3 Response Evaluable Set

The response evaluable set will include all randomized subjects who have measurable disease per RECIST v1.1 at baseline. Subjects will be analyzed according to the treatment arm assigned at randomization regardless of the actual treatment received. The response evaluable set will be used for the primary analysis of response related endpoints, e.g., ORR, DCR, and DOR.

5.4 Safety Analysis Set

The safety analysis set will include all enrolled subjects who receive any amount of study treatment. Subjects will be analyzed according to the actual treatment received. If a subject receives incorrect study treatment for part of the treatment period, the subject will be analyzed under the treatment with the greater number of cycles.

5.5 Patient Reported Outcomes Full Analysis Set (PRO FAS)

The PRO FAS will include all randomized subjects who have received any amount of study treatment and have completed at least one PRO assessment at baseline. Subjects will be analyzed according to the treatment arm assigned at randomization. The PRO FAS will be used for PRO analyses, unless otherwise specified.

5.6 Enfortumab Vedotin Pharmacokinetics Analysis Set

EV PK analysis set will include all enrolled subjects who receive any amount of EV and from whom at least one blood sample has been collected and assayed for EV and MMAE concentration. Corresponding records of the time of dosing and sample collection must also be available. The EV PK analysis set will be used for EV PK analyses.

5.7 DLT Evaluable Analysis Set

DLT evaluable analysis set will include all subjects who are enrolled for Japan-specific safety run-in and receive at least 75% of planned dose of EV and pembrolizumab during Cycle 1 and complete the DLT evaluation period or experience a DLT event during the DLT evaluation period.

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, maximum, 25^{th} and 75^{th} percentile for continuous variables, and the frequencies and percentages for categorical variables.

Unless otherwise specified, confidence intervals (CI) 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 6.4. Other statistical software, if used, will be described in the CSR.

6.2 Determination of Sample Size

The sample size of the study was determined to provide at least 90% power for each of the primary endpoints, at an initially allocated alpha of 0.005 (2-sided) for PFS and 0.045 (2-sided) for OS.

For the OS endpoint, 489 events are required to provide 93% power to detect a hazard ratio (HR) of 0.73 using a log-rank test at an alpha level of 0.045 (2-sided), taking into account one interim analysis at approximately 72.8% of the target number of events. Design assumptions for OS are the following: (1) OS curves follow piecewise exponential distribution with a reduced hazard rate (50% of the initial rate) starting from 24 months (Table 1); (2) hazard ratio of OS is 0.73 between Arm A and Arm B; (3) median OS for Arm B is 15.3 months; (4) an enrollment period of 30 months; (5) a yearly dropout rate of 5%.

Table 1: Hazard rates for OS curves

Period #	Starting at Time (months)	Hazard Rates	
		Arm B	Arm A
1	0	0.045	0.033
2	24	0.023	0.017

Under the above assumptions, approximately 860 subjects will be randomized in a 1:1 ratio to Arm A or Arm B. It is anticipated that 489 OS events will be observed approximately 17 months after the last subject randomized.

For the PFS endpoint, 526 events are required to provide 90% power to detect a hazard ratio of 0.7 at an alpha level of 0.005 (2-sided) under the following assumptions: (1) PFS curves follow piecewise exponential distribution with a reduced hazard rate (20% of the initial rate)

starting from 15 months (Table 2); (2) HR of PFS is 0.70 between Arm A and Arm B; (3) median PFS for Arm B is 7 months; (4) an enrollment period of 30 months; (5) a yearly dropout rate of 5%.

Table 2: Hazard rates for PFS curves

Period #	Starting at Time (months)	Hazard Rates Arm B	Hazard Rates Arm A
1	0	0.099	0.069
2	15	0.02	0.014

6.3 Randomization and Blinding

Subjects will be stratified based on the following stratification factors and randomized in a 1:1 ratio to Arm A or Arm B:

- Cisplatin eligibility (eligible or ineligible)
- PD-L1 expression (high or low)
- Liver metastases (present or absent)

Although this is an open-label study, until database lock and study unblinding for the pre-planned analyses, analyses or summaries by treatment arm will be limited to the IDMC monitoring or to support the information requests from regulatory authorities and be conducted by an external vendor that is independent from the study team. In addition, all imaging for response assessment will be centrally reviewed by the independent radiologists who do not have knowledge of treatment assignment.

6.4 Data Transformations and Derivations

6.4.1 General

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 or randomization for subjects without receiving any study treatment. For efficacy data, baseline values will be the assessments collected at baseline visit. For PRO analysis, baseline values will be the most recent non-missing assessments on or before Cycle 1 Day 1.

If multiple values are qualified 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 receive 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 are randomized but do not receive treatment, study day will be calculated relative to the 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 the EOT visit is performed or the date the decision is made to discontinue study treatment. If the EOT visit is not performed and date of decision made to discontinue study treatment is not available, then the EOT date will be 30 days after the last dose of study treatment or end of study (EOS) date, whichever is earlier.

6.4.2 Response Assessment Dates

At each response assessment time point, scans to evaluate target and non-target lesions can be performed on multiple dates. If the time point response is complete response (CR) or partial response (PR) or not evaluable (NE), then the latest date of all radiologic scans at the given response assessment visit will be the date of assessment. 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).

6.4.3 Adequate Response Assessment

An adequate tumor assessment must include a radiologic scan with the overall disease response of CR, PR, SD, non-CR/non-PD, or PD. Scans with the overall response not evaluable will not be considered an adequate response assessment for the purpose of PFS and DOR censoring.

6.5 Handling of Dropouts and Missing Data

With the exception of 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 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 Appendix B for treatment-emergent classification.

For death date, the start date of subsequent anticancer therapy, and the diagnosis date of locally advanced or metastatic disease, if month and year are present and only day is missing, date may be imputed. Imputation rules are presented in Appendix C

APPENDIX .

Censoring for time-to-event endpoints will be described in Section 7 with each planned analysis, as applicable.

Unless otherwise specified, lab values that are below the lower limit of quantification (LLOQ) will be assigned the value of LLOQ when a numeric value is required for the analysis (e.g., calculating the mean, assigning the CTCAE grade) and be listed as "<LLOQ" in the listings.

ATA titers that do not have numeric titer values (value of "<1") will be assigned to value of 1 when a numeric value is required for the analysis and be listed as "<1" in the ATA listing.

PK concentration values that are below the LLOQ will be assigned half (1/2) of the LLOQ when a numeric value is required for the analysis (e.g., calculating the mean) and be listed as "<LLOQ" in the listings. Summary statistics for a given timepoint will not be calculated when more than 50% of the records with value below LLOQ.

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

6.6 Multicenter Studies

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

6.7 Multiple Comparison/Multiplicity

The study is designed to assess the dual primary endpoints of PFS and OS. To maintain strong control of the family-wise type I error rate at 0.05 (2-sided), a graphical approach with group sequential testing outlined in Maurer and Bretz (Maurer 2013) is used. The initial alpha allocation is 0.005 to PFS and 0.045 to OS as shown in Figure 2. If one of the primary endpoints is statistically significant, the alpha initially assigned to that endpoint can be rolled over to the other endpoint.

The study will test PFS only once at the time of PFS final analysis. The efficacy boundaries at the initially allocated alpha of 0.005 and the updated alpha of 0.05 (if OS is statistically significant) are shown in Table 3.

Table 3: Efficacy boundaries for PFS analysis

	alpha=0.005		alpha=0.05	
Analysis	p-value	Approx. Obs. HR	p-value	Approx. Obs. HR
FA	0.005	0.783	0.05	0.843

The study is planned to have two OS analyses, one interim analysis at the time of PFS final analysis and one final analysis. The efficacy boundaries at the interim and final analyses will be determined using the Lan-DeMets spending function to approximate O'Brien-Fleming

boundaries. Table 4 shows the efficacy boundaries at the initially allocated alpha of 0.045 and the updated alpha of 0.05 (if PFS is statistically significant).

