NCT05551117

# STATISTICAL ANALYSIS PLAN
# VERSION 4.0

## CLINICAL STUDY PROTOCOL: CP-MGC018-03

## A Phase 2, Open-Label Study of Vobramitamab Duocarmazine in Participants with Metastatic Castration-Resistant Prostate Cancer and Other Solid Tumors (TAMARACK)

# TABLE OF CONTENTS

# LIST OF TABLES



# LIST OF FIGURES



# LIST OF APPENDICES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADA | anti-drug antibody |
| ADaM | analysis data model |
| ADT | androgen deprivation therapy |
| AE | adverse event |
| AESI | adverse event of special interest |
| ARAT | androgen receptor axis-targeted therapy |
| ATC | anatomical therapeutic chemical |
| BOR | best overall response |
| CDISC | Clinical Data Interchange Standards Consortium |
| CI | confidence interval |
| $C_{max}$ | maximum concentration |
| CR | complete response |
| eCRF | electronic case report form |
| CSR | clinical study report |
| DCR | disease control rate |
| DoR | duration of response |
| ECG | electrocardiogram |
| ECOG | eastern cooperative oncology group |
| EOTV | end of treatment visit |
| IDMC | Independent Data Monitoring Committee |
| IHC | immunohistochemistry |
| IRT | interactive response technology |
| ITT | intent-to-treat |
| IV | intravenous |
| LVEF | left ventricular ejection fraction |
| MedDRA | medical dictionary for regulatory activities |
| MUGA | multigated acquisition ventriculography scanning |

| | |
|---|---|
| NE | not evaluable |
| ORR | objective response rate |
| OS | overall survival |
| PCWG3 | prostate cancer clinical trials working group 3 |
| PD | progressive disease |
| PSA | prostate specific antigen |
| rPFS | radiographic progression-free survival |
| PFS | progression-free survival |
| PK | pharmacokinetics |
| PR | partial response |
| Q4W | every 4 weeks |
| SAE | serious adverse event |
| SAP | statistical analysis plan |
| SD | stable disease |
| SDTM | Study Data Tabulation Model |
| SOC | system organ class |
| SPP | statistical programming plan |
| SSE | symptomatic skeletal event |
| TEAE | treatment-emergent adverse event |

# 1 INTRODUCTION

This statistical analysis plan (SAP) provides a detailed and comprehensive description for the analysis of Study CP-MGC018-03 entitled "A Phase 2, Open-Label Study of Vobramitamab Duocarmazine in Participants with Metastatic Castration-Resistant Prostate Cancer and Other Solid Tumors (TAMARACK)". This SAP Version 4.0 applies to Protocol Amendment 4 (dated 22 January 2024) onwards and describes in detail the statistical methods to be used for analysis of the primary and secondary efficacy endpoints, the safety endpoints, and the pharmacokinetics (PK) parameters to be collected from this study.

# 2 STUDY OBJECTIVES

## 2.1 Primary Objectives

| Primary Objective | Estimand Description |
|---|---|
| Part 1: To evaluate the efficacy of vobramitamab duocarmazine in mCRPC at two dose levels as measured by rPFS by investigator assessment using PCWG3 criteria. | Population: Intent-to-treat |
| | Variable: rPFS |
| | Intercurrent events:<br>1) Subsequent anticancer therapy prior to rPFS event: censored<br>2) Missing ≥ 2 consecutive scheduled tumor assessments immediately followed by documented progression or death: censored<br>3) Study treatment discontinuation: ignored |
| | Summary measure: Kaplan-Meier estimates of rPFS |
| Part 2: To evaluate the efficacy of vobramitamab duocarmazine in select solid tumors as measured by investigator assessment of ORR per RECIST v1.1. | Population: Response evaluable |
| | Variable: ORR |
| | Intercurrent events:<br>1) Subsequent anticancer therapy: while not having started subsequent anticancer therapy<br>2) Disease progression: while not having evidence of progression<br>3) Death due to any reason: while being alive<br>4) Study treatment discontinuation: ignored |
| | Summary measure: Proportion of participants who achieve a BOR of confirmed CR or PR per RECIST v1.1 |

Abbreviations: BOR: best overall response; CR: complete response; ORR: objective response rate; PR: partial response; PCWG3: Prostate Cancer Clinical Trials Working Group 3; RECIST: Response Evaluation Criteria in Solid Tumors; rPFS: radiographic progression-free survival.

## 2.2 Secondary Objectives

| Secondary Objectives | Outcome or Endpoint |
|---|---|
| To characterize the frequency and severity of AEs and overall tolerability observed with vobramitamab duocarmazine. | Frequency and severity of AEs, SAEs, and AEs leading to study treatment discontinuation. |
| To evaluate the efficacy of vobramitamab duocarmazine as measured by PSA (Part 1 only). | PSA response rate per PCWG3 criteria; time to PSA progression per PCWG3 criteria, duration of PSA response per PCWG3 criteria.<br>PSA percent change over time and best PSA percent change. |
| To evaluate the efficacy of vobramitamab duocarmazine in participants with RECIST-evaluable disease as measured by ORR and DoR by investigator assessment using PCWG3 criteria (Part 1 only). | ORR and DoR per PCWG3 criteria. |

| Secondary Objectives | Outcome or Endpoint |
|---|---|
| To evaluate the effect of vobramitamab duocarmazine on SSEs (Part 1 only). | Time to first SSE. |
| To evaluate the efficacy of vobramitamab duocarmazine in participants with select solid tumors as measured by DoR and PFS by investigator assessment using RECIST v1.1 (Part 2 only). | DoR and PFS per RECIST v1.1. |

Abbreviations: AE: adverse event; ADA: anti-drug antibody; DoR: duration of response; IV: intravenous; ORR: objective response rate; PCWG3: Prostate Cancer Clinical Trials Working Group 3; PFS: progression-free survival; PK: pharmacokinetics; PSA: prostate specific antigen; Q4W: once every 4 weeks; RECIST: Response Evaluation Criteria in Solid Tumors; SAE: serious adverse event; SSE: symptomatic skeletal event.

## 2.3 Exploratory Objectives

| Exploratory Objectives | Outcome or Endpoint |
|---|---|
| To explore the effect of vobramitamab duocarmazine dose on OS. | OS |
| To explore the relationship between B7-H3 expression and efficacy. | Relationships between B7-H3 expression and tumor response and/or other endpoints. |
| To characterize the PK of vobramitamab duocarmazine. | Concentrations and PK parameters for vobramitamab duocarmazine (and metabolites as appropriate). |
| To characterize immunogenicity of vobramitamab duocarmazine. | Incidence of ADA formation against vobramitamab duocarmazine. |
| To evaluate the exposure-response relationship between vobramitamab duocarmazine exposure and select efficacy and safety variables. | Logistic regression curves relating vobramitamab duocarmazine indices of exposure (i.e., $C_{max}$ and AUC) to incidence of key efficacy (e.g., rPFS) and safety (AEs). |

Abbreviations: AE: adverse event; AUC: area under the concentration-time curve; $C_{max}$: maximum concentration; OS: overall survival; rPFS: radiographic progression-free survival.

Results of exploratory objectives may not be included in the clinical study report (CSR) or database lock unless they represent meaningful findings.

