Protocol B8011001

A Phase 1, Open-Label, Dose Escalation and Expansion Study of PF-06801591 in Patients with Locally Advanced or Metastatic Melanoma, Squamous Cell Head and Neck Cancer, Ovarian Carcinoma, Sarcoma, Non-Small Cell Lung Cancer, Urothelial Carcinoma or other Solid Tumors

Statistical Analysis
Plan (SAP)

Version: Amendment 2

Author: PPD (Early Clinical Development Statistics, Cambridge, MA, USA)

Date: 24-AUG-2020

[redacted]

TABLE OF CONTENTS

[redacted]

[redacted]

LIST OF TABLES

LIST OF FIGURES

[redacted]

1. VERSION HISTORY

Amendment 2	Section	Update
	5.5	A new section is added to address the analysis strategies on data that is impacted by the COVID-19 pandemic.
	Appendix 9.2	Removed the PFS window. The statistical programming standard for this study is switching from Pfizer Data Standard (PDS) to CaPS (CDISC and Pfizer Standard), to be consistent with the late stage studies of the program (B8011006 and B8011007) that are being conducted at the time of this SAP amendment. In CaPS, the PFS window is no longer required. Table 9.2.1 and 9.2.2 are updated to be consistent with CaPS requirements.
Amendment 1	Section	Update
	All	Texts taken directly from the protocol are made *italicized*
	2.1	Objectives are updated and reformatted per protocol amendment 3
	2.2	Study design is updated per protocol amendment 3
	3	Endpoints are updated per protocol amendment 3
	4	Analysis sets are updated per protocol amendment 3
	5.2.2.	Statistical decision rules are updated for Part 2 per protocol amendment 3
	5.3	General statistical methods are updated with additional details for programming personnel
	6	Details about unconfirmed and confirmed tumor response and progression for RECIST 1.1 and irRECIST added.
	6.1	Cycle delay is updated; Cycle skip is added
	6.3.1	Editorial changes were made to the efficacy endpoints analyses
	7	Informal data monitoring is added
	9.6	All lymphoma response assessment is removed per protocol amendment 3

[redacted]

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study B8011001. 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. This SAP amendment is based on the protocol amendment 3 dated June 3, 2017.

Note: in this document any text taken directly from the protocol is *italicized*.

2.1. Study Objectives

Part 1 (Dose-Escalation) Primary Objective

- To assess safety and tolerability of increasing dose levels of PF-06801591 in patients with locally advanced or metastatic melanoma, SCCHN, ovarian carcinoma, sarcoma, NSCLC, urothelial carcinoma or other solid tumor types with clinical evidence of response to anti PD-1 or PD-L1 agents to establish the MTD.

- To assess safety and tolerability of PF-06801591 administered SC in patients with locally advanced or metastatic melanoma, SCCHN, ovarian cancer, sarcoma, NSCLC, urothelial carcinoma or other solid tumor types with clinical evidence of response to anti PD-1 or PD-L1 agents.

Part 1 (Dose-Escalation) Secondary Objective

- To characterize the single-dose and multiple-dose PK of PF-06801591 following IV or SC administration.

- To evaluate the immunogenicity of PF-06801591 following repeated administration.

- To characterize PD-1 receptor occupancy (RO) in peripheral blood T cells following IV or SC PF-06801591 administration at each dose level.

- To evaluate preliminary anti-tumor activity of PF-06801591.

CCI

[redacted]

[redacted]

CCI

[redacted]

Part 2 (Dose-Expansion) Primary Objective

- *To further characterize the safety and tolerability of PF-06801591 following SC administration in NSCLC and urothelial carcinoma.*
- *To estimate clinical efficacy by overall response rate (ORR) of PF-06801591 following SC administration in NSCLC and urothelial carcinoma.*

Part 2 (Dose-Expansion) Secondary Objective

- To further evaluate preliminary anti-tumor activity of PF-06801591 following SC administration.
- *To evaluate overall survival (OS).*
- To collect PF-06801591 drug concentration data in patients following SC administration for evaluation of population PK.
- To evaluate the immunogenicity of PF-06801591 following repeated SC administration.

CCI

[redacted]

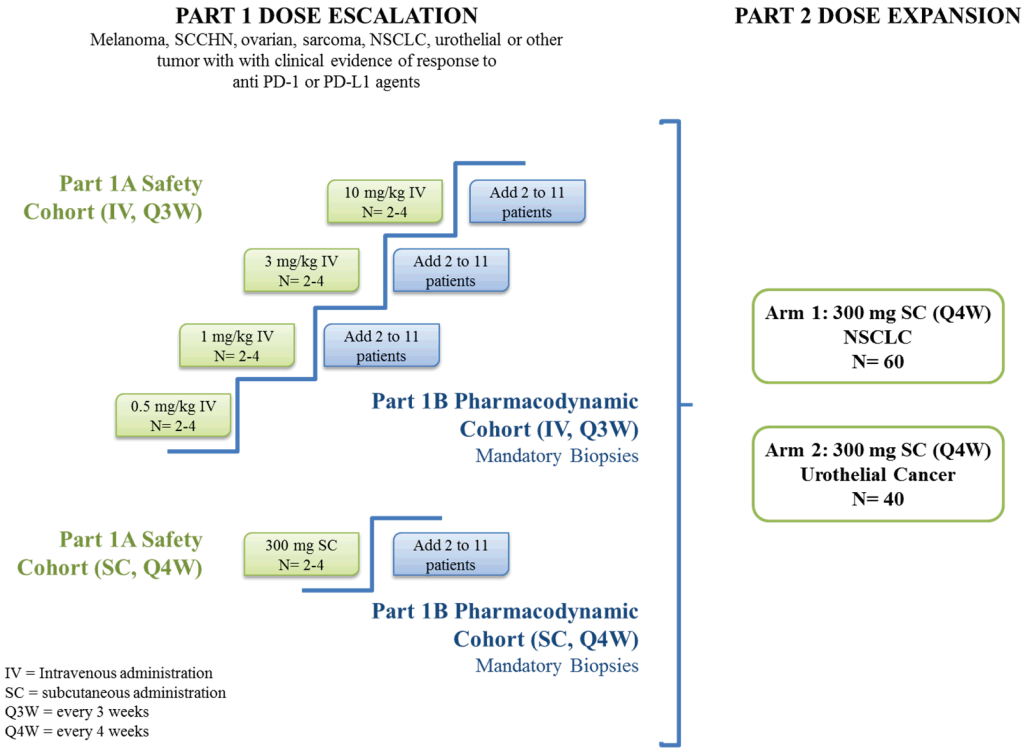
[redacted]

CCI

2.2. Study Design

This is a Phase 1, open-label, multi-center, multiple-dose, dose escalation and expansion, safety, PK, and PD study of PF-06801591 in previously treated patients with locally advanced or metastatic solid tumor types with clinical evidence of response to anti PD-1 or PD-L1 agents. This clinical trial will include 2 parts: Part 1 dose escalation and Part 2 dose expansion (Figure 1). A total of approximately 140 patients will be enrolled into this study.

Figure 1 B8011001 Study Schematic



Part 1 will enroll patients with locally advanced or metastatic melanoma, SCCHN, ovarian cancer, sarcoma, NSCLC, urothelial carcinoma, or other solid tumor types with clinical evidence of response to anti PD-1 or PD-L1 agents. Patients will receive 0.5, 1, 3, or 10 mg/kg PF-06801591 intravenously (IV) every 3 weeks (q3w), or 300 mg PF-06801591 subcutaneously every 4 weeks (q4w). The Part 1 dose escalation phase will enroll approximately 8 to 15 patients per dose level.

