NOVARTIS

Clinical Development

Spartalizumab (PDR001), LAG525, capmatinib (INC280), canakinumab (ACZ885)

CPDR001J2201 / NCT03484923

A randomized, open-label, phase II open platform study evaluating the efficacy and safety of novel spartalizumab (PDR001) combinations in previously treated unresectable or metastatic melanoma

Statistical Analysis Plan (SAP)

Author:	Trial Statistician, [redacted]

Document type:	SAP Documentation
Document status:	Final
Release date:	5-Oct-2018
Number of pages:	66

Template Document History – Changes compared to previous version

Date	Version number	Summary of changes
01-Jul-2015	1.0	N/A - First version

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
5-Oct-2018	Prior to DB lock	Creation of final version	N/A - First version	NA

Table of contents

List of abbreviations

ADA	Anti-Drug Antibodies
AE	Adverse event
AESI	Adverse Event of Special Interest
ALP	Alkaline Phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Classification
[redacted]	[redacted]
BID	bis in diem/twice a day
BOR	Best Overall Response
CD8	Cluster of differentiation 8
CI	Confidence Interval
CRS	Case Retriveal Strategy
CSR	Clinical Study report
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
DAR	Dosage Administration Record
DCR	Disease Control Rate
DI	Dose Intensity
DoR	Duration of Response
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
FAS	Full Analysis Set
HLGT	High Level Group Terms
HLT	High Level Terms
[redacted]	[redacted]
IRT	Interactive Response Technology
[redacted]	[redacted]
[redacted]	[redacted]
i.v.	Intravenous
LDH	Lactate Dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
NE	Not Evaluable

NMQ	Novartis MedDRA queries
ORR	Overall Response Rate
OS	Overall Survival
[redacted]	[redacted]
PDI	Planned Dose Intensity
PD-1	Programmed Death 1
PD-L1	Programmed Death-Ligand 1
PD-L2	Programmed Death-Ligand 2
pFBP	Proportion of subjects with a favorable biomarker profile
PFS	Progression-Free Survival
[redacted]	[redacted]
p.o.	Per os/by mouth/orally
PPS	Per-Protocol Set
PT	Preferred Term
Q4W	Every four weeks
RDI	Relative Dose Intensity
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
s.c.	Subcutaneous
SOC	System Organ Class
SMQ	Standardized MedDRA queries
TBL	Total Bilirubin
TA	Tumor Assessment
TIL	Tumor Infiltrating T cells
ULN	Upper Limit of Normal
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the clinical study report(s) (CSR) of study CPDR001J2201, a randomized, open-label, phase II open platform study evaluating and safety of novel spartalizumab (PDR001) combinations in previously treated unresectable or metastatic melanoma.

The content of this SAP is based on protocol CPDR001J2201 version 01. All decisions regarding final analysis, as defined in the SAP document, have been made prior to database lock and unblinding of the study data.

1.1 Study design

This is a randomized, open-label, multi-center, two-part, open platform phase II study to assess the efficacy and safety of the anti-PD-1 antibody spartalizumab in combination with novel agents in previously treated unresectable or metastatic melanoma.

This study consists of two parts:

- Part 1: Selection phase (refer to Section 1.1.1 for further details)
- Part 2 : Expansion phase (refer to Section 1.1.2 for further details)

At any time during the study, subjects deemed eligible after the completion of screening procedures will be randomized into any of the combination arms open to enrolment (in either part 1 or part 2) for which the subject meets the eligibility criteria as described in Section 5 of the protocol using an Interactive Response Technology (IRT) system. Randomization will be stratified by baseline level of lactate dehydrogenase (LDH): baseline LDH ≤ ULN (upper limit of normal) vs. baseline LDH > ULN.

Crossover or re-randomization of subjects to other combination arms, including better performing arms, is not allowed.

Overall response rate (ORR), as assessed by local review of tumor response and using Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 criteria, is the primary endpoint in this study. Duration of Response (DoR), as assessed by local review and using RECIST 1.1 criteria is the key secondary endpoint.

The primary analysis will be performed after all patients randomized in the expansion arm have been followed for at least 9 months or have discontinued study. Timing of the interim analyses is discussed further in Section 1.1.3.

1.1.1 Study design: Part 1: Selection phase

The primary objective of part 1 is to evaluate the preliminary ORR of spartalizumab in combination with novel agents in patients with previously treated unresectable or metastatic melanoma.

As dose regimens for additional new combinations with spartalizumab become established, new combination arms will be added under the same master protocol through protocol amendments. A maximum of 10 arms may be active (open to enrolment) at any given time under this master protocol, and any arms active in part 1 will continue to be randomized with equal probability.

At each interim analysis, it will be determined which arm(s) (1) is declared efficacious and will be expanded to part 2, (2) will continue enrolment in part 1 (up to a maximum of 45 subjects enrolled in a given arm) or (3) will be dropped for futility, taking into account all available efficacy, safety and biomarker data. If a combination arm can be expanded to part 2, the exact sample size for part 2 will be determined at that time when the decision to expand this arm is taken (refer to Section 10.8 of the protocol for further details on sample size considerations). The timing of interim analyses and the criteria used for decision-making at each interim analysis are detailed in Section 1.1.3 and Section 10 of the protocol.

Further discussion on the study design for part 1 can be found in Section 4.1 of the protocol.

1.1.2 Study design: Part 2: Expansion phase

The primary objective for part 2 is to further characterize ORR of spartalizumab in combination with novel agents with relevant preliminary activity, as determined in part 1, for patients with previously treated unresectable or metastatic melanoma.

Only combinations arms that have met the pre-specified criteria defined in Section 1.1.3 and Section 10 of the protocol will be opened to enrollment in part 2.

Exact randomization strategies in the situation when there is one (or more) arms open for enrollment in part 2 and there are still arms open for enrollment in part 1 will be determined at the time of the selection of the arm that will be expanded in part 2. Four factors will be considered at the time of opening an arm in part 2 to determine the exact randomization strategy and those are the following:

1) Number of subjects enrolled in part 1 for the arm that will graduate to part 2
2) Number of subjects that will need to be enrolled in part 2 to have sufficient predictive power and alpha control
3) Observed effect size for the arm that will graduate to part 2
4) Observed effect size of other arms which are open for enrolment in part 1

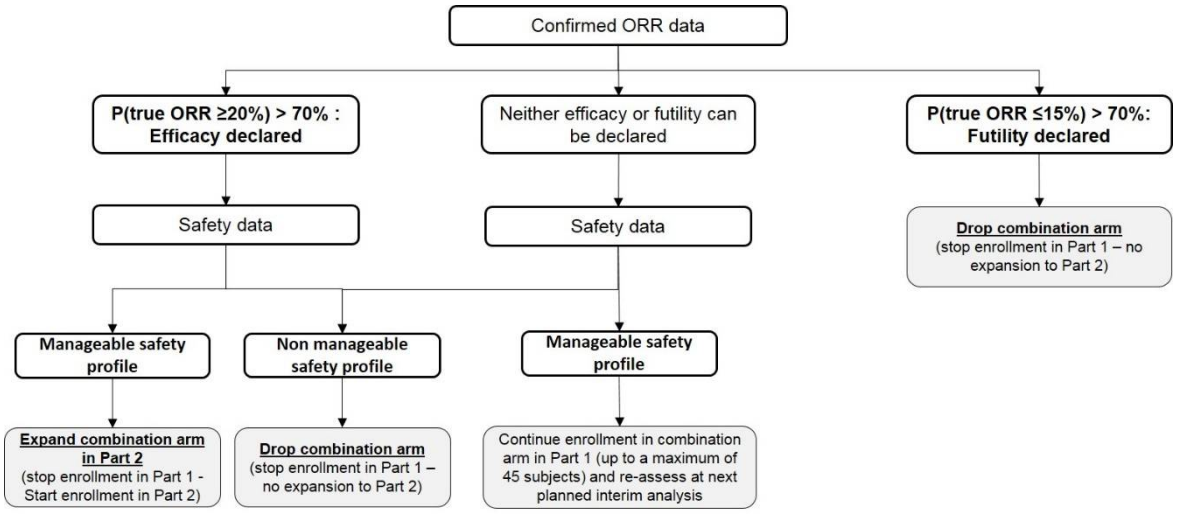
Further discussion on the study design for part 2 can be found in Section 4.1 of the protocol.

1.1.3 Timing of interim analyses and design adaptations

The first interim analysis will be conducted in part 1 after approximately 10 subjects have been enrolled in each of the first three arms and have completed the second post-baseline tumor assessment or have discontinued study prior to completing the second post-baseline tumor assessment. The requirement to have at least 10 subjects in each arm is based on simulations showing that 10 is the minimal number of subjects to be able to declare futility as per decision rules presented below and also in Section 10.4.2 of the protocol. If more combination arms are added to the study via future protocol amendments, those combination arms would also have their 1st interim analysis after approximately 10 subjects have been enrolled. Timing of subsequent interim analyses for part 1 and statistical thresholds are further discussed in Section 4.2 of the protocol and Section 10.4.2 of the protocol, respectively. _1362101091046446 2Statistical_hypoNo interim analysis is planned for part 2.

Figure 1-1 presents pre-specified criteria for decision-making in the selection phase.

Figure 1-1 Pre-specified criteria for decision making at each decision point (interim analysis) in part 1 to determine expansion in part 2 or drop of a combination arm



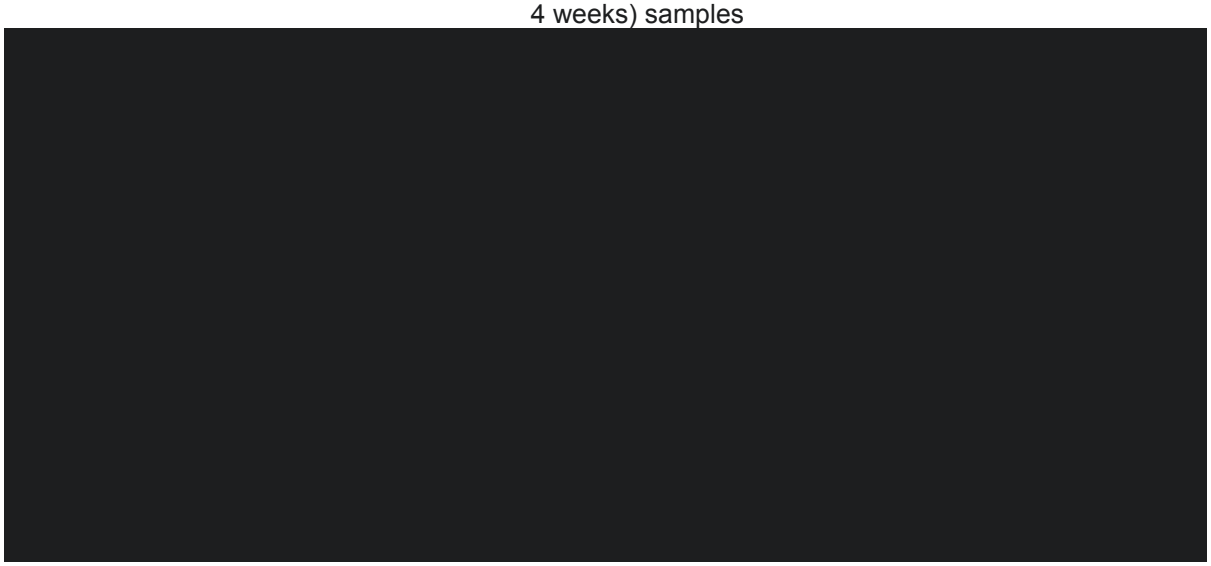
1.2 Study objectives and endpoints

Objectives and related endpoints are described in Table 1-1 below.

Table 1-1 Objectives and related endpoints

Primary objective	**Endpoint**
To evaluate the efficacy of each combination arm, as measured by confirmed objective response rate (ORR)	Confirmed ORR using RECIST v1.1, per local assessment
Key secondary objective	**Endpoint**
To evaluate the efficacy of each combination arm in terms of duration of response (DoR)	DoR using RECIST v1.1, per local assessment
Other secondary objectives	**Endpoint**
To evaluate the efficacy of each combination arm as measured by progression-free survival (PFS) and disease control rate (DCR)	PFS and DCR, assessed using RECIST v1.1, per local assessment
To evaluate the overall survival (OS) of each combination arm	Overall survival (OS)
To characterize the safety and tolerability of each combination arm	Safety: Incidence and severity of adverse event (AE) including changes in laboratory values, vital signs and cardiac assessment. Tolerability: Dose interruptions, reductions, and permanent discontinuations of study treatments

To characterize the prevalence and incidence of immunogenicity of spartalizumab, LAG525 and canakinumab in each combination arm	Anti-drug antibodies (ADA) prevalence at baseline and ADA incidence on treatment
To evaluate changes in levels and phenotype of T cell populations in the tumor and tumor microenvironment after treatment with combination therapies	Proportion of subjects with a favorable biomarker profile (pFBP), as defined by changes in numbers of cells expressing the $CD8^+$ T cell marker and/or T cell activation marker(s), T cell clonality and/or gene expression in tumor biopsy samples collected at baseline (screening) and compared to on treatment (3-4 weeks) samples



2 Statistical methods

2.1 Data analysis general information

All analysis will be performed by Novartis and/or a designated CRO. SAS version 9.4 or later software will be used to perform all data analyses and to generate tables, figures and listings. Analysis using Rshiny apps might also be used as needed.

Data included in the analysis

The analysis cut-off date for the primary analysis of study data will be established after all randomized patients in the expansion arm have either completed nine months of follow-up or have discontinued study. For the selection phase, interim analyses will be conducted once the 1st 10 subjects have been enrolled in each of the 1st three arms and then conducted approximately every 20 weeks thereafter. No formal interim analysis for efficacy will be conducted for combination arms in the expansion phase.

For each of the analyses, all statistical analyses will be performed using all data collected in the database up to the data cut-off date. All data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis. Any data collected beyond the cut-off date will not be included in the analysis and will not be used for any derivations.

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as 'ongoing'. The same rule will be applied to events starting before or on the cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these cases, the end date will not be imputed and therefore will not appear in the listings.

The analysis cutoff date for the final analysis of study data will be established at the end of the study defined as the earliest occurrence of one of the following:

- All subjects have died or discontinued from the study
- Another clinical study (i.e. rollover study) becomes available that can continue to provide the appropriate combination treatment in this subject population and all subjects ongoing are eligible to be transferred to that clinical study
- 52 weeks after data-cut for primary CSR. For subjects still receiving study treatment at the time of end of study, every effort will be made to continue provision of study treatment outside this study through an alternative setting (e.g. post-trial access programs) to subjects who in the opinion of the investigator are still deriving clinical benefit.

The timing of primary and final analyses will be updated, should new arms be added through protocol amendment or if more than one arm is expanded to the expansion phase.

General analysis conventions

Pooling of centers: Unless specified otherwise, data from all study centers will be pooled for the analysis. Due to expected small number of patients enrolled at each centers, no center effect will be assessed.

Qualitative data (e.g., gender, race, etc.) will be summarized by means of contingency tables; a missing category will be included as applicable. Percentages will be calculated using the number of patients in the relevant population or subgroup as the denominator.

Quantitative data (e.g., age, body weight, etc.) will be summarized by appropriate descriptive statistics (i.e. mean, standard deviation, median, minimum, and maximum).

2.1.1 General definitions

Investigational drug and study treatment

Investigational drug can refer to either PDR001 or any of the novel agents.

Study treatment will refer to PDR001 in combination with any of the novel agents.

The term investigational treatment may also be referred to as ***study treatment*** which is used throughout this document.

Date of first administration of investigational drug

The date of first administration of investigational drug is defined as the first date when a non-zero dose of investigational drug is administered and recorded on the Dosage Administration Record (DAR) Electronic Case Report Form (eCRF). The date of first administration of study drug will also be referred as start of investigational drug.

Date of last administration of investigational drug

The date of last administration of investigational drug is defined as the last date when a nonzero dose of investigational drug is administered and recorded on DAR eCRF. The date of last administration of investigational drug will also be referred as end of investigational drug.

Date of first administration of study treatment

The date of first administration of study treatment is derived as the first date when a nonzero dose of any component of study treatment was administered as per the Dosage Administration (e)CRF. (Example: if 1st dose of PDR001 is administered on 28-Jul-2018, and 1st dose of combination partner is administered on 26-Jul-2018, then the date of first administration of study treatment is on 26-Jul-2018). The date of first administration of study treatment will also be referred as ***start of study treatment***.

Date of last administration of study treatment

The date of last administration of study treatment is derived as the last date when a nonzero dose of any component of study treatment was administered as per Dose Administration (e)CRF. (Example: if the last PDR001 dose is administered on 26-Jul-2019, and the last dose of a combination partner is administered on 28-Jul-2019, then the date of last administration of study treatment is on 28-Jul-2019).

Study day

The study day, describes the day of the event or assessment date, relative to the reference start date.

The study day is defined as:

- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date + 1 if event is on or after the reference start date;
- The date of the event (visit date, onset date of an event, assessment date etc.) − reference start date if event precedes the reference start date.

The reference start date for safety assessments (e.g. adverse event onset, laboratory abnormality occurrence, vital sign measurement, dose interruption, [redacted] etc.) is the start of study treatment.

The reference start date for all other, non-safety assessments (i.e., tumor assessment, survival time, disease progression, tumor response, and Eastern Cooperative Oncology Group [ECOG]) performance status is the date of randomization.

The study day will be displayed in the data listings. If an event starts before the reference start date, the study day displayed on the listing will be negative.

Time unit

A year length is defined as 365.25 days. A month length is 30.4375 days (365.25/12). If duration is reported in months, duration in days will be divided by 30.4375. If duration is reported in years, duration in days will be divided by 365.25.

Baseline

For efficacy evaluations, the last non-missing assessment, including unscheduled assessments on or before the date of randomization is taken as "baseline" value or "baseline" assessment. In the context of baseline definition, the efficacy evaluations also include performance status.

For safety evaluations, the last available assessment is on or before the date of start of study treatment is taken or ''baseline" assessment.

In case time of assessment and time of treatment start is captured (e.g. pre-dose ECG), the last available assessment before the treatment start date/time is used for baseline.

For safety parameters (e.g ECGs or vital signs), where study requires multiple replicates per time point, the average of these measurements would be calculated for baseline (if not already available in the database).

In rare cases where multiple measurements meet the baseline definition, with no further flag or label that can identify the chronological order, then the following rule should be applied: If values are from central and local laboratories, the value from central assessment should be considered as baseline. If multiple values are from the same laboratory (local or central) or collected for ECGs or vital signs, then the last value should be considered as baseline.

If patients have no value as defined above, the baseline result will be missing.

On-treatment assessment/event and observation periods

For adverse event reporting the overall observation period will be divided into three mutually exclusive segments:

1. ***pre-treatment period***: from day of patient's informed consent to the day before first administration of study treatment
2. ***on-treatment period***: from date of first administration of study treatment to 30 days after date of last actual administration of any study treatment (including start and stop date)
3. ***post-treatment period***: starting at day 31 after last administration of study treatment.

Notes:

- If data on clock time is available in the clinical database (e.g. for time of blood/urine sample taken, ECG performed, etc. and first study treatment administration), a more precise distinction between pre-treatment and on-treatment periods is encouraged to be used.
- If dates are incomplete in a way that clear assignment to pre-, on-, post-treatment period cannot be made, then the respective data will be assigned to the on-treatment period.