# 3 STUDY DESIGN AND PLAN

## 3.1 Overall Study Design and Plan

Study CP-MGC018-03 is an open-label, two-part, Phase 2 study assessing the efficacy, safety, and tolerability of vobramitamab duocarmazine.

In Part 1 of the study, approximately 100 participants will be randomized 1:1 to one of two dose levels of vobramitamab duocarmazine (2.0 mg/kg Q4W and 2.7 mg/kg Q4W). Part 1 will enroll participants with mCRPC previously treated with one prior ARAT (abiraterone, enzalutamide, or apalutamide) for prostate cancer in either the metastatic or non-metastatic, castration-sensitive or castration-resistant setting. Participants may also have received up to one prior docetaxel-containing regimen for prostate cancer, and up to 3 total prior lines of therapy for mCRPC. Enrollment will use a central randomization scheme with stratified permuted blocks. Randomization will be stratified based on 1) presence of visceral disease (yes vs no), 2) prior taxane (yes vs no) and 3) region (US/Canada vs other).

In Part 2 of the study, participants with unresectable, locally advanced or metastatic anal SCC, HNSCC, melanoma, squamous NSCLC, or SCLC whose disease has progressed following at least one prior line of standard systemic therapy for advanced or metastatic disease will receive vobramitamab duocarmazine 2.7 mg/kg Q4W. Each disease-specific cohort will use a Simon's 2-stage design and enroll up to approximately 40 participants each if both stages are conducted.

Participants will receive vobramitamab duocarmazine at their assigned dose until protocol-specified progressive disease (PD), AE requiring discontinuation, physician decision, withdrawal of consent, allowed treatment duration (26 cycles or 2 years) is reached, or other treatment discontinuation criteria are met per protocol.

Tumor assessments are performed Q8W for the initial 24 weeks on study treatment then Q12W, regardless of dose delay, until PD. Tumor assessments are performed at the EOTV except if the prior assessment was ≤ 12 weeks of the EOTV and results are available for tumor response assessment.

Participants will be followed for safety throughout the study. Any participant who discontinues study treatment should complete an EOTV. The EOTV should be performed within 30 days following the last dose of study treatment whenever possible.

### 3.1.1 Part 1 Study Design

Part 1 of the study has pre-planned interim and final analyses. The interim analysis will be performed on each vobramitamab duocarmazine arm based on PSA response rate at 8 weeks when 20 PSA evaluable participants in each vobramitamab duocarmazine arm have been enrolled and followed for at least 8 weeks. The final rPFS analysis will be conducted when all participants on Part 1 have been randomized and followed for rPFS for at least 6 months.

An IDMC will provide study oversight and evaluate cumulative clinical data at regular intervals. If at any time during Part 1 of the study it becomes apparent that one or both doses of

vobramitamab duocarmazine have unacceptable toxicity, poor tolerability, or limited efficacy, that arm of the study may close to enrollment early while the other arm continues to enroll. If only one arm closes to enrollment, participants on that arm will stay on that arm and be analyzed according to their originally assigned arm, but be offered the opportunity to switch to the other dose.

Prior to Amendment 2, the study contained a control arm. Participants who enrolled prior to Amendment 2 and were randomized to the control arm may choose to stay on the control arm or discontinue study treatment. Participants who choose to discontinue study treatment may re-enroll on the study after implementation of Amendment 2 and be randomized to one of the vobramitamab duocarmazine arms if they meet eligibility criteria (including washout period). Participants assigned to one of the vobramitamab duocarmazine arms prior to Amendment 2 will stay on their assigned arm under Amendment 2 onwards with no changes to their study treatment.

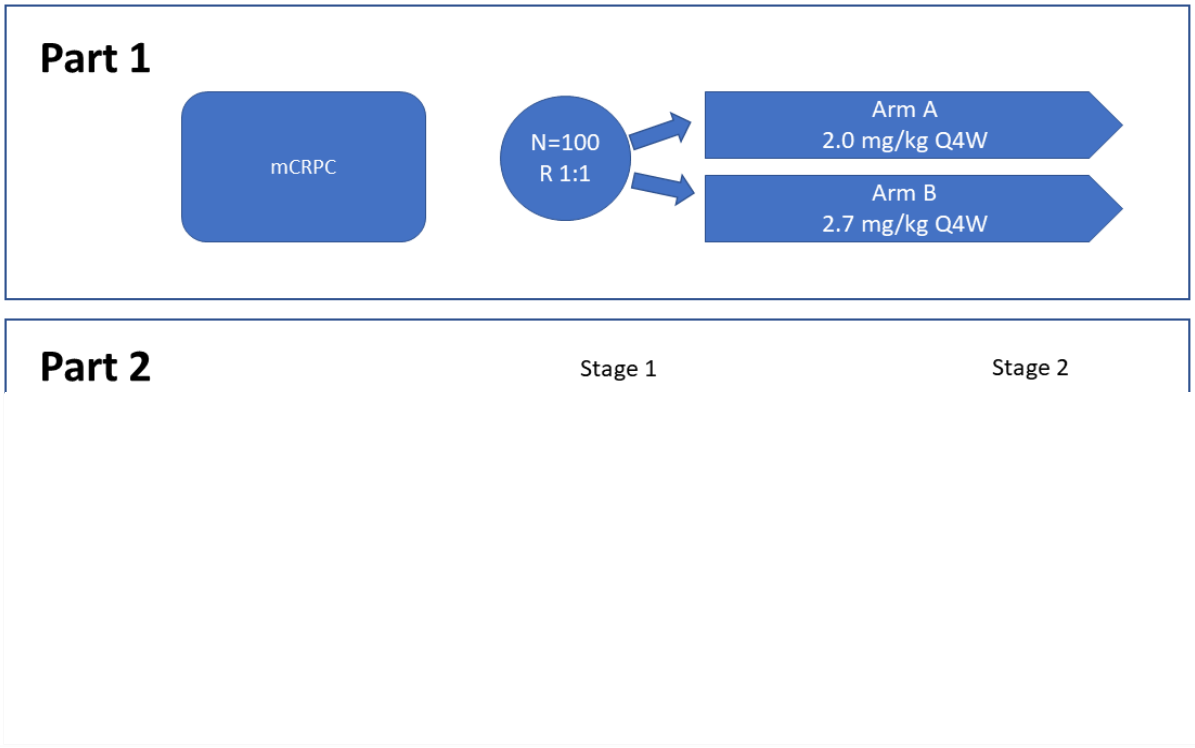
### 3.1.2 Part 2 Study Design

Part 2 will use Simon's 2-stage design for each tumor-specific expansion cohort. A futility assessment will be conducted at the end of Stage 1 on a pre-specified number of response-evaluable participants in each tumor-specific expansion cohort (**Section 3.4** and **Section 6.10.1**). Enrollment will be paused separately for each cohort while the analysis is conducted. Upon passing futility assessment, each cohort may enroll additional participants (up to a total of 37 to 42 response-evaluable participants per cohort; **Table 1**) to further evaluate antitumor activity, safety, tolerability, PK, and immunogenicity.

**Table 1 Part 2 Cohort Enrollment**

The study treatment schema is presented in **Figure 1**.