Table 4: Efficacy boundaries for OS analysis

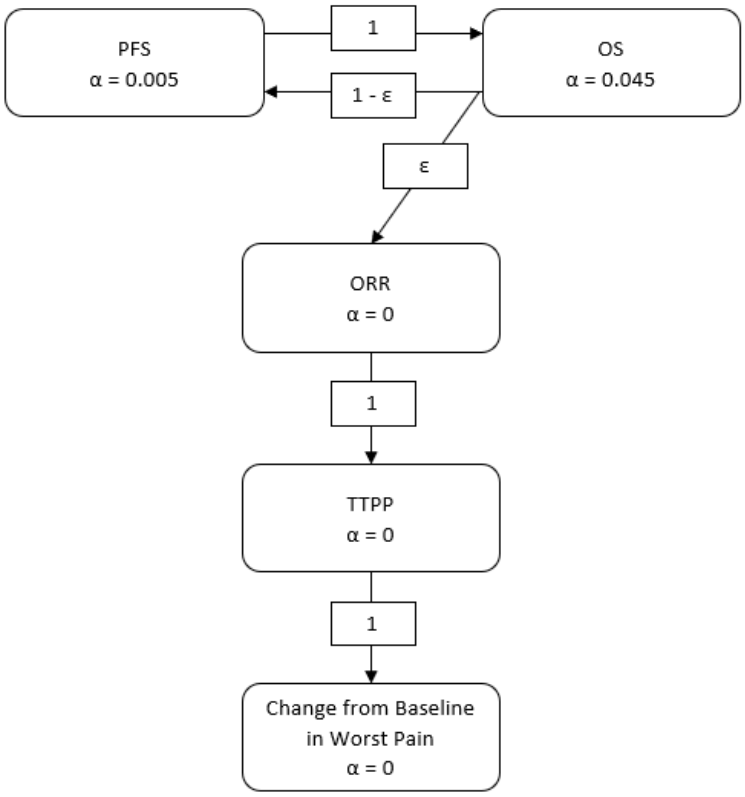
Analysis	alpha=0.045		alpha=0.05	
	p-value	Approx. Obs. HR	p-value	Approx. Obs. HR
IA (Information Fraction: 72.8%)	0.015	0.773	0.017	0.777
FA	0.040	0.831	0.045	0.834

If both PFS and OS are statistically significant, selected secondary endpoints will be tested in the following order using a gatekeeping testing strategy:

1. ORR by BICR
2. TTPP
3. Mean change from baseline in worst pain at Week 26

Each test will be at the 0.05 significance level (2-sided) as long as all preceding null hypotheses are rejected. These secondary endpoints will only be tested once at either the OS interim or final analysis after the null hypotheses for PFS and OS are both rejected, and the test statistics for inferential testing will be computed from the data at the time of OS interim analysis. In the event that only the null hypothesis for PFS is rejected and the superiority boundary for OS has not been crossed at the time of OS interim analysis, these secondary endpoints will not be tested and will only be analyzed descriptively.

Figure 2: Graphical Illustration of Multiplicity Adjustment



ε is a positive number close to zero (i.e., a negligible amount), indicating the potential to pass alpha to ORR only if both PFS and OS are statistically significant.

6.8 Examination of Subgroups

As supportive analyses, treatment effect for PFS, OS and ORR may be estimated and presented in forest plots for the following subgroup variables:

- Stratification factors:
    - Cisplatin eligibility (eligible, ineligible)
    - PD-L1 expression (low [CPS < 10], high [CPS ≥ 10])
    - Liver metastases status (present, absent)
- Age (<65, ≥65 years old)
- Region (North America, Europe, Rest of World)
- Sex (female, male)
- Race (white, non-white)
- ECOG performance status (PS) at baseline (0, 1-2)
- Visceral metastases vs. lymph nodes only metastases

- Primary disease site of origin (upper tract, lower tract)
- Renal function (normal, mild, moderate, severe)

Estimated treatment effect (with a nominal 95% CI) for PFS and OS will be provided for each subgroup using a stratified Cox proportional hazards regression model controlling stratification factors. If the subgroup is a stratification factor, then the stratified models will control for the rest of stratification factors. For ORR, the difference between treatment arms (with a nominal 95% CI) will be calculated for each subgroup. In addition, KM curves will be generated by arm for selected subgroups (e.g., cisplatin eligibility and PD-L1 expression).

A subgroup analysis may not be performed 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 may be considered when there is no sufficient sample size within one level.

6.9 Covariates

Stratified analyses will control for the stratification factors as recorded at randomization per Randomization and Trial Supply Management (RTSM).

6.10 Timing of Analyses

The PFS final analysis will be triggered when approximately 526 PFS events or 356 OS events (72.8% information fraction) in the ITT analysis set have occurred, whichever is later.

An OS IA is planned at the time of the PFS final analysis. At IA, the number of OS events will be approximately 356 (about 72.8% of information). The OS final analysis is triggered when approximately 489 OS events have occurred in the ITT analysis set. If OS is statistically significant at IA, this interim OS analysis will become the final OS analysis.

A summary of analyses timing at the planned PFS and OS analyses are presented in Table 5. The results are based on the initial alpha allocated to each endpoint before any alpha recycling: 0.005 to PFS and 0.045 to OS.

Table 5: Summary of timing of analyses

Analysis		2-sided alpha[d]	Est. Time after LPI	Planned # of Events	Power at Initial Alpha (0.005 to PFS, 0.045 to OS)	Power at Updated Alpha (0.05 to PFS and OS)
Analysis Time 1[a]	FA: PFS	0.005	7 months	526	90%	98%
	IA: OS	0.015	7 months	356 (72.8% info.)	70%[b]	72%[b]
Analysis Time 2	FA: OS	0.045	17 months	489	93%[c]	93%[c]

FA = Final Analysis; IA = Interim Analysis; LPI = Last Patient In

a Analysis Time 1 is triggered by the planned number of PFS events in total (526 events), or 356 OS events (72.8% information fraction), whichever is later.

b Probability to cross the efficacy boundary at the interim analysis

c Overall power at the final OS analysis

d Cumulative alpha assigned for each test without alpha roll-over. For OS IA, 0.015 is the estimated alpha spent at IA.

7 PLANNED ANALYSES

7.1 Disposition

Analysis set: ITT Analysis Set

A summary of study subjects by disposition will be tabulated by treatment arm and total. Reasons for discontinuation of treatment and study will be summarized. The number and percentage of subjects in follow up will be summarized.

The number and percentage of subjects who signed informed consent and the number of subjects in each analysis set will be summarized. The reasons for screen failure will be summarized and listed, if applicable.

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

7.2 Demographic and Baseline Characteristics

Analysis set: ITT Analysis Set

Demographics and baseline characteristics, including age, sex, ethnicity, race, baseline weight, body mass index, ECOG performance status, estimated creatinine clearance, hepatic function, hemoglobin A1c, smoking status, and Bajorin risk factors will be listed and summarized; summaries will be presented for each arm and the total.

Disease specific characteristics, including time from diagnosis of locally advanced or metastatic disease to randomization, histology, primary tumor location, disease status at randomization, disease stage, and sites of metastatic disease will be listed and summarized for each treatment arm and the total. Prior cancer treatment will be listed and summarized for each treatment arm and the 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 by treatment arm. Summary statistics for duration of treatment, number of cycles, and cumulative dose will be presented for the study treatment and for EV and pembrolizumab as applicable. Relative dose intensity (RDI) will be described for EV only. Listings may be presented as well.