Safety summaries (tables, figures) include only data from the on-treatment period with the exception of baseline data which will also be summarized where appropriate (e.g. change from baseline summaries). In addition, a separate summary for death including on treatment and post treatment deaths will be provided. In particular, summary tables for adverse events (AEs) will summarize only on-treatment events, with a start date during the on-treatment period (***treatment-emergent*** AEs). However, all safety data (including those from the post-treatment period) will be listed and those collected during the pre-treatment and post-treatment period will be flagged.

Additional summaries will be displayed to report deaths, all AEs, AEs related to study treatment, all Serious Adverse Events (SAEs) and SAEs related to study treatment collected up to 150 days after last administration of PDR001.

Note: If a patient starts a post-treatment antineoplastic therapy after the last administration of study treatment, only adverse events suspected to be related to study treatment will be collected up to 150 days after discontinuation of PDR001. All AEs/ all SAEs are planning to be reported up to 150 days after discontinuation of PDR001 acknowledging that complete collection of data may not be feasible for each patient.

However, all safety data (including those from the post-treatment period) will be listed and those collected during the pre-treatment and post-treatment period will be flagged.

Windows for multiple assessments

In order to summarize performance status (ECOG), [redacted] laboratory, and vital signs collected over time (including unscheduled visits), the assessments will be time slotted. The following general rule will be applied in creating the assessment windows: If more than one assessment is done within the same time window, the assessment performed closest to the target date will be used. If 2 assessments within a time window are equidistant from the target date, then the earlier of the 2 assessments will be used. If multiple assessments on the same date then the worst case will be used. Data from all assessments (scheduled and unscheduled), including multiple assessments, will be listed.

Table 2-1 Time window for ECOG PS assessments

Assessment	Target day of assessment	Time Interval
Baseline		*≤ Day 1*
Cycle 2 Day 1	*29*	*Day 15 to day 42*
Cycle 3 Day 1	*57*	*Day 43 to day 70*
Cycle 4 Day 1	*85*	*Day 71 to day 98*
Cycle k Day 1 (k≥5)	*d=(k-1)*28+1*	*Day d-14 to day d+13*
End of Treatment		*Assessment taken at the end of treatment visit*
30-day safety follow-up	*Post treatment study day 30*	*Assessment taken at the 30-day safety follow-up visit*

Table 2-2 Time window for Laboratory assessments

Assessment	Target day of assessment	Time Interval
Baseline		*≤ Day 1*
Cycle 1 Day 15	*15*	*Day 2 to day 21*
Cycle 2 Day 1	*29*	*Day 22 to day 42*
Cycle 3 Day 1	*57*	*Day 43 to day 63*
Cycle 3 Day 15	*71*	*Day 64 to day 77*
Cycle 4 Day 1	*85*	*Day 78 to day 98*
Cycle k Day 1 (k≥5)	*d=(k-1)*28+1*	*Day d-14 to day d+13*
End of Treatment		*Assessment taken at the end of treatment visit*
30-day safety follow-up	*Post treatment study day 30*	*Assessment taken at the 30-day safety follow-up visit*

Last contact date

The last contact date will be derived for patients not known to have died at the analysis cut-off using the last complete date among the following:

Table 2-3 Last contact date data sources

Source data	Conditions
Date of Randomization	No Condition
Last contact date/last date patient was known to be alive from Survival Follow-up page	Patient status is reported to be alive, lost to follow-up or unknown
Start/End dates from further antineoplastic therapy	Non-missing medication/procedure term.
Start/End dates from drug administration record	Non-missing dose. Doses of 0 are allowed.
End of treatment date from end of treatment page	No condition.
Tumor (RECIST) assessment date.	Evaluation is marked as ‘done’.
Verification for treatment beyond RECIST1.1 PD	At least one non missing parameter value
Laboratory [redacted] collection dates	Sample collection marked as ‘done’.
Vital signs date	At least one non-missing parameter value
Performance Status date	Non-missing performance status
Start/End dates of AE	Non-missing verbatim term

The last contact date is defined as the latest complete date from the above list on or before the data cut-off date. The cut-off date will not be used for last contact date, unless the patient was seen or contacted on that date. No date post cut-off date will be used. Completely imputed dates (e.g. the analysis cut-off date programmatically imputed to replace the missing end date of a dose administration record) will not be used to derive the last contact date. Partial date imputation is allowed for event (death)/censoring is coming from ‘Survival information’ eCRF.

The last contact date will be used for censoring of patients in the analysis of overall survival.

2.2 Analysis sets

All below defined analysis sets will be used to analyze data from both the selection phase (part 1) and the expansion phase (part 2).

Full Analysis Set

The Full Analysis Set (FAS) comprises all subjects to whom study treatment has been assigned by randomization. According to the intent to treat principle, subjects will be analyzed according to the treatment they have been assigned to during the randomization procedure regardless of whether or not treatment was administered. This population will be the primary population for efficacy analysis.

Per protocol set (PPS)

Not applicable.

Safety

The Safety Set includes all subjects who received at least one dose of study treatment. Subjects will be analyzed according to the study treatment received, where treatment received is defined as the randomized treatment if the subject took at least one dose of that treatment or the first treatment received if the randomized treatment was never received.

The actual treatment received corresponds to:

- the randomized treatment if patients took at least one dose of that treatment.
- the first treatment received if the randomized treatment was never received

Note: if a patient is assigned to a dose level for instance, PDR001 400 mg and just receives partial infusion of the treatment, the patient will still be analyzed under PDR001 400 mg treatment arm for safety analysis.

Biomarker analysis set

The biomarker analysis set, as it pertains to the secondary objective (Section 10.5.5 of the protocol), consists of all subjects in the safety set with an evaluable baseline tumor biopsy sample and at least one evaluable post-baseline tumor biopsy sample. [redacted]

Immunogenicity (IG) analysis sets

The *Immunogenicity prevalence set* includes all subjects in the Full analysis set with a determinant baseline IG sample **or** at least one determinant post-baseline IG sample.

The *Immunogenicity incidence set* includes all subjects in the Immunogenicity prevalence set with a determinant baseline IG sample **and** at least one determinant post-baseline IG sample.

See Section 2.10.1.1 for the definition of *determinant*.

Patient Classification:

Patients may be excluded from the analysis populations defined above based on the protocol deviations entered in the database and/or on specific subject classification rules defined in Table 2-4.

Table 2-4 Subject classification based on protocol deviations and non-PD criteria

Analysis set	Protocol deviations leading to exclusion	Non protocol deviation leading to exclusion
FAS	No written inform consent	Not applicable
Safety set	No written inform consent	No dose of study medication
Biomarker set	No written inform consent	No dose of study medication and no evaluable post-

		baseline tumor biopsy sample
[redacted]		
Immunogenicity Prevalence Set	No written informed consent	Not meeting the definition for inclusion in the Immunogenicity Prevalence set
Immunogenicity Incidence Set	No written informed consent	Not meeting the definition for inclusion in the Immunogenicity Incidence set

Withdrawal of Informed Consent

Any data collected in the clinical database after a subject withdraws informed consent from all further participation in the trial, will not be included in the analysis. The date on which a patient withdraws full consent is recorded in the eCRF.

Death events may be used in the analysis if captured from public records (registers), local law and subject informed consent pending.

Additional data for which there is a separate informed consent, e.g. [redacted] biomarker etc., collected in the clinical database without having obtained that consent will not be included in the analysis. These data will be excluded by the presence of the appropriate protocol deviation criterion.

2.2.1 Subgroup of interest

Efficacy

The primary efficacy endpoint will be summarized ***if there is sufficient number of patients*** in the following subgroups:

- Stratification factor (based on randomization data from IRT)
  - LDH (≤ ULN vs. > ULN)
- BRAFV600 mutation status (wild-type vs. mutant)

In addition, if a combination arm is expanded, the following subgroup analysis may also be performed ***if there is sufficient number of patients*** using all patients randomized within that combination arm:

- Age category (< 65 years, ≥ 65 years)
- ECOG PS (0 vs ≥0)
- PD-L1 expression (positive, negative) where a positive status is defined as having ≥ 1% expression and negative is defined as having <1% expression

No formal statistical test of hypotheses will be performed for the subgroups, only point estimate of the treatment effect and 95%-confidence intervals will be provided (see Section 2.5.2 and Section 5.4.1 for further analysis details). The objective of the efficacy subgroup analysis is to demonstrate homogeneity of treatment effect in the above subgroups, and to identify potential subgroups with differential efficacy. Summary tables will only be performed if at least 10% of patients or 10 patients are present in each subgroup.

Safety

Safety subgroup analyses will use the same method as for the analysis in the overall analysis set. Key safety analyses (e.g. deaths, SAEs, AESIs, AEs leading to discontinuation) will be repeated on safety set in the following subgroups:

- Sex
- Race
- Age category (< 65 years, ≥ 65 years)

The objective for carrying out these subgroup analyses is to identify potential safety issues that may be limited to a subgroup of patients, or safety issues that are more commonly observed in a subgroup of patients. Summary tables will only be performed if at least 10% of patients or 10 patients are present in each subgroup.

2.3 Patient disposition, demographics and other baseline characteristics

The Full Analysis Set (FAS) will be used for all baseline (including disease characteristics) and demographic summaries and listings unless otherwise specified. Summaries will be reported by treatment arm and for all patients and listings will be reported by treatment arm to assess baseline comparability. No inferential statistics will be provided.

Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed by treatment arm. Categorical data (e.g. gender, age groups: <65 and ≥ 65 years, race, ethnicity, etc...) will be summarized by frequency counts and percentages; the number and percentage of patients with missing data will be provided. Continuous data (e.g. age, weight, height, body mass index) will be summarized by descriptive statistics (N, mean, median, standard deviation,

minimum and maximum). BMI (kg/m2) will be calculated as weight[kg] / (height[m]2) using weight at Screening.

Baseline stratification factors

The number (%) of patients in each stratum (LDH (≤ ULN vs. > ULN)) based on data obtained from the IRT system will be summarized overall and by treatment arm for the FAS. Discordances between the stratum recorded in IRT at the time of randomization and the actual stratum recorded in the clinical database through the data collected on eCRF will be cross-tabulated and listed.

Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. This analysis will include the following: primary site of cancer, details of tumor histology/cytology, histological grade, stage at initial diagnosis, time since initial diagnosis, time from initial diagnosis to first recurrence/progression (in months), time since most recent relapse/progression to randomization (in months), stage at time of study entry, presence/absence of target and non-target lesions, number and type of metastatic sites involved, and number of tumor infiltrating T cells (TIL)). Note: Presence/absence of target and non-target lesions will be based on the data collected on RECIST target/non-target lesion assessment eCRF pages. Metastatic sites will be based on diagnosis page.

Medical history

Medical history and ongoing conditions, including cancer-related conditions and symptoms entered on (e) CRF will be summarized and listed by treatment arm. The summaries will be presented by primary system organ class (SOC), preferred term (PT) and treatment arm. Medical history and current medical conditions will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology. The MedDRA version used for reporting will be specified in the CSR and as a footnote in the applicable tables/listings.

Other

All data collected at baseline such as child bearing potential, [redacted] treatment beyond progression informed consent will be listed.

2.3.1 Patient disposition

Enrollment by country and center will be summarized for all screened patients and also by treatment arm using the FAS. The number (%) of randomized patients included in the FAS will be presented overall and by treatment group. The number (%) of screened and not-randomized patients and the reasons for screening failure will also be displayed. The number (%) of patients in the FAS who are still on treatment, who discontinued the study phases and the reason for discontinuation will be presented overall and by treatment group.

The following summaries will be provided (with % based on the total number of FAS patients):

- Number (%) of patients who were randomized (based on data from IRT system)

- Number (%) of patients who were randomized but not treated (based on 'DAR' eCRF page not completed for any study treatment component)
- Primary reason for not being treated (based on "End of Treatment Phase Completion" eCRF page)
- Number (%) of patients who were treated (based on 'DAR' eCRF pages of each study treatment component completed with non-zero dose administered)
- Number (%) of patients who are still on-treatment (based on the 'End of Treatment Phase' page not completed);
- Number (%) of patients who discontinued the study treatment phase (based on the 'End of Treatment Phase' page)
- Primary reason for study treatment phase discontinuation (based on the 'End of Treatment Phase' page)
- Number (%) of patients who have entered the post-treatment follow-up (based on the 'End of Treatment Phase' page);
- Number (%) of patients who have discontinued from the post-treatment follow-up (based on the End of Post-treatment follow-up page);
- Reasons for discontinuation from the post-treatment follow-up (based on End of Post-treatment follow-up page);
- Number (%) of patients who have entered the survival follow-up (based on the 'End of Treatment Phase' or 'End of Post-treatment follow-up' page).

Protocol deviations

The number (%) of patients in the FAS with any protocol deviation will be tabulated by deviation category (as specified in the study Data Handling Plan) overall and by treatment group for the FAS. Major protocol deviations leading to exclusion from analysis sets will be tabulated separately overall and by treatment group. All protocol deviations will be listed.

Analysis sets

The number (%) of patients in each analysis set (defined in Section 2.2) will be summarized by treatment group and stratum.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

The Safety set will be used for the analyses below. Categorical data will be summarized as frequencies and percentages. For continuous data, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum will be presented.

2.4.1 Study treatment / compliance

Duration of exposure, actual cumulative dose, dose intensity (DI) and relative dose intensity (RDI) will be summarized by combination arm, separately for each component of study treatment (PDR001 and appropriate novel agent). The duration of treatment will also be presented for each combination arm. Duration of exposure will be categorized into time intervals; frequency counts and percentages will be presented for the number (%) of subjects in each interval. The number (%) of subjects who have dose reductions or interruptions, and the reasons, will be summarized by combination arm.

Subject level listings of all doses administered on treatment along with dose change reasons will be produced.

The safety set will be used for all summaries and listings of study treatment.

Duration of exposure to study treatment

Duration of exposure to study treatment is considered by taking into account the duration of exposure to PDR001 and any novel agent partner for each combination arm:

Duration of exposure to study treatment (days) for each combination arm = (last date of exposure to study treatment) – (date of first administration of study treatment) + 1.

The last date of exposure to study treatment for each combination arm is the latest of the last dates of exposure to PDR001 or any novel agent partner (see Table 2-5).

Summary of duration of exposure of study treatment in appropriate time units will include categorical summaries and continuous summaries (i.e. mean, standard deviation etc.) using appropriate units of time.

Duration of exposure to PDR001 and novel agent combination partner

Duration of exposure to PDR001 (days) = (last date of exposure to PDR001) – (date of first administration of PDR001) + 1.

Table 2-5 Definition of last date of exposure of study drug

Scenario	Definition of last date of exposure of study drug	Example
PDR001 LAG525 ACZ885	The planned end date of the last cycle in which the last non-zero dose of the investigational drug was last administered (i.e last date of administration + (planned interval duration-1day)) Note : If the patient died or was lost to follow-up before the derived last date, the last date of exposure to investigational drug is the date of death or the date of last contact, respectively.	Example 1: If PDR001, LAG525, or ACZ885 is administered every four weeks, the last date of exposure is the date of administration in the last cycle + 27 days.

	If the derived last date of exposure goes beyond the data cutoff date, it should be truncated to the date of data cutoff.	
INC280	Date of last administration of a non - zero dose of the study drug.	Example 2: A patient had a permanent discontinuation of the study drug 06Jan2013 after being put on a temporary interruption since 01Jan2013. In this case the last date of exposure is- 31Dec2012.
Scenario 1: Novel agent with a cyclic administration other than Q4W	*The planned end date of the last cycle in which the last non-zero dose of the investigational drug was last administered.* *Note : If the patient died or was lost to follow-up before the derived last date, the last date of exposure to investigational drug is the date of death or the date of last contact, respectively.* *If the derived last date of exposure goes beyond the data cutoff date, it should be truncated to the date of data cutoff.*	*Example 1: In a once a week administration of an investigational drug, the last date of exposure is the date of administration in the last cycle + 6 days.* *Example 2: In a 21-day cycle with one or several infusions in the beginning of the cycle, the last date of exposure is the date of first infusion in the last cycle + 20 days.*

As more novel agents are combined with PDR001, the above table will be updated, and the novel agents added to the appropriate rows in the above table.

Summary of duration of exposure of PDR001 will include categorical summaries based on 28 day intervals and using descriptive statistics (mean, standard deviation etc).

The duration includes the periods of temporary interruption.

Cumulative dose

Cumulative dose of a study treatment is defined as the total dose given during the study treatment exposure and will be summarized for each of the study treatment components.

The **planned cumulative dose** for a study treatment component refers to the total planned dose as per the protocol up to the last date of investigational drug administration. The planned cumulative dose is not summarized or listed. It is used for relative dose intensity calculations.

The **actual cumulative dose** refers to the total actual dose administered, over the duration for which the subject is on the study treatment as documented in the Dose Administration eCRF.

For patients who did not take any drug the cumulative dose is by definition equal to zero.

For continuous dosing, the actual cumulative dose is the sum of the non-zero doses recorded over the dosing period and the planned cumulative dose is the planned starting dose summed over the same dosing period.

For intermittent dosing, the actual cumulative dose should be defined based on the days when the subject is assumed to have taken a non-zero dose during dosing periods.

Dose intensity and relative dose intensity

Dose intensity (DI) for patients with non-zero duration of exposure is defined as follows:

- ***For PDR001, LAG525, and Canakinumab (ACZ885):***
  - DI (mg / week) = Actual Cumulative dose (mg) / Duration of exposure to study treatment (week).
  - If for example, the duration of exposure to PDR001 is 8 weeks and the patient receives two injections of PDR001 at 400 mg :
  - DI (mg/week) = Actual Cumulative dose (mg) / Duration of exposure (week)

    = 800 (mg) / 8 (week) = 100 (mg/week)
- ***For Capmatinib (INC280):***
  - DI (mg/day)= Actual Cumulative dose (mg)/ Duration of exposure (in days)

For patients who did not take any drug the DI is by definition equal to zero.

Planned dose intensity (PDI) is defined as follows:

- ***For PDR001, LAG525, and Canakinumab (ACZ885):***
  - PDI (mg/week) = Planned Cumulative dose (mg) / Duration of exposure (week).
- ***For Capmatinib (INC280):***
  - PDI (mg/day)= Planned Cumulative dose (mg)/ Duration of exposure (in days).

Relative dose intensity (RDI) is defined as follows:

- ***For PDR001, LAG525, and Canakinumab (ACZ885):***
  - RDI = DI (mg /week) / PDI (mg /week).
- ***For Capmatinib (INC280):***
  - RDI = DI (mg/day) / PDI (mg /day).

For each PDR001 and novel agent combination, DI and RDI will be summarized separately for each of the study treatment components, but using the duration of exposure of each of the components.

Dose reductions, interruptions or permanent discontinuations

The number of subjects who have dose reductions, dose escalations, permanent discontinuations, administration stopped/paused during infusion or interruptions, and the reasons, will be summarized separately for each of the study treatment components. Dose reductions and dose escalations are only allowed for Capmatinib (INC280) (see Table 2-6).

'Dose changed', "Was drug administration stopped or paused" and 'Dose permanently discontinued' fields from the Dosage Administration CRF pages (DAR) will be used to determine the dose reductions, dose escalations, dose interruptions, administration stopped /paused and permanent discontinuations, respectively.

The corresponding fields 'Reason for dose change/dose interrupted', 'Reason for administration stopped or paused' and 'Reason for permanent discontinuation' will be used to summarize the reasons.