**Figure 1 Study Schema**




Abbreviations: HNSCC: head and neck squamous cell carcinoma; mCRPC: metastatic castration-resistant prostate cancer; NSCLC: non-small cell lung cancer; ORR: objective response rate; Q4W: once every 4 weeks; R: randomization; SCC: squamous cell carcinoma; SCLC: small cell lung cancer.

## 3.2 Description of Dosing Regimens

An overview of study treatments to be administered during this study is presented in **Table 2**.

**Table 2 Overview of Study Treatments**

| Product | Dose | Schedule | Route of Administration |
|---|---|---|---|
| Vobramitamab duocarmazine | 2.0 mg/kg [a] | Once every 4 weeks | IV |
| Vobramitamab duocarmazine | 2.7 mg/kg [a,b] | Once every 4 weeks | IV |

a Prednisone 40 mg PO twice daily will be administered on the first 3 days of each cycle **after** administration of vobramitamab duocarmazine.

b Part 2 dosing regimen.

Abbreviations: IV: intravenous

## 3.3 Randomization and Blinding

**Part 1**

Eligible participants will be randomized centrally using interactive response technology (IRT). Instructions for randomization will be provided via IRT. Every effort should be made to minimize the time between randomization and initiation of study treatment. It is recommended that subjects commence study treatment as soon as possible after randomization.

Enrollment will use a central randomization scheme with stratified permuted blocks. Note: for stratification, visceral disease is defined as liver, lung, or other non-nodal soft tissue metastases.

Participants will be randomized in a 1:1 ratio between the two treatment arms:

1. Experimental arm A: vobramitamab duocarmazine 2.0 mg/kg IV Q4W
2. Experimental arm B: vobramitamab duocarmazine 2.7 mg/kg IV Q4W

Randomization will use the following 3 stratification factors:

1. Presence of visceral disease (yes vs no)
2. Prior taxane (yes vs no)
3. Region (US/Canada vs other)

Participants assigned to the control arm prior to its removal in Amendment 2 who would like to receive vobramitamab duocarmazine treatment may discontinue the study and be re-randomized to one of the vobramitamab duocarmazine arms after a 4-week washout period if they meet eligibility criteria. Participants assigned to one of the vobramitamab duocarmazine arms prior to Amendment 2 will stay on their assigned arm under Amendment 2 onwards with no changes to their study treatment.

**Part 2**

Eligible participants will be assigned via IRT to receive vobramitamab duocarmazine at 2.7 mg/kg Q4W.

This is an open-label study. No blinding of site personnel or participants will be employed. Sponsor access to clinical data will be conducted according to the pre-specified study data access plan where only a limited number of sponsor employees will have access to patient-level and aggregated data; treatment assignment will be withheld unless required to help manage AEs or clinical supplies.

## 3.4 Sample Size

In total, up to approximately 382 participants may enroll in the study, including 182 participants in Part 1 and approximately 200 participants in Part 2. However, the number of participants enrolled cannot be determined precisely in advance and will be determined by the safety profile and early evidence of clinical benefit.

**Part 1**

Approximately 100 participants as planned per Amendment 2 will be randomized in a 1:1 ratio to each of the two vobramitamab duocarmazine arms. Randomization will be stratified by 3 stratification factors: 1) visceral disease (yes vs no), 2) prior taxane (yes vs no), and 3) region (United States/Canada vs other). The historic benchmarks for each vobramitamab duocarmazine arm to compare with are 25% PSA response rate and 8 months median rPFS observed on the control arm of the PRESIDE trial of docetaxel after enzalutamide (**10**).

Since the goal of this part of the study is to select one of two vobramitamab duocarmazine doses for further study in participants with mCRPC, Simon's randomized Phase 2 selection design (**11**) was used to determine the sample size for each arm. The planned sample size of 50 per arm would provide 76% probability of correctly selecting the best dose assuming 66% and 59% rPFS rates at 6 months (equivalent to 10 vs 8 months median rPFS) for the two vobramitamab duocarmazine arms, respectively. Sample size determination based on traditional hypothesis testing would be prohibitively large for a Phase 2 study. A sample size of 350 per arm would be required to have 70% power to detect 59% vs 66% rPFS rate between two arms at 1-sided 0.1 alpha level. Comparing each arm to a fixed benchmark would also require an unfeasibly large sample size of 160 per arm in order to have 70% power to reject a fixed null rPFS rate of 59% at a 1-sided 0.1 alpha level under an alternative rPFS rate of 66% for each arm, respectively. As both doses being tested fall within the range of doses that demonstrated biological activity based on CP-MGC018-01 study data, there is minimal risk of choosing a dose that is not biologically active. All available safety and efficacy data from this study will be used to justify the dose ultimately selected for further study (e.g., in a later registrational study).

For PSA response rate, the planned sample size of 50 per arm would provide 84% power to reject 25% null rate at 1-sided 0.1 alpha level for each vobramitamab duocarmazine arm at an alternative rate of 40%.

Total enrollment in Part 1 was planned be up to approximately 120 participants to account for participants originally enrolled on the control arm per Amendment 1. Due to unexpected rapid enrollment in Part 1, this part of the study enrolled a total of 182 participants.

**Part 2**

Part 2 may enroll up to approximately 200 response-evaluable participants, consisting of approximately 40 participants for each of 5 tumor-specific cohorts (**Table 3**). For each cohort, Simon's 2-stage design (**12**) is used to test null hypothesis $H_0$: $p \leq p_0$ vs alternative hypothesis $H_1$: $p = p_1$ at 1-sided significance level $\leq 0.1$ (alpha) with the power of approximately 80%, where $p_0$ is the benchmark response rate representing a lack of tumor response and $p_1$ is the target response rate representing evidence of sufficient tumor response worthy of further development. Specifically, for each cohort, Stage 1 will enroll $n_1$ response-evaluable participants. If the number of responses in a tumor-specific cohort enrolled in Stage 1 is less than or equal to the prespecified number of responses ($r_1$), then enrollment for the tumor-specific cohort will be stopped for futility, otherwise, cohort enrollment will continue to Stage 2 to enroll an additional $n_2$ response-evaluable participants. At the end of Stage 2, if the number of tumor responses out of all $n = n_1+n_2$ response-evaluable participants is at least r, then the null hypothesis $H_0$ will be rejected, and the cohort is considered to have sufficient efficacy for further development. Simon's 2-stage design parameters for each tumor-specific cohort are provided in **Table 3**.

**Table 3 Simon's Two-Stage Design Parameters**

1. Ott PA, et al. Safety and antitumor activity of the anti-PD-1 antibody pembrolizumab in patients with recurrent carcinoma of the anal canal. Ann Oncol. 2017; 28(5):1036-1041.

2. Marabelle A, et al. Pembrolizumab for previously treated advanced anal squamous cell carcinoma: results from the non-randomised, multicohort, multicentre, phase 2 KEYNOTE-158 study. Lancet Gastro & Hepa. 2022; 7(5):446-454.

3. Rao S, et al. A phase II study of retifanlimab (INCMGA00012) in patients with squamous carcinoma of the anal canal who have progressed following platinum-based chemotherapy (POD1UM-202). ESNO Open. 2022;7(4):100529.

