Protocol C3861001

A PHASE 1 DOSE ESCALATION AND EXPANSION STUDY EVALUATING THE SAFETY, TOLERABILITY, PHARMACOKINETICS, PHARMACODYNAMICS AND ANTI TUMOR ACTIVITY OF PF-07062119 IN PATIENTS WITH ADVANCED GASTROINTESTINAL TUMORS

Statistical Analysis Plan
(SAP)

Version: 1
Date: 14 Oct 2019

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

APPENDICES

1. VERSION HISTORY

This is the first version.

2. INTRODUCTION

PF-07062119 is an anti-Guanylyl Cyclase 2C (GUCY2c)/anti-CD3 bispecific Fc diabody targeting CD3 on T cells with one binding domain and GUCY2c with the other binding domain. PF-07062119 acts as a T cell redirecting bispecific for the treatment of gastrointestinal tumors including colorectal, gastric and esophageal adenocarcinomas.

This statistical analysis plan (SAP) provides the detailed methodology for summary and statistical analyses of the data collected in Study C3861001. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

2.1. Study Objectives, Endpoints and Estimands (part 1)

2.1.1. Primary Objectives

To assess safety and tolerability of increasing dose levels of PF-07062119 administered in patients with advanced gastrointestinal tumors, including colon, gastric and esophageal adenocarcinomas, for whom no standard therapy is available in order to estimate the Maximum Tolerated Dose (MTD) and select the Recommended Phase 2 Dose (RP2D) (as a monotherapy in Part 1A and in combination therapy in Part 1B with select tumor(s)).

2.1.2. Secondary Objectives

- *To characterize the single and multiple dose pharmacokinetics (PK) of PF-07062119.*
- *To evaluate the immunogenicity of PF-07062119.*
- *To evaluate the immunogenicity of PF-06801591 and bevacizumab-Pfizer given in combination with PF-07062119 (part 1B).*
- *To evaluate immune cells in archival biopsies and/or pre and post treatment biopsies (if available).*
- *To evaluate preliminary anti-tumor activity.*

2.1.3. Exploratory Objectives

- *To explore preliminary anti-tumor activity by standard and immune related response criteria and time to event endpoint.*
- *To explore the relationship between target expression and biomarkers in archival tumor biopsies and/or pre and post treatment tumor biopsies (if available), whole blood and serum samples in order to elucidate mechanism of action and response to therapy.*

- *To explore additional biomarkers related to clinical response, in order to elucidate the mechanism of action of PF-07062119, predict response to therapy and to understand resistance mechanisms that may predict escape from therapy.*

- *To evaluate T, B, and Natural Killer (NK) subtypes for immunophenotyping in addition to T cell proliferation, activation, and exhaustion markers in whole blood.*

- *To explore cytokines and chemokines as pharmacodynamic (PD) markers in pre and on treatment serum samples.*

- *To explore the tumor biomarker carcinoembryonic antigen (CEA) levels.*

- *To enable exploratory research through collection of banked biospecimens, unless prohibited by local regulations or ethics committee decision.*

2.2. Study Objectives, Endpoints, and Estimands (part 2)

2.2.1. Primary Objectives

- *To confirm safety and tolerability and explore preliminary evidence of anti-tumor activity of PF-07062119 at the RP2D in patients with advanced colorectal adenocarcinomas (as a monotherapy and in combination therapy).*

2.2.2. Secondary Objectives

- *To evaluate the PK of PF-07062119 at the RP2D.*

- *To evaluate the immunogenicity of PF-07062119.*

- *To evaluate the immunogenicity of PF-06801591 and bevacizumab-Pfizer given in combination with PF-07062119.*

- *To evaluate immune PD effects of PF-07062119 between pre-treatment and post treatment tumor biopsies.*

- *To evaluate preliminary anti-tumor activity through time to event endpoints.*

2.2.3. Exploratory Objectives

- *To document any anti-tumor activity as assessed by immune related response criteria.*

- *To explore the relationship between target expression and biomarkers in archival tumor biopsies and/or pre and post treatment tumor biopsies, whole blood and serum samples in order to elucidate mechanism of action and response to therapy.*
- *To explore additional biomarkers related to clinical response, in order to elucidate the mechanism of action of PF-07062119, predict response to therapy and to understand resistance mechanisms that may predict escape from therapy.*

- *To evaluate T, B and NK subtypes for immunophenotyping in addition to T cell proliferation, activation and exhaustion markers in whole blood.*

- *To explore cytokine and chemokine PD markers in pre and on treatment serum samples.*

- *To explore the tumor biomarker carcinoembryonic antigen (CEA) levels.*

- *To enable exploratory research through collection of banked biospecimens, unless prohibited by local regulations or ethics committee decision.*

2.3. Study Design

This is a phase I, open-label, multi-center, multiple-dose, safety, PK and PD study of PF-07062119. This study contains two parts, single agent dose escalation (Part 1A) and dose finding in combination with anti-PD-1 and in combination with anti-VEGF (Part 1B) followed by dose expansion (part 2).

Part 1A will estimate the MTD and/or RP2D of PF-07062119 as a monotherapy. In Part 1A, sequential cohorts of patients with advanced gastrointestinal tumors, including colorectal, gastric and esophageal adenocarcinomas, who are not candidates for regimens known to provide clinical benefit will receive escalating doses of PF-07062119 administered as subcutaneous (SC) injection as guided by the Bayesian Logistic Regression Model (BLRM).With the determination that a dose level is safe following a DLT observation period of 28 days and based on the discussion by the safety review team, dose escalation will proceed to the next dose level. The SC route has the potential to reduce the C_{max} which is believed to be associated with CRS and inflammatory responses, a common adverse event for bispecific antibodies. The starting SC dose is 45 µg Q2W. Maximum dose increases from the previous dose levels will be up to 200%, which is a consistent approach with biologic compounds including an anti CD3 bispecific (Saber et al, 2017),37 but will be adjusted to no more than 100% following the observation of a DLT or if there are two Grade 2 or higher clinically significant treatment related AEs. Cohort size will be approximately 3 participants. The first cohort of 45 µg Q2W in Part 1A will have at least 1 DLT-evaluable participant. All other cohorts in Part 1A will have at least 2 DLT evaluable participants. Part 1A is estimated to enroll approximately 30 participants, but the actual number of participants enrolled will depend on the tolerability of PF-07062119 and the number of dose levels required to identify the MTD, but will include at least 6 participants treated at the recommended RP2D/MTD.

In Part 1B, the dose finding evaluation with PF-07062119 in combination with anti-PD-1 [PF-06801591] and in combination with anti-VEGF [bevacizumab-Pfizer]) will occur in patients with colorectal, gastric and esophageal adenocarcinomas.