[redacted]

Part 1 will be further divided into Part 1A (safety cohort) CCI [redacted]. For both IV and SC administration portions, each safety cohort will enroll 2-4 patients per dose level. CCI [redacted]

All patients in Part 2 will receive 300 mg PF-06801591 SC q4w. Part 2 dose expansion will include 2 arms: Arm 1 will enroll approximately 60 patients with NSCLC who progressed on or were intolerant to systemic therapy or for whom systemic therapy was refused or unavailable but have not previously received anti-PD-1 or anti-PD-L1. Arm 2 will enroll approximately 40 patients with urothelial carcinoma who progressed on or were intolerant to systemic therapy or for whom systemic therapy was refused or unavailable but have not previously received anti-PD-1 or anti-PD-L1. Approximately 100 patients will be enrolled into Part 2.

2.2.1. MTD Determination (Part 1)

Part 1 will study sequential cohorts (0.5, 1, 3, and 10 mg/kg, or 300 mg SC) of PF-06801591 in adult patients with locally advanced or metastatic melanoma, SCCHN, ovarian cancer, sarcoma, NSCLC, urothelial carcinoma, or other solid tumor types with clinical evidence of response to anti PD-1 or PD-L1 agents who are unresponsive to currently available therapies or for whom no standard therapy is available. An mTPI method, targeting a dose limiting toxicity (DLT) rate of 27.5% will be utilized for dose escalation. *In the IV administration portion, if based on Part 1A information the dose level is deemed safe and well- tolerated, an additional 2 to 5 patients will enroll into the same dose level in Part 1B. In the SC administration portion, if based on Part 1A information the 300 mg dose level is deemed safe and well- tolerated, up to an additional 11 patients (approximately) will be enrolled into Part 1B. Up to approximately 9 patients may be enrolled into each dose level in the IV administration portion, and up to approximately 15 patients may be enrolled into the 300 mg SC administration cohort (Part 1A and Part 1B combined). The mTPI approach would be applied across Parts 1A and 1B to ensure that administered doses do not surpass the toxicity boundaries. Safety data from all patients in Parts 1A (safety cohort)* CCI [redacted] *will be used to determine the MTD.* The dose finding decision will be based on 1-cycle (21-day) DLT observation for patients enrolled into IV administration portion and 1-cycle (28-day) DLT observation for patients enrolled into subcutaneous dosing portion. A staggered start will be employed at all dose levels. A single patient will be dosed and observed for 48 hours. If no safety concerns arise during this 48-hour period, a second patient will be enrolled into the same dose level cohort.

Late immune-related DLTs are irAEs that meet the same grading criteria as DLT criteria but occur after the initial 21-day DLT period for IV administration, or 28 days for SC administration, and during the 120-day assessment period. Late immune-related DLTs will be added to the mTPI approach to reassess the dose-finding decisions.

Safety information, provided by additional patients enrolled into each dose level expansion will also be taken into account for MTD determination. If the DLT rate is estimated to reach >33% or more at any dose level, enrollment at that level and all higher levels will be temporarily stopped, and safety data will be analyzed. The decision to move forward with enrollment will follow the

[redacted]

same DLT target as described previously. Dose escalation will continue until an MTD has been established or a prespecified maximum dose level has been reached.

The mTPI method relies upon a statistical probability algorithm, calculated using all patients treated in the current dose level to determine one of the following dose-finding decisions: the subsequent dose should be escalated, maintained at the current dose, or de-escalated in the next cohort of 2 to 4 patients, or the trial should be terminated (see Table 1).

In principle, all patients must be evaluated for a minimum period of 21 days (q3w dosing interval) or 28 days for the subcutaneous cohort. If a patient withdraws from the study before Day 21 (or 28 days for the subcutaneous cohort) for reasons other than drug- related toxicity, another patient may be enrolled to replace that patient in the current cohort. However, if a patient discontinues close to Day 21 (or 28 days for the subcutaneous cohort) for reasons other than toxicity and due to an evident nondrug-related event, the patient may be deemed evaluable for safety if safety assessments have been unremarkable and the investigator and sponsor's medical monitor both agree that the patient is evaluable for DLT safety observation.

Table 1. Decision Rules

Number of Patients having DLT	Number of Patients Treated at a Dose Level													
	n=2	n=3	n=4	n=5	n=6	n=7	n=8	n=9	n=10	n=11	n=12	n=13	n=14	n=15
0	E	E	E	E	E	E	E	E	E	E	E	E	E	E
1	S	S	S	E	E	E	E	E	E	E	E	E	E	E
2	U	D	S	S	S	S	S	S	S	E	E	E	E	E
3		U	U	D	D	S	S	S	S	S	S	S	S	S
4			U	U	U	U	D	D	D	S	S	S	S	S
5				U	U	U	U	U	D	D	D	D	D	S
6					U	U	U	U	U	U	U	D	D	D
7						U	U	U	U	U	U	U	U	U

Actions to be taken:
D = De-escalate the dose; E: Escalate the dose; S: Stay at the dose.
U = Unacceptable toxicity.

The dose escalation Part 1 of the study will stop if any of the following criteria is met:

1. The maximum sample size has been achieved (approximately 40 patients in total);
2. 6 to 15 patients have been enrolled at a dose level that is predicted to be the MTD per the mTPI method;
3. All dose levels explored appear to be overly toxic, and the MTD cannot be determined;
4. All candidate dose levels have been tested and deemed safe.

[redacted]

2.2.2. Part 2 Dose Expansion

Part 2 dose expansion will include 2 arms: Arm 1 will enroll approximately 60 patients with NSCLC who progressed on or were intolerant to systemic therapy or for whom systemic therapy was refused or unavailable but have not previously received anti-PD-1 or anti-PD-L1. Arm 2 will enroll approximately 40 patients with urothelial carcinoma who progressed on or were intolerant to systemic therapy or for whom systemic therapy was refused or unavailable but have not previously received anti-PD-1 or anti-PD-L1. All patients in Part 2 will receive 300 mg PF-06801591 SC q4w. The 300 mg SC dose level in Part 2 was selected based upon safety, PK, PD, and preliminary anti-tumor activity observed in Part 1, as well as the maximum injection volume considered feasible with the current formulation.

After the first 30 NSCLC patients have been enrolled and have either completed their first tumor assessment (at approximately 8 weeks post treatment), or have discontinued from the study before their first scheduled tumor assessment, preliminary assessment of safety and efficacy data maybe completed. This preliminary data may provide guidance in the decision to the increase or decrease the sample size of each arm, to add other tumor types in a future amendment, or to initiate additional clinical studies. The specific details for this data look are described in Section 7.

3. ENDPOINTS AND COVARIATES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

Part 1:

- *DLTs at escalated doses of PF-06801591.*The specific definition of DLT is provided in the study protocol.

- Adverse Events (AEs) as characterized by type, frequency, severity (as graded by National Cancer Institute Common Terminology Criteria for Adverse Events [NCI CTCAE] version 4.03), timing, seriousness, and relationship to study therapy PF-06801591.

- Laboratory abnormalities as characterized by type, frequency, severity (as graded by NCI CTCAE version 4.03), and timing.

Part 2:

- AEs as characterized by type, frequency, severity (as graded by NCI CTCAE version 4.03), timing, seriousness and relationship to study therapy PF-06801591 *administered by SC administration*.

- Laboratory abnormalities as characterized by type, frequency, severity (as graded by NCI CTCAE version 4.03) and timing.

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

3.2. Secondary Endpoint(s)

Part 1:

- PK parameters of PF-06801591: Cycle 1 and Cycle 4 C_{max}, area under the concentration versus time curve (AUC) from time zero to the last quantifiable time point prior to the next dose (AUC_{last}), and if data permit, CL, V_d, volume of distribution at steady state (Vss), accumulation ratio (Rac) when feasible, and terminal elimination $t_{1/2}$.

- Incidence of anti-drug antibodies (ADA) and neutralizing antibodies (NAb) against PF-06801591.

- PD-1 receptor occupancy (RO) by PF-06801591, as assessed by measuring the levels of unbound (free) cell surface PD-1 on circulating T cells over time following PF-06801591 administration.

- Objective tumor response, as assessed using the Response Evaluation Criteria in Solid Tumor (RECIST) version 1.1 and immune-related RECIST (irRECIST); and proportion of subjects with PR (and irPR, as appropriate).

