

PROTOCOL A9421018

PHASE 1B/2 STUDY OF PF-04136309 IN COMBINATION WITH GEMCITABINE AND NAB-PACLITAXEL IN PATIENTS WITH PREVIOUSLY UNTREATED METASTATIC PANCREATIC DUCTAL ADENOCARCINOMA

STATISTICAL ANALYSIS PLAN (SAP)

Version: 1.0

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

Date: 8-APR-2016

TABLE OF CONTENTS

LIST OF TABLES.....	3
APPENDICES	4
1. VERSION HISTORY.....	5
2. INTRODUCTION	5
2.1. Study Objectives	5
2.2. Study Design.....	7
2.2.1. Phase 1b study: open label dose finding.....	7
2.2.2. Phase 2: a randomized double blinded placebo control study.....	8
3. ENDPOINTS AND COVARIATES: DEFINITIONS AND CONVENTIONS	9
3.1. Primary Endpoint(s).....	9
3.2. Secondary Endpoint(s).....	9
3.3. Exploratory Endpoints	10
3.4. Covariates	11
4. ANALYSIS SETS	11
4.1. Full Analysis Set.....	11
4.2. Safety Analysis Set	11
4.3. ‘PER PROTOCOL’ Analysis Set	11
4.4. PK Analysis Set	11
4.4.1. PK Concentration Set	11
4.4.2. PK Parameter Set.....	11
4.5. PD Biomarker Analysis Set	11
4.6. Modified Intent-to-Treat Set	12
4.7. Treatment Misallocations.....	12
4.8. Protocol Deviations.....	12
5. GENERAL METHODOLOGY AND CONVENTIONS	12
5.1. Statistical Hypotheses	12
5.2. Statistical Decision Rules	12
5.2.1. Part 1b.....	12
5.2.2. Phase 2.....	14
5.3. General Methods.....	15

5.3.1. Analyses for Time to Event Data	15
5.3.2. Analyses for Binary Data	16
5.3.3. Analyses for Continuous Data.....	16
5.4. Methods to Manage Missing Data	16
5.4.1. Missing Dates	16
5.4.2. Efficacy Analysis	16
5.4.3. Pharmacokinetics.....	16
5.4.4. Pharmacodynamic and health outcome parameters	17
6. ANALYSES AND SUMMARIES	17
6.1. Standard Analyses.....	17
6.2. Analysis of Primary Endpoints	19
6.2.1. DLT(Part 1)	19
6.2.2. Safety Endpoints.....	20
6.2.3. PFS (Phase 2)	20
6.3. Analysis for Other Endpoints.....	21
6.3.1. Other Efficacy Endpoints Analysis	21
6.3.2. Pharmacokinetics Analyses	22
6.3.3. Analysis of Pharmacodynamics.....	23
6.3.4. Population Pharmacokinetic Analysis or Pharmacokinetic/Pharmacodynamic (PK/PD) Modeling (Phase 1b and Phase 2).....	24
6.3.5. Analysis of Peripheral Neurological Adverse Events and EORTC QOL-C30 and QLQ-CIPN20 Questionnaire (Phase 2)	24
6.4. ECG and Vital Sign Data Analysis	25
7. INTERIM ANALYSES.....	26
8. REFERENCES	27
9. APPENDICES	28

LIST OF TABLES

Table 1.	13
Table 2. Progression Free Survival and Duration of Response.....	30
Table 3. Time to Progression with the double combination PF-04136309 and gemcitabine (maintenance therapy) after interruption of nab-paclitaxel	30

Table 4. Objective Response Status at each Evaluation.....36

Table 5. Objective Response Status at each Evaluation for Patients with Non-Target
Disease Only.....37

APPENDICES

Appendix 1. Details of Definitions of Endpoints28

Appendix 2. Time to Event Data Analysis Censoring Rules30

Appendix 3. Categorical Classes for ECG and Vital Signs.....32

Appendix 4. RECIST 1.1 Tumor Assessment Criteria.....33

090177e18f2f4f12\0.1\Draft\Versioned On:16-Aug-2018 13:12 (GMT)

1. VERSION HISTORY

This is the first version of statistical analysis plan, based on the protocol amendment dated Nov. 25, 2015 [1].