Upon determining the single agent PF-07062119 RP2D/MTD, the Part 1B cohort will be initiated to explore safety, tolerability and preliminary anti-tumor activity of PF-07062119 in combination with in combination with PF-06801591and bevacizumab-Pfizer according to the following considerations:

- *BLRM will be used to guide the dose escalation/de-escalation process for each combination (PF-07062119/PF-06801591 and PF-07062119/ bevacizumab-Pfizer) separately. The determination of the safety for a given dose level for each combination will be based on the discussion by the safety review team.*

- *For the PF-07062119 and /PF-06801591combination, the starting dose of PF-07062119 in combination with PF-06801591 will be one dose level lower than the PF-07062119 monotherapy MTD in order to account for any potential increase in CRS that might be associated with the PF-07062119/PF-06801591 combination. If the PF-07062119 RP2D from Part 1A is lower than the MTD, the RP2D will be used in combination with PF-06801591. Study participants will be evaluated for the safety and tolerability of the combination and for appropriate PF-07062119 dose adjustment; the PF-07062119 dose used in the next cohort may remain the same, escalated, or de-escalated as guided by emerging clinical data and the BLRM design. Maximum increases of PF-07062119 from the previous level will be 100%. However, PF-07062119 will not be dose escalated beyond the MTD/RP2D for PF-07062119 as a single agent in Part 1A. The dose of PF-06801591 will remain unchanged at 300 mg Q4W SC throughout the safety assessment period. Additionally, a less frequent PF-07062119 dosing regimen (eg, Q4W) or a staggered dosing approach (eg, PF-07062119 to be initiated alone on C1D1 and PF-06801591 to be initiated after 1or 2 weeks) may be attempted as guided by the clinical data.*

  The cohort size at each dose level of Part 1B for the PF-07062119/PF-06801591 and PF-07062119/bevacizumab-Pfizer combination will each be approximately 3 participants; all cohorts in Part 1B will each have at least 3 DLT evaluable participants. Part 1B is estimated to enroll approximately 20 participants (10 for each combination), but the actual number of participants enrolled will depend on the tolerability of PF-07062119 in combination with PF-06801591 and with bevacizumab-Pfizer and the number of dose levels required to identify the corresponding PF-07062119 MTDs in the combination setting, but will include at least 6 participants treated at the corresponding recommended RP2D/MTDs. As the RP2D/MTD is approached or the RP2D/MTD of PF-07062119 has been determined, an optional subset of participants may be enrolled (approximately 6-12 patients) with the requirement of providing mandatory paired fresh pre-treatment and on-treatment biopsy samples in Parts 1A and 1B to enable evaluation of tissue biomarker PD activity. The option to enroll the subset of participants for mandatory paired tumor samples will be based on evaluation of emerging clinical data including available safety/tolerability, PK, and PD findings by the sponsor.

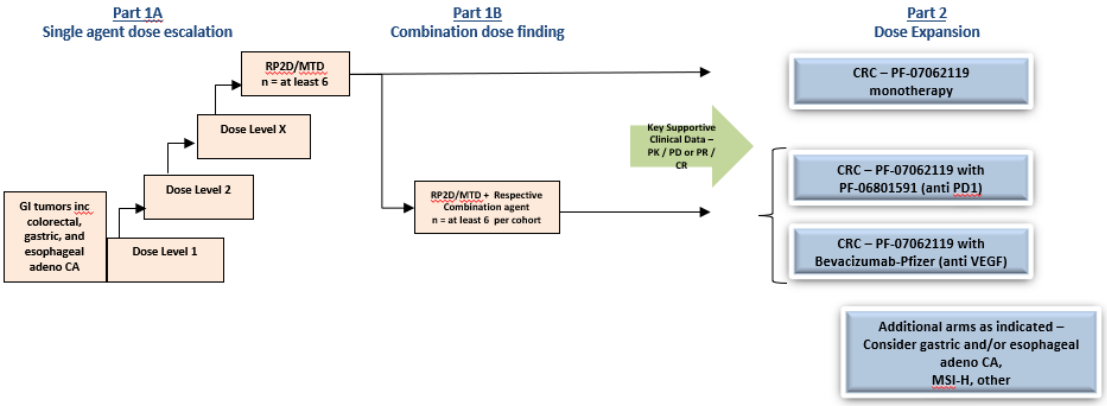
  In Part 2, the dose expansion phase will evaluate the RP2D of PF-07062119 as monotherapy and with a combination agent (anti-PD-1 [PF-06801591] and anti-VEGF [bevacizumab-Pfizer]) in patients with colorectal adenocarcinoma. In Part 2, approximately 60 patients will be enrolled; 20 patients in each monotherapy and combination cohorts. Responses in these initial dose expansion arms may lead to further evaluation of PF-07062119 in additional tumor indications and combination therapies.

Fresh pre-treatment biopsies in Part 2 will be required from all patients in all cohorts in order to establish relationship between target expression and efficacy observations. For a subset of patients in all cohorts (approximately 10 patients) mandatory fresh pre-treatment and on-treatment biopsy samples will be collected in order to confirm the mechanism of action and evaluate potential resistance mechanism during treatment.

***Number of Participants**: In the overall first -in-patient (FIP) study, up to approximately 110 patients are expected to be enrolled. Part 1 is estimated to enroll approximately 50 participants (30 participants in Part 1A and 10 patients in Part 1B for each of combination partners). In Part 2, approximately 20 patients will be enrolled in each monotherapy and combination cohorts.*

C3861001 Study Schema

Figure 1. Dose Escalation (BLRM design)



There will also be a minimum 72-hour interval between the first dose administered to each of the initial patients (ie, patients contributing to initial DLT evaluation) enrolled at a new dose level to allow the detection of possible infusion-site reactions or possible cytokine release syndrome. All patients will be observed in-patient for at least 48 hours after the first SC dose on Cycle 1 Day 1 (C1D1) for Part 1 and for at least 8 hours after the second SC dose on Cycle 1 Day 15 (C1D15). Patients should remain in house for observation for at least 1-hour post dose for all visits after C1D15 and may only be released after the investigator has confirmed the patient has not exhibited signs of cytokine reaction.

Treatment with study intervention will continue until either disease progression, patient refusal, or unacceptable toxicity occurs, whichever is earliest, unless the investigator and medical monitor agree to treatment beyond disease progression based on individual benefit/risk assessments.

The biomarker studies will be used to help understand the in vivo mechanism of action of the agent(s) studied as well as potential mechanisms of resistance. The studies may help in the future development of PF-07062119 as a single agent, or in combination with other

compounds, and may provide information on tumor sub-types that may respond to the Investigational Product.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS (PART 1)

3.1. Primary Endpoint(s)

- *First cycle Dose Limiting Toxicities (DLTs).*

  A participant is classified as dose-limiting toxicity (DLT)-evaluable if he/she experiences a DLT (irrespective of whether they received all of the planned doses of each investigational product and study intervention and scheduled safety assessments during the DLT window) or if he/she otherwise in the absence of a DLT receives all of the planned doses of each investigational product and study intervention and has received scheduled safety assessments during the DLT window. If a participant fails to meet these criteria, he/she may be replaced.

  Dose limiting toxicities (DLT) will be assessed through Cycle 1 (28-day cycle with Day 29 (eg, C2D1) laboratory assessments).

  For the purpose of dose escalation, the DLT observation period for Q2W SC regimen will be during the first cycle (28 days after the start of study treatment) in each participant.

  Significant adverse events considered to be related to the investigational product or treatment under investigation that occur after the DLT observation period will be reviewed in context of all safety data available. That review may result in re-evaluation of the dosing level or regimen.

  Severity of adverse events will be graded according to Common Terminology Criteria for Adverse Events (CTCAE) version 5.0.

  For the purpose of dose escalation, any of the following AEs with the associated parameters that occur in the first cycle of treatment or within 28 days after the start of the study treatment which are attributable to one, the other, or both agents in the combination will be classified as DLTs:

:

Hematological DLTs:

1. Grade 3 neutropenia must last >5 days.

2. Febrile neutropenia is a DLT defined as an absolute neutrophil count (ANC) <1.0 x 10^9/L with a single temperature of >38.3°C, or 101°F, or a sustained temperature of ≥38°C, or 100.4°F, for more than one hour.

3. Grade ≥3 neutropenia with infection is a DLT.

4. Grade 3 thrombocytopenia with Grade ≥2 (clinically significant) bleeding is a DLT.

5. Grade 4 thrombocytopenia is a DLT.

6. Anemia or thrombocytopenia requiring transfusion.

Non-hematological Exceptions:

1. Grade ≥3 fatigue lasting ≥7 days is a DLT.

2. For participants with liver, bone, or lung metastasis, an AST or ALT increase >8 x ULN or alkaline phosphatase >10 x ULN is a DLT.