4. Cohen EEW, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. Lancet. 2019; 393(10167):156-167.

5. Hodi FS, et al. Improved survival with Ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010; 363:711-723.

6. Robert C, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. N Engl J Med. 2015; 372:2521-2532.

7. Paz-Ares L, et al. Canakinumab in combination with docetaxel compared with docetaxel alone for the treatment of advanced non-small cell lung cancer following platinum-based doublet chemotherapy and immunotherapy (CANOPY-2): A multicenter, randomized, double-blind, phase 3 trial. Lung Cancer. 2024; 189:107451.

8. Trigo J, et al. Lurbinectedin as second-line treatment for patients with small-cell lung cancer: a single-arm, open-label, phase 2 basket trial. Lancet Oncol. 2020; 21(5):645-654.

9. Kaira K, et al. A phase II study of amrubicin, a synthetic 9-aminoanthracycline, in patients with previously treated lung cancer. Lung Cancer. 2010; 69(1):99-104.

10. Eckardt JR, et al. Phase III study of oral compared with intravenous topotecan as second-line therapy in small-cell lung cancer. Lung Cancer. 2007; 25(15):2086-2092.

## 3.5 Independent Data Monitoring Committee

An IDMC will oversee the ongoing monitoring and interpretation of safety data from all parts of this study (Part 1 and Part 2). An IDMC charter will be reviewed and approved by the IDMC prior to study initiation. The IDMC will evaluate aggregate safety data and efficacy data for the futility analysis in Part 1.

The first IDMC meeting will occur after 20 participants in each arm of Part 1 have been enrolled and completed at least 1 cycle of treatment or after approximately 6 months from first enrollment, whichever comes first. The IDMC may meet at any time if emergent safety data warrant a safety review. Reviews will continue in approximate 6-month intervals (± 1 month). If an IDMC meeting is scheduled to occur near the time of the planned futility analysis for Part 1, the timing of the meeting may be adjusted to align with the futility analysis.

The IDMC will be empowered to recommend changes to the study including discontinuation or modification of a study treatment arm, cohort, or overall termination of the study as appropriate.

The IDMC will be provided with clinical data tables and listings for these reviews. Additional details on the IDMC may be found in the IDMC charter.

# 4 ANALYSIS POPULATIONS

The study analyses will be performed on the following populations:

- **Safety Population:** All participants who received any dose of study treatment. This population will be used to summarize safety data. Participants will be analyzed according to the actual study treatment received. This population will also be used for summary of baseline data and analyses of PFS and OS for Part 2, and for supplementary analysis of rPFS and OS for Part 1.
- **Intent-to-Treat (ITT) Population:** All participants*. Participants will be analyzed according to the study treatment assigned. This population will be used for summary of baseline data and analyses of rPFS, OS, and other efficacy endpoints where applicable for Part 1.
- **Tumor Response Evaluable Population:** All participants* who received at least one dose of study treatment and had RECIST-evaluable disease per PCWG3 (Part 1) or RECIST v1.1 (Part 2) at baseline (i.e., participants with bone-only disease are excluded from this population). This population will be used for efficacy analyses related to tumor response and will be analyzed according to the study treatment assigned.
- **PSA Response Evaluable Population**: All participants in Part 1* who received at least one dose of study treatment, with a baseline PSA $\geq$ 2 ng/mL and at least one post-baseline PSA measurement. This population will be used to calculate and summarize PSA response rates according to the study treatment assigned.
- **PK Evaluable Population:** All participants who received at least one dose of study treatment, date and time of dose administration and relative PK sample collection are known, and have sufficient concentration data to derive at least one PK parameter. Participant data may be excluded from the PK analysis at the discretion of the pharmacokineticist.
- **ADA Evaluable Population:** All participants who received at least one dose of vobramitamab duocarmazine and have at least one post-baseline ADA sample.

* Note: participants in Part 1 who were assigned to the control arm prior to its removal in Amendment 2 who were re-randomized to a vobramitamab duocarmazine arm will be analyzed as a separate cohort for safety and efficacy endpoints.

# 5 ENDPOINTS

## 5.1 Efficacy Endpoints

### 5.1.1 Primary Efficacy Endpoints

#### 5.1.1.1 Radiographic Progression-Free Survival per PCWG3 (Part 1)

In Part 1, the primary efficacy endpoint is rPFS determined by the investigator. The rPFS is defined as the time from the date of randomization to the date of first documented PD per PCWG3 or death from any cause, whichever occurs first. Radiographic PD per PCWG3 is defined in **Section 11.1.1** and **Table 18** of the protocol as follows:

- Progression in soft tissue lesions per RECIST v1.1 (**5**)
- Progression in bone lesions per PCWG3 criteria:
  - At least 2 new lesions (compared to baseline) are observed on first post-baseline scan (Week 8 scan), and the next scan (confirmatory at least 6 weeks later) on Week 16 shows at least 2 additional new lesions
  - At least 2 new lesions (compared to Week 8 scan) are observed on a scan performed after Week 8, and the next scan (confirmatory at least 6 weeks later) show persistence or increase in number of new lesions

rPFS is calculated as:

rPFS (months) = (date of event [documented radiographic progression or death] or date of censoring – randomization date +1) / (365.25/12)

#### 5.1.1.2 Objective Response Rate per RECIST v1.1 (Part 2)

In Part 2, best overall response (BOR) will be categorized using RECIST v1.1 as CR, PR, SD, PD, or NE. To qualify as an objective response, CR and PR require confirmation at least 4 weeks after initial observation of CR or PR, and SD requires observation at least once after 6 weeks from the start of study treatment. BOR will be evaluated from the start of study treatment.

The number and percent of participants with BOR per RECIST v1.1 will be summarized. The ORR is estimated as the proportion of participants in the Tumor Response Evaluable Population who achieve a BOR of confirmed CR or PR per RECIST v1.1.

Disease control rate (DCR) is defined as the percentage of tumor response-evaluable participants who experienced response of CR, PR, or SD for at least 3 months.

## 5.1.2 Secondary Efficacy Endpoints

### 5.1.2.1 Prostate-Specific Antigen Response (Part 1)

PSA response is defined as a ≥ 50% decline in PSA from baseline with PSA confirmation ≥ 3 weeks after the first documented reduction in PSA of ≥ 50%. The PSA response rate will be calculated as the proportion of participants with a PSA response in the PSA Response Evaluable Population.

Duration of PSA response is defined as the time from the date of first documented PSA response to the earliest date of PSA progression. Duration of PSA response is calculated as:

Duration of PSA Response (months) = (date of the earliest PSA progression or date of censoring – date of initial PSA response +1) / (365.25/12)

### 5.1.2.2 Time to Prostate-Specific Antigen Progression (Part 1)

Per PCWG3, PSA progression is defined as follows:

- In participants with a decrease in PSA from baseline: ≥ 25% increase and an absolute increase of ≥ 2 ng/mL above the nadir value, which is confirmed by a consecutive second value obtained ≥ 3 weeks later.
- In participants with no decrease in PSA from baseline: ≥ 25% increase and an absolute increase of ≥ 2 ng/mL above the baseline value after 12 weeks.