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study A9421018. 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

Phase 1b (Dose Finding) Objectives

Primary Objectives

- To evaluate the safety and tolerability of PF-04136309 in combination with nab paclitaxel and gemcitabine in patients with metastatic pancreatic ductal carcinoma.
- To characterize the dose limiting toxicities (DLTs) and overall safety profile of escalated doses of PF-04136309 and the associated schedule.
- To determine the maximum tolerated dose (MTD) of PF-04136309 and select the Recommended Phase 2 Dose (RP2D).

Secondary Objectives

- To evaluate the pharmacokinetics (PK) of PF-04136309 when given in combination with nab-paclitaxel and gemcitabine in combination.
- To characterize the ex vivo inhibition of CCL2-induced ERK kinase phosphorylation as a measure of target engagement following treatment with PF-04136309 in combination with nab-paclitaxel + gemcitabine at each dose level.

CCI

C
C
I

[REDACTED]

I

[REDACTED]

I

[REDACTED]

I

[REDACTED]

Phase 2 (Randomized Double Blinded Placebo Control Study) Objectives**Primary Objectives**

- To assess the enhancement of efficacy of PF-04136309 in combination with nab- paclitaxel and gemcitabine versus nab-paclitaxel + gemcitabine + placebo in terms of progression free survival (PFS).

Secondary Objectives

- To assess the enhancement of efficacy of PF-04136309 in combination with nab- paclitaxel and gemcitabine versus nab-paclitaxel + gemcitabine + placebo in terms of Overall Survival (OS).
- To assess the enhancement of efficacy of PF-04136309 in combination with nab- paclitaxel and gemcitabine versus nab-paclitaxel + gemcitabine + placebo in terms of Objective Response Rate (ORR) and duration of response.
- To evaluate the safety and tolerability of PF-04136309 in combination with nab- paclitaxel + gemcitabine.
- To evaluate the population PK of PF-04136309 given in combination with nab- paclitaxel + gemcitabine.
- To validate the proof of mechanism of PF-04136309 in combination with nab- paclitaxel + gemcitabine compared to nab-paclitaxel + gemcitabine + placebo.
- To evaluate the improvement of peripheral neurotoxicity induced by nab- paclitaxel by the addition of PF-04136309 to the combination therapy of nab- paclitaxel plus gemcitabine.

CCI

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

2.2. Study Design

This is a multi-center Phase 1b/2 study in the first line treatment of patients with metastatic pancreatic ductal adenocarcinoma (mPDA).

The Phase 1b will be open label as patients will receive ascending doses of PF-04136309 in combination with nab-paclitaxel + gemcitabine. The Phase 2 will be controlled, randomized and double blinded as patients will be randomized to receive PF-04136309 in combination with nab-paclitaxel + gemcitabine (ARM A) versus nab-paclitaxel + gemcitabine + placebo (ARM B: Standard of Care).

2.2.1. Phase 1b study: open label dose finding.

The Phase 1b part of the study is a dose finding phase study. In the dose finding phase, patients will be enrolled in cohorts of 2 to 4 patients, starting with 750 mg BID of PF-04136309. A new cohort of n=2 to 4 patients exposed to the same dose 750 mg twice daily (BID) could be opened after safety has been established for at least the first cycle in all patients in the previous cohort. Escalation to a higher dose (eg 1000 mg BID) will require that at least n=6 patients are dosed at 750 mg BID and no more than 1 experienced a DLT.

DMB02-GSOP-RF02 3.0 STATISTICAL ANALYSIS PLAN TEMPLATE 30-Jun-2015

An MTD may not be identified, as DLTs may not be observed after administration of this investigational agent. Based on safety data, and, PK and PD data, dose escalation may continue up to a predefined Maximum Feasible Dose (ie, 1000 mg BID) or Phase 2 recommended dose.