3. Confirmed drug induced liver injury (DILI) meeting Hy's law criteria is a DLT.

4. Grade ≥3 vomiting or diarrhea lasting ≥3 days despite adequate treatment (eg, antiemetic or antidiarrheal medications, respectively) and other supportive care is a DLT.

5. Grade ≥3 CRS considered not to be due to an IRR, allergic reaction or anaphylaxis [unless i) not maximally treated or ii) events that are maximally treated, (eg, including the use of indicated immunosuppressive therapy and/or vasopressors), and resolve to ≤ Grade 2 within 72 hours] is a DLT.

6. Grade ≥3 QTcF prolongation irrespective of duration is a DLT.

7. Any death not clearly due to the underlying disease or extraneous causes is a DLT.

8. Clinically important or persistent toxicities (e.g., toxicities responsible for significant dose delay) that are not included in the above criteria may also be considered a DLT following review by the investigators and the Sponsor. All DLTs need to represent a clinically significant shift from baseline.

Otherwise, adverse events of CTCAE Grade greater than or equal to 3 will be considered to be DLTs, with the exception of the AE's listed below.

The following AEs will not be adjudicated as DLTs:

1. Isolated Grade 3 laboratory abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset.

2. Grade 3 amylase or lipase elevation not associated with clinical symptoms or clinical manifestations of pancreatitis.

3. Grade 3 infusion related reaction (IRR), allergic reaction or anaphylaxis will not be considered as DLTs but may be a reason for study discontinuation and will be reviewed with

the sponsor. The MTD is defined as a dose with probability of DLT from the target toxicity internal. The target interval for the DLT rate is defined as [0.16, 0.33].

- *Adverse Events as characterized by type, frequency, severity (as graded by National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE version [v.5.0], timing, seriousness, and relationship to study therapy. Detailed description is available in section 4.5.1.*

- *Laboratory abnormalities as characterized by type, frequency, severity (as graded by NCI CTCAE version [v.5.0]), and timing.*

3.2. Secondary Endpoint(s)

- *Pharmacokinetic parameters of PF-07062119:*

- *Cycle 1 and Cycle 4 PK parameters maximum concentration (C_{max}), time to achieve C_{max} (T_{max}), area under the concentration versus time curve from time zero to the last quantifiable concentration (AUC_{last}). If data permits, other PK parameters will be derived such as apparent clearance (CL/F), terminal half-life ($t^{1/2}$), and the area under the plasma concentration-time profile from time zero extrapolated to infinite time (AUC_{inf}).*

- *Pre-dose trough concentrations after multiple doses of PF-07062119.*

- *Incidence and titers of anti-drug antibodies (ADA) and neutralizing antibodies against PF-07062119.*

- *Incidence and titers of ADA and Nab against PF-06801591 and bevacizumab-Pfizer given in combination with PF-07062119 (Part 1B).*

- *Assessment of intra-tumor T cells (eg, CD8 IHC).*

- *Objective response, as assessed using the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1.*

3.3. Other Endpoint(s)

The exploratory endpoints are:

- *Progression free survival (PFS) and duration of response (DOR) by RECIST version 1.1.*

- *Overall Tumor assessment by the Immune Related Response Criteria in Solid Tumor (irRECIST) (criteria and immune-related progression free survival (irPFS).*

- *GUCY2c expression levels in archival and/or pre and on/post treatment tumor biopsies.*

- *Immunophenotyping of blood Immune cell subtypes frequency and activation.*

- *Intra-tumor immune pathway modulation, gene expression profiles, DNA mutations and immune cell biomarkers in pre- and post-dose tumor biopsies.*
- *Cytokine and chemokine PD markers in pre and on treatment serum samples.*
- *CEA levels in serum.*
- *Potential results from exploratory analysis of banked biospecimens (these results may or may not be generated in the context of the present study).*

Detailed methodology for summary and statistical analyses of the exploratory endpoints collected in this study will be further detailed in a biomarker statistical analysis plan (bSAP), which will be maintained by the sponsor.

4. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS (PART 2)

4.1. Primary Endpoint(s)

- *Adverse Events as characterized by type, frequency, severity (as graded by National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE version [v.5.0])), timing, seriousness, and relationship to study therapy.*
- *Laboratory abnormalities as characterized by type, frequency, severity (as graded by NCI CTCAE version [v.5.0]), and timing.*
- *Objective response rate (ORR) as determined by the Response Evaluation Criteria in Solid Tumor (RECIST) version 1.1 criteria.*

4.2. Secondary Endpoint(s)

- *Pharmacokinetic parameters and concentrations of PF-07062119.*
  - *Pre-dose trough concentrations after multiple doses of PF-07062119.*
- *Incidence and titers of ADA and Nab against PF-07062119.*
- *Incidence and titers of ADA and Nab against PF-06801591 and bevacizumab-Pfizer given in combination with PF-07062119.*
- *Assessment of increased intra-tumor T cells following PF-07062119 treatment (eg, CD8 IHC).*
- *PFS and DOR by RECIST version 1.1.*

4.3. Other Endpoint(s)

- *irRECIST and irPFS.*

- *GUCY2c expression levels in archival and/or pre- and on/post-treatment tumor biopsies.*
- *Immunophenotyping of blood Immune cell subtypes frequency and activation.*
- *Intra-tumor immune pathway modulation, gene expression profiles, DNA mutations and immune cell biomarkers in pre- and post-dose tumor biopsies.*
- *Cytokine and chemokine PD markers in pre- and on-treatment serum samples.*
- *Serum CEA levels.*
- *Potential results from exploratory analysis of banked biospecimens (these results may or may not be generated in the context of the present study).*

4.4. Baseline Variables

Baseline characteristics will be collected according to Schedule of Activities as specified in the protocol. No baseline variable will be used for stratification or as covariates in the statistical analysis. Unless otherwise specified, the baseline value is defined as the value collected at the time closest to, but prior to, the start of study drug administration in the first cycle.
Laboratory baseline will be the last predose measurement of before the first dose of any component of the study treatment.
Baseline for ECG is defined as Cycle 1 Day 1 predose. ECG measurements (an average of the triplicate measurements) will be used for the statistical analysis and all data presentations.

4.5. Safety Endpoints

4.5.1. Adverse Events

Severity of adverse events will be graded according to Common Terminology Criteria for Adverse Events (CTCAE) version 5.0.

For the purpose of dose escalation, any of the following AEs occurring in the first cycle of treatment or within 28 days after the start of the study treatment which are attributable to one, the other, or both agents in the combination will be classified as DLTs.

The definitions of an AE and an SAE can be found in Appendix 3 of the protocol.

AEs will be reported by the participant (or, when appropriate, by a caregiver, surrogate, or the participant's legally authorized representative).

The investigator and any qualified designees are responsible for detecting, documenting, and recording events that meet the definition of an AE or SAE and remain responsible to pursue and obtain adequate information both to determine the outcome and to assess whether it meets the criteria for classification as an SAE or that caused the participant to discontinue the study.

In addition, the investigator may be requested by Pfizer Safety to obtain specific follow up information in an expedited fashion.

4.5.2. Laboratory Data

Temperature (oral, tympanic, temporal or axillary), blood pressure (BP) and pulse rate will be assessed.

Further details of the laboratory tests can be found in Appendix 2 of the protocol.

5. ANALYSIS SETS (POPULATIONS FOR ANALYSIS)

Data for all participants will be assessed to determine if participants meet the criteria for inclusion in each analysis population prior releasing the database and classifications will be documented per standard operating procedures.