- Time-to-event endpoints based on RECIST and irRECIST and PFS (and irPFS as appropriate), Duration of stable disease (DOSD, or irDOSD as appropriate), and Duration of response (DOR, or irDOR as appropriate).

    - PFS is defined as the time from treatment start date to date of first documentation of progression or death due to any cause.

    - Duration of stable disease is defined from the start of the treatment until the criteria for progression are met, taking as reference the smallest tumor measurements recorded since the treatment started, including the baseline measurements. This endpoint is applicable to the subset of patients who achieved a best overall response of stable disease (SD), those patients whose best overall response is not SD will be excluded from this endpoint analysis. A minimum of 6 weeks (with a - 5 days window, which is essentially a minimum of 35 days) interval of two assessments is required for this endpoint.

    - 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.

[redacted]

Part 2:

- Time to event endpoints by PF-06801591 administered by SC based on RECIST and irRECIST, including time to response (TTR) and time to progression (TTP) as well as PFS (and irPFS as appropriate), DOSD (and irDOSD as appropriate), and DOR (and irDOR as appropriate). Time to response is defined as the time from the study treatment date to the first documentation of PR or CR. Time to progression is the time from start date to date of first documentation of objective progression.

- *Median time to death, proportion of patients alive at 6 months, 1 year, and 2 years.* Time to death (i.e. overall survival) is the time from treatment start date to date of death due to any cause.

- Trough PF-06801591 concentrations for selected cycles.

- Incidence of ADA and NAb against PF-06801591 administered by SC.

CCI

[redacted]

Detailed descriptions of Endpoints are provided in Appendice 9.1, 9.4, and 9.5.

[redacted]

CCI

[redacted]

4. ANALYSIS SETS

4.1. Full Analysis Set

The full analysis set includes all enrolled patients. This is equivalent to the ITT (intent-to- treat) population.

4.2. Safety Analysis Set

The safety analysis set includes all enrolled patients who receive at least one dose of study medication.

4.3. 'PER PROTOCOL' Analysis Set

The per protocol analysis set includes all enrolled patients who received at least one dose of study treatment and who did not have major treatment deviations during the Cycle 1. For the IV administration of PF-06801591, patients with major treatment deviations in the 21-day DLT observation period are not evaluable for the MTD assessment and will be replaced as needed to permit MTD estimation.

For the SC administration of PF-06801591, patients with major treatment deviations in the 28-day observation period will not be evaluable for DLT.

[redacted]

Major treatment deviations include:

- Administration of <50% of the planned dose of PF-06801591, provided that the reduction is not due to toxicity attributable to PF-06801591.
- Administration of >150% of the planned dose of PF-06801591.

4.4. PK Analysis Set

4.4.1. PK Concentration Set

The PK concentration population is defined as all patients who receive PF-06801591, have no protocol deviations affecting the PK assessment, and have at least 1 post-dose concentration measurement.

4.4.2. PK Parameter Set

The PK parameter analysis population is defined as all enrolled patients treated who have sufficient information to estimate at least 1 of the PK parameters of interest.

4.5. PD Biomarker Analysis Set

The biomarker analysis set includes all enrolled patients with at least one of the pharmacodynamic/biomarker parameters evaluated at pre- and/or post-dose.

4.6. Modified Intent-to-Treat Set

The modified intent-to-treat (mITT) population is defined as all the randomized subjects who have received at least 1 dose of study medication, have measurable disease baseline assessment (within 28 days prior to study entry) and at least 1 post baseline assessment or disease progression, global deterioration of health status, or death. The mITT population will be used for anti-tumor assessment.

4.7. Immunogenicity Analysis Set

The immunogenicity analysis set is defined as patients who receive at least 1 dose of study treatment and have at least 1 ADA or NAb sample collected.

4.8. Treatment Misallocations

Subjects who receive the wrong initial dose for whatever reason will be analyzed according to the initial dose actually received. Subjects who receive the wrong dose after the initial dose will be analyzed according to the initial dose received.

4.9. Protocol Deviations

The determination of protocol deviations (PDs) and important protocol deviations (IPDs) will follow Pfizer standard operating procedures. A full list of PDs, IPDs, and IPDs that are excluded from Per-protocol analysis will be determined prior to the database release and be included in the CSR.

[redacted]

5. GENERAL METHODOLOGY AND CONVENTIONS

This is an open-label dose escalation study and no interim analysis or blinding is planned for this study.

5.1. Statistical Hypotheses

There are no statistical hypotheses.

5.2. Statistical Decision Rules

5.2.1. Part 1 (MTD Finding)

Part 1 dose escalation phase of this study employs an mTPI design to estimate the MTD. The mTPI design employs a simple beta-binomial model with prior a conjugated prior beta (0.5, 0.5). Decision rules are based on calculating the unit probability mass (UPM) of 3 intervals corresponding to underdosing, proper dosing, and overdosing in terms of dose limiting toxicity. A proper dosing interval is centered at the target toxicity rate (pT) of 27.5% with 5% uncertainty (0.225 < pT <0.325). The under dosing interval is (0, 0.225), and the overdosing interval is (0.325, 1). The 3 dosing intervals are associated with 3 different dose escalation decisions. The underdosing interval corresponds to a dose escalation (E), overdosing corresponds to a dose de-escalation, and proper-dosing corresponds to staying at the same current dose. Given an interval and a probability distribution, the UPM of that interval is defined as the probability of the interval divided by the length of the interval. The mTPI design calculates the UPMs for the 3 dosing intervals, and the one with the largest UPM implies the corresponding dose-finding decision. That decision provides the dose level to be used for future patients. For example, if the underdosing interval has the largest UPM, decision E (to escalate) will be executed, and the next cohort of patients will be treated at the next-higher dose level. Under the mTPI design, a trial is terminated when either the lowest dose is above the MTD or a pre-specified maximum sample size for Part 1 (approximately 40) is reached.

The following table shows the probability of escalating to the next dose level for a range of underlying true DLT rates. For example, for a cohort size of n=3 and for a DLT that occurs in 10% of patients, there is a greater than 90% probability of escalating. Conversely, for a DLT that occurs with a rate of 70%, the probability of escalating is 3%. It is assumed that dose escalation occurs with either 0/3 or 1/6 patients with DLTs.

Probability of Escalating Dose									
True underlying DLT rate	10%	20%	30%	40%	50%	60%	70%	80%	90%
Probability of escalating	0.91	0.71	0.49	0.31	0.17	0.08	0.03	0.009	0.001

5.2.2. Part 2 (Dose Expansion)

Part 2 (N=100) of this study is intended to further characterize the safety, efficacy, PK, PD, and immunogenicity profiles of 300 mg of SC PF-06801591 in anti-PD-1 or anti-PD-L1 treatment naïve patients with NSCLC (N=60) and urothelial carcinoma (N=40) who have progressed on or were intolerant to systemic therapy or for whom systemic therapy was refused or unavailable.

[redacted]

Summary statistics will be provided for trough PF-06801591 concentrations, safety endpoints, immunogenicity, pharmacodynamic/biomarkers, and efficacy data.

When the first 30 NSCLC patients have been enrolled and have either completed their first tumor assessment (at approximately 8 weeks post treatment), or have discontinued from the study before their first scheduled tumor assessment, preliminary assessment of safety and efficacy data may be completed. This preliminary data may provide guidance in the decision to increase or decrease the size of each arm, to add other tumor types in a future amendment, or to initiate additional clinical studies.