Time to PSA progression is defined as the time from the date of randomization to the first documented PSA progression. Time to PSA progression is calculated as:

Time to PSA progression (months) = (date of first PSA progression or date of censoring – randomization date +1) / (365.25/12)

### 5.1.2.3 Progression-Free Survival per RECIST v1.1 (Part 2)

In Part 2, PFS is defined as the time from the first dose date to the date of first documented PD per RECIST v1.1 or death from any cause, whichever occurs first. PFS is calculated as:

PFS (months) = (date of event [documented progression or death] or date of censoring – randomization date +1) / (365.25/12)

### 5.1.2.4 Objective Response Rate per PCWG3 (Part 1)

In Part 1, BOR will be categorized using PCWG3 as CR, PR, SD, PD, or NE. To qualify as an objective response, CR and PR require confirmation at least 4 weeks after initial observation of CR or PR, and SD requires observation at least once after 6 weeks from the start of study treatment. BOR will be evaluated from the start of study treatment.

The number and percent of participants with BOR per PCWG3 criteria will be summarized. The ORR is estimated as the proportion of participants in the Tumor Response Evaluable Population who achieve a BOR of confirmed CR or PR without prior confirmed bone progression per PCWG3.

Disease control rate (DCR) is defined as the percentage of tumor response-evaluable participants who experienced response of CR, PR, or SD for at least 3 months.

#### 5.1.2.5 Duration of Response (Part 1 and Part 2)

The DoR will be calculated as the time from the date of initial tumor response (CR or PR) to the date of first documented PD or death from any cause, whichever occurs first. DoR is calculated only for participants with a tumor response (CR or PR) prior to initiation of new anticancer therapy as evaluated by investigators. DoR is calculated as:

DoR (months) = (date of event [documented progression or death] or date of censoring – date of initial response+1) / (365.25/12)

#### 5.1.2.6 Tumor Size Change Over Time (Part 1 and Part 2)

For participants with measurable disease at baseline, the tumor size is defined as the sum of diameters of the target lesions. Target lesions are identified and measured at baseline and post-baseline tumor assessment visits. All lesions identified at baseline should be measured on each post-baseline tumor assessment visit. If any lesion measurement is missing at a post-baseline tumor assessment visit, that visit will be excluded from the analysis for tumor size change. Percent change from baseline of Tumor Size (TS) is calculated as:

Percent change from baseline = 100 * (Post-baseline TS – Baseline TS)/Baseline TS

#### 5.1.2.7 Time to First Symptomatic Skeletal Event (Part 1)

An SSE is defined as any of the following events: new symptomatic pathological fracture, requirement for radiation therapy to relieve bone pain, spinal cord compression, or tumor-related orthopedic surgical intervention. The time to first SSE is defined as the time from the date of randomization to the first occurrence of SSE.

Time to first SSE is calculated as:

Time to SSE (months) = (date of first SSE or date of censoring – randomization date + 1) / (365.25/12)

### 5.1.3 Exploratory Efficacy Endpoints and Analyses

#### 5.1.3.1 Overall Survival (Part 1 and Part 2)

OS is defined as the time from the date of randomization (Part 1) or first dose (Part 2) to the date of death from any cause. OS is calculated as:

OS (months) = (date of death or last known alive – randomization or first dose date +1) / (365.25/12)

## 5.2 Safety Endpoints

### 5.2.1 Adverse Events

Safety will primarily be addressed by evaluations of the adverse events (AEs). An AE means any untoward medical occurrence associated with the use of a drug in humans. AEs may or may not be drug related. An AE is:

- Any unfavorable and unintended sign, symptom, or disease temporally associated with the use of a medicinal product.
- Any medical occurrence that is new or has increased in severity or frequency from the baseline condition.
- Any abnormal results of diagnostic procedures, including laboratory test abnormalities.

All AEs, whether serious or nonserious, will be captured starting with signing of the informed consent form until 30 days after the last dose of study treatment or until initiation of another anticancer therapy, whichever comes first. SAEs related to study treatment may be reported at any time, regardless temporal relationship to study treatment dosing. Verbatim terms will be coded to lower-level terms in the Medical Dictionary for Regulatory Activities (MedDRA). An assessment of severity grade will be made using NCI-CTCAE v5.0.

AEs and SAEs reported between the time the participant signs the informed consent form and the administration of the first dose of study treatment will be captured as concurrent medical history unless the events (AEs and SAEs) are attributed to protocol-specified procedures.

Only treatment emergent adverse events (TEAEs) will be summarized as safety endpoints. A TEAE is defined as any event that is newly occurring on or after the administration of study drug or an event that existed before but increased in severity on or after study drug administration.

### 5.2.2 Laboratory Evaluations

Clinical laboratory tests include hematology, serum chemistry, cardiac, coagulation, endocrine, and urinalysis. Both local and central laboratory testing will be used. Local laboratories should be used for all clinical decision-making, including but not limited to decisions regarding study treatment dosing and dose adjustments, if needed. Central laboratory testing will be used to determine participant eligibility and assess overall safety in the study, unless the sponsor medical monitor proactively approves the use of a local laboratory in place of an unanalyzable or unavailable central laboratory test result.

Clinically significant laboratory abnormalities are reported as AEs. A laboratory abnormality is considered clinically significant if:

- Associated with discontinuation of study treatment, study treatment dose reduction or delay, or required intervention.

- Suggestive of disease or organ toxicity.

### 5.2.3 Other Safety Endpoints

ECGs will be collected and analyzed for evidence of cardiac toxicity, including prolongation of QT interval.

LVEF will be evaluated by echocardiogram or MUGA and changes from baseline summarized.

Vital signs and weight will be summarized with descriptive statistics at each visit and time point collected. Shift tables may be produced.

# 6 STATISTICAL METHODOLOGY

## 6.1 General Considerations

Unless otherwise specified, the data for the stratification factors from IRT will be used in the stratified analysis. Discrepancy between IRT data and stratification factors collected on eCRF will be summarized.

Study Day 1 is defined as the first day of study drug administration. Every effort should be made to minimize the time between randomization and starting the study drug in Part 1. It is recommended that participants commence study drug as soon as possible after randomization in Part 1. Time to event endpoints (e.g., rPFS, OS) will be determined from date of randomization (Part 1) or first dose (Part 2) for all participants.

Unless otherwise noted, baseline values will be defined as the most recent value collected prior to the first dose of study treatment.

Categorical data will be summarized by the number and percent of participants falling within each category. Continuous variables will be summarized by descriptive statistics, including mean, standard deviation, median, minimum, and maximum.

P-values will be reported in 0.XXXX format where XXXX represent the value rounded to four decimal places. Any p-value <0.0001 will be presented as <0.0001.

Time to event endpoints will be summarized by the number and percent of the event, median time and corresponding 95% confidence interval (CI), and event free rate and corresponding 95% CI at the specified time points of interest.

All data summaries and tabulations will be conducted using SAS® software Version 9.4 or higher.

## 6.2 Missing Data

Data that are reported as missing will be treated as missing in all data summaries. Imputation rules for partially recorded dates, in case that the complete dates are required to carry out an analysis, will be provided in the statistical programming plan (SPP). In descriptive summaries for safety, observations that are spurious (extreme relative to the majority of the data) will not be altered or removed from the summary.