Dose levels may include de-escalation (500 mg BID) if a high number of the patients (ie, >33%) receiving 750 mg BID experiences a DLT related to PF-04136309. Based on safety and other results (eg, PK) from patients enrolled in the dose escalation cohorts, a dose level will be selected to be further evaluated as the Recommended Phase 2 Dose (RP2D). A minimum of 6 to 12 patients will be treated at this dose level to establish it as the RP2D.

2.2.2. Phase 2: a randomized double blinded placebo control study

In the Phase 2 part, ninety-two (92) patients will be randomized with a 1:1 to receive PF-04136309 in combination with nab-paclitaxel + gemcitabine (ARM A; n=46) versus nab-paclitaxel + gemcitabine + placebo (ARM B; n=46). The primary objective will be the enhancement of efficacy in terms of PFS.

Given the dynamic feature of this design, the exact number of dose arms that will be expanded cannot be prespecified, as it will depend on safety and anti-tumor activity.

In both studies, treatment may continue until progression of disease (PD) unacceptable toxicity or patient refusal, whichever comes first. After 6 months, in case of neurological or hematological chronic \geq grade 2 toxicity, nab-paclitaxel may be stopped and gemcitabine + PF-04136309 will be continued until disease progression or patient's or investigator's decision. If gemcitabine is not well tolerated, treatment may be continued with PF-04136309 alone until disease progression or patient's or investigator's decision.

The proof of mechanism of action of PF-04136309 will be assessed by the comparison of at least 12 paired (optional) biopsies, bone marrow aspirates and peripheral blood (pre- and post-treatment in the Phase 1b and Phase 2 study). Their optional collection will provide baseline, pre-dose fine-needle aspirate (FNA) or core needle biopsy, and/or bone marrow aspirates, and FNA or core needle biopsy, and/or bone marrow aspirates following the completion of Cycle 1 at C1D28 [± 7 days] or at C2D28 [± 7 days] if collection at Cycle 1 is not obtained.

Serial blood samples will be collected from patients enrolled in the Phase 1b cohorts to determine the multiple dose PK of PF-04136309 given in combination with nab-paclitaxel + gemcitabine. Sparse PK sampling will be performed for patients enrolled in the Phase 2 portion of the study to evaluate the population PK of PF-04136309.

3. ENDPOINTS AND COVARIATES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

Phase 1b

- Dose Limiting Toxicities (DLTs) in order to determine the Maximum Tolerated Dose (MTD) and Recommended Phase 2 Dose (RP2D).
- Adverse Events as characterized by type, frequency, severity (as graded by NCI CTCAE v 4.03) timing, seriousness and relationship to study therapy.
- Lab abnormalities as characterized by type, frequency, severity (as graded by NCI CTCAE v 4.03) and timing.

Phase 2

- Progression Free Survival (PFS).

3.2. Secondary Endpoint(s)

Phase 1b

- PK parameters of PF-04136309 after multiple dosing, including $C_{max,ss}$, T_{max} , $AUC_{tau,ss}$, $C_{min,ss}$, CL_{ss}/F , and as data permit, $t_{1/2}$ and V_{ss}/F .
- Ex vivo inhibition of CCL2-induced ERK kinase phosphorylation in the peripheral blood.
- Objective tumor response (OR) as assessed using the Response Evaluation Criteria in Solid Tumor (RECIST) version 1.1.

Phase 2

- Overall Survival (OS)
- Adverse events as characterized by type, frequency, severity (as graded by NCI CTCAE v.4.03), timing, seriousness and relationship to study treatment.
- Laboratory abnormalities as characterized by type, frequency, severity (as graded by NCI CTCAE v. 4.03) and timing.
- Objective response rate (ORR).
- Duration of Response. [The rate of progression-free survival at 1 year, the rate of disease control (confirmed response or stable disease for ≥ 16 weeks)].

- Time to progression with the double combination PF-04136309 and gemcitabine (maintenance therapy) after interruption of nab-paclitaxel.
- Trough PF-04136309 concentrations.
- PD markers in metastatic tumors: Pre biopsy – Post biopsy after 2 cycles of
- PF-04136309 + nab-paclitaxel + gemcitabine or nab-paclitaxel + gemcitabine + placebo
- Peripheral neurological adverse events as characterized by frequency, severity (as graded by NCI CTCAE v.4.03 and by self-assessment according to the EORTC QOL- CIPN20 questionnaire), and timing.