5.2.3. Sample Size Determination

The exact sample size for the dose escalation design in Part 1 cannot be specified in advance due to the dynamic features of mTPI. For the first dose level (0.5 mg/kg IV), in the absence of actual DLTs, a total of 2 patients will be enrolled: 2 patients in Part 1A and no patient in Part 1B. For subsequent dose levels it is anticipated that approximately 2 to 4 patients will be enrolled in Part 1A and approximately 2 to 5 patients in Part 1B with available pre- and on-treatment biopsies for IHC testing and up to 9 patients at dose levels 1, 3, and 10 mg/kg administered by IV in Part 1. A single cohort with SC administration at a dose level of 300 mg of PF-06801591 will be opened for enrollment shortly after the cohort at a dose level of 3 mg/kg IV in Part 1A will have completed enrollment and the 21-day DLT observation period. This cohort at a dose level of 300 mg SC will seek 2 to 4 patients in Part 1A, up to 11 patients in Part 1B with available pre- and on-treatment biopsies for IHC testing, and up to 15 patients in Part 1. The actual Part 1 sample size may be smaller, depending on the underlying dose toxicity profile, and the number of dose levels studied.

Part 2 of the study will enroll approximately 60 patients with NSCLC who progressed on or were intolerant to systemic therapy or for whom systemic therapy was refused or unavailable and are treatment naïve for anti-PD-1 or anti-PD-L1. NSCLC was selected for historical comparability to prior large anti-PD-1 studies and dosing of 300 mg SC in Part 2 has been selected based on safety, PK, available biomarker data and preliminary anti-tumor activity from Part 1. There is no hypothesis testing in Part 2. Estimation approach is used to characterize the precision of response data. The estimation of ORR using N=60 is described as follows. Suppose that the ORR estimate is 19% in Part 2 with N=60, then the 80% and 90% confidence intervals of the true ORR will be (13.3, 26.3%) and (12.1, 28.6%) respectively. Note that an ORR of 19% was observed in a clinical trial for nivolumab in previously treated NSCLC patients. In Part 2, approximately 40 urothelial cancer patients will also be enrolled, the sample size is based on clinical considerations of expanding the safety database.

5.3. General Methods

Whilst every effort has been made to pre-specify all analyses in this statistical analysis plan, if any additional exploratory analyses be found to be necessary, the analyses and the reasons for them will be detailed in the clinical study report (CSR).

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. All data will be categorized

based on the scheduled visit at which it was collected. These visit designators are predefined values that appear as part of the visit tab in the eCRF.

5.3.1. Analyses for Time to Event Data

Time to event endpoints in this study include progression-free survival (PFS), duration of response (DOR), duration of stable disease (DoSD), time to response (TTR), time to progression (TTP), and overall survival (OS). When appropriate and data permits, the immune-response version of these endpoints based on irRECIST will also be explored. The specific definitions of these endpoints are provided in Section 3. These endpoints will be summarized using the Kaplan-Meier method[2] and displayed graphically when appropriate. Median event times and 2-sided 95% confidence intervals for each time-to-event endpoint (Brookmeyer and Crowley, 1982)[3] will be provided.

5.3.2. Analyses for Binary Data

Binary endpoints in this study include ORR, complete response (CR), partial response (PR) based on RECIST 1.1. When appropriate and data permits, the immune-response version of these endpoints based on irRECIST will also be explored. If deemed necessary disease control rate (i.e. DCR, which is defined as the proportion of patients that achieved CR, or PR, or stable disease) may also be calculated for Part 2. Descriptive statistics along with the corresponding 2-sided 95% confidence intervals using an exact method will be provided for these endpoints.

5.3.3. Analyses for Continuous Data

Descriptive statistics, such as the mean, standard deviation, coefficient of variation, median, minimum, and maximum values, will be provided for continuous endpoints.

5.4. Methods to Manage Missing Data

5.4.1. 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.

5.4.2. Efficacy Analysis

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 9.2.

[redacted]

5.4.3. Pharmacokinetics

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 (i.e. lower limit of quantification) will be replaced with the value for the LLQ).

Deviations, missing concentrations and anomalous values

Patients who experience events that may affect their PK (eg, incomplete dosing) may be excluded from the PK analysis.

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

1. A concentration has been reported 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.

An anomalous concentration value is one that, after verification of bioanalytical validity, is grossly inconsistent with other concentration data from the same individual or from other subjects. For example, a BLQ concentration that is between quantifiable values from the same dose is considered as anomalous. Anomalous concentration values may be excluded from PK analysis at the discretion of the PK analyst.

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 not be presented for a particular treatment group if more than 50% of the data are NC. For statistical analyses, PK parameters coded as NC will also be set to missing.

If an individual subject has a known biased estimate of a PK parameter, this will be footnoted in summary tables and will not be included in the calculation of summary statistics or statistical analyses.

[redacted]

CCI

[redacted]

5.5. Statistical Considerations of COVID-19 Impacted Data

In March 2020, prior to the database release for this study, the World Health Organization (WHO) announced a global pandemic of the virus SARS-CoV-2 and the resulting disease COVID-19. During the conduct of this trial, if any participant's data is impacted by this pandemic, the following considerations will be given in the data analyses:

a. Death caused by COVID-19 is still considered as an "event" in the analysis of PFS (Parts 1 and 2) and OS (Part 2). If deemed necessary, a sensitivity analysis may be performed where COVID-19 driven death is censored at the death date.

b. If a scheduled tumor radiographic scan is *delayed* out of the Schedule of Activity allowable window, or is *missing* (i.e. participant skipped a scheduled tumor radiographic scan) due to any reasons related to the pandemic, this delay or missingness does not alter the censoring rules for PFS or TTP. A censoring reason of "COVID-19" may be added to the PFS or TTP summary if the specific reason of tumor scan delay or missing can be attributed to COVID-19. If deemed necessary, a sensitivity analysis may be performed where participants would be censored on the date of COVID-19 diagnosis.

   In the confirmed ORR analysis, as described in Section 6, if a response can't be confirmed by a subsequent tumor scan because of the pandemic (i.e. the subsequent tumor scan wasn't performed), then the initial response will be considered as unconfirmed. This is a conservative approach. No sensitivity analysis will be performed.

c. Any COVID-19 related symptoms are to captured as adverse events in the case report form. Those adverse events will be summarized in the same manner as other adverse events. If a label of COVID-19 can be identified in the investigator provided adverse event term, then a separate AE listing may be provided for just the COVID-19 related events.

d. If identifiable, the COVID-19 related data points, including missing data where the reason of missing is identified as COVID-19 related (site closure hence data could not captured; participants skipped a visit because of concern over the pandemic), protocol deviations driven by COVID-19, safety events caused by COVID-19 may be separately listed.

6. ANALYSES AND SUMMARIES

For efficacy related endpoints in this study, as both RECIST 1.1 and irRECIST will be used for tumor assessments, and there are different requirements in these criteria in terms of confirming tumor response or progression, the following analysis plan will be implemented:

[redacted]

Endpoint	Part 1		Part 2	
	RECIST 1.1	irRECIST	RECIST 1.1	irRECIST
ORR (or irORR)	Unconfirmed	Both	Both	Both
PFS (or irPFS)	Unconfirmed	Both	Unconfirmed*	Both
DOR (or irDOR)	Unconfirmed	Both	Both	Both
DOSD (or irDOSD)	Unconfirmed	Both	Both	Both
TTP	Unconfirmed	Both	Unconfirmed*	Both
TTR	Unconfirmed	Both	Both	Both

Note: "Both" in the above table represents both unconfirmed and confirmed tumor assessments.
*: Confirmation for progressed disease is not applicable in RECIST 1.1

In the "unconfirmed" analyses (i.e. tumor response or progression without confirmation required), all tumor assessments data will be included for analyses. Specifically, regardless if a patient's tumor response or progression is subsequently confirmed or not, the patient data will all be included in the "unconfirmed" analyses. This is a more comprehensive analysis where tumor confirmation is not required and not taken into account. For example in the ORR analysis, if a patient achieved CR or PR, regardless it is subsequently confirmed or not, the patient will be included in the numerator.