A dose change is either 'change in prescribed dose level' or 'dosing error' where actual dose administered/total daily dose is different from the prescribed dose.

Note: Dose interruptions are captured in the eCRF following the Oncology Data standard Reference manual and CRF Completions Guidelines. Depending on the dosing administration schedule (e.g. continuous dosing, interval dosing), data would be collected in different fashion. For instance, for continuous dosing, an interruption is defined as a dose of zero in a unit of time between two non-zero dosing records. In general, any rest period as part of the regimen/schedule is not considered as an interruption.

For the purpose of summarizing interruptions and reasons, in case multiple entries for interruption are entered on consecutive days/dose administrations with different reasons, separate interruptions will be counted. However, if the reason is the same in this mentioned multiple entries on consecutive days/dose administrations, then it will be counted as one interruption.

Reduction: A dose change where the prescribed dose level is lower than the previous prescribed dose level or where the actual dose administered/total daily dose is lower than the calculated dose amount based on the prescribed dose. Therefore any dose change to correct a dosing error will not be considered a dose reduction. Only dose change is collected in the CRF, number of reductions will be derived programmatically based on the change and the direction of the change.

Table 2-6 DAR record example for capmatinib (INC280)

DAR record number	Prescribed daily dose	Administered daily dose	Dose reduction Yes/No	Comment
1	800	800	No	
2	800	1000	No	Dosing error
3	800	800	No	Correcting dosing error
4	800	600	Yes	Due to AE
5	600	400	Yes	Dosing error
6	600	600	No	Correcting dosing error
7	400	600	Yes	Protocol-prescribed dose reduction (e.g. due to AE), dosing error
8	600	600	No	Dose increase
9	600	400	Yes	Dosing error

10	400	400	Yes	Dose reduction due to lower prescribed dose, although actual dose is not lower than before.

Note: It is assumed that if dose reduction is yes, that dose change was checked.

Escalation: For capmatinib, if the dose has been reduced per the mandatory dose modification, escalation to the previous dose level is permitted if the following criteria are met.

- A period of 4 weeks of treatment has passed since restarting dosing at the reduced dose level and there is no recurrence of the AE
- Subject is deriving clinical benefit.

Note: A dose rechallenge i.e when a subject goes back to the previous prescribed dose after an interruption, is not considered as an increase.

Missing data: If dose is recorded but regimen is missing or entered as 'none', it is assumed that the investigational drug was taken as per-protocol.

Treatment beyond RECIST progression

The number of patients who continue treatment beyond RECIST1.1 progression according to local investigators assessment based on protocol specified criteria will be summarized. It includes all patients who received any study treatment (i.e. at least one dose of any component of the study treatment) after RECIST 1.1 progression assessed by local investigators. Those patients will be identified using the field "Will the subject continue treatment beyond disease progression as per RECIST1.1?" on the Verification for Treatment beyond RECIST1.1 PD CRF pages.

2.4.2 Prior, concomitant and post therapies

Prior anti-cancer therapy

The number and percentage of patients who received any prior anti-neoplastic medications, prior checkpoint inhibitor therapy, prior anti-neoplastic radiotherapy prior anti-neoplastic surgery, or other therapies not defined by the previous categories, will be summarized by treatment arm. Prior anti-neoplastic medications will be summarized by therapy type (e.g. chemotherapy, hormonal therapy, immunotherapy etc.), setting (e.g. adjuvant, metastatic, etc.) and also by lowest Anatomical Therapeutic Classification (ATC) class, preferred term and treatment. Summaries will include total number of regimens, time from last treatment to progression for the last therapy, and type of therapy received as last therapy for unresectable and metastatic melanoma prior to enrollment (checkpoint inhibitor vs targeted therapy). Also, the prior checkpoint inhibitor therapy will be summarized by type of checkpoint inhibitor therapy (Anti-PD-1, Anti-CTLA-4, and combination therapy of anti-CTLA-4 and anti-PD-1). The medication therapy type of any combination therapy will be classified based on the following order: immunotherapy, chemotherapy, biologic therapy (other than immunotherapy), targeted therapy, hormonal therapy. For example, a combination therapy of chemotherapy and immunotherapy will be classified as 'immunotherapy'. For radiotherapy, time since last

radiotherapy, locations and setting of last therapy will be summarized. For prior surgery, time since last surgery and procedure will be summarized.

Separate listings will be produced for prior anti-neoplastic medications, radiotherapy, and surgery.

Anti-neoplastic medications will be coded using the World Health Organization (WHO) Drug Dictionary (WHO-DD); anti-neoplastic surgery will be coded using MedDRA. Details regarding MedDRA and WHO-DD version will be included in the footnote in the tables/listings.

The above analyses will be performed using the FAS.

Post treatment anti-cancer therapy

Anti-neoplastic therapies since discontinuation of study treatment will be listed and summarized by ATC class, preferred term, overall and by treatment group by means of frequency counts and percentages using FAS.

The number and percentages who reported taking at least one anti-neoplastic therapy since discontinuation of study treatment by category (see Table 2-7) will be summarized by treatment group.

Table 2-7 Categories of anti-neoplastic therapies since discontinuation

Category	Sub-category
Immunotherapy	Anti-CTLA-4 Anti-PD-1/Anti-PD-L1 Combination Anti-CTLA-4 + Anti-PD-1/Anti-PD-L1 IL-2 Interferon Other or investigational
Targeted therapy	
Chemotherapy	
Radiotherapy	
Surgery	
Other or investigational	

Concomitant medications

Concomitant therapy is defined as all interventions (therapeutic treatments and procedures) other than the study treatment administered to a patient coinciding with the study treatment period. Concomitant therapy include medications (other than study drugs) starting on or after the start date of study treatment or medications starting prior to the start date of study treatment and continuing after the start date of study treatment.

Concomitant medications will be coded using the World Health Organization (WHO) Drug Reference Listing (DRL) dictionary that employs the WHO Anatomical Therapeutic Chemical (ATC) classification system and summarized by lowest ATC class and preferred term using

frequency counts and percentages. Surgical and medical procedures will be coded using MedDRA and summarized by SOC and preferred term. Concomitant medications with immunosuppressive intent will be summarized by lowest ATC class and preferred term using frequency counts and percentages.

All reported systemic steroids will be summarized by lowest ATC class and preferred term using frequency counts and percentages. Reported cumulative dose and maximum cumulative dose, adjusted for relative anti-inflammatory potency [Schimmer and Parker 1996], will also be summarized along with route of administration.

Table 2-8 Systemic steroid anti-inflammatory potency relative to Hydrocortisone

Steroid Name	Anti-inflammatory potency, relative to Hydrocortisone
Hydrocortisone	1
Cortisone acetate	0.8
Prednisone	4
Prednisolone	4
Methylprednisolone	5
Dexamethasone	30

All summaries will include:

1. Medications starting on or after the start of study treatment but no later than 30 days after last dose of study treatment and
2. Medications starting prior to start of study treatment and continuing after the start of study treatment.

Additional summaries will be provided to report medications starting between 31 days after last dose of study treatment and 150 days after last dose of PDR001.

Note: If a patient starts a post-treatment antineoplastic therapy after the last administration of study treatment, only medications relative to the AEs/SAEs that are reported are collected up to 150 days after discontinuation of PDR001/placebo. Medications starting between 31 days after last dose of study treatment and 150 days after last dose of PDR001 will be summarized acknowledging that complete collection of data may not be feasible for each patient.

All reported concomitant therapies will be listed. Any concomitant therapies starting and ending prior to the start of study treatment or starting more than 150 days after the last dose of PDR001 or 30 days after last dose of study treatment whichever comes last, will be flagged in the listing. The safety set will be used for all concomitant medication tables and listings.

2.5 Analysis of the primary objective

The primary objective is to demonstrate the antitumor activity of spartalizumab in combination with novel agents as measured by ORR in previously treated unresectable or metastatic melanoma patients.

2.5.1 Primary endpoint

Confirmed ORR is defined as the proportion of patients with confirmed best overall response (BOR) of complete response (CR) or partial response (PR) according to RECIST 1.1 (see Appendix 2 of the protocol). For the primary analysis, ORR will be calculated based on the FAS using local investigators review of tumor assessment data. ORR will be summarized using descriptive statistics (N,%) along with 2-sided exact 95% confidence interval (CI) [Clopper and Pearson 1934]. Tumor assessments performed before the start of any further antineoplastic therapy (i.e. any additional secondary antineoplastic therapy or surgery) will be considered in the assessment of BOR. Limited-field palliative radiotherapy may be allowed after documented discussion with Novartis (see Section 6.4.4 of the protocol). For patients who receive limited-field palliative therapy, this therapy should not be considered in determining anti-neoplastic therapy usage.

2.5.2 Statistical hypothesis, model, and method of analysis

Part 1: Selection Phase

No formal hypothesis testing will be conducted in part 1. The objective of this part of the study is to explore the anti-tumor activity of the combination of spartalizumab with novel agents as measured by ORR per RECIST v1.1, and to select an arm(s) with promising efficacy that will be advanced to the expansion phase.

To determine whether arms have the potential to demonstrate clinically meaningful response rates, arms will need to show that they cross a specific efficacy probability threshold during the selection phase to advance to the expansion phase. Conversely, arms can also be shown to cross a specific futility probability threshold and consequently will be declared as futile and enrollment in those arms will be terminated. These thresholds and probabilities are highlighted in Table 2-9.

Table 2-9 Probabilities thresholds used to declare efficacy or futility for an arm

Decision rule	Action
P(ORR ≤15%) ≥ 70%	Drop arm for futility
P(ORR ≥20%) ≥ 70%	Efficacy declared; Advance arm to expansion phase (part 2)

The futility decision rule was determined based on simulations such that arms with response rate that is not considered clinically meaningful (i.e. ORR < 10%) will be dropped with high probability (as illustrated in Table 10-7 of the protocol). Similarly, the decision criterion for advancement to the expansion phase was determined so that arms with clinically meaningful response rate that warrants further development (i.e. ORR>25%) will be advanced with high probability ('power') (see also Table 10-7 of the protocol).

To calculate these probabilities, a beta-binomial posterior distribution will be used as shown below:

$$p_i \sim beta(y_i + 1, n_i - y_i + 1)$$

where y_i are the number of confirmed responses in a combination arm (i), and n_i are the number of subjects enrolled in a combination arm (i). In calculation of the posterior distribution, the uniformed prior, Beta[1,1] will be taken into account for each arm. The posterior probability will be calculated separately for each arm based on the observed ORR based on RECIST v1.1 criteria at each decision point, i.e. at each interim analysis as described in Section 10.7 of the protocol.

No formal comparisons will be made between the combination arms for ORR in this 'selection' phase; instead each of the arms will be evaluated against efficacy and futility thresholds as presented above.

The final decision on whether an arm will be selected for advancement to the expansion phase will primarily be based on the observed ORR using decision rule described above but other efficacy (in particular duration of response), safety, and biomarker data will be considered as well. Further considerations related to the design and execution of the selection phase will be captured in the [SAP Technical Appendix].

Part 2: Expansion Phase

For the arms which show promise for efficacy in part 1 (based on pre-defined criteria), these arms will be expanded. After completion of the expansion phase and within each arm separately, the following statistical hypothesis will be tested based on objective response rate (ORR):

- The null hypothesis is that the response rate is not clinically meaningful (≤ 10%).
- The alternative hypothesis is that the response rate is clinically meaningful (>10%). The target response rate that warrants further development is ≥ 25%

The null hypothesis will be formally rejected if the lower bound of exact 95% CI for ORR is above 10%. **The primary analysis of ORR at the end of part 2 will be based on a combined pool of part 1 and part 2 data for any given arm.**

For example, assuming 24 subjects enrolled in an arm of interest in part 1 (see Section 10.8 of the protocol for details), assuming this arm met criteria for expansion and finally assuming 50 subjects are enrolled in the expansion phase, then combined among 74 subjects enrolled in this arm in total and followed sufficiently, 14 or more responses need to be observed to reject the null hypothesis. With those 14 responders among 74 subjects, the point estimate for ORR will be 18.9% with an exact 95% confidence interval of (10.8%, 29.7%). Since the lower bound of the exact 95% CI will be > 10%, this leads to rejection of null hypothesis. The exact 95% CIs for potential observed ORR in 74 subjects are shown in Table 2-10:

Table 2-10 Exact binomial 95% confidence intervals around potential observed ORR for N=74

Number of responders	Observed ORR (%)	95% exact CI (%)
10	13.5	6.7, 23.5
11	14.9	7.7, 25.0
12	16.2	8.7, 26.6
13	17.6	9.7, 28.2
14	**18.9**	**10.8, 29.7**

Number of responders	Observed ORR (%)	95% exact CI (%)
15	20.3	11.8, 31.2
16	21.6	12.9, 32.7
17	23.0	14.0, 34.2
18	24.3	15.1, 35.7
19	**25.7**	**16.2, 37.2**
20	27.0	17.4, 38.6
21	28.4	18.5, 40.1
22	29.7	19.7, 41.5
23	31.1	20.8. 42.9
24	32.4	22.0, 44.3

Fifty subjects enrolled in the expansion phase considered in the example above also correspond to the number of subjects needed to ensure that the hypothesis to be tested after completion of part 2 has sufficient predictive power conditioned on the ORR observed in the selection phase (part 1) (see Section 10.8 of the protocol).

Alpha Control Considerations

For each arm (treatment combination cohort) included in the study, the potential for alpha inflation arises from the adaptive nature of the design, and in particular, from the presence of a selection process based on probabilistic decision rules described in Table 2-9.

To assess whether alpha is controlled for each combination treatment separately, simulations were conducted. Below in Table 2-11, simulation results for scenarios that include one or more arms with the true ORR of 0.10 corresponding to the null hypothesis threshold (see also Scenarios 1 and 2 in presented in Section 10.8 of the protocol), are presented. Scenario 2 is investigated to assess a potential impact of shrinkage estimator on type I error inflation. In simulations, the focus was on assessing type I error for each 'null hypothesis' arm separately defined as the proportion of significant results after completion of part 2 for given arm (successful arm). The technical details for the simulations are summarized in the [SAP Technical Appendix].

As shown in Table 2-11, for each of the 'null hypothesis' arms with the true ORR = 0.10, the probability is less than 0.025 (ranging between 0.0093 and 0.0126) and therefore alpha is well-controlled. Even though there no explicit alpha control measure used, alpha within each arm is still well-controlled due to the use of the futility criterion in the selection phase, as well as the adaptive nature of the sample size of part 2, based on the shrinkage estimator. Further simulation results for alpha for other scenarios and under a wide variety of samples sizes for the expansion arm will be presented in the [SAP Technical Appendix].

Table 2-11 Type I error assessment: Probability of a successful result for scenario where the true ORR equals 0.10 for each arm and sample size of the expansion arm is 50 subjects

Scenario	Arm	True ORR	Probability of significant result
1	**1**	0.10	0.0093
1	**2**	0.10	0.0114

1	**3**	0.10	0.0093
2	**1**	0.10	0.0126
2	**2**	0.25	NA*
2	**3**	0.25	NA*
* NA = not applicable in the context of type I error assessment; applicable only to the arms with true ORR = 0.10 corresponding to the null hypothesis.			

2.5.3 Handling of missing values/censoring/discontinuations

Patients with unknown or missing best overall response (BOR) will be counted as failures, i.e. non-responders. Patients who were non-responders before initiation of subsequent anti-cancer therapy will still be non-responders. If there is no baseline tumor assessment, all post-baseline overall lesion responses are expected to be 'Unknown'. If no valid post-baseline tumor assessments are available, the best overall response must be "Unknown" unless progression is reported. For the computation of ORR, these patients will be included in the FAS and will be counted as 'failures'.

2.5.4 Supportive analyses

As a supportive analysis, ORR based on RECIST v1.1 criteria as per blinded independent central review (BICR) will be analyzed with the same analysis conventions as the primary efficacy analysis, and the treatment effect will be summarized by the ORR with its 95% confidence interval using the Clopper-Pearson method.

Subgroup analyses for the primary endpoint

The primary endpoint of ORR will be summarized for the subgroups specified in Section 2.2.1 based on the primary analysis source (i.e., investigator assessment) and the same conventions as for the primary analysis. For each of the subgroups, the proportion of patients with objective response will be analyzed.

[redacted] Forest plot (n, odds ratio, 95% CI) will be produced to graphically depict the treatment effect estimates in different subgroups. No inferential statistics (p-values) will be produced for the subgroups.

Concordance analysis of BOR

An assessment of the concordance between BIRC assessment and local assessment of the Best Overall Response for each subject will be provided by treatment arm. The calculation will be based on the percent agreement (the proportion of response outcomes that agree or match across both Independent Reviewer and Investigator Assessments):

$$Percent\ Agreement = \frac{Number\ of\ matched\ responders + Number\ of\ matched\ non-responders}{Total\ Numbers\ of\ subjects\ assessed}$$

Reasons for "Unknown" BOR

Patients with 'unknown' BOR will be summarized by reason for having unknown status. The following reasons will be used:

- No valid post-baseline assessment
- All post-baseline assessments have overall lesion response UNK
- New anti-neoplastic therapy started before first post-baseline assessment
- SD or non-CR/non-PD too early
- PD too late

Note 1: A SD or Non-CR/Non-PD is considered as "SD too early"' if the SD or Non-CR/Non-PD is documented within first 11 weeks (takes into account the 1 week window for post-baseline assessments) after randomization date.

Note 2: A PD is considered as "PD too late" if the first documentation of PD is recorded more than 13 weeks after randomization date (takes into account the 1 week window for post-baseline assessments) with no qualifying CR, PR or SD or Non-CR/Non-PD in between.

Note 3: Special (and rare) cases where BOR is "unknown" due to both too early SD and too late PD will be classified as "SD too early".

Waterfall plot to depict anti-tumor activity

Waterfall graphs will be used to depict the anti-tumor activity based on local and central assessment separately. These plots will display the best percentage change from baseline in the sum of diameters of all target lesions for each patient. Only patients with measurable disease at baseline will be included in the waterfall graphs. Special consideration is needed for assessments where the target lesion response is CR, PR or SD, but the appearance of a new lesion or a worsening of non-target lesions results in an overall lesion response of PD. As a conservative approach, such assessments will not be considered for display as bars in the graph, since the percentage change in the sum of diameters of target lesions reflects the non-PD target lesion response, but the overall lesion response is PD. A patient with only such assessments will be represented by a special symbol (e.g. ✶) in the waterfall graph. Assessments with "unknown" target lesion response and assessments with unknown overall response will be excluded from the waterfall plots. Patients without any valid assessments will be completely excluded from the graphs.

The total number of patients displayed in the graph will be shown and this number will be used as the denominator for calculating the percentages of patients with tumor shrinkage and tumor growth. Footnote will explain the reason for excluding some patients (due to absence of any valid assessment).

All possible assessment scenarios are described in Table 2-12.

Table 2-12 Inclusion/exclusion of assessments used in waterfall graph

	Criteria for inclusion/exclusion	Possible sources of contradictions

case	Target response	Overall lesion	Include in waterfall?	Non-target	New lesion?
1	CR/PR/SD	PD	Yes but as " only	PD	any
2	CR/PR/SD	PD	Yes but as " only	any	Yes
3	UNK	UNK or PD	No	any	any
4	CR/PR/SD	UNK	No	UNK	No
5	CR/PR/SD	CR/PR/SD	Yes as a bar	SD/IR	No
6	PD	PD	Yes as a bar	any	any

Percentage change from baseline in the sum of diameters of all target lesions over the time will be displayed for individual patients per combination arm.