CCI

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]

CCI

- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]
- [REDACTED]



3.4. Covariates

Biomarker data and baseline characteristics may be considered as covariates in population PK, PK/PD and anti-tumor efficacy exploratory analyses. Covariates may also be considered when exploring QTc and PK relationship.

4. ANALYSIS SETS

Several analysis sets are defined and will be considered for this study.

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 important protocol deviations during the Cycle 1 that may impact DLT evaluations. Patients with the important protocol deviations during the first cycle of treatment that may impact DLT evaluations are not evaluable for the MTD or RP2D assessment.

4.4. PK Analysis Set

4.4.1. PK Concentration Set

The PK concentration population is defined as all patients who receive PF-04136309, 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 Phase 1b patients treated who have sufficient information to estimate at least 1 of the PK parameters of interest.

4.5. PD Biomarker Analysis Set

The PD/Biomarker analysis population is defined as all enrolled patients with at least 1 of the PD/Biomarkers evaluated (including both the patients with only pre-treatment data and those with both pre- and/or post-treatment data) .

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 and have measurable disease baseline assessment (within 28 days prior to study entry). The mITT population will be used for anti-tumor assessment.

4.7. 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.

4.8. 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.

5. GENERAL METHODOLOGY AND CONVENTIONS

The phase 1b part of the study is an open-label dose finding study and no interim analysis or blinding is planned for the Phase 1b part. The phase 2 part of the study is a double-blinded study and the final analysis for the phase 2 part will be conducted after the last subject last visit (LSLV) in the phase 2 part.

5.1. Statistical Hypotheses

Phase 1b (dose finding):

There is no statistical hypothesis.

Phase 2 Hypothesis:

H_0 : Hazard ratio $\lambda = 1$ vs. H_1 : $\lambda < 1$,

Where Hazard ratio λ is the hazard ratio for progression free survival between Arm A (PF-04136309 + nab-paclitaxel + gemcitabine) and Arm B (nab-paclitaxel + gemcitabine + placebo).

5.2. Statistical Decision Rules

5.2.1. Part 1b

Dose escalation and de-escalation

The modified toxicity probability interval (mTPI) design [2] uses a Bayesian statistics framework and a beta/binomial hierarchical model to compute the posterior probability of 3 dosing intervals that reflect the relative difference between the toxicity rate of each dose level

DMB02-GSOP-RF02 3.0 STATISTICAL ANALYSIS PLAN TEMPLATE 30-Jun-2015

to the target rate ($p_T = 0.275$). The prior distribution of DLT is set as a beta (0.5,0.5) and the threshold probability for early termination and dose exclusion is set to 0.95 as suggested in the mTPI design. Doses with an incidence of $DLT > 33\%$ (eg, 4 out of 10) cannot be selected as MTD although is allowed by the mTPI method. If the toxicity rate of the currently used dose level is far smaller than 27.5%, the mTPI will recommend escalating the dose level; if it is close to 27.5%, the mTPI will recommend continuing at the current dose; if it is far greater than 27.5%, the mTPI will recommend de-escalating the dose level. Being a model-based design, mTPI automatically and appropriately tailors dose-escalation and de-escalation decisions. Target size for each cohort will be $n = 2$ to 4 patients as described in mTPI design. All the dose-escalation decisions for a given trial can be pre-calculated under the mTPI design and presented in a two-way table. The decision rules to “dose escalate” (E), “no change in dose” (S), “dose de-escalate” (D) or “dose de-escalate, unacceptable toxicity” (U) are described in Table 1.

Cohorts of patients could receive doses already tested but a dose that is associated with decision “Dose de-escalate, unacceptable toxicity” cannot be revisited and no more patients should be treated at this dose or higher doses for the remainder of the trial.

Table 1.