Table 1. Analysis Sets

Population	Description
Full analysis set	*The full analysis set includes all enrolled participants.*
Per protocol analysis set (Evaluable for MTD)	*The per protocol analysis set includes all enrolled participants who had at least one dose of study treatment and either experienced DLT or do not have major treatment deviations during the DLT observation period.*
Safety set	*The safety analysis set includes all enrolled participants who receive at least one dose of study treatment. Unless otherwise specified the safety analysis set will be the default analysis set used for all analyses.*
Modified Intent to Treat (mITT)	*The modified intent to treat (mITT) is the analysis population that will follow the ITT principle and include participants receiving at least 1 dose of study medication with baseline assessment and at least 1 post baseline assessment, disease progression, or death before the first tumor assessment. The mITT population may be used for interim analysis and conference presentations when the study is still ongoing.*
PK analysis	*The PK parameter analysis population is defined as all enrolled participants treated who do not have protocol deviations influencing PK assessment, and have sufficient information to estimate at least 1 of the PK parameters of interest.* *The PK concentration population is defined as all enrolled participants who are treated and have at least 1 analyte*

Population	Description
	concentration.
Response Evaluable	*The response evaluable population will include all participants who received at least one dose of study treatment and had baseline disease and at least one post baseline disease assessment.*
PD/Biomarker	*The PD/Biomarker analysis population is defined as all enrolled participants with at least 1 of the PD/Biomarkers evaluated at pre and/or post dose.*
Immunogenicity analysis set	*The immunogenicity analysis set includes all enrolled participants who receive at least one dose of study treatment and have at least one sample tested for ADA.*

6. GENERAL METHODOLOGY AND CONVENTIONS

6.1. Hypotheses and Decision Rules

There will be no formal hypothesis testing in this study.

6.2. General Methods

The data are summarized by cohort defined by the initial dose of the study drug. DLT rates at the study dose levels will be presented via mean and medians and a Bayesian credible interval based on the posterior density from the full probability model and will be used for the dose escalation decision meetings.

6.2.1. Analyses for Continuous Data

Continuous data will be summarized with the mean, median, minimum, maximum, coefficient of variation and standard deviation. Missing values will be excluded from the analysis.

6.2.2. Analyses for Categorical Data

Categorical data will be summarized by number of unique patient incidence. Missing data will be excluded from the analysis.

6.2.3. Analyses for Binary endpoints

Binary data will be summarized using number of unique patient incidence, and Wilson's confidence interval for binomial proportions will be presented if warranted.

Binary endpoints in this study include ORR, complete response (CR), partial response (PR) based on RECIST 1.1. Descriptive statistics along with the corresponding 2-sided 95% confidence intervals using an exact method will be provided for these endpoints if the sample size per cohort permits.

6.2.4. Analyses for time-to-event data

The time-to-event data will be presented for individual patient and by cohort when applicable.

6.3. Methods to Manage Missing Data

For the analysis of safety endpoints, the sponsor data standard rules for imputation will be applied.

6.3.1. Missing Dates

In compliance with Pfizer standards, if the day of the month is missing for any date used in a calculation, the 1st of the month will be used to replace the missing date unless the calculation results in a negative time duration (eg, date of onset cannot be prior to day one date). In this case, the date resulting in 0 time duration will be used. Pfizer standards are also used if both month and day are missing (Jan 1 unless negative time duration). This excludes the pharmacokinetic, ECG, and pharmacodynamic analyses, which will only use the actual date collected or if date not available deem the data missing.

6.3.2. Efficacy Analysis

Response Evaluable Set will be used for all response related analyses including ORR, DOR, and PFS.

Tumor response will be presented in the form of participant data listings that include, but are not limited to tumor type, dose on Day 1, tumor response at each visit, and best overall response.

Part 1A and Part 1B: Progression date, death date, date of first response and last tumor assessment date and date of last contact will be listed.

Part 2: The Kaplan-Meier methods will be used to analyze all time to event endpoints. Median PFS (if reached).

PFS is defined the time from start date to date of first documentation of progression, or death due to any cause. Progression is defined as the appearance of local, regional or distant disease of the same type after complete response or progression of pre-existing lesions. It does not include second primary malignancies of unrelated types.

For the time-to-event endpoints, the missing data handling method will be censoring. Censoring rules for time-to-event endpoints are detailed in Appendix 2.

6.3.3. Pharmacokinetics

6.3.3.1. Concentrations Below the Limit of Quantification

In all data presentations (except listings), concentrations below the limit of quantification (BLQ) will be set to zero. (In listings BLQ values will be reported as "<LLQ", where LLQ will be replaced with the value for the lower limit of quantification.)

6.3.3.2. Deviations, Missing Concentrations and Anomalous Values

In summary tables and plots of median profiles, statistics will be calculated having set concentrations to missing if 1 of the following cases is true:

1. A concentration has been collected as ND (ie not done) or NS (ie no sample),

2. A deviation in sampling time is of sufficient concern or a concentration has been flagged anomalous by the pharmacokineticist.

Note that summary statistics will not be presented at a particular time point if more than 50% of the data are missing.

6.3.4. Pharmacokinetic Parameters

Actual PK sampling times will be used in the derivation of PK parameters.

If a PK parameter cannot be derived from a subject's concentration data, the parameter will be coded as NC (ie not calculated). (Note that NC values will not be generated beyond the day that a subject discontinues.)

In summary tables, statistics will be calculated by setting NC values to missing; and statistics will be presented for a particular dose with ≥3 evaluable measurements.

If an individual subject has a known biased estimate of a PK parameter (due for example to an unexpected event such as vomiting before all the compound is adequately absorbed in the body), this will be footnoted in summary tables and will not be included in the calculation of summary statistics or statistical analyses.

6.3.5. QTc

For the QTc analyses, no values will be imputed for missing data.

7. ANALYSES AND SUMMARIES

7.1. Primary Endpoint(s)

7.1.1. Dose Limiting Toxicities (DLTs)

- Analysis set: Per protocol analysis set. Dose limiting toxicities (DLT) will be assessed during Cycle 1 (the first 28 days) to inform dose escalation and determine the MTD.

- Analysis methodology:

Statistical Methods:

The dose escalation in the Part 1A and Part 1B of the study will be guided by a Bayesian analysis of dose limiting toxicity (DLT) data for PF-07062119. Toxicity is modelled using two-parameter logistic regression for the probability of a participant experiencing a DLT at the given dose.

Assessment of participant risk (the three dosing intervals boundaries can be modified):
After each cohort of participants, the posterior distribution for the risk of DLT for new participants at different doses of interest for PF-07062119 will be evaluated. The posterior distributions will be summarized to provide the posterior probability that the risk of DLT lies within the following intervals:

Under-dosing:	*[0, 0.16]*
Targeted dosing:	*[0.16, 0.33]*
Overdosing:	*[0.33, 1]*

The escalation with overdose control (EWOC) principle:
Dosing decisions are guided by the escalation with overdose control principal (Rogatko 2007). A dose may only be used for newly enrolled participants if the risk of excessive toxicity at that dose is less than 25%.

Prior distributions:

Weakly informative prior distributions based on pre-clinical/expert opinion information will be chosen for the logistic parameters for prior distribution in Part 1A, see Appendix 10.10.

A meta-analytic-predictive (MAP) approach will be used to derive the prior distribution for model parameters used in Part 1B based on the data collected in Part 1A. The MAP prior for the logistic model parameters for this study is the conditional distribution of the parameters given the historical data (see Spiegelhalter 2004, Neuenschwander 2010, Neuenschwander 2014). MAP priors are derived from hierarchical models, which take into account possible differences between the studies. A full description of the application of the MAP approach to derive the prior distributions of the model parameters is given in Appendix 10.9.

Starting dose:
The starting dose is 45 µg. For this dose the prior risk of overdosing is 9.7%, which satisfies the EWOC criterion. A full assessment of the prior risk to participants is given in Appendix 10.9.