In the "confirmed" analyses (i.e. tumor response or progression with confirmation required), only those tumor assessments data that's subsequently confirmed by a consecutive tumor scan will be considered as a "success" or "event" and be included for analyses as appropriate. For example, in the ORR analysis, a patient will only be counted in the numerator (the "success") if the tumor response (CR or PR) is subsequently confirmed. If a patient's tumor response is not subsequently confirmed, the patient will still be included in the denominator (the population set) but will not be included in the numerator (the "success"). In the PFS analysis, only a PD that's subsequently confirmed will be considered as an "event".

6.1. Standard Analyses

Study Conduct and Patient 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.

Baseline Characteristics

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.

[redacted]

Treatment Administration/Compliance

Listings and tables by dose level will be provided. Cycle length is 21 days (+/- 2 days) for IV infusion and 28 days (+/- 2 days) for subcutaneous administration. Day 1 of a cycle is the first date of dose within that cycle. The safety analysis set will be used.

Dose modifications may occur in the following ways:

- Cycle delay—Day 1 of current cycle starts later than 21 (+2) days from Day 1 of the previous cycle (only applies to cycle 2 and above). For subcutaneous administration, 28 (+2) days, rather than 21 (+2) days, will be used in determining cycle delay. For example, after cycle 1 ended for a patient in Part 2, a new cycle didn't start until 28 (+2) days after (but before 56 days after) cycle 1 day 1, the newly started cycle will be considered as cycle 2, and cycle 2 is considered delayed.

- Cycle skip – Day 1 of current cycle starts later than 42 (+4) days from Day 1 of the previous cycle (only applies to cycle 2 and above). For subcutaneous administration, 56 (+4) days, rather than 42 (+4) days, will be used in determining cycle skip. For example, After cycle 1 ended for a patient in Part 2, a new cycle didn't start until 56 (+4) days after cycle 1 day 1, the newly started cycle will be considered as cycle 3, and cycle 2 is considered skipped for this patient.

- Dose reduction— a decrease in the administered total daily dose (non-zero) compared to the planned total daily dose upon enrollment. If in the CRF the prescribed dose unit is mg/kg, but the actual dose is in mg the actual dose mg/kg should be calculated considering the body weight of the patient at that visit. Intrapatient dose reductions are not permitted during the study unless, in discussion with the sponsor, a dose level is deemed beyond the determined MTD.

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 subjects with cycle delays and cycle skips;

- 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) for cycle delays, dose interruptions and dose reductions;

[redacted]

- 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 delay (median, range);
- Number of cycles before 1st reduction (median, range);
- Number of cycles before 1st interruption (median, range);
- Number of cycles before 1st dose skip (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 percent of starting dose (compared to Day 1 dose of each cycle).

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 delay (yes/no), dose reduction (yes/no), and dose interruption (yes/no).

Prior, Concomitant, and Further Therapies

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.

6.2. Analysis of Primary Endpoint

6.2.1. DLT (Part 1)

Dose Limiting Toxicity is the primary endpoint of the dose escalation phase of the study, which will be summarized by dose level using the Per Protocol Analysis Set for patients in the dose escalation portion of the study. A listing of the DLTs will also be provided. If necessary, a summary and listing of the DLT by malignancy may be provided using the Per Protocol Analysis Set for patients in the MTD expansion portion of the study.

[redacted]

6.2.2. Safety Endpoints (Part 1 and Part 2)

In safety data tabulations, only data collected during the on-treatment period, defined as on or after the first dose of study treatment, and before the last dose of study treatment + 28 days, or the start of any other anti-cancer therapy, whichever is earlier, will be included. In safety data listings, all collected data will be listed.

Adverse Events

Adverse Events (AEs) will be graded by the investigator according to the CTCAE version 4.03 and coded using the MedDRA.4 The focus of AE summaries will be on Treatment Emergent Adverse Events, those with initial onset or increasing in severity after the first dose of study medication. The number and percentage of patients who experienced any AE, serious 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) for overall and each dose. The Safety Analysis Set will be used. Part 1 and Part 2 data will be summarized separately and will also be pooled together for analysis. Pfizer standard on safety data reporting will be followed.

Laboratory Tests Abnormalities

The number and percentage of patients who experienced laboratory test abnormalities will be summarized according to worst toxicity grade observed for each laboratory test for overall and each dose. The analyses will summarize laboratory tests both in the entire study period and by cycle (Cycle 1 and Cycles beyond 1). Shift tables will be provided to examine the distribution of laboratory abnormalities. The Safety Analysis Set will be used.

For laboratory tests without CTC grade definitions, results will be categorized as normal, abnormal high/low or not done.

6.2.3. ORR (Part 2)

ORR, as assessed using RECIST version 1.1 and irRECIST, is a primary efficacy endpoint for Part 2. The following analyses will be performed in the mITT population by tumor type (NSCLC or urothelial carcinoma) and pooled:

1. ORR by visit, separately for RECIST 1.1 and irRECIST: investigator provided tumor response CR, PR, and CR+PR will be presented by descriptive statistics (frequency and percentage) and 95% confidence interval.

2. Best overall response of CR, PR, and CR+PR, separately for RECIST 1.1 and irRECIST. This will include best overall response derivation without confirmation and with confirmation. Descriptive statistics (frequency and percentage) and 95% confidence interval will be provided.

[redacted]

6.3. Analysis for Secondary Endpoints

6.3.1. Efficacy Endpoints Analysis

Except ORR for Part 2, efficacy is a secondary objective. The efficacy analysis will be performed in the mITT population. Part 1 and Part 2 data will be summarized separately, and may also be pooled together for analysis if deemed necessary. Efficacy data assessed by RECIST will be analyzed separately from those assessed by irRECIST.

Tumor response data in Part 1, Part 2 and across the two Parts may be summarized with descriptive statistics (frequency and percentage) in the following groups by visit and then best overall response across all visits:

- Overall summary for all doses and all tumor types;
- By tumor type regardless of dose;
- By dose regardless of tumor type;
- By dose and tumor type if data permit.

Summary tables of best overall response rate, PFS, OS, DOSD, and DOR may be provided by the groups aforementioned when deemed necessary (e.g. if there are ≥5 patients in a specific group).

Efficacy listings (tumor measurements listings and tumor response listings) will be provided that include the investigator provided tumor measurement data, tumor response, best overall response, first CR/PR date, last date with CR or PR, most recent date without progression, progression date, death date, and last tumor assessment date, etc.

Swimmer plot for individual clinical response and time on treatment, waterfall plot for individual tumor size percent change from baseline, and spider plot for individual tumor size percent change from baseline over time will be presented for RECIST and irRECIST separately.

[redacted]

The following table provides an overview of the efficacy analysis.

Endpoint	Analysis Set	Statistical Method	Model/ Covariates/ Strata	Missing Data
Overall response	mITT	Exact CI	See aforementioned summary descriptions on data pooling across dose and tumor type	Observed case
Overall Survival	ITT	Kaplan-Meier	See aforementioned summary descriptions on data pooling across dose and tumor type	Censored at last visit
Progression Free Survival (PFS)	mITT	Kaplan-Meier	See aforementioned summary descriptions on data pooling 3721 across dose and tumor type	Censored per Appendix 9.2
Time to Progression (TTP) and Time to Response (TTR)	mITT	Kaplan-Meier	See aforementioned summary descriptions on data pooling across dose and tumor type	Censored per Appendix 9.2

[redacted]

Endpoint	Analysis Set	Statistical Method	Model/ Covariates/ Strata	Missing Data
Duration of Response (DOR)	mITT	Kaplan-Meier	See aformentioned summary descriptions on data pooling across dose and tumor type	Censored per Appendix 9.2
Duration of Stable Disease (DOSD)	mITT	Kaplan-Meier	See aformentioned summary descriptions on data pooling across dose and tumor type	Censored per Appendix 9.2

6.3.2. Pharmacokinetics Analyses

The concentration-time data of PF-06801591 will be summarized by descriptive statistics (n, mean, standard deviation, coefficient of variation, median, minimum, maximum, and geometric mean) according to dosing cohort and time for each part of the study. In addition, the concentration-time data from Part 2 will also be summarized by descriptive statistics according to tumor type.