2.6 Analysis of the key secondary objective

The key secondary objective of the study is to determine duration of response treatment with spartalizumab in combination with novel agents in patients with previously treated unresectable or metastatic melanoma.

2.6.1 Key secondary endpoint

The key secondary endpoint of the study is DoR, defined as the time from the date of first documented response of CR or PR (i.e., the start date of response, not the date when response was confirmed) to the date of the first documented progression or death due to underlying cancer. For the key secondary efficacy analysis, DoR will be based on local investigators review of tumor assessments and using RECIST 1.1 criteria (see Appendix 2 of the protocol). The key secondary analysis will be based on FAS and will include all data observed up-to the cut-off date. As a supportive analysis, DoR will also be summarized based on the BIRC review of tumor data.

If a subject has not had an event before leaving study or initiation of subsequent anti-cancer therapy, progression-free survival is censored at the date of last adequate tumor assessment. Censoring of subjects who have not had an event before initiation of subsequent anti-cancer therapy corresponds to option F(2) presented in Table 14-9 of the protocol and is used in this study in order to obtain DoR estimates that are not confounded by subsequent therapies.

Discontinuation due to disease progression (collected on the 'End of treatment' and 'End of post treatment follow up' disposition pages) without supporting objective evidence satisfying progression criteria per RECIST 1.1 will not be considered disease progression for DoR derivation. Clinical deterioration will not be considered as a qualifying event for progression for the key secondary analysis.

2.6.2 Statistical hypothesis, model, and method of analysis

No formal hypothesis testing will be conducted for the DoR analysis. The key secondary efficacy variable, DoR will be analyzed at the interim analyses for the selection phase, and the primary analysis for the expansion phase, based on the data observed in the FAS up-to the cut-off date, according to the treatment group and stratification assigned at randomization. The survival distribution of DoR will be estimated using the Kaplan-Meier method **from a**

combined pool of part 1 and part 2 data for any given arm. The results will be plotted graphically by treatment arm. The median 25th and 75th percentiles of DoR along with 95% confidence intervals will be presented by treatment group. The survival probabilities at 4, 6, 8 and 12 months, and the associated 95% confidence intervals will be summarized by treatment arm.

2.6.3 Handling of missing values/censoring/discontinuations

DoR will be censored at the date of the last adequate tumor assessment if no DoR event is observed prior to the analysis cut-off date or before the start of the new anticancer therapy date, whichever is earlier.

The date of last adequate tumor assessment is the date of the last tumor assessment with overall lesion response of CR, PR or SD before an event or a censoring reason occurred. In this case the last tumor evaluation date at that assessment will be used. If no post-baseline assessments are available (before an event or a censoring reason occurred) then the date of randomization/start date of treatment will be used.

In particular, DoR will be censored at the last adequate tumor assessment if one of the following occurs: absence of event; the event occurred after two or more missing tumor assessments; before the start of the new anti-cancer therapy. The term "missing adequate tumor assessment" is defined as a tumor assessment (TA) not performed or tumor assessment with overall lesion response of "UNK". The rule to determine number of missing TAs is based on the time interval between the date of last adequate tumor assessment and the date of an event. If the interval is greater than twice the protocol-specified interval between the TAs and 2 times the protocol-allowed time window around assessments, then the number of missing assessments will be 2 or more.

Refer to Table 2-13 for censoring and event date options and outcomes for DoR and PFS analysis (PFS analysis is further discussed in Section 2-7).

Table 2-13 Outcome and event/censor dates for DoR and PFS analysis

Situation	Date	Outcome
No baseline assessment	Date of randomization	Censored
Progression or death at or before next scheduled Assessment	Date of progression (or death)	Progressed
Progression or death after exactly one missing assessment	Date of progression (or death)	Progressed
Progression or death after two or more missing assessments	Date of last adequate assessment prior to missed assessment	Censored
No progression (or death)	Date of last adequate assessment	Censored
Treatment discontinuation due to 'Disease progression' without documented progression, i.e. clinical progression based on investigator claim	Date of last adequate assessment	Censored

Situation	Date	Outcome
New anticancer therapy given prior to protocol defined progression (*including patients who crossover from the control to the treatment arm.*)	Date of last adequate assessment on or prior to starting new anti-cancer therapy (or crossover treatment)	Censored
Death before first PD assessment	Date of death	Progressed

2.7 Analysis of secondary [redacted] efficacy objective(s)

The other secondary [redacted] efficacy objectives are to:

- Evaluate the efficacy of each combination arm, as measured by progression-free survival (PFS)
- Evaluate the efficacy of each combination arm, as measured by disease control rate (DCR)
- Evaluate the overall survival (OS) of each combination arm

[redacted]

2.7.1 Secondary [redacted] efficacy endpoints

Progression-Free Survival (PFS)

PFS is defined as the time from the date of randomization to the date of the first documented progression or death due to any cause. PFS will be based on both local investigator and central assessments of tumor assessments and using RECIST 1.1 criteria (see Appendix 2 of the protocol). The primary analysis will be based on FAS and will include all data observed up-to the cut-off date. PFS will be censored at the date of the last adequate tumor assessment if no PFS event is observed prior to the analysis cut-off date or before the start of the new anticancer therapy date, whichever is earlier. (See _Table_2-4_SubjectTable 2-13 and Section 2.7.3 for additional details regarding censoring rules and determination of date of last adequate tumor assessment). Discontinuation due to disease progression (collected on the 'End of treatment' and 'End of post treatment follow up' disposition pages) without supporting objective evidence satisfying progression criteria per RECIST 1.1 will not be considered disease progression for PFS derivation. Clinical deterioration will not be considered as a qualifying event for progression.

Disease Control Rate (DCR)

DCR is defined as the proportion of patients with a best overall response (BOR) of confirmed CR or PR, or SD lasting 12 weeks or longer, according to RECIST 1.1 criteria. A patient will be considered to have SD for 12 weeks or longer if a SD response is recorded at 11 weeks or later from randomization, allowing for the ±1 week visit window for tumor assessments. DCR will be calculated using the FAS based on both the local investigator and central tumor assessments.

Overall Survival (OS)

Overall Survival (OS) is defined as the time from date of randomization to date of death due to any cause. A cut-off date will be established for each analysis of OS. All deaths occurring on or before the cut-off date in the FAS will be used in the OS analysis.

If a patient is not known to have died at the time of analysis cut-off, OS will be censored at the date of last contact.

2.7.2 Statistical hypothesis, model, and method of analysis

There will be no hypothesis testing for any of the endpoints listed below.

Progression-Free Survival

The secondary efficacy variable, PFS, will be analyzed at the interim looks, the primary analysis, and the final analysis based on the data observed in the FAS up-to the cut-off date, according to the treatment group and strata assigned at randomization. The survival distribution of PFS will be estimated using the Kaplan-Meier method. The results will be plotted graphically by treatment arm. The median 25^{th} and 75^{th} percentile of PFS along with 95% confidence intervals will be presented by treatment group. The survival probabilities at 4, 6, 8 and 12 months, and

the associated 95% confidence intervals will be summarized by treatment arm. As a supportive analysis, PFS will also be summarized based on the BIRC review of tumor data.

Disease Control Rate

DCR will be summarized using descriptive statistics (N, %) by treatment arm along with two-sided exact binomial 95% CIs [Clopper and Pearson 1934]. As a supportive analysis, DCR will also be summarized using descriptive statistics (N, %) based on the BIRC review of tumor data.

Overall Survival

The survival distribution of OS distribution will be estimated using the Kaplan-Meier method, and Kaplan-Meier curves, medians and 95% confidence intervals [Brookmeyer and Crowley 1982] of the medians intervals along with the proportion of patients alive at 4, 6, 8 and 12 months and the associated 95% confidence intervals will be presented for each treatment group.

[redacted]

ECOG performance status

The ECOG PS scale (Table 2.14) will be used to assess physical health of patients, ranging from 0 (most active) to 5 (least active):

Table 2-14 ECOG Performance Scale

Score	Description
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair
5	Dead

Frequency counts and percentages of patients in each score category will be provided by treatment arm and time point based on the windows defined in Section 2.1.

2.7.3 Handling of missing values/censoring/discontinuations

Progression-Free Survival

The date of last adequate tumor assessment is the date of the last tumor assessment with overall lesion response of CR, PR or SD before an event or a censoring reason occurred. In this case the last tumor evaluation date at that assessment will be used. If no post-baseline assessments are available (before an event or a censoring reason occurred) then the date of randomization/start date of treatment will be used.

In particular, PFS will be censored at the last adequate tumor assessment if one of the following occurs: absence of event; the event occurred after a new anticancer therapy is administered; the event occurred after two or more missing tumor assessments. The term “missing adequate tumor assessment” is defined as a tumor assessment (TA) not performed or tumor assessment with overall lesion response of “UNK”. The rule to determine number of missing TAs is based on the time interval between the date of last adequate tumor assessment and the date of an event. If the interval is greater than twice the protocol-specified interval between the TAs and 2 times the protocol-allowed time window around assessments, then the number of missing assessments will be 2 or more.

Refer to Table 2-13 for censoring and event date options and outcomes for PFS.

Overall Survival

If a patient is not known to have died at the time of analysis cut-off, then OS will be censored at the date of last known date patient was alive, i.e., last contact date (see Section 2.4).

2.8 Safety analyses

For all safety analyses, the safety set will be used. All listings and tables will be presented by treatment group. In addition, pooled safety analysis based on all subjects will be produced.

The overall observation period will be divided into three mutually exclusive segments:

1. pre-treatment period: from day of subject's informed consent to the day before first dose of study medication
2. on-treatment period: from day of first dose of study medication to 30 days after last dose of study medication
3. post-treatment period: starting at day 31 after last dose of study medication.

If dates are incomplete in a way that clear assignment to pre-, on-, or post-treatment period cannot be made, then the respective data will be assigned to the on-treatment period.

Safety will be assessed based on all treated subjects; toxicities will be defined by CTCAE v4.03. Incidence and severity of adverse events, causality attribution to drug, time to event onset, duration of the event, its resolution and concomitant medications administered will be recorded. Additional safety assessments will include laboratory safety assessment, vital signs, and cardiac assessment.

2.8.1 Adverse events (AEs)

AE summaries will include all AEs occurring during on treatment period. Additional summaries will be displayed to report all AEs, AEs related to study treatment, all SAEs and SAEs related to study treatment collected up to 150 days after last administration of PDR001. All AEs collected in the AE (e)CRF page will be listed along with the information collected on those AEs e.g. AE relationship to study drug, AE outcome etc. AEs with start date outside of on-treatment period will be flagged in the listings.

AEs will be summarized by number and percentage of subjects having at least one AE, having at least one AE in each primary system organ class (SOC) and for each preferred term (PT) using MedDRA coding. A subject with multiple occurrences of an AE will be counted only once in the respective AE category. A subject with multiple Common Terminology Criteria for Adverse Events (CTCAE) grades for the same preferred term will be summarized under the maximum CTCAE grade recorded for the event. AE with missing CTCAE grade will be included in the 'All grades' column of the summary tables.

In AE summaries, the primary system organ class will be presented alphabetically and the preferred terms will be sorted within primary SOC in descending frequency. The sort order for the preferred term will be based on their frequency in the total column.

The following adverse event summaries will be produced by treatment arm; overview of adverse events and deaths (number and % of subjects who died, with any AE, any SAE, any dose reductions/interruptions, leading to treatment discontinuation), AEs by SOC and PT, summarized by relationship (all AEs and AEs related to study treatment), seriousness (SAEs and non-SAEs), leading to treatment discontinuation, leading to dose interruption/adjustment, requiring additional therapy, requiring immunosuppressive medication and leading to fatal outcome. In addition, a summary of serious adverse events with number of occurrences will be

produced (an occurrence is defined as >1 day between start and prior end date of record of same preferred term).

For the legal requirements of clinicaltrials.gov and EudraCT, two required tables on on-treatment adverse events which are not serious adverse events with an incidence greater than and equal to 5% and on on-treatment serious adverse events and SAE suspected to be related to study treatment will be provided by system organ class and preferred term on the safety set population.

If for a same patient, several consecutive AEs (irrespective of study treatment causality, seriousness and severity) occurred with the same SOC and PT:

- a single occurrence will be counted if there is ≤ 1 day gap between the end date of the preceding AE and the start date of the consecutive AE
- more than one occurrence will be counted if there is > 1 day gap between the end date of the preceding AE and the start date of the consecutive AE

For occurrence, the presence of at least one SAE / SAE suspected to be related to study treatment / non SAE has to be checked in a block e.g., among AE's in a ≤ 1 day gap block, if at least one SAE is occurring, then one occurrence is calculated for that SAE.

The number of deaths resulting from SAEs suspected to be related to study treatment and SAEs irrespective of study treatment relationship will be provided by SOC and PT.

2.8.1.1 Adverse events of special interest / grouping of AEs

Adverse events of special interest (AESI) during the on-treatment period will be tabulated. The list of AESI will also include relevant events for spartalizumab, and the novel agents.

Data analysis of AESIs

An adverse event of special interest (AESI) is a grouping of adverse events that are of scientific and medical concern specific to compound PDR001, the individual novel agents, as well as combination specific AESIs (PDR001 + novel agent). These groupings are defined using MedDRA terms, SMQs (standardized MedDRA queries), HGLTs (high level group terms), HLT (high level terms) and PTs (preferred terms). Customized SMQs (Novartis MedDRA queries, NMQ) may also be used. A NMQ is a customized group of search terms which defines a medical concept for which there is no official SMQ available or the available SMQ does not completely fit the need. It may include a combination of single terms and/or an existing SMQ, narrow or broad. All AESI terms for the individual compounds as well as for combination specific AESIs are captured in the case retrieval strategy (CRS).

For each specified AESI, number and percentage of patients with at least one event of the AESI occurring during on treatment period will be summarized.

Summaries of these AESIs will be provided by treatment arm, (specifying grade, SAE, relationship, leading to treatment discontinuation, leading to dose adjustment/interruption, hospitalization, death, requiring immunosuppressive medication etc.). If sufficient number of events occurred, analysis of time to first occurrence will be applied.

Additional summaries will be provided to report all AESIs, AESIs related to study treatment, all serious AESIs and serious AESIs related to study treatment collected during the on-treatment period and up to 150 days after last administration of PDR001.

A listing of all grouping levels down to the MedDRA preferred terms used to define each AESI will be generated. This listing should be included in section 16 of CSR.

2.8.2 Deaths

Separate summaries for on-treatment and all deaths will be produced by treatment arm, system organ class and preferred term. Additional summary will be displayed to report all deaths up to 150 days after last administration of PDR001.

All deaths will be listed; post treatment deaths will be flagged. A separate listing of deaths prior to starting treatment will be provided for all screened subjects.

Note: "Study indication" as primary reason of death should be coded using MedDRA terms based on the diagnosis eCRF field at start of study. If not coded accordingly in the database, it still must be included in the summary table. Coded reasons for deaths will then be summarized by category 'Study indication' and 'Other' (as selected by the investigator).

The death summaries cover subjects from the Safety Set. The count of deaths in safety analyses could be different from the number of deaths in the efficacy analyses.

2.8.3 Laboratory data

Grading of laboratory values will be assigned programmatically as per National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account.

CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher. Grade 5 will not be used.

For laboratory tests where grades are not defined by CTCAE v4.03, results will be categorized as low/normal/high based on laboratory normal ranges.

The following summaries will be generated separately for hematology, and biochemistry tests:

- Listing of all laboratory data with values flagged to show the corresponding CTCAE v4.03 grades if applicable and the classifications relative to the laboratory normal ranges

For laboratory tests where grades are defined by CTCAE v4.03

- Worst post-baseline CTCAE grade (regardless of the baseline status). Each subject will be counted only once for the worst grade observed post-baseline.
- Shift tables using CTCAE v4.03 grades to compare baseline to the worst on-treatment value

For laboratory tests where grades are not defined by CTCAE v4.03

- Shift tables using the low/normal/high/ (low and high) classification to compare baseline to the worst on-treatment value.

On analyzing laboratory, data from all sources (central and local laboratories (as applicable)) will be combined. The summaries will include all assessments available for the lab parameter collected no later than 30 days after the last study treatment administration date (see Section 2.1.1).

The following summaries will be produced for hematology and biochemistry laboratory data (by laboratory parameter and treatment):

- Worst post-baseline CTCAE grade (regardless of the baseline status). Each subject will be counted only for the worst grade observed post-baseline.
- Shift tables using CTCAE grades to compare baseline to the worst on-treatment value
- For laboratory tests where CTCAE grades are not defined, shift tables using the low/normal/high/(low and high) classification to compare baseline to the worst on-treatment value.
- Trends of lab parameter values over time (baseline and selected on-treatment timepoints) should be displayed via boxplots based on time windows and corresponding tables displaying the statistics used for the box plots by the selected time points.

The following listings will be produced for the laboratory data:

- Listings of all laboratory data, with CTCAE grades and classification relative to the laboratory normal range. Lab data collected during the post-treatment period will be flagged.
- Listing of all CTCAE grade 3 or 4 laboratory toxicities

Liver function parameters

Liver function parameters of interest are total bilirubin (TBL), alanine aminotransferase (ALT), aspartate aminotransferase (AST) and alkaline phosphatase (ALP). The number (%) of patients with worst post-baseline values as per Novartis Liver Toxicity guidelines will be summarized:

The following summaries will be produced:

- ALT or AST > 3xULN
- ALT or AST > 5xULN
- ALT or AST > 8xULN
- ALT or AST > 10xULN
- ALT or AST > 20xULN
    - TBL > 2xULN
    - TBL > 3xULN
    - ALT or AST > 3xULN & TBL > 2xULN

- ALT or AST > 3xULN & TBL > 2xULN & ALP < 2xULN (potential Hy's law)

Potential Hy's Law events are defined as those subjects with concurrent occurrence of AST or ALT > 3xULN and TBL > 2xULN and ALP < 2xULN in the same assessment sample during the on-treatment period. Further medical review has to be conducted to assess potential confounding factor such as, liver metastases, liver function at baseline etc.

2.8.4 Other safety data

2.8.4.1 ECG and cardiac imaging data

Categorical Analysis of QT/QTc interval data based on the number of subjects meeting or exceeding predefined limits in terms of absolute QT/QTc intervals or changes from baseline will be presented. In addition, a listing of these subjects will be produced by treatment group.

Data handling

In case the study requires ECG replicates at any assessment, the average of the ECG parameters at that assessment should be used in the analyses.

Data analysis

12-lead ECG including PR, QRS, QT, QTcF, and RR intervals will be obtained for each subject during the study. ECG data will be read and interpreted centrally.

The number and percentage of subjects with notable ECG values will be presented by treatment arm.

- QT, QTcF
  - New value of > 450 and ≤ 480 ms
  - New value of > 480 and ≤ 500 ms
  - New value of > 500 ms
  - Increase from Baseline of > 30 ms to ≤ 60ms
  - Increase from Baseline of > 60 ms
- HR
  - Increase from baseline >25% and to a value > 100 bpm
  - Decrease from baseline >25% and to a value < 50 bpm
- PR
  - Increase from baseline >25% and to a value > 200 ms
  - New value of > 200 ms
- QRS
  - Increase from baseline >25% and to a value > 120 ms
- New values of QRS > 120 ms

A listing of all ECG assessments will be produced by treatment arm and notable values will be flagged. A separate listing of only the subjects with notable ECG values may also be produced. In the listing, the assessments collected during the post-treatment period will be flagged.