The maximum number of participants in dose escalation part of the trial was set to 50. The trial will be stopped when the following criteria are met:

- *At least 6 participants have been treated at the recommended MTD/RP2D,* d'.
- *The dose* d' *satisfies one of the following conditions:*
- *The probability of target toxicity at dose* d' *exceeds 50%,*

  i.e. Pr(0.16 $\leq \pi_d <$ *0.33)*$\geq$*50%.*
- *A minimum of 15 participants have been treated in the trial.*

Intercurrent events and missing data: Cohort size will be approximately 3 participants, with at least 1 DLT-evaluable participant per cohort in the first dose level of Part 1A and at least 2 DLT evaluable participants per cohort in the remaining cohorts of Part 1A and at least 3 DLT-evaluable participants per cohort in Part 1B. The actual number of patients enrolled will depend on the tolerability of PF-07062119 and the number of dose levels required to identify the MTD. Missing values will not be imputed.

7.1.2. Adverse Events

- Analysis set: Safety analysis set.
- Analysis methodology: characterized by type, frequency, severity (as graded by National Cancer Institute Common Terminology Criteria for Adverse Events [NCI CTCAE version 5.0]) timing, seriousness, and relationship to study therapy. Further description is given in section 7.4.1. The treatment emergent adverse events will be defined according to Pfizer's standard definition.
- Intercurrent events and missing data: intermediate missing values will not be imputed.

7.1.3. Laboratory abnormalities

- Analysis set: safety analysis set.
- Analysis methodology: Presented as tables and characterized by type, frequency, severity (as graded by NCI CTCAE version 5.0), and timing. The number and percentage of participants who experienced laboratory test abnormalities will be summarized according to worst toxicity grade observed for each laboratory assay. The analyses will summarize laboratory tests both on the entire study period and by cycle (Cycle 1 and Cycles beyond 1). For laboratory tests without CTCAE grade definitions, results will be categorized as normal, abnormal, or not done. Shift tables or figures may be created for select laboratory abnormalities. Results meeting the Hy's Law as defined in the protocol may also be presented as tables or listings.
- Intercurrent events and missing data: intermediate missing values will not be imputed.

7.2. Secondary Endpoint(s)

7.2.1. Pharmacokinetic Analysis

The concentrations of PF-07062119 will be summarized by descriptive statistics (n, mean, standard deviation, coefficient of variation, median, minimum maximum, and geometric mean) by dosing cohort, cycle, and nominal time. Individual patient and median profiles of the concentration-time data will be plotted by dosing cohort and cycle using nominal times. Median profiles will be presented on both linear-linear and log-linear scales.

The individual concentration-time data of PF-07062119 following the Cycle 1 Day 1 dose and Cycle 4 Day 1 dose will be analyzed separately using non-compartmental analysis to estimate the PK parameters. The PK parameters estimated will include C_{max}, time to maximum concentration (T_{max}), and concentration versus time curve (AUC_{last}). If data permit or if considered appropriate, other PK parameters including terminal elimination half-life ($t_{1/2}$), clearance (CL or CL/F), and volume of distribution at steady state (V_{ss} or $V_{ss/F}$) may be determined. Actual sample collection times will be used for the parameter calculations.

The concentrations of PF-06801591 or bevacizumab-Pfizer from Part 1B and Part 2 will be summarized by descriptive statistics (n, mean, standard deviation, coefficient of variation, median, minimum maximum, and geometric mean) by dosing cohort, cycle, and nominal time.

Pharmacokinetic/Pharmacodynamic (PK/PD) Correlation

PK and data from this study may be analyzed using population modeling approaches and may also be pooled with data from other future studies to investigate any association between PF-07062119 exposure and biomarkers or selected safety and/or efficacy endpoints. The results of these analyses, if performed, may be reported separately from the clinical study report.

7.2.2. *Pharmacodynamics Analysis*

For biopsy samples, summary statistics (eg, the mean and standard deviation, median, and minimum/maximum levels of continuous, and frequencies and percentages of categorical biomarker measures) will be determined at baseline and post-treatment.

7.2.3. *Immunogenicity Data Analysis*

For the immunogenicity data, the percentage of patients with positive ADA will be summarized. Listings and summary tabulations of the ADA data at baseline and post-randomization will be generated. Samples may also be analyzed for the presence of neutralizing antibodies (NAb), and any data will be similarly summarized. For patients with positive ADA or NAb, the magnitude (titer), time of onset, and duration of ADA or NAb response will also be described, if data permit. The potential impact of immunogenicity on PK and clinical response including pharmacodynamic markers, safety/tolerability and efficacy will be explored, if warranted by the data.

7.2.4. Tumor Response

- Overall Response Rate (ORR).

- ORR as assessed using RECIST version 1.1. ORR is defined as the proportion of patients who achieved completed response (CR) or partial response (PR) per RECIST 1.1.

- Duration of Response (DOR).

- Duration of response is defined as the time from start date (which is the date of first documentation of PR or CR) to date of first documentation of objective progression or death. DOR is only applicable to those patients with an objective response.

- Analysis set: Response Evaluable.

- Intercurrent events and missing data: Data after study drug discontinuation and rescue will be excluded; intermediate missing values will not be imputed.

7.3. Baseline and Other Summaries and Analyses

7.3.1. Baseline Summaries

Baseline characteristics such as demographics, prior medication, medical history, ECOG performance status, and primary diagnosis will be tabulated and listed. For ECOG performance status a shift table (worst post-baseline vs baseline) may be produced. The Safety Analysis Set will be used.

7.3.2. Study Conduct and Participant Disposition

An accounting of the study patients will be tabulated. The subject evaluation groups will be listed. The Full Analysis Set will be used.

Subject discontinuation from treatment and study will be tabulated and listed separately with their reason for discontinuation. The Safety Analysis Set will be used.

7.3.3. Study Treatment Exposure

The safety analysis set will be used.

Dose modifications are described in the protocol. The following will be summarized by subject for overall and each dose level:

- Number of subjects per dose level;

- Median and range of number of cycles started per subject;

- Number (%) of subjects starting a cycle (1, 2, 3…);

- Number (%) of dose interruptions (include both known and unknown dates);

- Number (%) of subjects with dose reductions;

- Number (%) of each reason (drug related AE vs AE vs. Other) dose interruptions and dose reductions;
- Time on treatment (median, range).

The following will be summarized by cycle received for overall and each dose level:

- Total number of cycles started;
- Number of cycles started per subject (median, range);
- Number of cycles before 1st reduction (median, range);
- Number of cycles before 1st interruption (median, range).

The following will be summarized for cumulative dose by dose level and cycle:

- Summary statistics (mean, median, standard deviation and range) of cumulative dose and cumulative percent of administered dose (compared to planned dose) by cycle and overall.

Listings by subject (ordered by dose level): start date and stop date of each dosing period within each cycle (including records with 0 mg), administered total daily dose for each period, any missed doses with unknown dates (Y/N), number of missed doses with unknown dates, reason for any dosing changes.

Listings by subject and each cycle (ordered by dose level): cycle length, total planned dose, administered total dose, percentage of planned dose, dose reduction (yes/no), and dose interruption (yes/no).

7.3.4. Concomitant Medications and Nondrug Treatments

Prior, concomitant, and further therapies (drug and non-drug treatments) will be coded by the World Health Organization (WHO) medical dictionary. Listings of prior, concomitant, and further therapies will be provided separately.

7.4. Safety Summaries and Analyses

All safety analyses will be performed on the safety population.

Summaries and analyses of safety parameters will include all participants in the safety analysis set.