	Number of Patient Treated at a Dose Level													
Number of Patients Having DLT	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	S	E	E	E	E
3		U	U	D	D	S	S	S	S	S	S	S	S	E
4			U	U	U	U	D	D	D	D	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

D: De-escalate the dose; E: Escalate the dose; S: Stay at the dose; U: Unacceptable toxicity

Note: Starting dose is 750 mg BID. Escalation to a higher dose (eg 1000 mg BID) will require that at least 6 patients are dosed at 750 mg BID and no more than 1 experienced a DLT.

Estimating the MTD and RP2D (Phase 1b)

An MTD may not be identified, as DLTs may not be frequently observed after administration of this investigational agent. Dose escalation may continue up to a predefined Maximum Feasible Dose (MFD) (eg 1000 mg) and the escalation is halted without a maximum tolerated dose (MTD) estimate. The study will continue accruing until one of the three stopping conditions below is triggered.

The algorithm will stop if any of the following criteria is met:

1. MTD or RP2D has been identified with sufficient accuracy: at least 6 to 12 patients have been accumulated on a dose that is currently estimated to be the MTD or RP2D; or
2. All doses explored appear to be overly toxic.

Due to binomial data variability in small samples, DLTs may be observed in a first cohort(s) by chance even when the true Probability (DLT) is fairly low. This could result in the estimated posterior DLT rate to exceed the targeted very early in the trial, triggering an early stop when very few patients (eg, 3) have been treated. To prevent stopping the trial prematurely in such cases, a step-down option with a lower dose of 500 mg added to the dose grid.

Sample Size (Phase 1b)

The number of patients to be enrolled in the Phase 1b part will depend upon the observed safety profile, which will determine the number of patients at each dose level and the number of dose levels explored. It is envisaged that the maximum sample size at the RP2D would be N=12 (see Protocol Section 9.2.3). To further evaluate safety and pharmacodynamics, the number of patients enrolled during this part of the study (Phase 1b) may be N up to 20.

5.2.2. Phase 2

Sample Size (Phase 2)

Ninety-nine (92) patients will be randomized in a 1:1 ratio Arm A versus Arm B. Median PFS in patients enrolled in the control arm (Arm B) is expected to be approximately 5.5 months. A total of 92 patients and 57 events are required for a log-rank test to have an overall 1-sided significance level of 0.1 ($\alpha=0.2$) and power of 0.80. This assumes an 82% improvement in median PFS from 5.5 months to 10 months and an HR=0.55. Predicted accrual is approximately n=7 patients per month and the predicted dropout rate is approximately 15% at 1 year. This sample size is also based on a planned non-binding interim analysis that will be conducted when 60% of the PFS events are observed, by using Lan-DeMets O'Brien-Fleming error spending function for both Type I and Type II errors.

Decision Rules at the interim and the final analyses (Phase 2)

An interim Analysis (IA) will be conducted based on PFS after 34 patients have had an event, which is 60% of the total required number of events (ie, 57 events).

The table below displays the decision rules at the interim and at the end so that the overall Type I error is maintained at one-sided 0.10 and power is approximately 80%. Efficacy in terms of HR would be declared at the time of the IA if the HR surpasses the IA boundary for efficacy that is $HR < 0.533$ (or $p\text{-value} < 0.033$). Futility, in terms of HR, could be declared at the time of IA if the HR surpasses the IA boundary for futility that is $HR > 0.863$.

Look	Information fraction	Efficacy Boundary			Futility Boundary	
		HR scale	Z-score scale	p-value scale	Z-score scale	HR scale
Interim	60% (34 events)	0.533	-1.836	0.033	-0.429	0.863
Final	100% (57 events)	0.701	-1.339	0.067	-1.339	0.701

The study will not stop for efficacy regardless whether or not a statistical significance is found in the interim analysis. The final analysis will be performed when all randomized subjects have either completed study participation period or withdrawn early or should the study be stopped prematurely due to safety reasons. If statistical significance is claimed at the interim analysis, the p-values at the interim will be considered as the primary p-values for the study and the efficacy data after the interim analysis will be used for summary, supportive purpose. If statistical significance is not claimed at the interim analysis, the efficacy boundaries for the final analysis will apply as the inferential base. In the example that an interim analysis is done and no statistical significance is detected at the interim, the statistical significance will be claimed at the final analysis if the p-value is less than or equal to 0.067.