2.8.4.2 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The following parameters were collected: height (cm), weight (kg), body temperature (°C), heart rate (beats per minute), systolic and diastolic blood pressure (mmHg).

Data handling

Vital signs collected on treatment will be summarized. Values measured outside of on treatment period will be flagged in the listings.

Data analysis

For analysis of vital signs the clinically notable vital sign criteria are provided in Table 2-15 below.

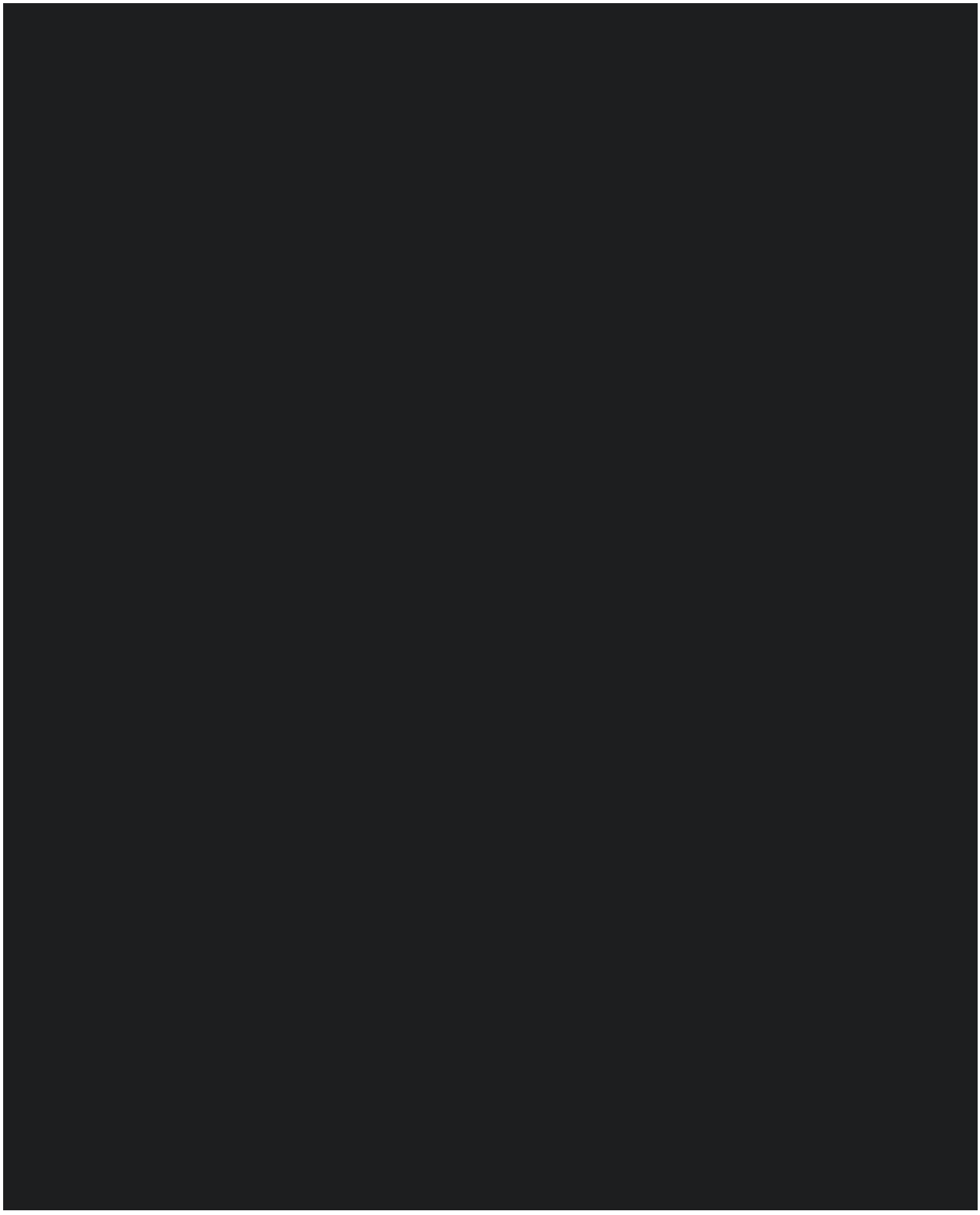
Table 2-15 Clinically notable changes in vital signs

Vital sign (unit)	Clinically notable criteria	
	above normal value	**below normal value**
Weight (kg)	increase > 10% from Baseline	decrease > 10% from Baseline
Systolic blood pressure (mmHg)	>=180 with increase from baseline of >=20	<=90 with decrease from baseline of >=20
Diastolic blood pressure (mmHg)	>=105 with increase from baseline of >=15	<=50 with decrease from baseline of >=15
Pulse rate (bpm)	>=100 with increase from baseline of >25%	<=50 with decrease from baseline of > 25%
Body temperature	>= 39.1	-

The number and percentage of subjects with notable vital sign values (high/low) will be presented by treatment arm.

A listing of all vital sign assessments will be produced by treatment arm and notable values will be flagged. A separate listing of only the subjects with notable vital sign values may also be produced. In the listing, the assessments collected outside of on-treatment period will be flagged.

[redacted]



2.10 PD and [redacted]PD analyses

2.10.1 Immunogenicity

2.10.1.1 Sample ADA Status

Each IG sample is assessed in a three tiered anti-drug anti-body (ADA) testing approach. All IG samples are analyzed in the initial screening assay (first tier). Samples testing negative in the screening assay are not subject to a confirmatory assay. Samples testing positive in the screening assay are then subjected to the confirmatory assay to demonstrate that ADA are specific for the therapeutic protein product (second tier). The titer of confirmatory positive samples will be subsequently determined in the titration assay (third tier). Samples identified as positive in the confirmatory assay are considered ADA positive and are further characterized in the neutralization assay to indicate the presence of neutralizing antibodies (NAb). Samples can test negative in either the screening or confirmatory assay but for analysis purposes they are not differentiated. The following properties of each sample will be provided in the source data:

- Result of assay according to pre-specified confirmatory cut point: ADA positive (yes) or ADA negative (no)
- Titer (for positive samples): numerical representation of the magnitude of ADA response
- Presence of NAb (for positive samples, if NAb assay results are available): yes or no
- Drug tolerance level: highest drug concentration that does not interfere in the ADA detection method
- Fold titer change (i.e. x-fold): threshold for determining treatment boosted

- Sample ADA status is determined based on the following definitions:
  - *ADA-inconclusive sample*: Sample where assay is ADA negative [redacted] at the time of IG sample collection is greater than or equal to the drug tolerance level or missing.
  - *Unevaluable sample*: Sample where assay is not available.
  - *Determinant sample:* Sample that is neither ADA-inconclusive nor unevaluable.

The following definitions apply only to determinant samples:

- *ADA-negative sample*: Determinant sample where assay is ADA negative [redacted] at the time of IG sample collection is less than the drug tolerance level.
- *ADA-positive sample*: Determinant sample where assay is ADA positive.
- *ADA-positive NAb sample*: Determinant sample where assay is ADA positive and presence of NAb = yes.

The following definitions apply only to post-baseline ADA-positive samples with a corresponding determinant baseline sample. To be classified as *treatment-boosted* or *treatment-unaffected*, both the post-baseline and baseline titer must be non-missing:

  - *treatment-induced ADA-positive sample:* ADA-positive sample post-baseline with ADA-negative sample at baseline.

- *treatment-boosted ADA-positive sample:* ADA-positive sample post-baseline with titer that is at least *the fold titer change* greater than the ADA-positive baseline titer.
- *treatment-unaffected ADA-positive sample*: ADA-positive sample post-baseline with titer that is less than *the fold titer change* greater than the ADA-positive baseline titer

[redacted]

The following summaries of ADA sample status (n and %) will be provided using *Immunogenicity prevalence set*:

- ADA-positive samples (i.e. ADA prevalence) and ADA-positive NAb samples, both overall and by time point (including baseline). For summaries by time point, the denominator is the number of subjects at that time point with a determinant sample.

Listings will be provided of sample ADA status (including titer for positive samples).

2.10.1.2 Subject ADA status

Any IG sample collected after 150 days of the last dose of PDR001, LAG525, or Canakinumab will not be used for summaries or derivations and will only be included in the listing.

Subject ADA status is defined as follows:

- *Treatment-induced ADA-positive subject*: subject with ADA-negative sample at baseline and at least one treatment-induced ADA-positive sample.
- *Treatment-boosted ADA-positive subject*: subject with ADA-positive sample at baseline and at least one treatment-boosted ADA-positive sample.
- *Treatment-unaffected ADA-positive subject*: subject with ADA-positive sample at baseline, no treatment-boosted ADA-positive samples, and at least one treatment-unaffected ADA-positive sample.
- *Treatment-reduced ADA-positive subject*: subject with ADA-positive sample at baseline and at least one post baseline determinant sample, all of which are ADA-negative samples.
- *ADA-negative subject*: subject with ADA-negative sample at baseline and at least one post baseline determinant sample, all of which are ADA-negative samples.
- *Inconclusive subject*: subject who does not qualify as treatment-induced ADA-positive, treatment-boosted ADA-positive, treatment-unaffected ADA-positive, treatment-reduced ADA-positive, or ADA-negative.

The following summaries of ADA subject status (n and %) will be provided using *Immunogenicity incidence set*:

- Treatment-boosted ADA-positive subjects; denominator is the number of subjects with ADA-positive sample at baseline.

- Treatment-induced ADA-positive subjects; denominator is the number of subjects with ADA-negative sample at baseline.
- ADA-negative subjects: denominator is the number of subjects in *Immunogenicity incidence set*.
- ADA-positive subjects (i.e. ADA incidence): calculated as the number of treatment-boosted ADA-positive and treatment-induced ADA-positive subjects; denominator is the number of subjects in *Immunogenicity incidence set*.

Listings will be provided of subject ADA status.

[redacted]

2.11 Patient-reported outcomes

Not applicable.

2.12 Biomarkers

2.12.1 Introduction

As a project standard, Novartis Oncology Biostatistics will analyze only biomarkers collected in the clinical database. The biomarker analysis covered within the SAP pertains mainly to the secondary objective endpoint of determining the proportion of subjects with a favorable biomarker profile (pFBP, Table 1-1 and Table 2-18).

[redacted] For example, there may be inadequate sample number, issues related to the quality of the sample or issues related to the assay that preclude analysis. Alternatively, there may be insufficient efficacy to allow for correlative analyses. Therefore, depending on the results obtained during the study, collection/analysis of some samples may be omitted.

[redacted]

2.12.2 Secondary biomarker objectives

Changes in phenotype and/or activation of T cell populations in the tumor and tumor microenvironment will be analyzed on an on-going basis as part of the secondary objectives (Table 1-1 and Table 2-18). Every effort will be made to make this data available for internal decision making purposes. Such analyses may include, but are not limited to number of $CD8^+$ T cell infiltration, T cell activation level by IHC, gene expression analysis, and T cell repertoire/clonality by TCR-sequencing (Table 2-18).

Table 2-18 Parameter definitions for categorization of biomarker outcome at subject level

Biomarker parameter	Analysis	Criteria for favorable biomarker result
Number of tumor infiltrating T cells (TILs)	Changes of $CD8^+$ T cell numbers as assessed by IHC in tumor/ tumor microenvironment with treatment	Increase
Activation level of TILs	Change of T cell activation marker level(s) as assessed by IHC in tumor/tumor microenvironment with treatment	Increase
TIL repertoire/specificity of T cell response to tumor	TCR clonality changes as assessed by TCR-sequencing in tumor/tumor microenvironment with treatment	Increase
Changes in immune response gene expression signatures	Changes in expression of genes and/or gene signatures as assessed by mRNA analysis in tumor/tumor microenvironment with treatment	Defined per gene/gene signature

2.12.3 Biomarker data

As a project standard, Novartis Oncology Biostatistics will analyze only biomarkers collected in the clinical database. The biomarker analysis covered within the SAP pertains mainly to the secondary objective endpoint of determining the proportion of subjects with a favorable biomarker profile (pFBP, Table 1-1 and Table 2-18). [redacted]

The biomarker analysis set, as it pertains to the secondary objectives (Table 1-1) as well as the analysis of the change in LAG-3 expression levels, consists of all subjects in the safety set with an evaluable baseline tumor biopsy sample and at least one evaluable post-baseline tumor biopsy sample.

Table 2-19 summarizes the biomarker collection schedules and sample types which are in scope with this SAP.

Table 2-19 Sample biomarker summary table

Biomarker parameter	Biomarker analysis measurement	Time point	Sample	Method	Dataset
Number of TILs	Percent of tumoral CD8+ cells	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	IHC	B1

Biomarker parameter	Biomarker analysis measurement	Time point	Sample	Method	Dataset
Activation level of TILs	Percent of tumoral GzmB+ within CD8+cells	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	FIHC/AQUA	B1
Activation level of TILs	Percent of tumoral CD8+ within Ki67+ cells	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	FIHC/AQUA	B1
TIL repertoire/ specificity of T cell response to tumor	TCR clonality	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	TCR sequencing	B1
Changes in immune response gene expression signatures	Gene signature expression levels	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	Nanostring	B1

[redacted]

2.12.4 General data handling and preprocessing

The mean of all pre dose assessments (i.e. screening and Cycle 1 Day 1 assessments) will be used as the baseline value. The on-treatment assessment will be the result from the tumor biopsy taken between C1D21 and C2D1.

When more than one biomarker data value are available for a subject at any time point, the mean of the replicate values will be used for all statistical analyses.

2.12.5 Secondary biomarker objectives

The secondary objective biomarker data will be listed for each subject for all time points (baseline and on treatment) and ordered by the treatment arm. Absolute and relative change (percent change) from baseline of each biomarker parameter will be calculated for each patient and summarized for each treatment arm. Graphical presentations of the change and percentage change from baseline will be produced via spaghetti plots and boxplots and ordered by the treatment arm.

2.12.5.1 IHC for quantifying tumor infiltrating T cells (TILs)

The number of TILs at each timepoint will be assessed by measuring % CD8+ cells within the invasive margin, tumor and stroma area (as defined by Melamask staining), using the following variables provided in the raw data:

- Intraepithelial tumor (IE) CD8+ density (% of cells containing CD8)
- Intrastroma (IST) CD8+ density
- IST CD8+ density 0 to 10 μm (% of CD8+ IST cells within the 0 to 10 μm region of the IE-IST border)
- IST CD8+ density >10 to 20 μm
- IST CD8+ density >20 to 30 μm
- IST CD8+ density >30 to 40 μm
- IST CD8+ density >40 to 50 μm

For all CD8 parameters above, the baseline value, on-treatment (C1D21-C2D1) value, and the difference from baseline to on-treatment will be listed for each individual patient by treatment arm. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all patients (regardless of treatment arms) will be reported for baseline and change from baseline to on-treatment by combination arm. Inter-quartile range is the number of data points between the 25th and 75th percentile.

The proportion of GzmB+ T cells and Ki67+ T cells will be assessed using FIHC and AQUA at each timepoint. The % of GzmB+CD8+ and Ki67+CD8+ cells within the invasive margin, tumor and stroma area (as defined by Melamask staining) will be calculated using the following variables provided in the raw data:

- IE CD8+Ki67+ density (% of CD8+cells containing Ki67)
- IST CD8+Ki67+ density
- IE CD8+GzmB+ density (% of CD8+cells containing GzmB)
- IST CD8+GzmB+ density

The GzmB+ density is expected to be low but unique to T cells. In instances where the density is <10%, the following raw data variables may be employed in lieu of the one above for all samples within a particular analysis:

- IE GzmB+ density (% of cells containing GzmB)
- IST GzmB+ density (% of cells containing GzmB)

For all parameters above, the baseline value, on-treatment value, and the difference from baseline to 3-4 weeks will be listed for each individual patient by treatment arm. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all patients (regardless of treatment arms) will be reported for baseline and change from baseline to on-treatment (C1D21-C2D1) by combination arm. Inter-quartile range is the number of data points between the 25th and 75th percentile.

2.12.5.2 TCR sequencing for quantifying T cell clonality

T cell clonality is calculated as follows:

$$1 - (normalized\ Shannon\ entropy)$$

Using the following variable provided in the raw data:

- Normalized Shannon entropy.

2.12.5.3 Nanostring for quantifying gene expression signatures

Gene expression data will be calculated using the following variables provided in the raw data:

- raw gene expression
- housekeeper-gene normalized gene expression

The gene signatures that will be scored for expression are defined in Table 2-20. The score of each gene signature will be calculated using the simple average of log expression values for all genes. The aggregate score for the three gene signatures, calculated as average of the score of each gene signature will be used as one of the criteria for favorable biomarker result. Additional genes may be considered for gene expression analyses pertaining to biomarker outcomes.

Table 2-20 Immune response gene expression signatures

Signature	Gene names
T-cell inflamed [Ayers, 2017]	CD3D IDO1 CIITA CD3E CCL5 GZMK CD2 HLA-DRA CXCL13 IL2RG NKG7 HLA-E CXCR6 LAG3 TAGAP CXCL10 STAT1 GZMB
CD8 T-cell	CCL5 CTSW FASLG CD8B ZNF683 GZMA XCL2 CD7 KLRC1 CD8A XCL1 TRDC NKG7 KLRK1 GNLY PRF1 GZMB GZMH LAG3 KLRD1
IFN-gamma	IFNG CXCL9 CXCL10 CXCL11 GBP1

For all parameters above, the baseline value, on-treatment (C1D21-C2D1) value, and the difference from baseline to 3-4 weeks will be listed for each individual patient by treatment arm. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all patients (regardless of treatment arms) will be reported for baseline and change from baseline to on-treatment by combination arm. Inter-quartile range is the number of data points between the 25th and 75th percentile. Calculation of proportion of patients with favorable biomarker profile

The biomarkers listed in Table 2-20 will be used to calculate the proportion of subjects with a favorable biomarker profile (pFBP).

- As for the parameter of number of tumor infiltrating T cells (TIL), an increase from baseline to on-treatment (C1D21-C2D1) in the IE CD8+ density is considered a favorable change. Whether such increases are accompanied by a decrease in IST CD8+ density may be taken into consideration as a signal for CD8+ T cell migration into the tumor. Furthermore increases in CD8+ density in IST areas closer to the IE-IST border (e.g. an increase in 0-10 µm and >10-20 µm regions compared to more distal regions, >30 µm) may be taken into consideration as another signal for CD8+ T cell migration into the tumor
- As for the activation level of TIL, an increase from baseline to on-treatment in percentage of GzmB+ T cells is considered favorable. In the presence or absence of an increase in GzmB+ T cells, an increase in percentage of Ki67+ T cells is also considered favorable. If decreases in both GzmB + T cells and Ki67+ T cell are observed, this parameter would not be considered favorable.

- As for the TIL repertoire/specificity of T cell response to tumor, an increase from baseline to on-treatment in TCR clonality is considered favorable.
- An increase from baseline to on-treatment in the aggregate gene signature score would be considered favorable.

For each biomarker involved in the evaluation of pFBP as stated above, the baseline, 3-4 weeks biomarker values, and the difference from baseline to 3-4 weeks will be listed for each individual patient by combination arm. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all patients (regardless of treatment arms) will be reported for baseline and change from baseline to 3-4 weeks by combination arm. Inter-quartile range is the number of data points between the 25^{th} and 75^{th} percentile.