AEs, ECGs, BP, PR, continuous cardiac monitoring, and safety laboratory data will be reviewed and summarized on an ongoing basis during the study to evaluate the safety of participants. Any clinical laboratory, ECG, BP, and PR abnormalities of potential clinical concern will be described. Safety data will be presented in tabular and/or graphical format and summarized descriptively, where appropriate.

Medical history and physical examination and neurological examination information, as applicable, collected during the course of the study will be considered source data and will not be required to be reported, unless otherwise noted. However, any untoward findings identified on physical/or neurological examinations examinations conducted during the active collection period will be captured as AEs, if those findings meet the definition of an AE. Data collected at screening that are used for inclusion/exclusion criteria, such as laboratory data, ECGs, and vital signs, will be considered source data, and will not be required to be reported, unless otherwise noted. Demographic data collected at screening will be reported.

7.4.1. Adverse Events

AEs will graded by the investigator according to the CTCAE version 5.0 and coded using the Medical Dictionary for Regulatory Activities (MedDRA). Adverse event data will be reported in tables and listings. Summaries of adverse event by mapped terms, appropriate thesaurus level, toxicity grade, and seriousness and relationship to study treatment will be presented, as well as summaries of adverse events leading to death and premature withdrawal from study treatment. The number and percentage of participants who experienced any AE, SAE, treatment related AE, and treatment related SAE will be summarized according to worst toxicity grades. The summaries will present AEs both on the entire study period and by cycle (Cycle 1 and Cycles beyond 1). Listings of DLTs and deaths will be provided.

7.4.2. Laboratory Data

Safety laboratory tests will be performed as described in the protocol.

To determine if there are any clinically significant laboratory abnormalities, the hematological, clinical chemistry (serum) and urinalysis safety tests will be assessed against the criteria specified in the sponsor reporting standards. The assessment will take into account whether each subject's baseline test result is within or outside the laboratory reference range for the particular laboratory parameter.

The data will be summarized with respect to demographic and baseline characteristics, efficacy observations and measurements, safety observations and measurements, pharmacokinetic and biomarker measurements.

7.4.3. Vital Signs

Vital signs will be measured with the participant in a semi-supine position after 5 minutes of rest and will include temperature, systolic and diastolic blood pressure, and pulse rate.

8. ELECTROCARDIOGRAMS

The analysis of ECG results will be based on participants in the safety analysis set with baseline and on-treatment ECG data. Baseline is defined as a Cycle 1 Day 1 pre-dose.

Standard 12-Lead ECGs utilizing limb leads (with a 10-second rhythm strip) should be collected at times specified in the SoA section of this protocol using an ECG machine that automatically calculates the heart rate and measures PR, QT, and QTc intervals and QRS

complex. Alternative lead placement methodology using torso leads (e.g. Mason-Likar) is not recommended given the potential risk of discrepancies with ECGs acquired using standard limb lead placement. All scheduled ECGs should be performed after the participant has rested quietly for at least 10 minutes in a supine position.

At each time point, 3 consecutive ECGs will be performed at approximately [1-4] minutes apart to determine the mean QTc interval. To ensure safety of the participants, a qualified individual at the investigator site will make comparisons to baseline measurements. Additional ECG monitoring will occur if a) the mean value from the triplicate measurements for any postdose QTc interval is increased by ≥60 msec from the baseline ***and*** *is >450 msec; or b) an absolute QTc value is ≥500 msec for any scheduled ECG. If either of these conditions occurs, then a single ECG measurement must be repeated at least hourly until QTc values from 2 successive ECGs fall below the threshold value that triggered the repeat measurement. In addition, if verified QTc values continue to exceed the criteria above, immediate correction for reversible causes including electrolyte abnormalities, hypoxia and concomitant medications for drugs with the potential to prolong the QTcF interval should be performed.*

If the QTc interval reverts to less than the threshold criteria listed above, and in the judgment of the investigator(s) and sponsor, it is determined that the cause(s) of QTc prolongation is something other than investigational product, treatment may be continued with regular ECG monitoring. If in that timeframe the QTcF intervals rise above the threshold values, the investigational product will be held until the QTc interval decreases to below the threshold values. Participants will then restart the investigational product at the next lowest dose level. If the QTc interval has still not decreased to 501 msec after 2 weeks, or if at any time a participant has a QTcF interval >515 msec or becomes symptomatic, the participant will be removed from the study. Additional triplicate ECGs may be performed as clinically indicated.

In some cases, it may be appropriate to repeat abnormal ECGs to rule out improper lead placement as contributing to the ECG abnormality. It is important that leads be placed in the same positions each time in order to achieve precise ECG recordings. If a machine-read QTc value is prolonged, as defined above, repeat measurements may not be necessary if a qualified medical provider's interpretation determines that the QTc values are in the acceptable range.

If a participant experiences a cardiac or neurologic AE (specifically syncope, dizziness, seizures, or stroke), an ECG (triplicate) should be obtained at the time of the event.

For each patient and by treatment, the maximum change from baseline will be calculated as well as the maximum post-baseline interval across time points. Categorical analysis will be conducted for the maximum change from baseline in corrected QT and the maximum post-baseline QT interval.

Shift tables will be provided for baseline vs worst on treatment corrected QT (one or more correction methods will be used) using maximum absolute QTc results. Shift tables will also be provided for ECG abnormality at baseline vs. on treatment (yes, no, not done: (n, %).

If more than 1 ECG is collected at a nominal time after dose administration (for example, triplicate ECGs), the mean of the replicate measurements will be used to represent a single observation at that time point. If any of the 3 individual ECG tracings has a QTc value >500 msec, but the mean of the triplicates is not >500 msec, the data from the participant's individual tracing will be described in a safety section of the clinical study report (CSR) in order to place the >500-msec value in appropriate clinical context. However, values from individual tracings within triplicate measurements that are >500 msec will not be included in the categorical analysis unless the average from the triplicate measurements is also >500 msec. Changes from baseline will be defined as the change between the postdose QTc value and the average of the time-matched baseline triplicate values on Day -1, or the average of the predose triplicate values on Day 1.

In addition, an attempt will be made to explore and characterize the relationship between plasma concentration and QT interval length using a PK/PD modeling approach. If a PK/PD relationship is found, the impact of participant factors (covariates) on the relationship will be examined.

ECG measurements (an average of the triplicate measurements) will be used for the statistical analysis and all data presentations. Any data obtained from ECGs repeated for safety reasons after the nominal time-points will not be averaged along with the preceding triplicates. Interval measurements from repeated ECGs will be included in the outlier analysis (categorical analysis) as individual values obtained at unscheduled time points.

QT intervals will be corrected for heart rate (HR) (QTc) using standard correction factors (ie, Fridericia's (default correction), Bazett's, and possibly a study-specific factor, as appropriate). Data will be summarized and listed for QT, HR, response rate (RR), partial response (PR), QRS, QTcF (and other correction factors, eg, QTcB as appropriate) by study arm and presented by time and dose. Individual QT (all evaluated corrections) intervals will be listed by study, time and dose. The most appropriate correction factor will be selected and used for the following analyses of central tendency and outliers and used for the study conclusions. Descriptive statistics (n, mean, median, standard deviation, minimum, and maximum) will be used to summarize the absolute corrected QT interval and changes from baseline in corrected QT after treatment by dose and time point.

Changes from baseline for the ECG parameters QT interval, heart rate, QTc interval, PR interval, and QRS complex will be summarized by treatment and time.

The number (%) of participants with maximum postdose QTc values and maximum increases from baseline in the following categories will be tabulated by treatment:

Table 2. Safety QTc Assessment

Degree of Prolongation	***Mild (msec)***	***Moderate (msec)***	***Severe (msec)***
Absolute value	*>450-480*	*>480-500*	*>500*
Increase from baseline		*30-60*	*>60*

In addition, the number of participants with uncorrected QT values >500 msec will be summarized.