Duration of treatment (except when calculating exposure such as RDI) is defined as the time from first dose date to the earliest of either:

- Day 21 of the last treatment cycle (i.e., Day 1 of the last treatment cycle + 20)
- Date of death or end of study date
- Start of any subsequent anticancer therapy
- Analysis data cutoff date if the subject is still on treatment at the time of data cutoff for an analysis

Cumulative Dose is the sum of the actual dose levels a subject received across all cycles.

Intended Dose Intensity (IDI)

- **EV:** the planned dose is 1.25 mg/kg on Day 1 and Day 8 of every 3-week cycle. Thus, the IDI is 2.50 mg/kg per 3-week cycle.

Absolute Dose Intensity (ADI) = sum of dose strength / duration of treatment (week) * 3 weeks/cycle. ADI will be calculated for EV only, where ADI (mg/kg per 3-week cycle) = Sum of (actual dose in mg / weight in kg) / duration of treatment (in weeks) * 3 weeks/cycle. For ADI calculation, duration of treatment is defined as the time from the first dose of study drug to Day 21 of the last treatment cycle, regardless of if death occurs before the end of the cycle.

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

7.5 Efficacy Analyses

Efficacy endpoints will be summarized using the ITT analysis set or response evaluable set. Efficacy analyses will be performed for Arms A and B. For Arm C, only ORR and DCR will be summarized due to small sample size. Efficacy data of all enrolled subjects will be provided in listings.

For the primary analysis of PFS and OS, the randomization stratification factors per RTSM will be used as strata in stratified analysis. If the total number of subjects in a stratum of a stratification factor is less than 10% of the total population, that stratification factor may be excluded in the stratified analysis. In the case of >5% stratification inconsistency between RTSM and eCRF, as a sensitivity analysis, the hazard ratio and its 95% CI will be estimated using a stratified Cox proportional hazards regression model controlling for the stratification factors based on actual strata information collected at baseline eCRF. In addition, hazard ratios and 95% confidence intervals of PFS and OS will be estimated for each of the subgroups specified in Section 6.7.

Only the adequate tumor assessments on or before the start date for subsequent anticancer therapy will be considered for the primary efficacy analysis. For efficacy analysis, maintenance therapy (e.g., avelumab) that is received after discontinuation or completion of

chemotherapy in Arm B and local therapy (e.g., surgery, palliative radiotherapy) per medical adjudication will not be considered as subsequent anticancer therapy.

7.5.1 Primary Endpoints

The primary endpoints of this study are PFS per RECIST v1.1 by BICR and OS. The study can be considered positive if either PFS by BICR or OS comparison is statistically significant between Arm A and Arm B. See Section 6.7 for multiplicity adjustment for primary hypotheses.

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

7.5.1.1 Progression-free Survival by BICR

Analysis set: ITT Analysis Set

PFS is defined as the time from the date of randomization to the first documented disease progression (as determined by BICR 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 6.

Table 6: Censoring scheme for primary analysis of PFS per RECIST v1.1 by BICR

Scenario	Progression/Censor Date	Outcome
PD or death before subsequent anticancer treatment and not after two or more consecutive missed assessments	Earliest date of PD or death	Event
PD or death immediately after two or more consecutive missed/NE tumor assessments	Date of last adequate tumor assessment prior to the missed visits or date of randomization if no post-baseline tumor assessment	Censored
No PD or death and no post-baseline tumor assessments	Date of randomization	Censored
No PD or death and no subsequent anticancer treatment	Date of last adequate tumor assessment	Censored
Subsequent anticancer treatment started before PD or death	Date of last adequate tumor assessment on or prior to start of subsequent anticancer treatment	Censored

Note: maintenance therapy (e.g., avelumab) that is received after discontinuation or completion of chemotherapy in Arm B and local therapy per medical adjudication will not be considered as subsequent anticancer therapy.

The two arms will be compared for PFS using a stratified log-rank test controlling for the stratification factors. The initial 2-sided alpha level for PFS is 0.005. The alpha recycling algorithm and the efficacy boundaries are described in Section 6.7. KM curves will be generated for each arm. KM estimates of the median PFS and its 2-sided 95% CI will be calculated for each arm. In addition, KM estimates of the 25th and 75th percentiles of PFS and the observed minimum and maximum PFS will be reported.

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

```
** pfstime = PFS time;
** censor = Censor indicator (1 = censored)
** Trt = treatment arm (A vs. B)

** SF1 = Stratification Factor 1

** SF2 = Stratification Factor 2

** SF3 = Stratification Factor 3

ODS OUTPUT HomTests=chisq;

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

To describe the treatment effect, a hazard ratio between Arm A and Arm B 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 = pfsdata;
        CLASS Trt(ref='B');
                STRATA SF1 SF2 SF3;
                MODEL pfstime*censor(1) = Trt / TIES=EFRON RL;
                HAZARDRATIO Trt;
RUN;
```

Proportional hazards (PH) assumption for treatment arms may be examined with the following approach: (other methods may also be applied to check the PH assumption if needed)

- Plot of Schoenfeld residuals for the treatment variable versus time. Schoenfeld residuals including LOESS curve may be plotted in SAS using the PHREG procedure via the OUTPUT statement (keyword=RESSCH). With proportional hazards the LOESS curve is expected to be parallel to the x-axis. Residuals that do not show any trend indicate the PH assumption is met.

Sensitivity Analyses for PFS

Analysis sets: ITT Analysis Set

1. **Unstratified Analysis:** PFS analysis by unstratified log-rank test. Additionally, an unstratified Cox regression model will be used to estimate the hazard ratio and the corresponding 95% CI for the treatment effect. The primary censoring method (Table 6) will be used.

2. **Subsequent therapy before PD/death:** For subjects who received subsequent anticancer therapy before PD or death, a sensitivity analysis without considering subsequent anticancer therapies as a censoring reason will be performed.

3. **Missing Tumor Assessments:** To explore the potential impact of missing tumor assessments on PFS, subjects who miss two or more consecutive scheduled assessments before death or PD are considered to have an event on the date of death or progression.

4. **Mis-stratification:** In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the stratification factors collected in eCRF will be used as strata in both stratified log-rank test and stratified Cox proportional hazards regression model. A listing and summary of the stratification factor values per RTSM and eCRF will be provided.

5. **Non-proportional hazard:** In the case that the PH assumption is violated, PFS 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 PFS event time for Arm A, largest observed PFS event time for Arm B). The difference in RMST between the treatment arms will be analyzed adjusting for the randomization stratification factors. RMST and its 95% CI will be estimated for each treatment arm. The difference in RMST between 2 arms will be provided with 95% CI.

7.5.1.2 Overall Survival

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, OS will be censored at the date the subject is last known to be alive or at the analysis cutoff date, whichever is earlier.

OS analysis will be performed using the same statistical methods as described for PFS primary analyses in Section 7.5.1.1. The Lan-DeMets O'Brien-Fleming approximation spending function will be used to obtain the efficacy boundaries at the IA and final OS analyses. The initial 2-sided alpha level is 0.045 for OS. The alpha recycling algorithm and efficacy boundaries are described in Section 6.7.

PH assumption may be examined using the same method as described in Section 7.5.1.1.

Sensitivity Analyses for OS

1. **Unstratified Analysis:** OS analysis by unstratified log-rank test. Additionally, an unstratified Cox regression model will be used to estimate the hazard ratio and the corresponding 95% CI for the treatment effect.

2. **Mis-stratification:** In the case of >5% inconsistency in stratification factors between RTSM and eCRF, the stratification factors collected in eCRF will be used as strata in both stratified log-rank test and stratified Cox proportional hazards regression model.
3. **Initiation of Subsequent Anticancer Therapy:** The use of subsequent anticancer 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, sensitivity analysis of OS using inverse probability of censoring weights (IPCW) method [Robins and Finkelstein, 2000] will be performed. Subjects who start subsequent therapy will be censored at the time of subsequent anticancer therapy, and subjects are weighted according to their probability to take subsequent therapy. This analysis will be performed at OS interim analysis if the OS primary analysis is statistically significant or at OS final analysis.
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 Arm A, largest observed OS event time for Arm B). The difference in RMST between the treatment arms will be analyzed adjusting for the randomization stratification factors. RMST and its 95% CI will be estimated for each treatment arm. The difference in RMST between 2 arms will be provided with 95% CI.