If an individual subject has two or more favorable biomarker parameters from Table 2-16 then that subject will be considered to have a favorable biomarker profile.

For each combination arm, the proportion of subjects with a favorable biomarker profile (pFBP) are then defined as the number of subjects with favorable biomarker profiles divided by the number of subjects enrolled into the arm which are included in the biomarker analysis set for that particular interim analysis and will be summarized.

[redacted]

2.14 Interim analysis

Primary endpoint: Objective Response Rate (ORR)

Part 1: Selection phase

The 1st interim analysis is planned once at least ten patients have been randomized into each of the three initial treatment combination arms, and have completed at least 20 weeks of treatment (corresponding to two post-baseline assessments) or discontinued the study. It is possible to terminate one or more combination arms at the 1st interim analyses for lack of adequate efficacy, i.e. futility. It is also possible for advance one or more combination arms at the 1st interim analysis to the expansion phase if those arms cross a specific efficacy probability threshold.

Assessment of futility and efficacy will be based on the calculated Bayesian predictive probabilities of occurrence, using the following criteria highlighted in Table 2-9.

The (posterior) predictive distribution of future responders conditioned on the observed data at the interim is a beta-binomial distribution. More details on the (posterior) predictive distribution used for this study are highlighted in Section 2.5.2.

Enrollment and assessment at subsequent interim analyses for each treatment group will continue at an assessment period as defined in Section 4.2 of the protocol, or 45 subjects have been randomized into a treatment group rules presented in Section 2.5.2.

Note: If new combination arms are added to the study, those combination arms will have their interim analyses at the same time-points as the first three combination arms on this study (after 10 patients have been randomized and have sufficient data for analysis, and then if needed, at an assessment period as defined in Section 4.2 of the protocol, or 45 patients have been randomized into each arm).

Part 2: Expansion phase

No formal interim analysis is planned for this part of the trial. The primary analysis will be performed once the expanded arm has fully enrolled and also after all patients randomized in the expansion phase had at least 9 months of follow-up. This timing would allow for a potential minimum duration of response of 6 months for all subjects randomized. Formal

hypothesis testing of the primary endpoint (as described in Section 2.5.2) will be performed at the primary analysis. The final analysis will be performed at the end of the study as described in Section 2.5.1.

3 Sample size calculation

In this section, sample size will be presented based on specific potential scenarios for true ORR. In particular, for selection phase (part 1), since one can consider a wide variety of potential true ORR scenarios, the sample size will not be based on an exact calculation leading to one fixed number, but will instead be characterized by simulation based on average number of subjects enrolled into these arms under each individual scenario for true ORR. It is important to emphasize that the true underlying efficacy activity (in terms of ORR) of the three initial arms will have a substantial bearing on the total sample size needed in part 1. Subsequently, and given the adaptive nature of the study design, the sample size needed in expansion phase (part 2) to ensure sufficient power will depend on the ORR observed in part 1 and in particular on the rate observed in the arm that qualifies for expansion to part 2.

The sample size calculation for parts 1 and 2 were performed using SAS 9.4 software, based on the probabilities of first determining a efficacious treatment and a futile (non-efficacious) treatment in the selection phase (part 1), and then determining the number of subjects needed to reject the null hypothesis (part 2). Assumptions and sample size calculations for both the selection and expansion phases are summarized in Section 10.8 of the protocol. Additional simulation results for both the selection and expansion phases will be summarized in the [SAP Technical Appendix].

4 Change to protocol specified analyses

The only changes to protocol specified analysis in the SAP is the inclusion of an IG analysis set for immunogenicity analysis, as well as specific sub-group analyses for efficacy and safety analysis.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

The following rule should be used for the imputation of the dose end date for a given study treatment component:

Scenario 1: If the dose end date is completely missing and there is no EOT page and no death date, the patient is considered as on-going:

The patient should be treated as on-going and the cut-off date should be used as the dose end date.

Scenario 1 should not applicable for final CSR. All patients should have EOT page complete before the Database lock for Final CSR

Scenario 2: If the dose end date is completely or partially missing and the EOT page is available:

Case 1: The dose end date is completely missing, and the EOT completion date is complete, then this latter date should be used.

Case 2: Only Year(yyyy) of the dose end date is available and yyyy < the year of EOT date:

- **Use Dec31yyyy**

Case 3: Only Year(yyyy) of the dose end date is available and yyyy = the year of EOT date:

- **Use EOT date**

Case 4: Both Year(yyyy) and Month (mm) are available for dose end date, and yyyy = the year of EOT date and mm < the month of EOT date:

- **Use last day of the Month (mm)**

All other cases should be considered as a data issue and the statistician should contact the data manager of the study.

After imputation, compare the imputed date with start date of treatment, if the imputed date is start date of treatment:

- **Use the treatment start date**

Patients with missing start dates are to be considered missing for all study treatment component related calculations and no imputation will be made. If start date is missing then end-date should not be imputed.

5.1.2 AE, ConMeds and safety assessment date imputation

Table 5-1 Imputation of start dates (AE, CM) and assessments (LB, EG, VS)

Missing Element	Rule
day, month, and year	• No imputation will be done for completely missing dates

day, month	• If available year = year of study treatment start date then
	  ○ If stop date contains a full date and stop date is earlier than study treatment start date then set start date = 01JanYYYY
	  ○ Else set start date = study treatment start date.
	• If available year > year of study treatment start date then 01JanYYYY
	• If available year < year of study treatment start date then 01JulYYYY

Missing Element	**Rule**
day	• If available month and year = month and year of study treatment start date then o If stop date contains a full date and stop date is earlier than study treatment start date then set start date= 01MONYYYY. o Else set start date = study treatment start date. • If available month and year > month and year of study treatment start date then 01MONYYYY • If available month and year < month year of study treatment start date then 15MONYYYY

Table 5-2 Imputation of end dates (AE, CM)

Missing Element	Rule (*=last treatment date plus 30 days not > (death date, cut-off date, withdrawal of consent date))
day, month, and year	• Completely missing end dates (incl. ongoing events) will be imputed by the end date of the on-treatment period*
day, month	• If partial end date contains year only, set end date = earliest of 31DecYYYY or end date of the on-treatment period *
day	• If partial end date contains month and year, set end date = earliest of last day of the month or end date of the on-treatment period*

Any AEs and ConMeds with partial/missing dates will be displayed as such in the data listings.

Any AEs and ConMeds which are continuing as per data cut-off will be shown as 'ongoing' rather than the end date provided.

The above imputations are only used for analyses of time to and duration of AEs and concomitant medications.

5.1.2.1 Other imputations

Incomplete date of initial diagnosis of cancer and date of most recent recurrence

Missing day is defaulted to the 15th of the month and missing month and day is defaulted to 01-Jan.

Incomplete assessment dates for tumor assessment

All investigation dates (e.g. MRI scan, CT scan) must be completed with day, month and year. If one or more assessment dates are incomplete but other investigation dates are available, this/these incomplete date(s) are not considered for calculation of the assessment date and assessment date is calculated as the latest of all investigation dates (e.g. MRI scan, CT scan) if the overall response at that assessment is CR/PR/SD/UNK. Otherwise – if overall response is progression – the assessment date is calculated as the earliest date of all investigation dates at that evaluation number. If all measurement dates have no day recorded, the 1st of the month is used. If the month is not completed, for any of the investigations, the respective assessment will be considered to be at the date which is exactly between previous and following assessment. If a previous and following assessment is not available, this assessment will not be used for any calculation.

Applying the cut-off to tumor assessment

For tumor related assessments, if an evaluation has some assessments done prior to cut-off date and others after the cut-off date , then the evaluation is considered post-cut-off date and will be excluded from analysis.

5.2 AEs coding/grading

Adverse events are coded using the Medical dictionary for regulatory activities (MedDRA) terminology.

Note: The latest available MedDRA version at the time of the analyses should be used. The MedDRA version used should be specified in the footnote of relevant tables.

AEs will be assessed according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03.

The CTCAE represents a comprehensive grading system for reporting the acute and late effects of cancer treatments. CTCAE grading is by definition a 5-point scale generally corresponding to mild, moderate, severe, life threatening, and death. This grading system inherently places a value on the importance of an event; although there is not necessarily proportionality among grades (a grade 2 is not necessarily twice as bad as a grade 1).

5.3 Laboratory parameters derivations

Grade categorization of lab values will be assigned programmatically as per NCI Common Terminology Criteria for Adverse Events (CTCAE) version 4.03. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account. The criteria to assign CTCAE grades are given in Novartis internal criteria for CTCAE grading of laboratory parameters. The latest available version of the document based on the underlying CTCAE version 4.03 at the time of analysis will be used. The Novartis internal CTC grading document should be added as appendix to the SAP and to Section 16.1.9 of the CSR. For laboratory tests where grades are not defined by CTCAE version 4.03, results will be graded by the low/normal/high (or other project-specific ranges, if more suitable) classifications based on laboratory normal ranges.

A severity grade of 0 will be assigned for all non-missing lab values not graded as 1 or higher. Grade 5 will not be used. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia.

Imputation Rules

Common Toxicity Criteria (CTC) grading for blood differentials is based on absolute values. However, this data may not be reported as absolute counts but rather as percentage of WBC.

If laboratory values are provided as '<X' (i.e. below limit of detection) or '>X', prior to conversion of laboratory values to SI unit, these numeric values are set to X.

The following rules will be applied to derive the WBC differential counts when only percentages are available for an xxx differential

$$xxx\ count = (WBC\ count) * \left(\frac{xxx\ \%\ value}{100}\right)$$

Further derivation of laboratory parameters might be required for CTCAE grading. For instance, corrected calcium can be derived using the reported total calcium value and albumin at the same assessment using the following formula:

$$Corrected\ Calcium\ \left(\frac{mg}{dL}\right) = Calcium\left(\frac{mg}{dl}\right) - 0.8\left[Albumin\ \left(\frac{g}{dL}\right) - 4\right]$$

In order to apply the above formula, albumin values in g/L will be converted to g/dL by multiplying by 0.1), calcium values in mmol/L will be converted to mg/dL by dividing by 0.2495. For calculation of laboratory CTC grades 0 and 1, the normal range for derived corrected calcium is set to the same limits (in mg/dL) as for calcium.

CTC grades for the derived absolute WBC differential counts (neutrophils, lymphocytes) and corrected calcium will be assigned as described above for grading

5.4 Statistical models

5.4.1 Primary analysis

Please refer to Section 10 of the protocol, Section 2.5.2, and the [SAP Technical Appendix] for extended details on the modeling and SAS procedures used for both the selection and expansion phases of the trial. Other details for the primary analysis are shown below

Analysis of Binary Data

Single arm study

For the selection phase, the following statistical hypothesis will be tested:

H_0: ORR ≤ 10 % will be tested vs H_1: ORR > 10 %

using a one-sided test with α=0.025 based on the exact binomial distribution.

Responses will be summarized in terms of percentage rates with 95% CIs. An exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement for one-way tables) will be calculated [Clopper and Pearson 1934]

SAS procedure FREQ will be used to estimate the proportion of responders (binary outcome = 1 or "Yes"), along with the associated 95% (=100 × (1 – *two-sided alpha level*)) two-sided Pearson-Clopper CI and exact one-sided p-value for the hypothesis test of the *null proportion* (0.10).

Confidence interval for response rate

Responses will be summarized in terms of percentage rates with 100(1 − α)% confidence interval using exact binomial confidence interval (implemented using SAS procedure FREQ with EXACT statement for one-way table [Clopper and Pearson 1934].

5.4.2 Key secondary analysis

Kaplan-Meier estimates

For Duration of Response, an estimate of the survival function in each treatment group will be constructed using Kaplan-Meier (product-limit) method as implemented in PROC LIFETEST with METHOD=KM option. The PROC LIFETEST statement will use the option CONFTYPE=LOGLOG.

Median survival for each treatment group will be obtained along with 95% confidence intervals calculated from PROC LIFETEST output using the method of [Brookmeyer and

Crowley 1982]. Kaplan-Meier estimates of the survival function with 95% confidence intervals at specific time points will be summarized. The standard error of the Kaplan-Meier estimate will be calculated using Greenwood's formula [Collett 1994].

Treatment of ties

The ties = EFRON statement in LIFETEST procedure will be used to analyze time to event data with ties.

5.4.3 Audit-based BIRC assessment of PFS

Not applicable.

5.5 Rule of exclusion criteria of analysis sets

Not applicable.

6 Reference

1. Amit O, Mannino F, Stone AM, Bushnell W, Denne J, Helterbrand J and Burger HU. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. European Journal of Cancer 2011; 47: 1772-1778.
2. Armitage P, McPherson CK and Rowe BC (1969). Repeated significance tests on accumulating data. J .R. Statist. Soc. A, 132, 232-244.
3. Ayers M, Lunceford J, Nebozhyn M, et al (2017). IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest, 127(8), 2930-2940.
4. Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time. Biometrics, 38, 29 - 41.
5. Clopper CJ and Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrical, 26, 404-413.
6. Collet D (1994). Modelling survival data in medical research. London, Chapman & Hall.
7. Daud A, Wolchok J, Robert C et al (2016). Programmed death ligand 1 expression and response to the anti-programmed death 1 antibody pembrolizumab in melanoma. J Clin Oncol, 34:4102-4109.
8. Dodd EL, Korn LE, Freidlin B, Gray R, Bhattacharya S. An audit strategy for progression-free survival. Biometrics, 2011, 67 (3): 1092-99.
9. FDA briefing Document Oncologic Drugs Advisory Committee Meeting, July 24, 2012, Evaluation of radiologic review of progression-free survival in Non-hematologic malignancies. http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/oncologicdrugsadvisorycommittee/ucm311141.pdf

10. Fleming TR, Harrington DP and O'Brien PC (1984). Designs for group sequential tests. Controlled Clinical Trials, 5, 348-361.
11. Greenland S, Lanes S and Jara M (2008). Estimating effects from randomized trials with discontinuations: the need for intent-to-treat design and G-estimation. Clinical Trials, 5, 5–13.
12. Glimm E, Maurer W and Bretz F (2010). Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine, 29, 219-228.
13. Hwang IK, Shih WJ and De Cani JS (1990). Group sequential designs using a family of type 1 error probability spending functions. Statistics in medicine, 9(12), 1439-1445.
14. Kalbfleisch, J. D., and Prentice, R. L. (1981). Estimation of the Average Hazard Ratio. Biometrika, 68, 105-112
15. Kalbfleisch JD and Prentice RL. The Statistical Analysis of Failure Time Data (2002). Wiley Series in Probability and Statistics.
16. Klein J.P., Logan B., Harhoff M and Andersen P K, Analyzing survival curves at a fixed point in time, Statistics in Medicine, 2007, 26, 4505-4519.
17. Korhonen P, Laird NM and Palmgren J. Correcting for non-compliance in randomized trials: an application to the ATBC study (1999). Statistics in Medicine, 18, 2879-2897.
18. Lan KKG, DeMets DL (1983). Discrete sequential boundaries for clinical trials. Biometrika, 70, 649-53.
19. Mantel, N. and Joseph L Fleisś (1980). Minimum expected cell size requirements for the Mantel-Haenszel one degree-of-freedom Chi-Square Test and a related rapid procedure. American Journal of Epidemiology, 112, 129-134.
20. Neuenschwander B, Branson M and Gsponer T (2008). Critical aspects of the Bayesian approach to Phase I cancer trials. Statistics in Medicine, 27, 2420-2439.
21. O'Brien PC and Fleming TR (1979). A multiple testing procedure for clinical trials. Biometrics, 35, 549-53.
22. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J and Smith PG (1976). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. British Journal of Cancer, 34, 585-607.
23. Royston and Parmar (2013). Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Medical Research Methodology 2013, 13:152.
24. Robins JM and Finkelstein DM (2000). Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics, 56(3),779-88.
25. Robins JM and Tsiatis AA (1991). Correcting for non-compliance in randomized trials using rank preserving structural failure time models (RPSFTM). Commun. Statist.-Theory Meth., 20(8), 2609-31.
26. Rup B, et al. (2015). Standardizing terms, definitions and concepts for describing and interpreting unwanted immunogenicity of biopharmaceuticals: recommendations of the

Innovative Medicines Initiative ABIRISK consortium. Clinical and Experimental Immunology 181:385-400.

27. Schimmer BP and Parker KL (1996). Adrenocorticotropic hormone; adrenocortical steroids and their synthetic analogs; inhibitors of the synthesis and actions of adrenocortical hormones. In Goodman & Gilman's The Pharmacological Basis of Therapeutics (9th Edition), Published by McGraw-Hill, New York: 1459-1485.
28. Shankar G, et al. (2014). Assessment and reporting of the clinical immunogenicity of therapeutic proteins and peptides – harmonized terminology and tactical recommendations. The AAPS Journal 16:658-673.
29. Smith B, Vandenhende FR, DeSante KA, Farid NA, Welch PA, Callaghan JT and Forgue ST (2000). Confidence Interval Criteria for Assessment of Dose Proportionality. Pharmaceutical Research, 17(10), 1278-83.
30. Spiegelhalter DJ, Abrams KR and Myles JP (2004). Bayesian Approaches to Clinical Trials and Health-care Evaluation. Chichester, Wiley.
31. Zhang JJ, Chen H, He K, Tang S, Justice R, Keegan P, Pazdur R and Sridhara R. Evaluation of blinded independent central review of tumor progression in oncology clinical trials: a meta-analysis. Therapeutic innovation & regulatory science: 2012, 47(2): 167-174.

NOVARTIS

Clinical Development

Spartalizumab (PDR001), LAG525

CPDR001J2201 / NCT03484923

A randomized, open-label, phase II open platform study evaluating the efficacy and safety of novel spartalizumab (PDR001) combinations in previously treated unresectable or metastatic melanoma

Statistical Analysis Plan (SAP) for non-randomized section (single-arm 1A)

Author:	Trial Statistician, [redacted]

Document type:	SAP Documentation
Document status:	Final
Release date:	12-Nov-2020
Number of pages:	26

Document History – Changes compared to previous final version of SAP

Date	Time point	Reason for update	Outcome for update	Section and title impacted (Current)
12-Nov-2020	Prior to FPFV of Arm 1A	Initial version	N/A - First version	NA

Table of contents

List of abbreviations

AE	Adverse event
AESI	Adverse Event of Special Interest
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BIRC	Blinded Independent Central Review
BOR	Best Overall Response
CI	Confidence Interval
CR	Complete Response
CSR	Clinical Study report
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease Control Rate
DoR	Duration of Response
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
FAS	Full Analysis Set
IA	Interim Analyses
[redacted]	[redacted]
IHC	Immunohistochemistry
[redacted]	[redacted]
LDH	Lactate Dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
NMQ	Novartis MedDRA queries
ORR	Overall Response Rate
OS	Overall Survival
PD-1	Programmed Death 1
PD-L1	Programmed Death-Ligand 1
pFBP	Proportion of subjects with a favorable biomarker profile
PFS	Progression-Free Survival
PR	Partial Response
PT	Preferred Term
RECIST	Response Evaluation Criteria in Solid Tumors
SAP	Statistical Analysis Plan
SD	Stable Disease
SMQ	Standardized MedDRA queries
TIL	Tumor Infiltrating T cells
ULN	Upper Limit of Normal
WHO	World Health Organization
WHO-DD	World Health Organization Drug Dictionary

1 Introduction

This statistical analysis plan (SAP) describes all planned analyses for the non-randomized section of the clinical study report (CSR), including the non-randomized single-arm 1A of CPDR001J2201. CPDR001J2201 is a randomized, open-label, phase II open platform study evaluating efficacy and safety of novel spartalizumab (PDR001) combinations in previously treated unresectable or metastatic melanoma. All planned analyses for the randomized section of the study are described in a separate analysis plan, referred as "main SAP for the randomized part of the study" in this document.