Patients experiencing clinically-relevant morphological ECG changes will be summarized (including frequency and percentage).

Categorical data analysis will follow Appendix 3.

8.1.1. Physical Examination

Participants will have a physical examination to include weight, vital signs, assessment of ECOG performance status and height; height will be measured at screening only.

A complete physical examination will include, at a minimum, assessments of the cardiovascular, respiratory, gastrointestinal, and neurological systems.

A brief physical examination will include, at a minimum, assessments of the skin, lungs, cardiovascular system, and abdomen (liver and spleen).

Investigators should pay special attention to clinical signs related to previous serious illnesses.

9. INTERIM ANALYSES

No formal interim analysis will be conducted for this study. As this is an open-label study, the sponsor may conduct unblinded reviews of the data during the course of the study for the purpose of safety assessment, facilitating dose-escalation decisions, facilitating PK/PD modeling, and/or supporting clinical development.

10. REFERENCES

11. APPENDICES

Appendix 1. Summary of Efficacy Analyses

Endpoint	Analysis Type	Population	Data Inclusion and Rules for Handling Intercurrent Events and Missing Data	Analysis Model
Tumor response: Objective response rate (ORR) and duration of response (DoR), as assessed using the Response Evaluation Criteria in Solid Tumors (RECIST Appendix 4) version 1.1, PFS	Summary	Response Evaluable	Separately for observed data.	N/A

Appendix 2. Time to Event Data Analysis Censoring Rules

Table 3. Progression Free Survival and Duration of Response

Situation	Date of Progression/Censoring[1]	Outcome
Inadequate baseline assessment	First dosing date in Cycle 1	Censored
No on-study assessments	First dosing date in Cycle 1	Censored
Alive, on treatment2 and no Progression	Date of last objective tumor assessment	Censored
Progression Documented on or between scheduled tumor assessments prior to treatment discontinuation2	Date of first objective tumor assessment showing objective progression	Progressed (Event)
Treatment discontinuation for undocumented progression	Date of last objective tumor assessment prior to discontinuation[2]	Censored
Treatment discontinuation due to toxicity or other reason	Date of last objective tumor assessment prior to discontinuation[2]	Censored
Death prior to first planned tumor assessment	Date of death	Death (Event)
Death without objective progression prior to treatment discontinuation[2]	Date of death	Death (Event)
Death or progression after 2 or more missed tumor assessments	Date of last objective tumor assessment prior to the event	Censored

1. For date of censorship, if a tumor assessment takes place over a number of days (eg, superficial lesions one day, scans another), the last date is used as the assessment date.
2. or within 28 days of discontinuation of treatment.

Table 4. Time to Progression

Situation	Date of Progression/Censoring1	Outcome
Inadequate baseline assessment	First dosing date in Cycle 1	Censored
No on-study assessments	First dosing date in Cycle 1	Censored
Alive, on treatment2 and no Progression	Date of last objective tumor assessment	Censored
Progression Documented on or between scheduled tumor assessments prior to treatment discontinuation2	Date of first objective tumor assessment showing objective progression	Progressed (Event)
Treatment discontinuation for undocumented progression	Date of last objective tumor assessment prior to discontinuation2	Censored
Treatment discontinuation due	Date of last objective tumor	Censored

to toxicity or other reason	assessment prior to discontinuation	
New anticancer treatment <28 days after discontinuation of treatment without progression	Date of last objective tumor assessment prior to new anticancer treatment	Censored
Death prior to first planned tumor assessment	Start date (C1D1)	Censored
Death without objective progression prior to treatment discontinuation2	Date of last objective tumor assessment prior to death	Censored
Progression after 2 or more missed tumor assessments	Date of last objective tumor assessment prior to the event	Censored

3. For censoring date, if a tumor assessment takes place over a number of days (eg, superficial lesions one day, scans another), the last date is used as the assessment date.
4. or within 28 days of discontinuation of treatment.

DOSD and DOR

Censoring rules for DOSD and DOR will be the same as for PFS.

Appendix 3. Categorical Classes for ECG and Vital Signs

Categories for QTcB and QTcF

QTcB/QTcF (ms)	max. ≤450	450<max. ≤480	480<max.≤500	max. >500
QTcB/QTcF (ms) increase from baseline	max. <30	30≤max. <60	max. ≥60	

Categories for PR and QRS

PR (ms)	max ≥300	
PR (ms) increase from baseline	Baseline >200 and max. ≥25% increase	Baseline ≤200 and max. ≥50% increase
QRS (ms)	Max. ≥200	
QRS (ms) increase from baseline	Baseline >100 and max. ≥25% increase	Baseline ≤100 and max. ≥50% increase

Categories for Vital Signs

Systolic BP (mm Hg)	min. <90	
Systolic BP (mm Hg) change from baseline	max. decrease ≥30	max. increase ≥30
Diastolic BP (mm Hg)	min. <50	
Diastolic BP (mm Hg) change from baseline	max. decrease ≥20	max. increase ≥20
Supine pulse rate (bpm)	min. <40	max. >120

Measurements that fulfil these criteria are to be listed in the study report.

Appendix 4. RECIST 1.1 Tumor Assessment Criteria

Adapted from E.A. Eisenhauer, P. Therasseb, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, J. Verweij: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). European Journal of Cancer 45 (2009) 228–247.

At baseline, individual tumor lesions will be categorized by the investigator as either measurable or not, according to the criteria summarized below:

Measurable Lesions

Lesions that can be accurately measured in at least one dimension.

- Lesions with longest diameter twice the slice thickness and at least 10 mm or greater when assessed by CT or MRI (slice thickness 5-8 mm).
- Lesions with longest diameter at least 20 mm when assessed by Chest X-ray.
- Superficial lesions with longest diameter 10 mm or greater when assessed by caliper.
- Malignant lymph nodes with the short axis 15 mm or greater when assessed by CT.

NOTE: The shortest axis is used as the diameter for malignant lymph nodes, longest axis for all other measurable lesions

Non-measurable Lesions

- Non-measurable disease includes lesions too small to be considered measurable

(including nodes with short axis between 10 and 14.9 mm) and truly non-measurable disease such as pleural or pericardial effusions, ascites, inflammatory breast disease, leptomeningeal disease, lymphangitic involvement of skin or lung, clinical lesions that cannot be accurately measured with calipers, abdominal masses identified by physical exam that are not measurable by reproducible imaging techniques.

- Bone disease: Bone disease is non-measurable with the exception of soft tissue components that can be evaluated by CT or MRI and meet the definition of measurability at baseline.

- Previous local treatment: A previously irradiated lesion (or lesion patientive to other local treatment) is non-measurable unless it has progressed since completion of treatment.

Normal sites:

- Cystic lesions: Simple cysts should not be considered as malignant lesions and should not be recorded either as target or non-target disease. Cystic lesions thought to represent cystic metastases can be measurable lesions, if they meet the specific definition above. If non-cystic lesions are also present, these are preferred as target lesions.

- Normal nodes: Nodes with short axis <10 mm are considered normal and should not be recorded or followed either as measurable or non-measurable disease.

Recording Tumor Assessments

All sites of disease must be assessed at baseline. Baseline assessments should be done asclose as possible prior to study start. For an adequate baseline assessment, all required scans must be done within 28 days prior to treatment and all disease must be documented appropriately. If baseline assessment is inadequate, subsequent statuses generally should beindeterminate.

Note: For the patient population being evaluated in this protocol, the baseline assessment may be completed within 6 weeks prior to randomization.

Target Lesions

Response in target lesions is defined as follows:

- **Complete Response (CR):** Complete disappearance of all target lesions with theexception of nodal disease. All target nodes must decrease to normal size (short axis<10 mm). All target lesions must be assessed.