7.5.2 Secondary Efficacy Endpoints

7.5.2.1 Objective Response Rate

Analysis set: Response Evaluable Set

ORR is defined as the proportion of subjects achieving a confirmed CR or PR per RECIST v1.1. ORR by BICR and by investigator will be analyzed as separate endpoints. Only response assessments before the first documented PD or subsequent anticancer 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. For subjects who do not have confirmed CR or PR, the best overall response is defined as SD if the subject has at least one response assessment of CR/PR/SD (or non-CR/non-PD) at least 6 weeks after the randomization date.

The ORR and its 2-sided 95% exact CI will be calculated for each treatment arm using the Clopper-Pearson method (Clopper 1934). The statistical comparison of ORR by BICR between 2 arms will be performed only if both PFS and OS are statistically significant (Section 6.7), and the comparison will be evaluated at the 0.05 significance level (2-sided) by a Cochran-Mantel-Haenszel (CMH) test controlling for the stratification factors.

The ORR by investigator assessment will not be tested statistically. The nominal p-value from the stratified CMH test will be reported. In addition, the concordance between BICR and investigator assessed responders and non-responders will be summarized.

Time to response will be summarized for responders only. Time to response is defined as the time from the date of randomization to the first documentation of objective response (CR or PR that was subsequently confirmed).

7.5.2.2 Duration of Response

Analysis set: Response Evaluable Set (Responders only)

DOR is defined as the time from the first objective response (CR or PR that is subsequently confirmed) to the first documented PD per RECIST v1.1 or death from any cause, whichever occurs first. DOR will only include subjects with a confirmed response (CR or PR per RECIST v1.1). The same censoring rules as described in Table 6 for primary PFS analysis will be applied for DOR. DOR per the assessment by BICR and by investigator will be analyzed as separate endpoints.

DOR will be displayed by KM plots. KM estimates of median DOR with corresponding 95% CIs for each treatment arm will be provided. In addition, KM estimates of the 25^{th} and 75^{th} percentiles of DOR and the observed minimum and maximum DOR will be reported.

7.5.2.3 Disease Control Rate

Analysis set: Response Evaluable Set

DCR is defined as the proportion of patients with confirmed response (CR or PR) or SD per RECIST v1.1. Subjects who have no post-baseline response assessments will be considered as non-responders for calculating the DCR. Only response assessments before the first documented PD or subsequent anticancer therapies will be considered. DCR per the assessment by BICR and by investigator will be analyzed as separate endpoints.

The DCR and its 95% CIs will be provided using Clopper-Pearson methodology for each treatment arm (Clopper 1934). Comparison of DCR between 2 arms will be evaluated by a 2-sided CMH test controlling for the stratification factors. The nominal p-value from the stratified CMH test will be reported.

7.5.2.4 Progression-Free Survival by Investigator Assessment

Analysis set: ITT Analysis Set

PFS by investigator assessment is defined as the time from the date of randomization to the first documented disease progression per RECIST v1.1 by investigator assessment or death from any cause, whichever occurs first. The same censoring rules as described in Table 6 will be applied for PFS by investigator assessment.

PFS by investigator assessment will be analyzed using the same method as described in Section 7.5.1.1 for the primary analysis of PFS by BICR. Subgroup analyses and sensitivity analyses will not be conducted for PFS by investigator assessment.

In addition, the concordance between BICR and investigator assessed PFS events will be summarized.

7.5.3 Exploratory Efficacy Endpoints

Tumor assessment by iRECIST will be collected for Arm A. Subjects may continue to be treated beyond PD per RECIST v1.1 until confirmed PD per iRECIST guidance (Seymour 2017) assessed by the investigator. All collected iRECIST assessments will be used for iRECIST related analyses, including data after the 1st unconfirmed progression event (iUPD). The following iRECIST endpoints will be derived and analyzed for Arm A.

- PFS per iRECIST (iPFS) by investigator assessment
- ORR per iRECIST (iORR) by investigator assessment
- DOR per iRECIST (iDOR) by investigator assessment

7.5.3.1 PFS per iRECIST

Analysis set: ITT Analysis Set (Arm A only)

PFS by iRECIST is defined as the time from date of randomization to the date of the first iUPD which is subsequently confirmed per iRECIST (iCPD), or to death due to any cause, whichever comes first. Specifically,

iPFS = Date of Event per iRECIST or Censoring date – Date of randomization +1

Following the iRECIST guidance, the first iUPD occurs when the first PD per RECIST v1.1 is declared. The event and censoring rules are defined in Table 7.

iPFS will be analyzed using the KM method. For Arm A, the median will be provided with the 2-sided 95% CI using the complementary log-log transformation method (Collett 1994). In addition, KM estimates of the 25th and 75th percentiles of iPFS and the observed minimum and maximum iPFS will be reported.

Table 7: Event/censoring scheme for analysis of PFS per iRECIST v1.1 (iPFS)

Situation	Date of iPFS Event or Censoring	Outcome
No post-baseline tumor assessments and No death	Date of randomization	Censored
Death before first documented iUPD and subsequent anticancer therapy and not after two or more consecutive missed assessments	Date of death	Event
Documented iCPD (confirmed progression) before subsequent anticancer therapy and not after two or more consecutive missed assessments	Date of the initial iUPD that is subsequently confirmed	Event
Documented iUPD and the subjects is not clinically stable to continue on treatment	Date of iUPD	Event
Documented iUPD without subsequent scans before subsequent anticancer therapy and the subject is no longer in response follow-up	Date of iUPD	Event
Documented iUPD and all subsequent assessments are iUPD(s) or NE(s) and the subject is no longer in response follow-up	Date of iUPD	Event

Situation	Date of iPFS Event or Censoring	Outcome
None of the above and subsequent anticancer therapy started	Date of last disease assessment prior to start of subsequent anticancer therapy	Censored
Death or documented iUPD right after two or more consecutive missed disease assessments	Date of last disease assessment prior to the missed visits	Censored
None of the above	Date of last disease assessment	Censored

Note: maintenance therapy (e.g., avelumab) that is received after discontinuation or completion of chemotherapy in Arm B and local therapy will not be considered as subsequent anticancer therapy.

For illustration purpose, some examples of iPD date determinations are provided below.

Case 1: iPD date = TP1.

Time Point (TP)	**TP1**	TP2
iRECIST Response	**iUPD**	iCPD

Case 2: iPD date = TP1.

Time Point (TP)	**TP1**	TP2	TP3
iRECIST Response	**iUPD**	iUPD	iCPD

Case 3: iPD date = TP4.

Time Point (TP)	TP1	TP2	TP3	**TP4**	TP5
iRECIST Response	iUPD	iPR	iPR	**iUPD**	iCPD

Case 4: iPD date = TP1; no confirming scan is done due to EOS.

Time Point (TP)	**TP1**	TP2
iRECIST Response	**iUPD**	Not Done (due to EOS)

Case 5: iPD date = TP1; neither iCR/iPR/iSD followed the iUPD nor confirmed progression and the subject is off study.

Time Point (TP)	**TP1**	TP2	TP3
iRECIST Response	**iUPD**	iUPD or NE	iUPD or NE, and no subsequent TP due to EOS

Case 6: no iPD is observed, iPFS will be censored at TP3 with iPR

Time Point (TP)	TP1	TP2	**TP3**
iRECIST Response	iUPD	iPR	**iPR**

Case 7: no iPD is observed, iPFS will be censored at TP3 with iUPD if the subject is still on study.