The content of this SAP is based on CPDR001J2201 protocol amendment version 05. All decisions regarding final analysis, as defined in the SAP document, have been made prior to database lock of the study data.

1.1 Study design

This is a randomized, open-label, multi-center, two-part, open platform phase II study to assess the efficacy and safety of the anti-PD-1 antibody spartalizumab in combination with novel agents in previously treated unresectable or metastatic melanoma.

A non-randomized single-arm (Arm 1A) has been added with protocol amendment 5 to assess the efficacy and safety of spartalizumab in combination with LAG525 in subjects with previously treated unresectable or metastatic LAG-3 positive melanoma, based on encouraging results from arm 1 following interim analysis 2 of the randomized section of the study.

This study consists of two parts, applicable for both randomized and non-randomized sections of the study :

- Part 1 : Selection phase
- Part 2 : Expansion phase

For more details on the design of the randomized part, refer to protocol or corresponding statistical analysis plan.

In the non-randomized section, arm 1A will enroll up to approximately 100 subjects in total with LAG-3 positive melanoma: 20 subjects in part 1 and up to approximately 80 subjects in part 2.

Study treatment in Arm 1A will be: LAG525 600 mg i.v. Q4W and spartalizumab 400 mg i.v. Q4W.

Two interim analyses for efficacy and futility will be conducted for arm1A: based on data from 20 subjects treated in Part 1 and from 20 additional subjects treated in part 2 (i.e. 40 subjects in total combined with part 1 subjects). The primary analysis will occur after all 100 subjects have been enrolled (combining part 1 and part 2) and followed for at least 9 months. The final analysis will be performed at the end of study (as defined in the protocol).

1.2 Study objectives and endpoints

The same objectives than for the randomized section of the study have been defined for the non-randomized section of the study. Refer to the main SAP for the randomized part of the study

for more details. [redacted]

2 Statistical methods

2.1 Data analysis general information

All analysis will be performed by Novartis. SAS version 9.4 or later software will be used to perform all data analyses and to generate tables, figures and listings.

Data included in the analysis

For each of the analyses, all statistical analyses will be performed using all data collected in the database up to the data cut-off date. All data with an assessment date or event start date (e.g. vital sign assessment date or start date of an adverse event) prior to or on the cut-off date will be included in the analysis. Any data collected beyond the cut-off date will not be included in the analysis and will not be used for any derivations.

All events with start date before or on the cut-off date and end date after the cut-off date will be reported as 'ongoing'. The same rule will be applied to events starting before or on the cut-off date and not having documented end date. This approach applies, in particular, to adverse event and concomitant medication reports. For these cases, the end date will not be imputed and therefore will not appear in the listings.

For Part 1, the interim analysis will be conducted once 20 patients enrolled in arm 1A have either completed two post-baseline tumor assessments or have discontinued the study treatment. For Part 2, if the arm 1A is expanded after interim analysis in part 1, the 2nd interim analysis will be performed once 40 patients (e.g. 20 in part 1 and 20 in part 2) have either completed two post-baseline tumor assessments or have discontinued the study treatment.

The cut-off date of the primary analysis of this non-randomized section will be determined once all patients from Part 2 are enrolled and followed for at least 9 months or have discontinued the study. The analysis cut-off for the final analysis will be established at the end of the study as defined in the protocol.

General analysis conventions

For information on general analysis conventions, see the main SAP for the randomized part of the study.

2.1.1 General definitions

For information on general definitions, see the main SAP for the randomized part of the study except for the ones below.

Study day

For arm 1A, the reference start date for non-safety assessments is the date of start of study treatment.

Baseline

For arm 1A, for efficacy evaluations, the last non-missing assessment, including unscheduled assessments on or before the start of study treatment is taken as "baseline" value or "baseline" assessment. In the context of baseline definition, the efficacy evaluations also include performance status.

Extended follow-up

Additional summaries will be displayed to report deaths, all AEs, AEs related to study treatment, all serious adverse events (SAEs), SAEs related to study treatment and AESI collected up to 150 days after last administration of PDR001/LAG525.

2.2 Analysis sets

All below defined analysis sets will be used to analyze data from both the selection (Part 1) and expansion (Part 2) phases. Note, the immunogenicity (IG) analysis sets are outside the scope of the interim analyses (IA).

Full Analysis Set

The Full Analysis Set (FAS) comprises all subjects who received at least one dose of study treatment. This population will be the primary population for efficacy analysis.

Safety

The Safety Set is identical to the FAS.

Note: Even if the subject receives partial infusion of the treatment or only one component of the treatment, the subject will still be analyzed under PDR001+LAG525 treatment arm.

Biomarker analysis set

The biomarker analysis set, as it pertains to the secondary objective (Section 10.5.5 of the protocol), consists of all subjects in the safety set with an evaluable baseline tumor biopsy sample and at least one evaluable post-baseline tumor biopsy sample.

Evaluability for the analysis of the secondary objective will consist of the following:

- Having both tumor biopsy samples as stated above,
- Having quantifiable results for at least two assays to calculate the secondary endpoint as described in Section 2.12.2.

Immunogenicity (IG) analysis sets

The immunogenicity prevalence set includes all subjects in the Full analysis set with a determinant baseline IG sample **or** at least one determinant post-baseline IG sample.

The immunogenicity incidence set includes all subjects in the Immunogenicity prevalence set with a determinant baseline IG sample **and** at least one determinant post-baseline IG sample.

See Section 2.10.1.1 of the main SAP for the randomized part of the study for the definition of determinant.

For more information on the additional analyses sets as well as subject classification and withdrawal of informed consent information, see the main SAP for the randomized part of the study.

2.2.1 Subgroup of interest

Subgroup analysis as described in the main SAP for the randomized part of the study may be only performed at the primary analysis (e.g. once all patients from part 2 have been enrolled and followed for at least 9 months) but not at interim analyses. Of note, LDH subgroup defined in the main SAP for the randomized part of the study as a stratification factor is not a stratification factor for this non-randomized section of the study. In addition, subgroup linked to ECOG performance status will be defined as 0 versus 1.

2.3 Subject disposition, demographics and other baseline characteristics

The Full Analysis Set (FAS) will be used for all baseline (including disease characteristics) and demographic summaries unless otherwise specified. Summaries will be reported for all subjects to assess baseline comparability. No inferential statistics will be provided.

Basic demographic and background data

All demographic and baseline disease characteristics data will be summarized and listed. Categorical data (e.*g.* gender, age groups: < 65 and ≥ 65 years, race, ethnicity, ECOG performance status, etc...) will be summarized by frequency counts and percentages; the number and percentage of subjects with missing data will be provided. Continuous data (e.g. age, weight, height, body mass index) will be summarized by descriptive statistics (N, mean, median, standard deviation, minimum and maximum). BMI (kg/m2) will be calculated as weight[kg] / (height[m]2) using weight at Screening.

Diagnosis and extent of cancer

Summary statistics will be tabulated for diagnosis and extent of cancer. This analysis will include the following: primary site of cancer at initial diagnosis, diagnosis of disease, BRAF mutation status, details of tumor histology/cytology, stage at study entry per AJCC Edition 8, time since initial diagnosis, time from initial diagnosis to first recurrence/progression (in months), time since most recent relapse/progression to randomization (in months), presence/absence of target and non-target lesions, sum of lesion diameter for target lesions at baseline, number and type of metastatic sites involved, and number of metastatic sites. Note: Presence/absence of target and non-target lesions will be based on the data collected on RECIST target/non-target lesion assessment eCRF pages. Metastatic sites will be based on diagnosis page. Tumor (T) stage at study entry, lymph nodes (N) stage at study entry, and metastases (M) at study entry will be further listed in the Subject disease history listing.

Medical history

Medical history and ongoing conditions will be presented identically to the content in the main SAP for the randomized part of the study.

2.3.1 Subject disposition

The number (%) of enrolled Arm 1A subjects included in the FAS will be presented overall. The number (%) of screened and not-treated subjects and the reasons for screening failure will also be displayed. The number (%) of subjects in the FAS who are still on treatment, who discontinued the study phases and the reason for discontinuation will be presented.

The remaining summaries for subject disposition are identical to those in the main SAP for the randomized part of the study.
The following summaries will be provided (with % based on the total number of FAS patients):

- Number (%) of patients who completed the screening phase but not treated,
- Primary reason for not being treated,
- Number (%) of patients who were treated,
- Number (%) of patients who are still on-treatment,
- Number (%) of patients who discontinued the study treatment phase,
- Primary reason for study treatment phase discontinuation,
- Number (%) of patients who have entered the post-treatment follow-up,
- Number (%) of patients who have discontinued from the post-treatment follow-up,
- Reasons for discontinuation from the post-treatment follow-up,
- Number (%) of patients who have entered the survival follow-up.

Protocol deviations

Summaries and listing for protocol deviations are identical to those in the main SAP for the randomized part of the study.

Analysis sets

The number (%) of subjects in each analysis set (defined in Section 2.2) will be summarized.

2.4 Treatments (study treatment, rescue medication, concomitant therapies, compliance)

The Safety set will be used for the analyses below. Categorical data will be summarized as frequencies and percentages. For continuous data, mean, standard deviation, median, 25th and 75th percentiles, minimum, and maximum will be presented.

2.4.1 Study treatment / compliance

Study treatment and compliance information will be presented identically for the IAs, the primary analysis and the final analysis. Therefore, more details on study treatment and compliance information can be found in the main SAP for the randomized part of the study.

For PDR001 and LAG525, for the calculation of the dose intensity, planned dose intensity and relative dose intensity, the full theoretical treatment cycle duration (e.g. 28 days) for the last cycle will be considered : the modified duration of exposure will use the last date of exposure of study drug as defined in the main SAP for the randomized part of the study, but without cutting at the date of death, last contact date or cut-off date in case the patient died, was lost to follow-up, before the end of the cycle or the last exposure date is going beyond the cut-off date, respectively. Table 2-1 below reflects the definition of the last date of exposure of study drug to be used for the modified duration of exposure in the dose intensity calculation.

Duration of exposure to study treatment, dose intensity and planned dose intensity

Duration of exposure to study treatment (days) = (last date of exposure to study treatment) – (date of first administration of study treatment) + 1.

The last date of exposure to study treatment is the latest of the last dates of exposure to PDR001 and to LAG525 (see Table 2-1 below).

Table 2-1 Definition of last date of exposure of study drug

Scenario	Definition of last date of exposure of study drug	Example
PDR001 LAG525	The planned end date of the last cycle in which the last non-zero dose of the investigational drug was last administered (i.e last date of administration + (planned interval duration-1day)) The note below will not be applicable for dose intensity calculation : Note : If the patient died or was lost to follow-up before the derived last date, the last date of exposure to investigational drug is the date of death or the date of last contact, respectively	Example 1: If PDR001 or LAG525, is administered every four weeks, the last date of exposure is the date of administration in the last cycle + 27 days.

	If the derived last date of exposure goes beyond the data cutoff date, it should be truncated to the date of data cutoff.	

Dose intensity and planned dose intensity

For drugs administered every 4 weeks (PDR001 and LAG525), the dose intensity and planned dose intensity will be reported in mg/cycle, defined as follows :

- DI (mg / cycle) = (Actual Cumulative dose (mg) / modified duration of exposure to study treatment (days)) * 28, where modified duration of exposure (days) = last date of exposure of study drug (as defined in Table 2-1 above) - (date of first administration of study drug) + 1.

  If for example, the duration of exposure to PDR001 is 6 days because patient died 6 days after first day of dosing, the modified duration of exposure to PDR001 used in dose intensity calculation will be 28 days.

- PDI (mg/cycle) is 400 mg/cycle and 600 mg/cycle, for PDR001 and LAG525, respectively.

Dose interruptions or permanent discontinuations

The number of subjects who have dose permanent discontinuations, dose interruptions, and the reasons, will be summarized separately for each study drug, as described in the main SAP for the randomized part of the study.

Treatment beyond RECIST progression

The number of subjects who continue treatment beyond RECIST1.1 progression according to local investigators assessment based on protocol specified criteria will be summarized. It includes all subjects who received any study treatment (i.e. at least one dose of any component of the study treatment) after RECIST 1.1 progression assessed by local investigators. Those subjects will be identified using the field “Will the subject continue treatment beyond disease progression as per RECIST1.1?” on the Verification for Treatment beyond RECIST1.1 PD CRF pages.

2.4.2 Prior, concomitant and post therapies

Prior anti-cancer therapy

The number and percentage of subjects who received any prior anti-neoplastic medications, prior checkpoint inhibitor therapy, prior anti-neoplastic radiotherapy, prior anti-neoplastic surgery, or other therapies not defined by the previous categories, will be summarized. Prior anti-neoplastic medications will be summarized by therapy type (e.g. chemotherapy, hormonal therapy, immunotherapy etc.) and by setting (e.g. adjuvant, metastatic, etc.). Summaries will include total number of regimens, time from last treatment to progression for the last therapy, type of therapy received at any time or as last systemic therapy for unresectable and metastatic melanoma prior to enrollment (checkpoint inhibitor vs targeted therapy vs Other) excluding

radiotherapy and surgery, best response to last systemic therapy, to anti Anti-PD-1/Anti-PD-L1 therapy and best response to any immunotherapy, based on RECIST v1.1.

For radiotherapy, time since last radiotherapy, locations and setting of last therapy will be summarized. For prior surgery, time since last surgery and procedure will be summarized.

The prior checkpoint inhibitor therapy will be further summarized by type of checkpoint inhibitor therapy (Anti-PD-1, Anti-CTLA-4, and combination therapy of anti-CTLA-4 and anti-PD-1). The prior targeted therapy will be further summarized by type of targeted therapy (BRAF Inhibitor and MEK Inhibitor). The medication therapy type of any combination therapy will be classified based on the types of treatment taken at each regimen. For example, a combination therapy of targeted therapy and immunotherapy will be classified in both categories with appropriate sub-types.

Swimmer plot will be performed to present the subject's response according treatment duration. The different type of checkpoint inhibitor therapy (Anti-PD-1, Anti-CTLA-4, and combination therapy of anti-CTLA-4 and anti-PD-1, …) received by the patient will be presented in parallel on the same graph, with the duration of the different prior therapy and best response to prior therapy. This plot will be produced at time of interim analyses, if number of patients is appropriate to keep the plot readable, for patients who responded (BOR in (CR,PR)) or with a BOR as SD.

Separate listings will be produced for prior anti-neoplastic medications, radiotherapy, and surgery.

Anti-neoplastic medications will be coded using the World Health Organization (WHO) Drug Dictionary (WHO-DD); anti-neoplastic surgery will be coded using MedDRA. Details regarding MedDRA and WHO-DD version will be included in the footnote in the tables/listings.

The above analyses will be performed using the FAS.

Post treatment anti-cancer therapy

Anti-neoplastic therapies since discontinuation of study treatment will be listed and summarized by ATC class, preferred term, by means of frequency count and percentages using FAS at primary and final analysis. The number and percentage of patients who reported at least one anti-neoplastic therapy since discontinuation of study treatment by category (see main SAP for the randomized part of the study Table 2-7) will be reported.

Concomitant medications

Summaries and listings as described in the main SAP for the randomized part of the study will be provided at primary and final analysis.

2.5 Analysis of the primary objective

The primary objective is to demonstrate the antitumor activity of spartalizumab in combination with LAG525 as measured by overall response rate (ORR) in previously treated unresectable or metastatic melanoma LAG-3 positive subjects. In particular, for the IAs, the primary objective will be to determine if investigator assessed ORR is sufficient to allow for the non-randomized arm 1A to be expanded or not.

2.5.1 Primary endpoint

Confirmed ORR is defined as the proportion of subjects with confirmed best overall response (BOR) of complete response (CR) or partial response (PR) according to RECIST 1.1.

Of note, BOR of SD will be assigned if at least one SD assessment (or better) occurred > 11 weeks after randomization (and not qualifying for CR or PR).

Tumor assessments performed after the start of any further antineoplastic therapy (i.e. any additional secondary antineoplastic therapy, radiotherapy or surgery, including limited-field palliative radiotherapy) will not be considered in the assessment of BOR. Further antineoplastic therapies will be identified via protocol deviations and from the data collected on "antineoplastic therapies since discontinuation of study treatment" (e)CRF pages. Of note, "biopsy" will not be considered as a new antineoplastic surgery.

For additional information on the primary endpoint, see the main SAP for the randomized part of the study.

2.5.2 Statistical hypothesis, model, and method of analysis

Part 1: Selection Phase

No formal hypothesis testing will be conducted in part 1. The objective of this part of the study is to explore the anti-tumor activity of the combination of spartalizumab with LAG525 agent in subjects with LAG-3 positive melanoma as measured by ORR per RECIST v1.1, and to expand this new arm if it shows promising efficacy.

To determine whether the arm has the potential to demonstrate clinically meaningful response rate, a specific efficacy probabilistic threshold should be met during the selection phase to advance to the expansion phase. If not, the arm will be declared as futile and won't be expanded. These thresholds and probabilities are highlighted in Table 2-2.

Table 2-2 Probabilities thresholds used to declare efficacy or futility for arm 1A

Decision rule	Action
P(ORR ≥20%) ≤ 70%	Drop arm for futility
P(ORR ≥20%) > 70%	Efficacy signal - Consider advancing arm to expansion phase (part 2)

To calculate the posterior probabilities, a beta-binomial posterior distribution will be used as shown below:

$$p \sim beta(y+1, n-y+1)$$

Where y is the number of confirmed responses in arm 1A and n is the number of subjects treated. In calculation of the posterior distribution, uninformed prior Beta[1,1] will be taken into account. The posterior probability will be calculated based on the observed ORR using RECIST v1.1 criteria at interim analysis.

The final decision on whether the arm will be selected for advancement to the expansion phase will primarily be based on the observed ORR using decision rule described above but other efficacy (in particular duration of response), safety, and biomarker data will be considered as well. Novartis may decide to not expand the arm despite the efficacy probability threshold crossed.

With 20 treated patients, 5 or more responders need to be observed to meet the efficacy probabilistic threshold.

Part 2: Expansion Phase

If the pre-defined efficacy probabilistic threshold is met after the selection part, the arm 1A might be expanded.

A second interim analysis will be conducted including 20 additional patients treated in the expansion phase, i.e. based on the total of 40 patients. The objective of this second interim analysis is to support and inform decision-making on whether arm 1A will be: (1) further explored (2) stopped for futility . To determine whether arm 1A has the potential to demonstrate clinically meaningful ORR, an efficacy probabilistic threshold same as in part 1 will need to be crossed (see Table 2-2). If not, the arm 1A will be declared futile and further exploration/enrollment will be stopped in this combination. With 40 subjects at time of 2nd interim analysis treated in arm 1A (combining part 1 and part 2), 9 responders will have to be observed to cross the efficacy probabilistic threshold.

If efficacy probabilistic threshold is met at the 2nd interim analysis enrollment may continue in expansion up to approximately 80 subjects (i.e. 100 subjects in total adding the subjects in part 1).