The futility analysis is non-blinding, ie, the decision whether to stop the trial for futility is not bound to the futility boundary in the table. Even the HR surpasses the IA boundary for futility that is $HR > 0.863$, the trial may not stop. Other related data including other efficacy (such as ORR and OS etc.), biomarker data and safety data may be weighed into the decision of stopping the trial for futility.

Both central review and investigator assessment efficacy data will be analyzed in the interim analysis. If the two analyses lead to inconsistent conclusions, the claim of efficacy or futility may not be made and the trial may continue as planned.

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, Phase 1b and Phase 2 data will be analyzed separately in all analyses. 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 (DOR, PFS, TTP and OS etc) curves between the two treatment groups will be compared with a log-rank test at significance level of 0.1 (one-sided). Cox model will be used to estimate the hazard ratio as well as 2-sided 80% and 95% confidence intervals, and may be used to explore the potential influences of the baseline factors (eg, age, gender and ethnic origin) on the time to event data. The time to event data will also be summarized using the

DMB02-GSOP-RF02 3.0 STATISTICAL ANALYSIS PLAN TEMPLATE 30-Jun-2015

Kaplan-Meier method [2]. The Kaplan-Meier estimates for the 1-y PFS and OS rates, and the 2-sided 80% and 95% confidence intervals of the rates using the Greenwood's formula [4] will be reported.

5.3.2. Analyses for Binary Data

Point estimates of binary endpoints for each treatment arm will be provided along with the corresponding 2-sided 95% confidence intervals using an exact method (Blyth and Still (1983) [5]). The difference between the two treatments in Phase 2 will be estimated and the corresponding 80% and 95% intervals will be constructed using an exact method (Chan and Zhang (1999) [6]). Logistic regression may be employed to compare the two treatments in Phase 2.

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. Linear or non-linear models may be employed to analyze the continuous data.

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 2](#).

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 will be replaced with the value for the lower limit of quantification.).

Deviations, missing concentrations and anomalous values

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

In drug concentration summary tables and plots of median PK profiles (plots of PK profiles will only be done for Phase 1b patients), 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 (due for example to an unexpected event such as vomiting before all the drug is 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.

5.4.4. Pharmacodynamic and health outcome parameters

Missing data for the pharmacodynamic and health outcome parameters will be treated as such and no imputed values will be derived.

6. ANALYSES AND SUMMARIES

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.

DMB02-GSOP-RF02 3.0 STATISTICAL ANALYSIS PLAN TEMPLATE 30-Jun-2015

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.

Treatment Administration/Compliance

Unless otherwise stated, listings and tables by treatment will be provided and within each treatment, the summary for each drug, ie, PF-04136309, nab-paclitaxel, or gemcitabine respectively within treatment, will be presented. Cycle length is 28 days. Day 1 of a cycle is the first date of dose within that cycle. The safety analysis set will be used.

One may refer to Protocol Section 5.4.5 and 5.4.6 for dosing interruptions, delays and reductions for PF-04136309, nab-paclitaxel, or gemcitabine respectively.

The following will be summarized for each treatment and within treatment for each drug (PF-04136309, nab-paclitaxel, or gemcitabine respectively). For phase 1b data the overall summary across treatment will also be presented.

- Number of subjects
- 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
- 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
- Time (in months) on treatment (median, range)

The following will be summarized for each treatment and within treatment for each drug (PF-04136309, nab-paclitaxel, or gemcitabine respectively). For phase 1b data the overall summary across treatment will also be presented.

- 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)

The following will be summarized by cycle. For phase 1b data the overall summary across treatment will also be presented.

- For nab-paclitaxel and gemcitabine respectively, summary statistics (mean, median, standard deviation and range) of cumulative prescribed dose, cumulative actual dose, relative dose intensity (%), and percent of compliance will be provided by cycle, where relative dose intensity (%) = cumulative prescribed dose / (planned dose * planned # of treatment administrations) * 100%, and percent of compliance (%) = cumulative actual dose / cumulative prescribed dose * 100%.
- For PF-04136309, the percent of compliance by cycle is calculated as percent of compliance (%) = (# of days on PF-04136309) / (# of days should be on PF-04136309) * 100%

Listings by subject (ordered by treatment, and within treatment by drug): 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 treatment, and within treatment by drug): 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 Endpoints

6.2.1. DLT(Part 1)

Dose Limiting Toxicity (DLT) is the primary endpoint of the dose finding component of the study, which will be summarized by treatment (PF dose level) using the Per Protocol Analysis Set for patients in the dose finding 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.