Time Point (TP)	TP1	TP2	**TP3**
iRECIST Response	iUPD	iUPD or NE	**iUPD**

7.5.3.2 ORR per iRECIST

Analysis set: Response Evaluable Set (Arm A)

Objective response per iRECIST is defined as achieving a confirmed CR or PR per iRECIST as assessed by investigator. The iORR and its 2-sided 95% exact CI will be calculated using the Clopper-Pearson method (Clopper 1934).

7.5.3.3 DOR per iRECIST

Analysis set: Response Evaluable Set (Responders in Arm A)

DOR per iRECIST is defined as the time from first documented objective response of iCR or iPR (that is subsequently confirmed) to the first documented PD per iRECIST or to death due to any cause, whichever comes first. iDOR will only be calculated for the subjects in Arm A with a confirmed iCR or iPR. The same censoring rules as described in Table 7 for iPFS will be applied for iDOR.

The KM estimate of median iDOR with corresponding 95% CIs will be estimated. In addition, KM estimates of the 25th and 75th percentiles of iDOR and the observed minimum and maximum iDOR will be reported.

7.6 Safety Analyses

Analyses of safety will be performed by treatment arm using the safety analysis set. If notable imbalances in one or more baseline characteristics are observed between treatment arms, subgroup analyses may be performed to evaluate their potential impact on safety.

Adverse event will be coded using Medical Dictionary for Regulatory Activities (MedDRA, version 23.0 or higher).

Laboratory values will be graded using the National Cancer Institute's Common Toxicity Criteria for Adverse Events (CTCAE, version 4.03 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

Adverse events will be summarized by MedDRA preferred terms (PT) in descending frequency of occurrence in Arm A 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 adverse event (TEAE) is defined as a newly occurring or worsening AE after the first dose of study treatment through 30 days after the last dose of study treatment, or through 90 days after the last dose of study treatment for SAE in arms utilizing pembrolizumab. See Appendix B for details regarding treatment-emergent classification.

Treatment-related adverse events are events assessed by the investigator as 'related' to any of the study drugs: EV, pembrolizumab, cisplatin, carboplatin, or gemcitabine.

An overall summary of TEAEs and treatment-related TEAEs will be provided. Summaries of TEAEs and treatment-related TEAEs by PT will be provided for the following:

- TEAEs
- Grade 3 or higher TEAEs
- Serious TEAEs
- TEAEs leading to interruption of EV
- TEAEs leading to interruption of pembrolizumab
- TEAEs leading to interruption of any study drugs
- TEAEs leading to dose reduction of any study drugs
- TEAEs leading to discontinuation of EV
- TEAEs leading to discontinuation of pembrolizumab
- TEAEs leading to discontinuation of any study drugs
- TEAEs leading to death

In addition, the summary of TEAEs by SOC, PT and maximum severity will be provided.

Exposure-adjusted analyses may be performed to account for imbalanced treatment exposure between Arm A and B as appropriate. The exposure-adjusted TEAEs will be calculated as event rate per patient-year of exposure, and exposure is defined as the time from first dose date to the earliest of last dose date + 30 days, EOS date, death date, or data cutoff date.

All adverse events, grade 3 or higher adverse events, serious adverse events, adverse events leading to treatment discontinuation, and adverse events leading to death will be listed.

7.6.1.1 Adverse Events of Special Interest

Enfortumab Vedotin Adverse Events of Special Interest (AESI): AESI for enfortumab vedotin include peripheral neuropathy, ocular disorders, hyperglycemia, skin reactions, and infusion related reactions (IRR). AESI are medical concepts of composite terms based on the

search criteria (standard MedDRA query [SMQ] or sponsor specified query [SSQ]). The search criteria for AESI will be maintained in a separate document and finalized prior to database lock for the primary analyses.

Treatment-emergent AESI will be summarized by PT and maximum severity. In addition, serious treatment-emergent AESIs, treatment-related treatment-emergent AESI, and AESI leading to dose interruption, reduction and discontinuation will be summarized in the same format as listed in Section 7.6.1 for TEAEs.

For peripheral neuropathy (PN), an analysis by pre-existing condition of PN will also be performed. Hyperglycemia by pre-existing condition of hyperglycemia, baseline BMI (< 30; ≥30 kg/m^2), and baseline HbA1c (<5.7%, ≥5.7% to <6.5%, and ≥6.5%) will be summarized.

For selected AESIs, time to onset, event outcome, and time to resolution will be summarized, as appropriate.

- Time to onset of a specific AESI is calculated as time from the date of first dose to start date of first treatment emergent event that meets the respective search criteria.
- For events with an outcome of 'recovered/resolved' or 'recovered/resolved with sequelae', time to resolution will be calculated as time from the event start date to event end date.

Time to onset will be summarized at the subject level and time to resolution will be summarized at the event level.

Separate listings of AESI will be presented.

Pembrolizumab Adverse Events of Special Interest: This includes immune related adverse events and infusion reactions, following the Keytruda Adverse Events of Special Interest (AEOSI) List (v19.0 or higher).

Overall summary of pembrolizumab AEOSI, and the summary by PT and maximum severity will be provided. In addition, time to onset of first pembrolizumab AEOSI, pembrolizumab AEOSI outcome, and the concomitant systemic corticosteroid use for pembrolizumab AEOSI will be summarized.

7.6.2 Vital Signs

Analysis set: Safety Analysis Set

Vital sign measurements (systolic and diastolic blood pressure, heart rate, and body temperature) will be listed by subject and scheduled visit.

7.6.3 Clinical Laboratory Parameters

Analysis set: Safety Analysis Set

Clinical laboratory results (serum chemistry and hematology) to be collected up to the end of treatment visit will be presented in standardized units.

Median, 25th and 75th percentile of laboratory results will be plotted by treatment arm and scheduled visit up to Cycle 17 (i.e., 1 year of treatment) and EOT. If multiple post-baseline records are in a given scheduled visit for a laboratory parameter, the latest non-missing record will be selected for by visit analysis. Shift from baseline to maximum post-baseline NCI CTCAE grade will be summarized for each lab test by treatment arm with number and percentage of subjects with Grade 1, 2, 3, 4. Treatment-emergent laboratory abnormalities, defined as post-baseline laboratory values of a higher CTCAE grade than the baseline grade, will also be summarized.

Laboratory results and NCI CTCAE grades for hematology and serum chemistry will be presented 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.

7.6.3.1 Liver Safety Assessment

Analysis set: Safety Analysis Set

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 criteria post-baseline will be summarized for each treatment arm.

- ALT: >3xULN, >5xULN, >10xULN, >20xULN
- AST: >3xULN, >5xULN, >10xULN, >20xULN
- ALT or AST: >3xULN, >5xULN, >10xULN, >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.

7.6.4 ECOG Performance Status

Analysis set: Safety Analysis Set

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

7.6.5 Electrocardiogram

Analysis set: Safety Analysis Set

Electrocardiograms will be collected at baseline. A by-subject listing will be generated.

7.6.6 Pregnancy

Positive pregnancy tests will be listed by subject.

7.6.7 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 peripheral neuropathy, skin reactions and hyperglycemia will be presented in separate listings and summarized in tables.

7.6.8 Deaths

Analysis set: Safety Analysis Set

Death information will be listed by subjects. Cause of death due to AE will be summarized by descending MedDRA preferred term (unless otherwise specified).

7.7 Patient Reported Outcomes

Electronic PRO (ePRO) assessments will include the EQ-5D-5L, EORTC QLQ-C30, BPI-SF, and HRU questionnaires. ePROs will be administered at Cycle 1 Day 1 before study treatment, once weekly for the first 12 weeks and once every 3 weeks for the remainder of the study through progression and survival follow-up. Visit windows for the analysis timepoints and selection of analysis record are defined in Appendix E. Selected analysis records will be summarized by analysis timepoint for EQ-5D-5L, EORTC QLQ-C30 and BPI-SF. All records are included in the calculation of cumulative incidence of HRU by analysis timepoint.