After completion of the expansion phase, the following statistical hypothesis will be tested based on objective response rate (ORR) at time of primary analysis:

- H_0: ORR $\leq$ 15%
- H_1: ORR >15%

The null hypothesis will be formally rejected if the lower bound of exact 95% CI for ORR is above 15%. The primary analysis of ORR at the end of part 2 will be based on a combined pool of part 1 and part 2 data.

23 or more responses need to be observed among the 100 subjects to reject the null hypothesis. With those 23 responders among 100 subjects, the point estimate for ORR will be 23% with an exact 95% confidence interval of (15.2, 32.5).

2.5.3 Handling of missing values/censoring/discontinuations

For information on the handling of missing values/censoring/discontinuations, see the main SAP for the randomized part of the study.

2.5.4 Supportive analyses

Same supportive analysis on ORR using BIRC than the one described in the main SAP for the randomized part of the study will be performed for Arm 1A at 2nd interim analysis and primary analysis.

Concordance analysis of BOR

Refer to the main SAP for the randomized part of the study for details on the concordance analysis between BIRC assessment and local assessment.

Reasons for "Unknown" BOR

Reasons for "Unknown" BOR will be presented identically for the IAs and primary analysis. Therefore, more details on reasons for "Unknown" BOR can be found in the main SAP for the randomized part of the study.

Waterfall plot to depict anti-tumor activity

Waterfall plot to depict anti-tumor activity will be presented identically for the IAs and primary analysis. Therefore, more details on reasons for "Unknown" BOR can be found in the main SAP for the randomized part of the study.

Subgroup analyses

At primary analysis, data from part 2 will be combined with part 1 to assess ORR in the following subgroups :

- LDH (≤ ULN vs. > ULN),
- ECOG PS (0 vs. 1).

2.6 Analysis of the key secondary objective

The key secondary objective of the study is to determine duration of response (DoR) treatment with spartalizumab in combination with LAG525 in subjects with previously treated unresectable or metastatic LAG3-positive melanoma. It will be analyzed at time of IAs, primary and final analysis.

When the main SAP for the randomized part of the study referred to the date of randomization for endpoints definition, the date of first study treatment dose is rather used for this arm 1A analysis.

2.6.1 Key secondary endpoint

For information on the key secondary endpoint, see the main SAP for the randomized part of the study. Note, due to small sample size at 1st interim analysis, only listings will be generated for DoR.

2.6.2 Statistical hypothesis, model, and method of analysis

For information on the statistical hypothesis, model, and method of analysis for the key secondary endpoint, see the main SAP for the randomized part of the study.

2.6.3 Handling of missing values/censoring/discontinuations

For information on the handling of the missing values/censoring/discontinuations for the key secondary endpoint, see the main SAP for the randomized part of the study.

2.6.4 Duration of follow-up

Study follow-up will be summarized as well as patient follow-up and will be defined as follows :

- Duration between the date of first dose of study treatment in arm 1A and data cut-off date = (Cut-off date – Date of first dose of study treatment in arm 1A+ 1) / 30.4375 (months),
- Duration between the date of first dose of study treatment in arm 1A and the last contact date = (Date of last contact/death date (equivalent to the date used for OS time) – Date of first dose of study treatment in arm 1A+ 1) /30.4375 (months).

2.7 Analysis of secondary [redacted] efficacy objective(s)

The other secondary objectives are the following:

- Evaluate the efficacy of arm 1A, as measured by progression-free survival (PFS)
- Evaluate the efficacy of arm 1A, as measured by disease control rate (DCR)
- Evaluate the overall survival (OS) of arm 1A

[redacted]

2.7.1 Secondary [redacted] efficacy endpoints

For secondary endpoints definition, see the main SAP for the randomized part of the study.

Note, due to small sample size at 1st interim analysis, only listings for the following secondary endpoints PFS and OS will be provided. Full analyses using Kaplan Meier method will be performed at 2nd interim analysis, primary and final analyses.

[redacted]

2.7.2 Statistical hypothesis, model, and method of analysis

There will be no hypothesis testing for any of the endpoints listed above.

For the method of analysis of secondary [redacted] endpoints, refer to the main SAP for the randomized part of the study.

2.7.3 Handling of missing values/censoring/discontinuations

For handling of missing values/censoring/discontinuations for PFS and OS endpoints, please see the main SAP for the randomized part of the study.

2.8 Safety analyses

For all safety analyses, the safety set will be used. All listings and tables will include only arm 1A.

The overall obervation period, incomplete dates, and safety assessments will be treated similarly in the IA analyses, the primary analysis and the final analysis. More details on these components of the safety analyses can be found in the main SAP for the randomized part of the study.

2.8.1 Adverse events (AEs)

Adverse event summaries and details will be presented identically for the IAs, the primary and the final analysis. Therefore, more details on AEs can be found in the main SAP for the randomized part of the study. Of note, the additional summaries to report all AEs, AEs related to study treatment, all SAEs, and SAEs related to study treatment as described in the main SAP for the randomized part of the study as "collected up to 150 days after last administration of PDR001/LAG525" will rather use the extended follow-up period definition (section 2.1.1).

2.8.1.1 Adverse events of special interest / grouping of AEs

Adverse events of special interest (AESI) during the on-treatment period will be tabulated. The list of AESI will also include relevant events for spartalizumab, and LAG525 agent.

Refer to the main SAP for the randomized part of the study for details on the analysis of AESIs. Of note, the additional summaries to report all AESIs as described in the main SAP for the randomized part of the study as "collected up to 150 days after last administration of PDR001/LAG525" will rather use the extended follow-up period definition (section 2.1.1). In addition, analysis of time to first occurrence summaries will not be generated.

2.8.2 Deaths

Deaths will be presented identically for the IAs, the primary and the final analysis. Therefore, more details on deaths can be found in the main SAP for the randomized part of the study. The additional summary reported all deaths "up to 150 days after last administration of PDR001/LAG525", as described in the main SAP for the randomized part of the study, will rather used the extended follow-up definition (section 2.1.1).

2.8.3 Laboratory data

Grading of laboratory will be assigned programmatically as per National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v5.0. The calculation of CTCAE grades will be based on the observed laboratory values only, clinical assessments will not be taken into account.

CTCAE Grade 0 will be assigned for all non-missing values not graded as 1 or higher. Grade 5 will not be used.

For laboratory tests where grades are not defined by CTCAE v5.0, results will be categorized as low/normal/high based on laboratory normal ranges.

Laboratory data summaries and listings will be presented identically for the IAs, the primary and the final analysis. More details on laboratory data summaries and listings can be found in the main SAP for the randomized part of the study.

For laboratory figures, only the eDish plot examining the relationship between peak total bilirubin (measured as the maximum post-baseline total bilirubin and maximum post-baseline alanine aminotransferase (ALT) (both in reference to the ULN) or maximum post-baseline aspartate aminotransferase (AST) will be produced at IAs, primary and final analysis.

Liver function parameters

The following liver function parameters are of interest : total bilirubin (TBL), alanine aminotransferase (ALT), aspartate aminotransferase (AST) and alkaline phosphatase (ALP). The number (%) of subjects with worst post-baseline values as per Novartis Liver Toxicity guidelines will be summarized.

The following summaries on peak post-baseline values will be produced:

- ALT >3xULN, ALT>5x ULN, ALT>10xULN, ALT>20xULN
- AST >3xULN, AST>5xULN, AST>10xULN, AST>20xULN
- ALT or AST > 3xULN, ALT or AST > 5xULN, ALT or AST > 8xULN, ALT or AST > 10xULN, ALT or AST > 20xULN
- TBL > 2xULN, TBL > 3xULN

For patients with AST and ALT ≤ ULN at baseline:

- ALT or AST > 3xULN & TBL > 2xULN
- ALT or AST > 3xULN & TBL > 2xULN & (ALP < 2xULN or ALP ≥2xULN)

For patients with ALT or AST > ULN at baseline

- Elevated ALT or AST (>3xULN if ≤ULN at baseline, or (>3xBsl or 8xULN) if >ULN at baseline)
    - and TBIL (>2xBsl and 2xULN)
    - and TBIL (>2xBsl and 2xULN) and ALP ≥2xULN
    - and TBIL (>2xBsl and 2xULN) and ALP < 2xULN

Potential Hy's Law events are defined as occurrence of AST or ALT > 3xULN and TBL > 2xULN, or occurrence of (ALT or AST >3xbaseline) OR (ALT or AST > 8.0xULN), whichever is lower, combined with (TBIL>2xbaseline AND >2.0xULN), for patients with abnormal ALT or AST baseline values; where combined elevations post-baseline are based on the peak values at any post-baseline time for a subject.

Further medical review has to be conducted to assess potential confounding factor such as, liver metastases, liver function at baseline etc.

2.8.4 Other safety data

2.8.4.1 ECG and cardiac imaging data

Categorical analysis of QT/QTc interval data based on the number of subjects meeting or exceeding predefined limits in terms of absolute QT/QTc intervals or changes from baseline will be presented. In addition, a listing of these subjects will be produced.

Data handling

In case the study requires ECG replicates at any assessment, the average of the ECG parameters at that assessment should be used in the analyses.

Data analysis

Data analysis for ECG and cardiac imaging data will be presented identically for the IAs, the primary and the final analysis. More details on the data analysis of ECG and cardiac imaging data can be found in the main SAP for the randomized part of the study. Only notable ECG values will be listed.

2.8.4.2 Vital signs

Vital sign assessments are performed in order to characterize basic body function. The following parameters were collected: height (cm), weight (kg), body temperature (°C), heart rate (beats per minute), systolic and diastolic blood pressure (mmHg).

Data handling

Vital signs collected on treatment will be summarized at primary and final analyses (but not for IAs purpose). Values measured outside of the on-treatment period will be flagged in the listings.

[redacted]

2.10 PD and [redacted]PD analyses

Same analyses than in the main SAP for the randomized part of the study will be performed for immunogenicity.

[redacted]

2.11 Subject-reported outcomes

Not applicable

2.12 Biomarkers

2.12.1 Introduction

As a project standard, Novartis Oncology Biostatistics will analyze only biomarkers collected in the clinical database. The biomarker analysis covered within the SAP pertains mainly to the secondary objective endpoint of determining the proportion of subjects with a favorable biomarker profile (pFBP). More details on the pFBP can be found in the main SAP for the randomized part of the study.

[redacted]

2.12.2 Secondary biomarker objectives

Changes in phenotype and/or activation of T cell populations in the tumor and tumor microenvironment will be analyzed on an on-going basis as part of the secondary objective. Every effort will be made to make this data available for internal decision making purposes. Such analyses may include, but are not limited to number of $CD8^+$ T cell infiltration, T cell activation level by IHC and gene expression analysis. T cell repertoire/clonality by TCR-sequencing is excluded from all interim analyses as well as beyond due to low assay success. Further discussions of the parameter definitions to support the biomarker objectives can be found in Table 2-3 below.

Table 2-3 Parameter definitions for categorization of biomarker outcome at subject level

Biomarker parameter	Analysis	Criteria for favorable biomarker result
Number of tumor infiltrating T cells (TILs)	Changes of $CD8^+$ T cell numbers as assessed by IHC in tumor/ tumor microenvironment with treatment	Increase
Activation level of TILs*	Change of T cell activation marker level(s) as assessed by IHC in tumor/tumor microenvironment with treatment	Increase
Changes in immune response gene expression signatures	Changes in expression of genes and/or gene signatures as assessed by mRNA analysis in tumor/tumor microenvironment with treatment	Defined per gene/gene signature

*Of note, the activation level of TILs by IHC (AQUA analysis) initially reported in the protocol may be excluded due to sample limitation associations with arm 1A, created by the need to retain tissue for potential further LAG3 test validation in support of future diagnostic development.

2.12.3 Biomarker data

As a project standard, Novartis Oncology Biostatistics will analyze only biomarkers collected in the clinical database. The biomarker analysis covered within the SAP pertains mainly to the secondary objective endpoint of determining the proportion of subjects with a favorable biomarker profile (pFBP). [redacted]

The biomarker analysis set, as it pertains to the secondary objectives consists of all subjects in the safety set with an evaluable baseline tumor biopsy sample. When evaluating the pFBP, the analysis set should include all subjects in the safety set with an evaluable baseline tumor biopsy sample and at least one post-baseline tumor biopsy sample with evaluable results for 2 or more biomarker parameters listed in Table 2-3 (paired biomarker analysis set).

Table 2-4 summarizes the biomarker collection schedules and sample types.

Table 2-4 Sample biomarker summary table

Biomarker parameter	Biomarker analysis measurement	Time point	Sample	Method	Dataset
Number of TILs	Percent of tumoral CD8+ cells	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	IHC	B1
Changes gene expression signatures	T-cell Inflammed Signature (TIS) levels	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	Nanostring	B1
Baseline LAG-3 expression (LAG525 treatment arm only)	Percent of tumoral LAG-3+ cells	Screening Btw, C1D1 and C2D1*	Newly obtained tumor biopsy	IHC	B1

2.12.4 General data handling and preprocessing

The mean of all pre dose assessments (i.e. screening and Cycle 1 Day 1 assessments) will be used as the baseline value. The on-treatment assessment will be the result from the tumor biopsy taken between C1D21 and C2D1.

When more than one biomarker data value are available for a subject at any time point, the mean of the replicate values will be used for all statistical analyses.

2.12.5 Secondary biomarker objectives

The secondary objective biomarker data will be listed for each subject for all time points (baseline and on treatment). Absolute and relative change (percent change or fold change, depending on biomarker) from baseline of each biomarker parameter will be calculated for each subject and summarized for arm 1A.

2.12.5.1 IHC for quantifying tumor infiltrating T cells (TILs)

The number of TILs at each timepoint will be assessed by measuring % CD8+ cells within the invasive margin, tumor and stroma area (as defined by Melamask staining).

For all CD8 parameters above, the baseline value, on-treatment (C1D21-C2D1) value, and the absolute difference from baseline to on-treatment will be listed for each individual subject for arm 1A. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all subjects (regardless of treatment arms) will be reported for baseline and change from baseline to on-treatment by combination arm. Inter-quartile range is the number of data points between the 25th and 75th percentile.

The summary statistics : mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range for baseline %CD8+ cells will be summarized for patients with best overall response of CR or PR (e.g. responders) versus others.

For all parameters above, the baseline value, on-treatment value, and the absolute difference from baseline to 3-4 weeks will be listed for each individual subject. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all subjects (regardless of treatment arms) will be reported for baseline and change from baseline to on-treatment (C1D21-C2D1). Inter-quartile range is the number of data points between the 25th and 75th percentile.

2.12.5.2 Nanostring for quantifying gene expression signatures

The Tcell inflamed signature will be assess by the T-cell Inflammed Signature (TIS) score for Nanostring data.

For TIS score, the baseline value, on-treatment (C1D21-C2D1) value, and the fold change (on-treatment / baseline) will be listed for each individual subject. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all subjects (regardless of treatment arms) will be reported for baseline and fold change (on-treatment / baseline) . Inter-quartile range is the number of data points between the 25th and 75th percentile.

The summary statistics : mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range for baseline TIS score will be summarized for patients with best overall response of CR or PR (e.g. responders) versus others.

2.12.5.3 Calculation of pFBP

The biomarkers listed in Table 2-3 will be used to calculate the proportion of subjects with a favorable biomarker profile (pFBP).

For each biomarker involved in the evaluation of pFBP as stated above, the baseline, 3-4 weeks biomarker values, and the absolute difference from baseline to 3-4 weeks will be listed for each individual subject by combination arm. The mean, standard deviation, %CV, median, minimum, maximum, inter-quartile range and number for all subjects (regardless of treatment arms) will be reported for baseline and change from baseline to 3-4 weeks by combination arm. Inter-quartile range is the number of data points between the 25th and 75th percentile.

If an individual subject has two favorable biomarker parameters from Table 2-3 then that subject will be considered to have a favorable biomarker profile. A favorable biomarker will be defined as follows:

Number of tumor infiltrating T cells (TIL) as measured by CD8 IHC:

- An increase from baseline to on-treatment (C1D21-C2D1) in the IE CD8+ density is considered a favorable change as follows:
  - IE CD8 increase which is ≥1% in magnitude (i.e. [on-treatment C1D21-C2D1 %IE value] – [baseline IE value = ≥1%])

And

  - Percent fold change in IE is ≥30% (i.e. ([on-treatment C1D21-C2D1 %IE value] – [baseline IE value])/[baseline IE value] = ≥30%])

Changes in immune response gene expression signature as measured by Nanostring:

- For baseline to on-treatment changes in the TIS score, the following would be considered favorable:
  - ≥1.5 fold change (i.e. on-treatment aggregate gene signature score]/[baseline aggregate gene signature score])

For each combination arm, the proportion of subjects with a favorable biomarker profile (pFBP) are then defined as the number of subjects with favorable biomarker profiles divided by the number of subjects enrolled into the arm which are included in the paired biomarker analysis set for that particular interim analysis and will be summarized.

[redacted]

2.13 Other Exploratory analyses

Not applicable

2.14 Interim analysis

Part 1 : Selection phase

As described in protocol amendment v05 Section 4.2.2, there will be one interim analysis planned for ORR once all patients treated in arm 1A have either completed two post-baseline tumor assessments or have discontinued study treatment.

Part 2: Expansion phase

A 2nd interim analysis will be conducted once 40 patients have been treated within selection and expansion phases and have either completed two post-baseline tumor assessments or have discontinued study treatment.The primary analysis of this non-randomized section of the study will be performed once all patients treated in arm 1A had at least 9 months of follow-up. This timing would allow for a potential minimum duration of response of 6 months for all subjects enrolled. Formal hypothesis testing of the primary endpoint (as described in Section 2.6.2) will be performed at this primary analysis. The final analysis will be performed at the end of the study as described in protocol Section 4.3.

3 Sample size calculation

Further discussion on sample size calculations can be found in Section 10.8.2 of the protocol.

4 Change to protocol specified analyses

Not applicable.

5 Appendix

5.1 Imputation rules

5.1.1 Study drug

Imputation rules for study drugs are identical for the IAs,the primary analysis and the final analysis. Therefore, more details concerning the imputation rules for study drug can be found in the main SAP for the randomized part of the study.

5.1.2 AE date imputation

Imputation rules for AE date are identical for all analyses. Therefore, more details concerning the imputation rules for AE date imputation can be found in the main SAP for the randomized part of the study.

5.1.3 Concomitant medication date imputation

Imputation rules for concomitant medication date are identical for all analyses. Therefore, more details concerning the imputation rules for concomitant medication date imputation can be found in the main SAP for the randomized part of the study.

5.2 AEs coding/grading

AEs coding/grading are identical for all analyses. Therefore, more details concerning AEs coding/grading can be found in the main SAP for the randomized part of the study.

5.3 Laboratory parameters derivations

Laboratory parameter definitons are identical for all analyses. Therefore, more details concerning laboratory parameter definitions can be found in the main SAP for the randomized part of the study.

5.4 Statistical models

5.4.1 Primary analysis

Please refer to Section 10 of the protocol, Section 2.5.2 of the current SAP for extended details on the modeling and SAS procedures used.

5.4.2 Key secondary analysis

Kaplan-Meier estimates

For information on the Kaplan-Meier estimates and dealing with ties, see the main SAP for the randomized part of the study.

5.4.3 Audit-based BIRC assessment of PFS

Not applicable.

5.5 Rule of exclusion criteria of analysis sets

Not applicable.

[redacted]