## 6.3 Participant Disposition and Baseline Characteristics

### 6.3.1 Participant Disposition

For participant disposition, the number and percentage of participants who reach various study milestones are summarized. All screened participants are broken down by screen failures (with reasons if collected) and enrolled. Then the category of enrolled is broken down by never treated (with reasons if collected) and treated. The category of treated will further be broken down by treatment ongoing and treatment discontinuation (with reasons for discontinuation, which also

include protocol-defined treatment completion, if any). The end of study status for all enrolled participants will also be included.

### 6.3.2 Participant Demographics and Baseline Characteristics

Participant demographics, baseline characteristics, cancer disease history, prior anticancer therapy, and medical history will be summarized using descriptive statistics. For classification of metastatic sites at baseline (e.g. bone ± lymph node, visceral [lung/liver], lymph node only) for participants in Part 1, see **Appendix 1**.

## 6.4 Study Drug Exposures and Concomitant Medications

Study treatment exposures will be summarized by descriptive statistics using the Safety Population.

The summary of study treatment exposure will include descriptive statistics as well as frequency counts for the number of doses or cycles received, the total dose actually administered as well as the total dose intended, and the dose intensity which is calculated as percentage of total dose actually administered divided by total dose intended during whole treatment period. Duration of study treatment exposure will also be summarized. Duration of study treatment in months is calculated as:

(end of treatment visit (EOTV) date – first dose date + 1)/(365.25/12) for participants who have discontinued treatment;

(data cutoff date– first dose date + 1)/(365.25/12) for participants whose treatment is ongoing.

Prior and concomitant medications are coded using the World Health Organization Drug Dictionary. Prior medication is defined as any medication administered from 30 days prior to study treatment administration that was stopped prior to the first dose of study treatment. All concomitant medications, including ADT, prophylactic pre-infusion medications, and blood products administered during the study until 30 days following the last dose of study treatment or until initiation of another anticancer therapy, if earlier, will be recorded. The number and percentage of participants who received prior and concomitant medications will be summarized by Anatomical Therapeutic Chemical (ATC) level 4 and Preferred Term using the ITT population for Part 1 and using Safety Population for Part 2, respectively.

## 6.5 Subsequent Anticancer Therapy

The definition of another anticancer treatment after discontinuation of study treatment includes, but is not limited to, chemotherapy, hormonal therapy, gene therapy, immunotherapy, cryotherapy, biologics, systemic radiation, and surgery for prostate cancer. Number and percentage of participants who received at least one subsequent anticancer therapy will be summarized based on the ITT population for Part 1 and Safety Population for Part 2, respectively. All anticancer therapies will be provided in a data listing.

## 6.6 Protocol Deviations

Major protocol deviations will be identified prior to database lock for final analysis and will be listed and summarized.

## 6.7 Efficacy Endpoint Analyses

### 6.7.1 Primary Efficacy Endpoints Analyses

#### 6.7.1.1 Radiographic Progression-Free Survival per PCWG3 (Part 1)

For participants without an event of PD or death at the time of data cutoff for rPFS analysis, the rPFS will be censored at the date of the last tumor assessment. Specifically, the censoring rules described in **Table 4** will apply to the analysis of rPFS.

**Table 4 Censoring Rules for Primary Analysis of Radiographic Progression-Free Survival**

| Situation | Date | Outcome |
|---|---|---|
| No baseline tumor assessment | Randomization date | Censored |
| Death prior to first scheduled tumor assessment | Date of death | Progressed |
| No post-baseline tumor assessment in absence of death prior to first scheduled tumor assessment | Randomization date | Censored |
| Documented progressive disease | Date of progressive disease | Progressed |
| Initiation of subsequent anticancer therapy in absence of documented progressive disease | Date of last tumor assessment prior to initiation of subsequent anticancer therapy | Censored |
| Death or documented progression immediately after missing two or more consecutive scheduled tumor assessments | Date of last tumor assessment prior to missed tumor assessments | Censored |

##### 6.7.1.1.1 Primary Analysis of rPFS

The Kaplan-Meier method will be used to generate rPFS curves and estimate the median rPFS along with 95% confidence intervals (CIs) for each study treatment arm, based on ITT population. The method of Brookmeyer and Crowley (**2**) will be used to construct 95% CI for median rPFS. The 95% CIs for rPFS rates at 6, 8, and 12 months will be calculated by normal approximation after log(-log) transformation.

The analysis of rPFS will occur when all participants in Part 1 have been enrolled to the vobramitamab duocarmazine arms and followed for rPFS for at least 6 months. The 2.0 mg/kg Q4W dose of vobramitamab duocarmazine will be selected for further study if it demonstrates promising efficacy, unless there is compelling additional activity at the higher dose and an acceptable safety and tolerability profile. Efficacy, safety, and exposure-response variables will inform the selection decision.

#### 6.7.1.1.2 Sensitivity Analysis of rPFS

A sensitivity analysis will be performed that includes documented PD or death as a rPFS event regardless of when it occurs during the study. Specifically, death prior to first scheduled tumor assessment or death immediately after missing two or more consecutive scheduled tumor assessments will be considered as having an event of death. A documented progression immediately after missing two or more consecutive scheduled tumor assessments will be considered as having an event of PD. Other sensitivity analyses may be performed if necessary. Supplementary analyses may be performed based on the Safety Population.

#### 6.7.1.1.3 Subgroup Analysis of rPFS

Subgroup analysis will be performed in the participant subgroups defined by randomization stratification factors, i.e. region, prior cabazitaxel, and presence of visceral disease. The other prognostic factors may also be used to perform subgroup analysis if necessary. The subgroup analysis will be skipped if there are less than 5 events in any subgroup defined by a factor.

### 6.7.1.2 Objective Response Rate per RECIST v1.1 (Part 2)

The number and percent of participants with BOR per RECIST v1.1 will be summarized. ORR and DCR will be summarized, and the 2-sided 95% exact binomial CI will be calculated using the Clopper-Pearson method. For the computation of ORR, participants without any post-baseline measurements are considered as non-responders.

## 6.7.2 Secondary Efficacy Endpoints Analyses

### 6.7.2.1 Prostate-specific Antigen (PSA) Response (Part 1)

PSA response rates at 8 weeks, 12 weeks, and at any time will be summarized. The 2-sided 95% exact binomial CI for PSA response rate will be calculated.

For duration of PSA response, participants without PSA progression at the time of analysis will be censored to the date of their last PSA assessment. Kaplan-Meier method will be applied to estimate duration of PSA response.

The percent change in PSA from baseline over time will be summarized and presented by spider plot. The best PSA percent change from baseline will be presented by waterfall plot.

### 6.7.2.2 Time to Prostate-Specific Antigen Progression (Part 1)

Time to PSA Progression is defined as the time from the date of randomization to the first documented PSA progression. Participants without PSA progression at the time of analysis will be censored to the date of their last PSA assessment. Kaplan-Meier method will be applied to estimate time to PSA progression.

#### 6.7.2.3 Progression-Free Survival per RECIST v1.1 (Part 2)

The same censoring rules and analysis method used for rPFS will be used for PFS analysis (detailed in **Section 6.7.1.1**), but based on Safety Population.

#### 6.7.2.4 Objective Response Rate (Part 1)

BOR, ORR and DCR will be analyzed using the same analysis method as described in **Section 6.7.1.2**.