Selected PRO-based secondary endpoints on pain will be hierarchically tested after OS, PFS, and ORR are all statistically significant as outlined in Section 7.7.1.1 and 7.7.1.2. See Section 6.7 for further details on hierarchical testing.

For QLQ-C30, EQ-5D-5L, and BPI-SF, completion and compliance rates will be summarized for each scheduled assessment by treatment arm using the ITT analysis set for Arm A and Arm B. Completion rate is defined as the proportion of subjects who completed

the instrument among the ITT analysis set. Compliance rate is defined as the proportion of subjects who completed the instrument among those who are expected to complete at a given visit.

Average score and change from baseline will be reported at each visit for all QLQ-C30 scales and items, BPI-SF domains, and VAS and utility score of EQ-5D-5L (see section 7.7.1, 7.7.2, and 7.7.3). Summary statistics will be calculated for analysis timepoints with at least 20% completion rate in either treatment arm. Additional analyses of these PROs will be described in a supplemental PRO SAP.

7.7.1 BPI-SF

Analysis set: PRO FAS

Descriptively, summary of the worst, least, and average pain experienced in the last 24 hours as well as pain right now and number of pain locations will be provided for each treatment arm. In addition, the summary of the pain severity of the seven interference items to measure pain interference will be provided. All scoring will be completed according to the BPI-SF user manual (Cleeland 2009). At each post-baseline assessment, change in worst/least/average pain, as well as interference scores from baseline will be calculated.

7.7.1.1 Time to Pain Progression

TTPP is defined as the time from the date of randomization to date of pain progression. Pain progression is defined as a subject reporting either of the following, whichever comes first:

1. Increase of 2 or more points from baseline on BPI-SF Question 3 (i.e., pain at its worst in the last 24 hours) maintained for at least two consecutive assessments
2. Initiation of new opioid medication from baseline for pain with usage maintained for at least two consecutive assessments as recorded in BPI-SF Question 7

This definition assumes that an increase of 2 or more points from baseline on the BPI-SF Question 3 constitutes meaningful within-patient change (MWPC).

Censoring scheme for the primary analysis of TTPP is described below in Table 8. For analysis of TTPP, maintenance therapy (e.g., avelumab) that is received after discontinuation or completion of chemotherapy in Arm B will not be considered as subsequent anticancer therapy.

Table 8: Censoring scheme for TTPP analysis

Scenario	Event/Censor Date	Outcome
Pain progression	Date of the first assessment that qualified as pain progression	Event
No post-baseline BPI-SF assessments	Date of randomization	Censored
BPI-SF Question 3 score of 9 or 10 at baseline	Date of randomization	Censored
Death before pain progression	Date of last BPI-SF assessment	Censored

No recorded pain progression and no subsequent anticancer treatment	Date of last BPI-SF assessment	Censored
Subsequent anticancer treatment started before pain progression	Date of last BPI-SF assessment on or prior to start of subsequent anticancer treatment	Censored

Maintenance therapy (e.g., avelumab) that is received after discontinuation or completion of chemotherapy in Arm B will not be considered as subsequent anticancer therapy.

The two arms will be compared for TTPP using a stratified log-rank test controlling for age (<65 or ≥65 years old), sex, region, and randomization stratification factors at 2-sided alpha level of 0.05 if PFS by BICR, OS, and ORR by BICR are all statistically significant. KM curves will be generated for each arm. KM estimates of the median TTPP and its 2-sided 95% CI will be calculated. In addition, KM estimates of the 25th and 75th percentiles of TTPP and the observed minimum and maximum TTPP will be reported. The hazard ratio between 2 arms and its 95% CI will be estimated using a Cox proportional hazards regression model controlling for age (<65 or ≥65 years old), sex, region, and randomization stratification factors.

A sensitivity analysis for TTPP will be performed using the same censoring rule as specified in Table 8 except that subjects who are taking opioid pain medication at baseline will be censored at randomization date.

7.7.1.2 Mean Change from Baseline in Worst Pain at Week 26

Using the BPI-SF Question 3, mean change from baseline in worst pain will be calculated for each post-baseline assessment timepoint for Arm A and Arm B. A mixed model for repeated measures (MMRM) will be used to estimate marginal mean change scores and differences in mean scores between Arm A and Arm B. The difference in mean change from baseline between treatment arms at Week 26 will be evaluated for statistical significance at 2-sided alpha level of 0.05 if PFS by BICR, OS, ORR by BICR, and TTPP are all statistically significant. The MMRM will control for timepoint (categorical week), age (<65 or ≥65 years old), sex, region, baseline worst pain score, baseline use of opioid pain medication (yes or no), randomization stratification factors, and a treatment-by-timepoint interaction. The estimated MMRM will use an unstructured covariance matrix. If the model does not converge, a compound symmetry covariance matrix will be used as appropriate.

7.7.2 EORTC-QLQ-C30

Analysis set: PRO FAS

The EORTC QLQ-C30 scale scores will be calculated using the EORTC QLQ-C30 Scoring Manual (Fayers 2001) for Global health status / Quality of Life, functional scales, and symptom scales, symptom items and Financial Impact. The following summary will be provided for EORTC QLQ-C30 by treatment arms and by scheduled assessments:

- Completion rate for EORTC QLQ-C30 assessment (ITT analysis set)
- Compliance rate for EORTC QLQ-C30 assessment (ITT analysis set)

- Summary statistics of the actual value and change from baseline for global health status/QoL and each functional and symptom scale.
- Change in symptoms/functioning/global health status from baseline will be categorized as improvement using the threshold of 10.

Although not included in the original scoring manual (Fayers 2001), it has been suggested in the literature that a single summary score can also be calculated for this instrument (Giesinger 2016):

- Physical Functioning + Role Functioning + Social Functioning + Emotional Functioning + Cognitive Functioning + (100-Fatigue) + (100-Pain) + 100-(Nausea and Vomiting) + (100-Dyspnoea) + (100-Sleeping Disturbances) + (100-Appetite Loss) + (100-Constipation) + (100-Diarrhoea))/13

If at least half of the items in a scale are presented for a timepoint, then the score will be calculated using the average of all items answered; otherwise the score will be set to missing. For single measures, if the item is missing the scale score is set to missing (Fayers 2001).

7.7.3 EuroQol-5D-5L

Analysis set: PRO FAS

The EQ-5D-5L 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 (EQ-5D-5L User Guide, 2019). 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. EQ-5D-5L index score using UK specific weights will be presented. 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) visual analogue scale (VAS).

The following summary will be provided for EQ-5D-5L by treatment arm and by scheduled assessments for Arm A and Arm B:

- Completion rate for EQ-5D-5L assessment (ITT analysis set)
- Compliance rate for EQ-5D-5L assessment (ITT analysis set)
- Summary statistics of the actual value and change from baseline for VAS and Health State Index Score
- Summary of Dimension Scores

7.7.4 Health Care Resource Utilization

Analysis set: ITT Analysis Set

HRU information will be collected using a questionnaire designed to collect data from the subject perspective summarizing unplanned use of health resources outside of the clinical trial. Cumulative incidence of HRU including length of stay, hospitalizations, and emergency room (ER) visits will be summarized by treatment arm and timepoint.

7.8 Pharmacokinetics and Immunogenicity Endpoints

7.8.1 Enfortumab Vedotin Pharmacokinetics

Analysis set: EV PK Analysis Set

Concentrations of enfortumab vedotin antibody-drug conjugate (ADC) and MMAE will be summarized using descriptive statistics (including geometric mean and coefficient of variation) at each PK sampling timepoint for the PK analysis set.