The actual time of sample collection will be used in PK parameter calculation. In the event that the actual sampling time is not available, the nominal time may be used if there is no evidence that the actual sampling time deviates substantially from the nominal time.

Presentation of PF-06801591 concentration-time data

The concentration-time data of PF-06801591 will be presented as below:

- a listing of all concentrations by cohort, subject ID and nominal time. The concentration listing will also include the actual times. Deviations from the nominal time will be given in a separate listing.

- a summary of concentrations by cohort and nominal time, where the set of statistics will include n, mean, standard deviation, median, coefficient of variation (cv), minimum, maximum and the number of concentrations above the lower limit of quantification.

- for the concentration-time data after the 1st and 4th dose, median concentration-time plots (on both linear and semi-log scales) against nominal time postdose by cohort (all cohorts on the same plot per scale, based on the summary of concentrations by cohort and time postdose).

- for the concentration-time data after the 1st and 4th dose, mean concentration-time plots (on both linear and semi-log scales) against nominal time postdose by cohort (all cohorts on the same plot per scale, based on the summary of concentrations by cohort and time postdose).

[redacted]

For drug concentration summary statistics, median and mean plots by sampling time, the nominal PK sampling time will be used; for individual subject plots by time, the actual PK sampling time will be used, with the pre-dose time set to zero.

Calculation of PF-06801591 PK parameters

For patients enrolled in Part 1 of the study, the individual concentration-time data of PF-06801591 during Cycle 1 and Cycle 4 will be analyzed separately by non-compartmental methods to estimate the PK parameters. The PK parameters estimated will include C_{max}, Tmax, and AUC_{last} (AUC_{tau} at steady state). If data permit or if considered appropriate, $t_{1/2}$, CL (or CL/F for SC cohort), V_d (or V_d/F for SC cohort; V_{ss} at steady state), and accumulation ratio (R_{ac}) will also be estimated for Cycle 1 and Cycle 4. For IV cohorts in Part 1, individual Cycle 6 pre-dose concentrations will be used in lieu of Cycle 5 pre-dose concentrations while calculating Cycle 4 PK parameters. The PK parameters will be summarized descriptively by dose level and cycle.

Additionally, for Part 1, dose-normalized AUC_{last} and C_{max} will be plotted against dose (using a logarithmic scale) by cycle. These plots will include individual patient values and the geometric means for each dose. These plots will be used to help understand the relationship between the PK parameters and dose.

For patients enrolled in Part 2 of the study, trough concentrations of PF-06801591 will be summarized descriptively by cycle and by tumor type.

PK parameters will be calculated using standard non-compartmental methods:

Parameter	*State*	*Method of Determination*
AUC_{tau}	sd, ss	Linear/Log trapezoidal method
AUC_{last}	sd,ss	Linear/Log trapezoidal method
AUC_{inf} [a]	sd	AUClast + (Clast*/kel), where Clast* is the predicted serum concentration at the last quantifiable time point estimated from the log-linear regression analysis.
C_{max}	sd, ss	Observed directly from data
T_{max}	sd, ss	Observed directly from data as time of first occurrence
CL and CL/F	sd, ss	Dose/AUC_{inf} for sd[a] Dose/AUC_{tau} for ss
Vd/F	sd, ss	Dose/AUC_{inf} for sd Dose/(AUC_{τ} / kel) for ss
Vss [a]	sd, ss	CL * MRT, where MRT is the mean

[redacted]

		residence time adjusted for the duration of infusion
t½[a]	sd	Loge(2)/kel, where kel is the terminal phase rate constant calculated by a linear regression of the log-linear concentration time curve.

[a] if data permit.

6.3.3. Immunogenicity Assessment

For the immunogenicity data, the percentage of patients with positive ADAs and neutralizing antibodies will be summarized by overall, tumor type, and dose within tumor type as data permit. For patients with positive ADAs, the magnitude (titer), time of onset, and duration of ADA response will also be described, if data permit. In addition, efforts will be made if data permit, as appropriate, to examine possible correlations of the ADA response with clinical data on the PK, safety and/or efficacy.

6.3.4. RO of PD-1 by PF-06801591 in Circulating T Cells

Descriptive statistics of PD-1 receptor occupancy by PF-06801591 evaluated from the percentage of free (ie unbound) PD-1 on the surface of circulating T cells will be summarized by overall, tumor type, and dose within tumor type as data permit.

CCI

[redacted]

6.5. Population PK and PK/PD Modeling

Pharmacokinetic and PD data from this study may be analyzed using modeling approaches and may also be pooled with data from other studies to investigate any association between PF-06801591 exposure and biomarkers or significant safety endpoints. The results of these analyses, if performed, may be reported separately.

[redacted]

6.6. ECG and Vital Sign Data Analysis

The analysis of ECG results will be based on patients in the Safety Analysis Set with baseline and on-treatment ECG data, and will follow the ICH E14 guidance on the clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs.5

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 HR (QTc) using standard correction factors [ie, Fridericia's (default correction), Bazett's, and possibly a study specific factor, as appropriate]. QTcF interval will be calculated using the Friderica formula, as follows:

$$QTcF = \frac{QT}{\sqrt[3]{RR}}$$

Data will be summarized and listed for QT, HR, response rate (RR), PR, QRS, QTcF (and/or QTcB if deemed appropriate by overall, and dose. Individual QT` (all evaluated corrections) intervals will be listed by study arm 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 study arm dose and time point. 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 CTCAE version 4.03 Grade. Shift tables will also be provided for ECG abnormality at baseline vs. on treatment (yes, no, not done: (n, %).

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

The effect of drug concentrations on corrected QT change from baseline may be explored graphically. Additional concentration-corrected QT analyses may be performed. Data may be pooled with other study results and/or explored further with PK/PD models.

Changes from baseline for the ECG parameters QT interval, heart rate (HR), QTc interval, PR interval and QRS interval will be summarized by treatment and visit. Categorical data analysis will follow Appendix 9.3.

[redacted]

If more than one ECG is collected at a nominal time post dose (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 three individual ECG tracings has a QTc value ≥500 msec, but the mean of the triplicates is not ≥500 msec, the data from the subject's individual tracing will be described in a safety section of the study report 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 QTc post dose from the time-matched average of the pre-dose 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 subject factors (covariates) on the relationship will be examined.

7. INTERIM ANALYSES

There is no formal interim analysis planned in this study. In Part 2, a Bayesian approach will be used aiming to detect early sign of efficacy. The following analyses will be performed:

1. Approximately 8 weeks after the first 30 NSCLC patients have enrolled, a data snapshot will be taken for an early data assessment. Descriptive statistics (frequency and percentage) for tumor response (CR, PR, PD, stable disease [SD] etc.) from the first tumor assessment; ORR, the proportion of patients who achieved unconfirmed CR or PR, and the 95% confidence interval for ORR; DCR (disease control rate), the proportion of patients who achieved unconfirmed CR or PR or SD, and the 95% confidence interval for the DCR will be presented. If tumor response data are collected based on both RECIST 1.1 and irRECIST, separate presentations will be generated for each response criterion. In this data snapshot, some patients may have gone through their second tumor assessment as of the data snapshot, however only data through the first tumor assessment will be used for this analysis. The above described analyses may be repeated, if deemed necessary, when the first 40 or 45 NSCLC patients are enrolled and have gone through their first tumor assessment. This preliminary data may provide guidance in the decision to the increase or decrease the sample size of each arm, to add other tumor types in a future amendment, or to initiate additional clinical studies.

2. In the early data assessment, a Bayesian approach will be used to calculate the posterior probability that the true ORR is greater than or equal to the minimal expected ORR. For example, the minimal expected ORR if all patients received only one tumor assessment is approximately 10% but may be adjusted based on how many patients had assessments beyond 8 weeks and the PD-L1 status of patients which will be tested at baseline. A beta prior (0.235, 1) is chosen, where the parameter beta is set to be 1, and parameter alpha 0.235 was calculated from a reference product pembrolizumab on NSCLC patients where a 19% ORR was observed. For example, if 4 responders are observed out of the first 30 patients as of the first tumor assessments, the posterior probability that the true ORR is greater than a minimal expected ORR of 10% is approximately 70%.