#### 6.7.2.5 Duration of Response (Part 1 and Part 2)

For participants without an event of PD or death at the time of data cutoff for DoR analysis, the DoR will be censored at the date of the last tumor assessment. The same censoring rules and analysis method used for rPFS will be used for duration of response (detailed in **Section 6.7.1.1**).

DoR = rPFS/PFS event or censoring date – Date of initial tumor response (PR or CR) + 1

The Kaplan-Meier method will be used to generate DoR curves and estimate the median DoR along with 95% CIs. The DoR analyses will be performed only if there are enough responders to render the analyses meaningful.

#### 6.7.2.6 Tumor Size Change Over Time (Part 1 and Part 2)

Descriptive statistics for the best percent change from baseline in Tumor Size (TS) will be provided. A waterfall plot of the best percent change in the TS will be presented. A spider plot of the percentage change in TS versus time for each subject will be presented. Tumor assessments collected after the first PD or after the first next line anti-cancer therapy will be excluded.

#### 6.7.2.7 Time to First Symptomatic Skeletal Event (Part 1)

Participants without an event of Symptomatic Skeletal Event (SSE) at the time of analysis will be censored at the last SSE assessment. The Kaplan-Meier method will be applied to estimate time to first SSE.

### 6.7.3 Exploratory Efficacy Endpoints Analyses

#### 6.7.3.1 Overall Survival (Part 1 and Part 2)

The analysis of OS will be based on the ITT population for Part 1 and Safety Population for Part 2, respectively. For participants without an event of death at the time of data cutoff for OS analysis, OS will be censored at the time the participant is last known to be alive.

The Kaplan-Meier method will be used to generate OS curves and estimate the median OS along with 95% CIs. The method of Brookmeyer and Crowley (**2**) will be used to construct 95% CI for median OS. The 95% CIs for OS rates at 12,18 and 24 months will be calculated by normal approximation after log(-log) transformation. These time points may be adjusted according to

actual data observed in the study without amendment to this SAP. Supplementary analyses may be performed based on the Safety Population.

OS follow-up time will be calculated as the time from randomization (Part 1) or first dose (Part 2) to the date of last known alive or date of death and will be summarized descriptively for all participants and for censored participants, respectively.

Additional supplementary analyses (e.g. stratified by certain prognostic factors, or within different participant population) may be performed for some or all the efficacy endpoints.

## 6.8 Safety Endpoints Analyses

### 6.8.1 Adverse Events Analyses

Treatment-emergent AEs will be summarized in tables and listings. Tables will display the number and percent of participants that experience the given event. AEs prior to the first dose of study treatment (e.g., due to study-related procedures) will be presented in listings only.

AEs will be coded to the Medical Dictionary for Regulatory Activities (MedDRA) dictionary. AEs will be summarized by system organ class (SOC) and preferred term (PT), relationship to study treatment(s), and highest severity. Similar PTs may be combined if necessary for safety topic characterization. High level terms (HLTs) or standardized MedDRA queries (SMQs) may also be used to summarize certain safety topics or AESIs as appropriate. The following AEs will be summarized:

- All treatment-emergent AEs
- Treatment-emergent AEs with CTCAE severity ≥ Grade 3
- Treatment-related AEs
- Treatment-related AEs with CTCAE severity ≥ Grade 3
- SAEs
- Treatment-related sAEs
- AEs resulting in discontinuation of study drug
- AEs leading to interruption of study drug
- Fatal AEs
- Adverse event of special interest (AESIs)

An overall summary of treatment-emergent AEs will display the number and percent of participants who experience at least one event of each of the above types.

### 6.8.2 Laboratory Values Analyses

Summaries of laboratory values will display descriptive statistics for numerically quantified labs. Summaries will be grouped by laboratory panel (hematology, serum chemistry, urinalysis,

endocrine, coagulation and cardiac) and will be displayed by visit for each laboratory parameter. A listing of abnormal values will be provided. A shift table based on lab normal ranges may be produced. Graphs of mean values over time may be generated.

### 6.8.3 Other Safety Endpoints Analyses

ECGs will be collected and analyzed for evidence of cardiac toxicity, especially prolongation of QT interval. The following categories for QTcF interval and maximum post dose change from baseline QTcF interval (ΔQTcF) will be used in summary and shift tables:

QTcF: ≤ 450 msec, >450 to 480 msec, >480 to 500 msec, and >500 msec

ΔQTcF: ≤ 30 msec, >30 to 60 msec, and >60 msec

Vital signs, weight, and ECOG performance status will be summarized with descriptive statistics at each visit and time point where they are collected. Shift tables may be produced.

LVEF will be evaluated by echocardiogram or MUGA and changes from baseline summarized. A graph of LVEF change from baseline will be generated. In addition, time to >15% reduction in LVEF value may be summarized using the Kaplan-Meier method in the Safety Population.

Additional supplementary analyses (e.g. stratified by certain prognostic factors, or within different participant population) may be performed for some or all the safety endpoints.

## 6.9 Pharmacokinetic, Exposure-Response, Immunogenicity, and Biomarker Parameter Endpoint Analyses

**PK Analysis:** Concentrations and PK parameters for vobramitamab duocarmazine will be summarized by dose cohort using descriptive statistics including number of participants, arithmetic mean, standard deviation, minimum, median, maximum, geometric mean, and coefficient of variation (CV%) of the mean and geometric mean. Time-course of drug concentrations will be plotted as appropriate.

Population PK analyses may be conducted using nonlinear mixed effects methodology. Data from this study may be pooled with data from other studies for analysis. A covariate analysis may be performed to relate the effect of intrinsic and extrinsic participant factors to exposure.

**Exposure-Response Analyses**: An exposure-response model to assess the relationship between vobramitamab duocarmazine exposure indices derived from population PK analyses and select efficacy and AEs may be evaluated, e.g., logistic regression curves relating indices of drug exposure (i.e., $C_{max}$ and AUC) to incidence of key efficacy (e.g., rPFS) and safety (AEs).

**Immunogenicity Analysis:** Incidence of ADA, and treatment-emergent and treatment-boosted ADA, will be assessed as absolute occurrence (n) and percent (%) of participants and will include ADA titer level. For participants who have a treatment-emergent ADA positive result, time to first treatment-emergent ADA positive result and the time course of ADA positivity will be summarized.

**Biomarker Analysis:** Summary statistics for biomarker parameters will be summarized. Analysis of B7-H3 expression by IHC on archival or fresh pre-treatment tumor specimens (if no archival specimen is available) may be performed. Relationships between B7-H3 expression and efficacy and other endpoints may be assessed.

## 6.10 Timing of Analyses

### 6.10.1 Interim Analysis

In Part 1, one futility analysis will be performed on each vobramitamab duocarmazine arm based on PSA response rate at 8 weeks when 20 PSA evaluable participants in a vobramitamab duocarmazine arm. At the futility analysis, if the probability of true PSA response rate for a vobramitamab duocarmazine arm greater than or equal to 40% is <10%, that is, if ≤ 25% (5/20) PSA response rate is observed, enrollment to that arm may stop unless other data warrant continuing enrollment.