If performed, population PK and exploratory PK/pharmacodynamic (PD), PK/safety, or PK/efficacy analyses will be described in a separate analysis plan and the results reported separately from the CSR.

If performed, pembrolizumab related PK analyses will be provided in a separate document.

7.8.2 Antitherapeutic Antibody

Analysis set: Safety Analysis Set

The ATA incidence rate is defined as the proportion of subjects that develop ATA at any time during the study. EV related ATA incidence will be summarized using the safety analysis set in Arm A and Arm C. If performed, pembrolizumab related ATA incidence will be provided in a separate document.

7.8.3 Pharmacodynamic and Pre-treatment Biomarkers

The H-score of Nectin-4 expression at baseline will be summarized with descriptive statistics. Baseline H-score will be displayed in a boxplot by best overall response (responders vs. non-responders) as assessed by BICR.

The number and percentage of subjects with baseline PD-L1 CPS<10 vs. ≥10 will be summarized.

Additional analyses for pharmacodynamic biomarkers and for pre-treatment biomarkers that are potentially predictive of subject outcomes may be described in a separate analysis plan and the results may be reported separately from the CSR. For example, efficacy data may be summarized using different Nectin-4 H-score thresholds based on its distribution and/or historical data.

7.9 Additional Analyses

7.9.1 Subsequent Anticancer Therapy

Analysis set: ITT Analysis Set

The number and percentage of subjects who receive subsequent anticancer therapies, including palliative radiotherapy, systemic therapy for disease progression, maintenance therapy, and surgical procedure, will be summarized for Arm A and Arm B. In addition, number of therapies and time from last dose of study treatment to the first subsequent anticancer therapy for progressive disease will be summarized by treatment arm. Subsequent anticancer therapy will be listed.

7.9.2 COVID-19 Impacts

Analysis set: ITT Analysis Set

COVID-19 impacts will be summarized and listed. Subjects with suspected or confirmed case of COVID-19 will be listed.

7.9.3 Japan-Specific Safety Run-In

Subjects in Japan-specific safety run-in are not part of the randomized portion of the study. Their data will be included in listings without separate summary due to small sample size. DLT will be listed.

7.9.4 Arm C (EV + Pembrolizumab + Cisplatin or Carboplatin)

Subjects in Arm C will be summarized in separate tables as appropriate. This includes summary for disposition, baseline demographics and disease characteristics, exposure, concomitant medication, ORR, DCR, and key safety. Data for Arm C subjects will be included in listings.

8 INTERIM ANALYSIS

No IAs are planned for the primary endpoint of PFS. One IA for superiority is planned for the primary endpoint of OS at the time of the PFS final analysis. For the OS analysis, O'Brien-Fleming boundaries will be used to calculate the efficacy boundaries based on the actual information fraction observed. Timing and efficacy boundaries are described in Sections 6.10 and 6.7, respectively.

9 CHANGES FROM PLANNED ANALYSES

9.1 Changes from the Original Protocol

Not Applicable.

9.2 Changes from the Original SAP

Not Applicable.

9.3 Changes from the SAP version 1.0

The SAP has been updated to reflect changes in Protocol Amendment 6. The following changes have been made:

- Signature page has been updated

- Section 3.3 has been updated to align with Protocol Amendment 6
- Section 4: increase number of subjects to 860 and update Figure 1
- Section 5: add clarification for analysis sets
- Section 5.3 and 5.4: add clarification for PRO FAS and PK analysis set
- Sections 6.2, 6.7 and 6.10 have been updated to reflect the following study design changes:
    - Sample size increases from 760 to 860
    - Initial alpha split changes to 0.005 for PFS and 0.045 for OS
    - Median OS for control arm changes to 15.3 months
    - PFS and OS events, power, efficacy boundaries and timing of analyses are updated accordingly
- Section 7.2: add body mass index, hepatic function, hemoglobin A1c, and Bajorin risk factors in baseline characteristics summary. update disease specific characteristics summary
- Section 7.4: remove IDI and RDI for pembrolizumab, cisplatin, gemcitabine. Add end of study date in the duration of treatment definition
- Section 7.5:
    - Add a paragraph to clarify for the purpose of efficacy analysis, maintenance therapy (e.g., avelumab) are not considered as subsequent new anticancer therapy.
    - Update Table 6 for clarification
    - Update sensitivity analyses for OS and PFS.
    - Add clarification for ORR analysis
    - Remove SD maintained for >= 24 weeks requirement for DCR
    - Update Table 7 for clarification
- Section 7.6.1: update the definition of TEAE and list of analyses for TEAE and adverse events of special interest
- Move Section of Deaths from Section 7.9 to Section 7.6
- Add section 7.9.2 COVID-19 impacts
- Update the references for PRO in Section 7.7 and 10

- Add imputation rule for locally advanced disease dates and metastatic disease dates in Appendix C. Move imputation rule for partial subsequent anticancer therapy date from Appendix F to C.
- Change Appendix G to F
- Remove Appendix H
- Minor edits and clarification throughout the document

9.4 Changes from the SAP version 2.0

The SAP has been updated to reflect changes in Protocol Amendment 7. The following changes have been made:

- Separate ORR by BICR and by investigator assessment, and add TTPP and mean change from baseline in worst pain at Week 24 as secondary objectives in Section 2.2
- Separate ORR by BICR and by investigator assessment, and add TTPP and mean change from baseline in worst pain at Week 24 as secondary endpoints in Section 3.2
- Update exploratory objectives and endpoints in Section 2.3 and 3.3
- Add the response evaluable set in Section 5
- Update Section 5.4 to clarify that subjects will be analyzed according to the treatment arm assigned at randomization.
- Add 25^{th} and 75^{th} percentile for continuous variables, and remove the language on additional analysis for PRO endpoints in Section 6.1
- Add statistical testing for ORR and selected PRO endpoints in Section 6.7 and update Section 7.5.2 accordingly.
- Add renal function in subgroups, ORR may be analyzed for subgroups and clarifications in Section 6.8
- Add clarification and update Table 5 using planned 526 PFS events for PFS FA in Section 6.10
- Remove summary of dose modification in Section 7.4 as summary of TEAE leading to dose reduction and dose interruption are summarized in Section 7.6.1 and add clarifications in Section 7.4
- Update analysis set from ITT analysis set to response evaluable set in Sections 7.5, 7.5.2.1, 7.5.2.2, 7.5.2.3, 7.5.3.2, and 7.5.3.3
- Add local therapy per medical adjudication will not be considered as subsequent anticancer therapy for efficacy analysis in Section 7.5
- Update Table 6 for clarity in Section 7.5.1.1

- Update Section 7.5.1.2 to clarify censoring for OS
- Add time to response and statistical comparison for ORR by BICR in Section 7.5.2.1
- Remove KM plot for iDOR in Section 7.5.3.3
- Add an analysis for pembrolizumab AEOSI outcome and update the definition for event resolution and remove the definition of improvement and time to improvement in Section 7.6.1.1
- Add clarification for timepoint and laboratory results will be plotted in Section 7.6.3
- Update Section 7.7 to reflect hierarchical testing for selected PRO endpoints and subsections are re-arranged accordingly
- Add new Sections 7.7.1.1 and 7.7.1.2 for TTPP and mean change from baseline in worst pain at Week 24
- Update Section 7.7.2 to describe QLQ-C30 and remove the analysis using threshold developed by Cocks 2012 and Appendix F
- Change to ITT analysis set and add clarification in Section 7.7.4
- Add clarification in Sections 7.8.1 and 7.9.1
- Update Appendix C to include imputation rule for partial start date of concomitant medication
- Minor edits and clarification throughout the document