[redacted]

A final analysis will be performed when all 60 NSCLC patients and 40 urothelial cancer patients are enrolled and have completed their scheduled tumor assessments.

This is an open label study, the Pfizer study team will review safety, immunogenicity, pharmacodynamics, [redacted] and other data throughout the study.

[redacted]

8. REFERENCES

1. B8011011 Protocol Amendment 1, Dec. 10, 2015; Protocol Amendment 2, July 21, 2016; Protocol Amendment 3, June 3, 2017

2. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958; 53:457-481.

3. Brookmeyer R, Crowley JJ. A confidence interval for the median survival time. Biometrics 1982; 38:29-41.

4. NCI Common Terminology Criteria for Adverse Events v4.03 (CTCAE). http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf.

5. ICH E14 - The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. CHMP/ICH/2/04.

6. 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.

7. Nishino M, Gargano M, Suda M, Ramaiya NH, Hodi FS. Optimizing immune-related tumor response assessment: does reducing the number of lesions impact response assessment in melanoma patients treated with ipilimumab? J Immunother Cancer 2014;2:17.

8. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. J Clin Oncol 2007; 25: 579-86. Responses must last for at least 4 weeks off treatment.

9. Guha M:The new era of immune checkpoint inhibitors; Pharm J Volume 293, 18 NOV 2014, No 7837/8 online URL 200671127.

[redacted]

9. APPENDICES

9.1. Details of Definitions of Endpoints

DLT Definitions

Previous anti-PD1 mAbs were administered in >1000 patients and were associated in the clinic with inflammatory adverse reactions resulting from increased or excessive immune activity (immune-mediated adverse reactions), likely to be related to the mechanism of action. Immune-mediated adverse reactions, which can be severe, may involve the gastrointestinal, skin, liver, endocrine, respiratory, renal or other organ systems.[9]

Severity of AEs will be graded according to CTCAE version 4.03. For the purpose of a dose finding decision, any of the following drug-related AEs occurring during the first cycle of treatment (21 days, or 28 days for subcutaneous administration) will be classified as DLTs following review by the investigators and the sponsor:

- Grade 5 AE.
- Hematologic toxicity:
  - Any Grade 4 hematologic AE, with the following clarifications.
  - Grade 4 neutropenia lasting >5 days from initiation of granulocyte-colony stimulating factor.
  - Grade 4 thrombocytopenia with bleeding.
  - Platelet transfusion requirement or a platelet count <10,000/uL.
- Non-Hematologic Toxicity:
  - Grade 4 non-hematologic AE.
  - Grade 3 AE lasting >7 days despite optimal supportive care.
  - Grade 3 central nervous system AE regardless of duration.
  - Concurrent aspartate aminotransferase (AST) or alanine aminotransferase (ALT) >3X upper limit of normal (ULN) and total bilirubin >2X ULN (potential Hy's law case see section Potential Cases of Drug-Induced Liver Injury).

The following AEs will not be adjudicated as DLTs:

- Any Grade 3 endocrinopathy that is adequately controlled by hormonal replacement.

[redacted]

- Grade 3 AE of tumor flare (defined as local pain, irritation, or rash localized at sites of known or suspected tumor).

- Isolated Grade 3-4 laboratory abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset.

- Grade ≥3 infusion reactions and allergic reactions will not be considered dose limiting as they are unlikely to be dose related, but all available information on these events will be collected. If Grade ≥3 infusion reactions occur in ≥2 of the first 10 patients at any dose level, or if the occurrence is ≥5% thereafter, a mandatory pre- treatment regimen for all new patients will be implemented. The incidence of Grade 1 and 2 reactions will also be taken into account. If a total rate of >10% all-grade infusion or allergic reactions is observed, a mandatory pre-treatment regimen for all new patients will be implemented.

Information regarding the DLT observation period can be found in the Criteria for Dose Finding (See protocol Section 3.1.4). For dose escalation, DLT observation is required for 21 days for IV infusion and 28 days for subcutaneous administration. However, DLT observation will continue for at least 120 days from first dose (or completion of 5 cycles, if still on treatment) to assess late immune-related dose-limiting AEs, and will be taken into account for MTD determination.

Late immune-related DLTs

Late immune-related DLTs are irAEs (see protocol Appendix 9) that meet the same grading criteria as DLT criteria but occur after the initial 21-Day DLT period for IV or 28-day DLT period for SC and during the first 120-day assessment period. Late immune-related DLTs will be added to the mTPI approach to reassess the dose-finding decisions.

For any patient being treated at dose levels that are subsequently considered to exceed the MTD, the option to reduce their dose will be discussed between the investigator and the sponsor's medical monitor. If the patient tolerated the above-MTD dose level well and is deriving clinical benefit, continuation of treatment at the above-MTD dose level will require re-consenting.

Maximum tolerated dose (MTD)

The maximum tolerated dose (MTD) is defined as the highest dose with true toxicity probabilities in the equivalence interval (EI) where the EI is defined as [22.5%-32.5%].

In practice, the MTD will be the highest dose associated with the occurrence of DLTs ≤33% (eg, ≤3 of 9 evaluable patients experience a DLT during at least [90 days after the first dose]) or the first 5 treatment cycles if the patient remains on treatment.

Recommended Phase 2 Dose (RP2D) Definition

The recommended Phase 2 dose (RP2D) is the dose chosen for further study combining the MTD and OBD, based on Phase 1 study results, from the primary and secondary endpoints of the study. If the MTD proves to be clinically feasible for long-term administration in a reasonable number of patients, then this dose usually becomes the RP2D. Further experience with the MTD may result in a RP2D lower than the MTD. Next to safety assessment, careful consideration will be given to immunomodulatory effects, PK information, and preliminary anti-tumor activity. If an OBD can be determined, this will be a key factor in the determination of the RP2D.

[redacted]

9.2. Time to Event Data Analysis Censoring Rules

Table 9.2.1: 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 over the course of the study and no progression	Date of last objective tumor assessment	Censored
Progression or death documented on or between scheduled tumor assessments	Date of first objective tumor assessment showing objective progression or death date	Progressed/Death (Event)
Treatment discontinuation for undocumented progression	Date of last objective tumor assessment	Censored
Treatment discontinuation due to toxicity or other reason	Date of last objective tumor assessment	Censored
Progression/Death prior to second planned tumor	Date of death	Progressed/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. The first two reasons (Inadequate baseline assessment; No on-study assessments) are not applicable for Duration of Response, as that endpoint is only for the subset of responders.

Table 9.2.2: Time to Progression

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 over the course of the study and no progressionProgression	Date of last objective tumor assessment	Censored
Progression Documented on or between scheduled tumor assessments	Date of first objective tumor assessment showing objective progression	Progressed (Event)
Treatment discontinuation for undocumented progression	Date of last objective tumor assessment	Censored

[redacted]

Situation	Date of Progression/Censoring[1]	Outcome
Treatment discontinuation due to toxicity or other reason	Date of last objective tumor assessment	Censored
New anticancer treatment prior to documented 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	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.

[redacted]

9.3. 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

[redacted]

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.

[redacted]

9.4. 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.[6]

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 (longest diameter in the plane of measurement is to be recorded) with a minimum size of:

- 10 mm for lesions other than lymph nodes and assessed by CT scan (CT scan slice thickness no greater than 5 mm).
- 10 mm for lesions assessed clinically by caliper measurement (lesions which cannot be accurately measured with calipers should be recorded as non-measurable).
- 20 mm for lesions assessed by chest X-ray.
- 15 mm in short axis for lymph nodes when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm).