In Part 2, interim analyses will be performed independently for each disease-specific cohort according to Simon's 2-stage design (see **Section 3.4**) and will be based on confirmed and unconfirmed (preliminary) responses. Enrollment to each cohort will pause after Stage 1 is fully enrolled to await response data. A cohort may proceed to enroll Stage 2 as soon as the futility threshold for that cohort (see **Table 3**) is passed. If all participants in Stage 1 have had at least 2 post-baseline scans or come off study treatment for any reason and the number of responders is less than or equal to the futility threshold indicated in **Table 3**, then the cohort will not proceed to Stage 2 unless other data warrant continuing enrollment.

### 6.10.2 Primary Analysis of rPFS (Part 1)

The analysis of rPFS will occur when all participants have been enrolled to the vobramitamab duocarmazine arms and followed for rPFS for at least 6 months.

## 6.11 Data Standards

Clinical Data Interchange Standards Consortium (CDISC) standards will be used. The latest version of Study Data Tabulation Model (SDTM) will be used for data tabulations of the eCRF data, and the latest version of Analysis Dataset Model (ADaM) will be used for the analysis datasets.

# 7 LIST OF TABLES, LISTINGS AND FIGURES

The list of tables, listings, and figures (TLFs) and associated shells planned for CSR based on the analyses described in this SAP will be provided in a separate statistical programming plan (SPP), which will also include data reporting conventions and programming specifications for the development of these TLFs.

# 8 REFERENCES

1. **Bauer P, Posch M.** Letter to the Editor, Modification of the sample size and the schedule of interim analyses in survival trials based on data inspections. Statistics in Medicine 2004; 23:1333-1335.

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

3. **Carreras M, Gutjahr G, Brannath W.** Adaptive seamless designs with interim treatment selection: a case study in oncology. Statistics in Medicine 2015; 34:1317-1333.

4. **de Wit R, de Bono J, Sternberg CN, Fizazi K, Tombal B, Wulfing C, et al.** Cabazitaxel versus abiraterone or enzalutamide in metastatic prostate cancer. N Engl J Med. 2019; 381:2506-2518.

5. **Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al.** New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009; 45:228-247.

6. **Friede T, Parsons N, Stallard N, Todd S, Valdes Marquez E, Chataway J, et al.** Designing a seamless phase II/III clinical trial using early outcomes for treatment selection: an application in multiple sclerosis. Statistics in Medicine 2011; 30:1528-1540.

7. **Jenkins M, Stone A, Jennison C.** An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. Pharmaceutical Statistics 2011; 10:347-356.

8. **Kenward, M. G., and Roger, J. H.** Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997; 53:983-997.

9. **Scher HI, Morris MJ, Stadler WM, Higano C, Basch E, Fizazi K, et al.** Trial design and objectives for castration-resistant prostate cancer: updated recommendations from the prostate cancer clinical trials working group 3. J Clin Oncol. 2016; 34:1402-1418.

10. **Merseburger AS, Attard G, Åström L, Matveev VB, Bracarda S, Esen A, et al.** Continuous enzalutamide after progression of metastatic castration-resistant prostate cancer treated with docetaxel (PRESIDE): an internati35andomizedomised, phase 3b study. Lancet Oncol. 2022; 23(11):1398-408.

11. **Simon R, Wittes RE, Ellenberg SS.** Randomized phase II clinical trials. Cancer Treat Rep. 1985; 69(12):1375-81.

12. **Simon R,** Optimal two-stage designs for phase II clinical trials. Control Clin Trials, 1989; 10(1):1-10.

13. **Ott PA, Piha-Paul SA, Munster P, Pishvaian MJ, van Brummelen EMJ, et al.** Safety and antitumor activity of the anti-PD-1 antibody pembrolizumab in patients with recurrent carcinoma of the anal canal. Ann Oncol. 2017; 28(5):1036-1041.

14. **Marabelle A, Cassier PA, Fakih M, Kao S, Nielsen D, Italiano A, et al.** Pembrolizumab for previously treated advanced anal squamous cell carcinoma: results from the non-randomised, multicohort, multicentre, phase 2 KEYNOTE-158 study. Lancet Gastro & Hepa. 2022; 7(5):446-454.

15. **Rao S, Anandappa G, Capdevila J, Dahan L, Evesque L, Kim S, Saunders MP, et al.** A phase II study of retifanlimab (INCMGA00012) in patients with squamous carcinoma of the anal canal who have progressed following platinum-based chemotherapy (POD1UM-202). ESNO Open. 2022;7(4):100529.

16. **Cohen EEW, Soulières D, Tourneau CL, Dinis J, Licitra L, Ahn MJ, et al.** Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-0436andomizedomised, open-label, phase 3 study. Lancet. 2019; 393(10167):156-167.

17. **Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al.** Improved survival with Ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010; 363:711-723.

18. **Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, Daud A, et al.** Pembrolizumab versus Ipilimumab in Advanced Melanoma. N Engl J Med. 2015; 372:2521-2532.

19. **Paz-Ares L, Goto Y, Lim DWT, Halmos B, Cho BC, Cobo M, et al.** Canakinumab in combination with docetaxel compared with docetaxel alone for the treatment of advanced non-small cell lung cancer following platinum-based doublet chemotherapy and immunotherapy (CANOPY-2): A multicenter, randomized, double-blind, phase 3 trial. Lung Cancer. 2024; 189:107451.

20. **Trigo J, Subbiah V, Besse B, Moreno V, López R, Sala MA, et al.** Lurbinectedin as second-line treatment for patients with small-cell lung cancer: a single-arm, open-label, phase 2 basket trial. Lancet Oncol. 2020; 21(5):645-654.

21. **Kaira K, Sunaga N, Tomizawa Y, Yanagitani N, Shimizu K, Imai H, et al.** A phase II study of amrubicin, a synthetic 9-aminoanthracycline, in patients with previously treated lung cancer. Lung Cancer. 2010; 69(1):99-104.

22. **Eckardt JR, Pawel JV, Pujol JL, Papai Z, Quoix E, Ardizzoni A, et al.** Phase III study of oral compared with intravenous topotecan as second-line therapy in small-cell lung cancer. Lung Cancer. 2007; 25(15):2086-2092.

# Appendix 1 Classification of Site of Metastases at Baseline by Mutually Exclusive Categories

| Site of Metastases | Presence of LN Metastases [1] | Presence of Bone Metastases [2] | Presence of Visceral Metastases [3] | Category |
|---|---|---|---|---|
| LN only | Yes | No | No | LN |
| Bone only | No | Yes | No | Bone |
| Bone with LN involvement | Yes | Yes | No | Bone |
| Liver | No/Yes | No/Yes | Yes | Liver [4] |
| Lung | No/Yes | No/Yes | Yes | Lung [4] |
| Other | No/Yes | No/Yes | Yes | Other [4] |

Abbreviations: LN: lymph node.

1 Presence of LN metastases defined as Yes for lesion location is lymph nodes (target and non-target lesion at baseline).

2 Presence of bone metastases defined as Yes for lesion location bone at baseline.

3 Presence of visceral metastases defined as Yes for all lesion locations except lymph node and bone (target and non-target lesion at baseline).

4 Any participant with liver metastases is categorized as having liver metastases even if they had other metastatic sites; participants with lung metastases are denoted as having lung metastases, unless they also had liver metastases； participants with visceral metastases are denoted as having other metastases, unless they also had liver or lung metastases (**10**).