9.5 Changes from the SAP version 3.0

- Update endpoint of mean change from baseline in worst pain in Sections 3.2 and 7.7.1.2
- Add description for Japan-specific safety run-in in Section 4
- Add clarifications in Sections 5, 6.3, 6.4.1 and 6.5
- Add KM curves for selected subgroups in Section 6.5
- Update Section 7.5 to add clarification and Case 7
- Update the analyses for TEAE leading to dose interruption and discontinuation in Section 7.6.1
- Update analyses for laboratory parameters in Section 7.6.3
- Add visit windows and selection of analysis records in Section 7.7
- Update the definition of TTPP and censoring rule in Section 7.7.1

- Add 2 new Section 7.9.3 and 7.9.4 for analyses related to Japan-specific safety run-in and Arm C respectively
- Add imputation for death date in Appendix C
- Drop Appendix D and E
- Add Appendix D for Japan specific analysis
- Add Appendix E for definition of analysis windows
- Minor edits and clarification throughout the document

10 REFERENCES

Cleeland CS. The Brief Pain Inventory User Guide. https://www.mdanderson.org/research/departments-labs-institutes/departments-divisions/symptom-research/symptom-assessment-tools/brief-pain-inventory.html. 2009.

Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934;26(4):404-13.

Cocks K, King MT, Velikova G, et al. Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. Eur J Cancer. 2012;48(11):1713-21.

Collett D. Interval-censored survival data. Modelling survival data in medical research. London, Chapman & Hall. 1994:237-51.

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

Giesinger JM, Kieffer JM, Fayers PM, et al. Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. J Clin Epidemiol. 2016;69:79-88.

Maurer W, Bretz F. Multiple testing in group sequential trials using graphical approaches. Stat Biopharm Res. 2013;5(4):311-20.

Royston P, Parmar MK. 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. 2011;30(19):2409-21.

Seymour L, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. Lancet Oncol. 2017;18(3):e143-52.

Uno H, Claggett B, Tian L, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol. 2014;32(22):2380-5.

van Reenen M, Janssen B, Stolk E, et al. EuroQol Research Foundation. EQ-5D-5L User Guide, 2019. Version 3.0. https://euroqol.org/publications/user-guides. 2019.

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:
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:
AE start date will be imputed as the first day of the month

AE day and month are missing, or month only is missing

If the year is the same as the year of first dose of any study treatment:
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:
AE start date will be imputed as January 1st

AE day, month and year are missing, or year only is 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.

* only use condition end date if known and full end date is available.

Incomplete AE End Date:

If AE outcome is "not recovered/resolved", "unknown", or blank: AE condition end date will not be imputed.

If AE outcome is "recovering/resolving", "recovered/resolved", "recovered/resolved with sequelae", or "fatal" apply the following:

AE day only is missing

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

AE day and month are missing, or month only is missing

If the year is equal to the year of the last dose date:
AE condition end date will be imputed as the minimum of (last dose date + 30, death date, 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:

AE condition end date will be imputed as the minimum of (death date, data extraction date, December 31st of the end date year, EOS date)

AE day, month and year are missing, or year only is 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.

Example

AE Number 4: Condition/Event NAUSEA
First dose date 02APR2012

Prior to imputation

Log Line	Start date	Condition end date	Severity	Outcome
1	25APR2012	UNAPR2012	2	recovering/resolving
2	UNAPR2012	04MAY2012	1	recovered/resolved

Post imputation

Log Line	Start date	Condition end date	Severity	Outcome
1	25APR2012	**30**APR2012	2	recovering/resolving
2	02APR2012	04MAY2012	1	recovered/resolved

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:

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 discontinued the treatment,

   AE Start Date ≤ 30 days after last dose of study treatment

   OR

   SAE=Yes and AE Start Date ≤ 90 days after last dose of study treatment in pembrolizumab containing arm.

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 DATES

1. 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. If both month and day are missing, disease diagnosis dates will not be imputed.

Missing locally advanced disease dates will be imputed as the maximum of (the first day of the month 01mmmyyyy, disease diagnosis dates) if both month and year are present and only day is missing. If both month and day are missing, locally advanced disease dates will not be imputed.

Missing metastatic disease dates will be imputed as the maximum of (the first day of the month 01mmmyyyy, disease diagnosis dates) if both month and year are present and only day is missing. If both month and day are missing, metastatic disease dates will not be imputed.

2. Subsequent anticancer 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 the end-of-study (EOS) date or 30 days after the last dose of any study drug, whichever is earlier.

3. Concomitant medications

Concomitant medication start date will be imputed as follows if the indication is for adverse events.

- If only the day is missing, concomitant medication start date will be imputed as the first day of the month or the start date of the adverse event that is linked to the medication, whichever is later.
- If both the day and month are missing, concomitant medication start date will be imputed as January 1st of the year or the start date of the adverse event that is linked to the medication, whichever is later.
- If concomitant medication start date is completely missing, the start date will not be imputed.

Concomitant medication start date will be imputed as follows if the indication is not for adverse events.

- If 'Started prior to first dose of any study treatment' is No and only the day is missing, concomitant medication start date will be imputed as the first day of the month or the first dose date of any study treatment, whichever is later.
- Else if only the day is missing, concomitant medication start date will be imputed as the first day of the month.
- If month and/or year is missing, concomitant medication start date will not be imputed.

4. Death Date

If both month and year are present and only day is missing, death date will be imputed as the first day of the month or the last known alive date, whichever is later.

APPENDIX D: JAPAN SPECIFIC ANALYSIS

This appendix describes the planned regional specific analyses in addition to the analyses described in the main body of this SAP. The regional analyses specified in this appendix are for the PMDA submission only.

The following regional specific analyses of Japan will be performed as appropriate:

- Selected analyses as described in the main body of this SAP will be repeated for the subgroups of Japan vs Non-Japan or subgroup of Japan only.
- Selected listings will be sorted by sites with Japan sites followed by Non-Japan sites.
- Efficacy endpoints (PFS, OS, ORR, DCR) and TTPP will be evaluated using unstratified analysis due to the small numbers of Japan patients.
- MMRM for mean change from baseline in worst pain as described in Section 7.7.1.2 will be modified and not control for stratification factors due to the small numbers of Japan patients.

Subgroup of Japan includes subjects who are enrolled in Japan sites regardless of race.
Subgroup of Non-Japan includes subjects who are enrolled in Non-Japan sites regardless of race.

APPENDIX E: DEFINITION OF ANALYSIS WINDOWS AND SELECTION OF ANALYSIS RECORD

For PRO, analysis window is defined as listed in Table G1. For EQ-5D-5L, EORTC QLQ-C30 and BPI-SF, if there are multiple records in the same analysis window, select the record closest to the target day. If two records are of the same distance to the target day, select the record with the earlier date.

Table G1: Analysis Windows for PRO Assessments

Analysis Visit	Target Day (Study Day)	Time Window (Study Days Range)
Baseline	1	<=1
Week 1	5	[2, 6]
Week 2	12	[7, 13]
Week 3	19	[14, 20]
Week 4	26	[21, 27]
Week 5	33	[28, 34]
…	(x-1)*7 + 5	[(x-1)*7, x*7-1]
Week 12	82	[77, 83]
Week 14	96	[84, 97]
Week 17	117	[98, 118]
Week 20	138	[119. 139]
….	(x-1)*7 + 5	[(x-3)*7, x*7-1]
EOT	Day of EOT	

Signature Page for SGN-CLIN-023739 v2.0

eSignature Approval	[redacted] 3 17:06:26 GMT+0000
eSignature Approval	[redacted] GMT+0000
eSignature Approval	[redacted] 3 17:18:04 GMT+0000
eSignature Approval	[redacted] 3 17:23:58 GMT+0000
eSignature Approval	[redacted] GMT+0000

Signature Page for SGN-CLIN-023739 v2.0