- **Partial Response (PR):** Greater than or equal to 30% decrease under baseline of the sum of diameters of all target measurable lesions. The short diameter is used in the sum for target nodes, while the longest diameter is used in the sum for all other target lesions. All target lesions must be assessed.

- **Stable Disease (SD):** Does not qualify for CR, PR or Progression. All target lesions must be assessed. Stable can follow PR only in the rare case that the sum increases by less than 20% from the nadir, but enough that a previously documented 30% decrease no longer holds.

- **Objective Progression (PD):** 20% increase in the sum of diameters of target measurable lesions above the smallest sum observed (over baseline if no decrease in the sum is observed during therapy), with a minimum absolute increase of 5 mm.

- **Indeterminate**: Progression has not been documented, and one or more target measurable lesions have not been assessed; or assessment methods used were inconsistent with those used at baseline; or one or more target lesions cannot be measured accurately (eg, poorly visible unless due to being too small to measure); or one or more target lesions were excised or irradiated and have not reappeared or increased.

Non-Target Lesions

While some non-target lesions may actually be measurable, they need not be measured and instead should be assessed only qualitatively at the time points specified in the protocol.

Response in non-target lesions is defined as follows:

- **Complete Response (CR):** Disappearance of all non-target lesions and normalization of tumor marker level. All lymph nodes must be non-pathological in size (<10 mm short axis).

- **Non-CR/Non-PD**: Persistence of any non-target lesions and/or tumor marker level above the normal limits.

- **PD:** Unequivocal progression of pre-existing lesions. Generally, the overall tumor burden must increase sufficiently to merit discontinuation of therapy. In the presence of SD or PR in target disease, progression due to unequivocal increase in non-target disease should be rare.

- **Indeterminate**: Progression has not been determined and one or more non-target sites were not assessed or assessment methods were inconsistent with those used at baseline.

New Lesions

The appearance of any new unequivocal malignant lesion indicates PD. If a new lesion isequivocal, for example due to its small size, continued assessment will clarify the etiology. If repeat assessments confirm the lesion, then progression should be recorded on the date of the initial assessment. A lesion identified in an area not previously scanned will be considered a new lesion.

Supplemental Investigations

- If CR determination depends on a residual lesion that decreased in size but did not disappear completely, it is recommended the residual lesion be investigated with biopsy or fine needle aspirate. If no disease is identified, objective status is CR.

- If progression determination depends on a lesion with an increase possibly due tonecrosis, the lesion may be investigated with biopsy or fine needle aspirate to clarify status.

Subjective Progression

Patients requiring discontinuation of treatment without objective evidence of disease progression should not be reported as PD on tumor assessment CRFs. This should be indicated on the end of treatment CRF as off treatment due to Global Deterioration of Health Status. Every effort should be made to document objective progression even afterdiscontinuation of treatment.

Determination of Overall Response by the RECIST 1.1 Criteria

When both target and non-target lesions are present, individual assessments will be recorded separately. The overall assessment of response will involve all parameters as depicted in the following table.

Table 5. Response Evaluation Criteria in Solid Tumors by RECIST 1.1

Target lesions	Non-target lesions	New Lesions	Overall response
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
CR	Indeterminate or Missing	No	PR
PR	Non-CR/Non-PD, Indeterminate or Missing	No	PR
SD	Non-CR/Non-PD, Indeterminate or Missing	No	SD
Indeterminate or Missing	Non-PD	No	Indeterminate
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD
CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease			

Table 6. Objective Response Status at Each Evaluation for Patients with Non Target Disease Only

Non-target Disease	New Lesions	Objective Status
CR	No	CR
Non-CR/Non-PD	No	Non-CR/Non-PD
Indeterminate	No	Indeterminate
Unequivocal progression	Yes or No	PD

Any	Yes	PD
CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease		

Appendix 5. Immune-Related Response Criteria Derived from RECIST 1.1 (irRECIST)

Increasing clinical experience indicates that traditional response criteria may not be sufficient to fully characterize activity in this new era of targeted therapies and/or biologics.

This is particularly true for immunotherapeutic agents such as anti-CTLA4 and anti-PD-1/anti-PD-L1 antibodies which exert the antitumor activity by augmenting activation and proliferation of T cells, thus leading to tumor infiltration by T cells and tumor regression rather than direct cytotoxic effects.18,19 Clinical observations of patients with advanced melanoma treated with ipilimumab, for example, suggested that conventional response assessment criteria such as Response Evaluation Criteria in Solid Tumors (RECIST) and WHO criteria are not sufficient to fully characterize patterns of tumor response to immunotherapy because tumors treated with immunotherapeutic agents may show additional response patterns that are not described in these conventional criteria.

Furthermore, the conventional tumor assessment criteria (RECIST and WHO criteria) havebeen reported as not capturing the existence of a subset of patients who have an OS similar to those who have experienced CR or PR but were flagged as PD by WHO criteria.

On these grounds, a tumor assessment system has been developed that incorporates thesedelayed or flare-type responses into the RECIST v1.1 (irRECIST).

For irRECIST, only target and measurable lesions are taken into account. In contrast to RECIST v1.1, irRECIST:

- Requires confirmation of both progression and response by imaging at least 4 weeks from the date first documented, and

- Does not necessarily score the appearance of new lesions as progressive disease if thesum of lesion diameters of target lesions (minimum of 10 mm longest diameter per non-nodal lesion and 15 mm shortest diameter per nodal lesion, maximum of 5 targetlesions, maximum of 2 per organ) and measurable new lesions does not increase by ≥20%.

The same method of assessment and the same technique should be used to characterize each identified and reported target lesion(s) at baseline and throughout the trial. irRECIST is defined as follows:

1. Overall immune-related complete response (irCR): Complete disappearance of all lesions (whether measurable or not) and no new lesions. All measurable lymph nodes also must have a reduction in short axis to <10 mm.

2. Overall immune-related partial response (irPR): Sum of the diameters (longest for non-nodal lesions, shortest for nodal lesions) of target and new measurable lesions decreases ≥30%.

3. Overall immune-related stable disease (irSD): Sum of the diameters (longest for non-nodal lesions, shortest for nodal lesions) of target and new measurable lesions is neither irCR, irPR, (compared to baseline) or immune-related progressive disease (irPD, compared to nadir).

4. Overall immune-related progressive disease (irPD): Sum of the diameters (longest for non-nodal lesions, shortest for nodal lesions) of target and new measurable lesions increases ≥20% (compared to nadir), confirmed by a repeat, consecutive observation at least 4 weeks from the date first documented.

New measurable lesions: Incorporated into tumor burden (ie, added to the target lesion measurements). A lymph node has to be ≥15 mm in short axis to be a measurable new lesion and its short axis measurement is included in the sum. Up to 2 new lesions per organ and up to 5 new lesions in total can be added to the measurements.

New non-measurable lesions: Do not define progression but preclude irCR.

Table 7. Overall Response Derived from Changes in Index, Non-index and New Lesions

Index and New Measurable Lesions[1]	Non-Index Lesions	Measurable Lesions	Overall response using irRECIST[2]
Decrease 100%	CR	No	CR
Decrease 100%	Non-CR/Non-PD	No	PR
Decrease 100%	Indeterminate or Missing	No	PR
Decrease ≥30%	Non-CR/Non-PD, Indeterminate or Missing	No	PR
Decrease ≥30%	Non-CR/Non-PD, Indeterminate or Missing	No	SD
Decrease <30% and increase <20%	Non-PD	No	Indeterminate
Decrease <30% and increase <20%	Any	Yes or No	PD
Increase ≥20%	PD	Yes or No	PD

[1] Decrease assessed relative to baseline.

[2] Response (irCR and irPR) and progression (irPD) must be confirmed by a second consecutive assessment at least 4 weeks apart.