Non-measurable Lesions

Non-measurable lesions include small lesions (longest diameter <10 mm or pathological lymph nodes with a ≥10 but <15 mm short axis) as well as truly non-measurable lesions. Truly non-measurable lesions include: leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses identified by physical exam and not measurable by reproducible imaging techniques.

[redacted]

Nodes that have a short axis <10 mm are considered non-pathological and should not be recorded or followed.

Special Considerations Regarding Specific Lesions

Bone lesions:

- Bone scan, PET scan or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions.

- Lytic bone lesions or mixed lytic-blastic lesions, with identifiable soft tissue components, that can be evaluated by cross sectional imaging techniques such as CT or MRI can be considered as measurable lesions if the soft tissue component meets the definition of measurability described above.

- Blastic bone lesions are non-measurable.

Cystic lesions:

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor non-measurable) since they are, by definition, simple cysts.

- 'Cystic lesions' thought to represent cystic metastases can be considered as measurable lesions, if they meet the definition of measurability described above. However, if non-cystic lesions are present in the same patient, these are preferred for selection as target lesions.

Lesions with prior local treatment:

- Tumor lesions situated in a previously irradiated area, or in an area subjected to other loco-regional therapy, are usually not considered measurable unless there has been demonstrated progression in the lesion.

Solitary lesions:

If a measurable disease is restricted to a solitary lesion, its neoplastic nature should be confirmed by cytology/histology.

[redacted]

Recording Tumor Measurements

All measurable lesions up to a maximum of 2 lesions per organ and up to 5 in total and representative of all involved organs should be identified as target lesions and measured and recorded at baseline and at the stipulated intervals during treatment. Target lesions should be selected on the basis of their size (lesions with the longest diameters) and their suitability for accurate repetitive measurements (either by imaging techniques or clinically).

The longest diameter will be recorded for each target lesion. The sum of the longest diameter of all target lesions will be calculated and recorded as the baseline sum diameter to be used as reference to further characterize the objective tumor response of the measurable dimension of the disease during treatment.

One exception to the above described approach is related to pathological lymph nodes. Pathological lymph nodes are defined as measurable lesions and may be identified as target lesions if the criterion of a short axis of ≥ 15 mm by CT scan is met. Only the short axis of these nodes will contribute to the baseline sum. Nodal size is normally reported as two dimensions in the plane in which the image is obtained (for CT scan this is almost always the axial plane; for MRI the plane of acquisition may be axial, sagittal or coronal). The smaller of these measures is the short axis.

A sum of the diameters (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and reported as the baseline sum diameters. The baseline sum diameters will be used as reference to further characterize any objective tumor regression in the measurable dimension of the disease.

All other lesions (or sites of disease) should be identified as non-target lesions and should also be recorded at baseline. Measurements are not required and these lesions should be followed as 'present', 'absent', or in rare cases 'unequivocal progression'. In addition, it is possible to record multiple non-target lesions involving the same organ as a single item on the case record form (eg, 'multiple enlarged pelvic lymph nodes' or 'multiple liver metastases').

Definition of Tumor Response

Target Lesions

Response in target lesions is defined as follows:

Complete Response (CR): disappearance of all target lesions.

- **Partial Response (PR):** at least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.

[redacted]

- **Progressive Disease (PD):** at least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. The appearance of one or more new lesions is also considered a sign of progression.

- **Stable Disease (SD):** neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.

When nodal disease is included in the sum of target lesions and the nodes decrease to 'normal' size (<10 mm), they may still have a measurement reported on scans. This measurement should be recorded even though the nodes are normal in order not to overstate progression should it be based on increase in size of the nodes. As noted earlier, this means that patients with CR may not have a total sum of 'zero' on the CRF.

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 one or more non-target lesion(s) and/or maintenance of tumor marker level above the normal limits.

- Progressive **Disease (PD):** Unequivocal progression of existing non-target lesions. (Note: the appearance of one or more new lesions is also considered progression).

Cytology, histology

These techniques can be used to differentiate between PR and CR in rare cases if required by protocol (for example, residual lesions in germ cell tumors). When effusions are known to be a potential adverse effect of treatment (eg, taxane compounds or angiogenesis inhibitors), the cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment can be considered if the measurable tumor has met criteria for response or stable disease in order to differentiate between response or stable disease and progressive disease.

For patients having effusions or ascites, only cases having cytological proof of malignancy should be recorded on the CRF. Effusions that have not been evaluated using cytology or were found to be non-malignant should not be recorded on the CRF.

[redacted]

New Lesions

The appearance of new malignant lesions indicates PD. New lesion should be unequivocal (eg, not attributable to differences in imaging technique, or change in imaging modality or findings not attributable to tumor). If a new lesion is equivocal, for example due to its small size, continued therapy and follow-up assessment will clarify the etiology of the disease. If repeat scans confirm there is definitely a new lesion, then progression should be declared using the date of the initial scan.

The use of FDG-PET is sometimes reasonable to complement a CT scan assessment of a PD (particularly for possible 'new' disease). New lesions on the basis of FDG-PET imaging can be identified according to the following algorithm:

- Negative FDG-PET at baseline, with a positive FDG-PET at follow-up.
- No FDG-PET at baseline and a positive FDG-PET at follow- up: if the positive FDG-PET at follow-up corresponds to a new site of disease confirmed by CT, this is PD.

If the positive FDG-PET at follow-up is not confirmed as a new site of disease on CT, additional follow-up CT scans are needed to determine if there is truly progression occurring at that site (if so, the date of PD will be the date of the initial abnormal FDG-PET scan).

If the positive FDG-PET at follow-up corresponds to a pre-existing site of disease on CT that is not progressing on the basis of the anatomic images, this is not PD.

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 10.4.1: 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	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD

[redacted]

Target lesions	Non-target lesions	New Lesions	Overall response
Any	Any	Yes	PD
CR = complete response, PR = partial response, SD = stable disease,			
PD = progressive disease, and NE = inevaluable.			

Best overall response

The best overall response is defined according to the tumor response along the study. Complete or partial responses may be claimed only if the criteria for each are met at a following time point as specified in the protocol (generally 4 weeks later). In this circumstance, the best overall response can be interpreted as the following table.

Table 10.4.2 Best Overall Response When Confirmation of CR and PR Required

Overall response First time point	Overall response Subsequent time point	BEST overall response
CR	CR	CR
CR	PR	SD, PD or PRa
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE
CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable.		

[a] If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

[redacted]

Patients with a global deterioration of health status requiring discontinuation of treatment without objective evidence of disease progression at that time should be reported as 'symptomatic deterioration'. Every effort should be made to document objective progression even after discontinuation of treatment. Symptomatic deterioration is not a descriptor of an objective response: it is a reason for stopping study therapy. The objective response status of such patients is to be determined by evaluation of target and non-target lesions.

In some circumstances it may be difficult to distinguish residual disease from normal tissue. When the evaluation of CR depends upon this determination, it is recommended that the residual lesion be investigated (fine needle aspirate/biopsy) before assigning a status of complete response. FDG-PET may be used to upgrade a response to a CR in a manner similar to a biopsy in cases where a residual radiographic abnormality is thought to represent fibrosis or scarring. The use of FDG-PET in this circumstance should be prospectively described in the protocol and supported by disease specific medical literature for the indication. However, it must be acknowledged that both approaches may lead to false positive CR due to limitations of FDG-PET and biopsy resolution/sensitivity.

[redacted]

9.5. Immune-related RECIST (irRECIST) Tumor Assessment Criteria

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-cytotoxic T lymphocyte-associated protein 4 (CTLA4) and anti PD-1/anti-PD-L1 antibodies which exert the antitumor activity by augmenting activation and proliferation of T cells, thus leading totumor infiltration by T cells and tumor regression rather than direct cytotoxic effects. 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) have been 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 these delayed or flare type responses into the RECIST v1.1 (irRECIST).[7]

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 the sum 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 target lesions, 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:

- 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.

- 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%.

[redacted]

- 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).

- 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) with a minimum absolute increase of 5 mm, 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.

[redacted]