6.2.2. Safety Endpoints

Adverse Events

Adverse Events (AEs) will be graded by the investigator according to the CTCAE version 4.03 and coded using the MedDRA [7]. 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) by treatment in Phase 1b and Phase 2 respectively, and for Phase 1b the overall summary will also be presented. Note that AEs are the primary endpoints in phase 1b, but the secondary endpoints in phase 2. However the analysis methods are the same for both so we put them together here. 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 each treatment for Phase 1b the overall summary will also be presented. 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. Note that laboratory test abnormalities are the primary endpoints in phase 1b, but the secondary endpoints in phase 2. However the analysis methods are the same for both so we put them together here. Pfizer standard on safety data reporting will be followed.

6.2.3. PFS (Phase 2)

PFS is defined as the time from randomization to first documentation of objective tumor progression or to death due to any cause, whichever occurs first. Censoring of the PFS data is described in [Appendix 2](#). PFS will be compared between the 2 treatment arms using a one-sided log-rank test with a significance level of 0.1 (two-sided $\alpha=0.2$). Cox model will be used to estimate the hazard ratio as well as 2-sided 80% and 95% confidence intervals, and may be used to explore the potential influences of the baseline factors (eg, age, gender and ethnic origin) on PFS. PFS will also be summarized using the Kaplan-Meier method. The Kaplan-Meier estimates for the 1-y PFS rates, and the 2-sided 80% and 95% confidence intervals of the rates using the Greenwood's formula will be reported. Primary analysis of PFS will be based on central review data and sensitivity analyses will also be conducted using investigational site tumor assessments. PFS analysis will be conducted in the mITT population.

6.3. Analysis for Other Endpoints

6.3.1. Other Efficacy Endpoints Analysis

Summary tables of the efficacy data including Objective Response Rate (ORR), Progression Free Survival (PFS), and Overall Survival (OS), and Duration of Response (DOR) will be provided by the groups aforementioned. Efficacy listings will be provided that include best response, first CR/PR date, last date with CR or PR, most recent date without progression, progression date, death date, date of first response 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.

Objective Response Rate (ORR), Complete Response (CR) and Partial Response (PR)

ORR is defined as the proportion of patients with confirmed complete response (CR) or confirmed partial response (PR) according to RECIST, relative to all randomized patients who have baseline measurable disease. Confirmed responses are those that persist on repeat imaging study ≥ 4 weeks after initial documentation of response. Patients who do not have on-study radiographic tumor re-evaluation or who die, progress, or drop out for any reason prior to reaching a CR or PR will be counted as non-responders in the assessment of ORR. A patient, who initially meets the criteria for a PR and then subsequently becomes a confirmed CR, will be assigned a best response of CR. ORR, CR and PR point estimates for each treatment arm will be provided along with the corresponding 2-sided 95% confidence intervals using an exact method (Blyth and Still (1983)). The difference between the two treatments in Phase 2 will be estimated and the corresponding 80% and 95% confidence intervals will be constructed using an exact method (Chan and Zhang (1999)).

Overall Survival (OS)

Time to death is defined as the time from date of randomization to death due to any cause. Patients will be censored for this endpoint on the date of the last tumor assessment if they do not die at that time. A log-rank test will be used to compare the treatment groups with respect to PFS at a 2-sided alpha level of 0.20. Overall survival will also be summarized using the Kaplan-Meier method. The Kaplan-Meier estimates for the 1-year OS rates, and the 2-sided 80% and 95% confidence intervals of the rates using the Greenwood's formula will be reported. In addition, Cox model will be used to estimate the hazard ratio and its 80% and 95% CIs for the treatment effect. The analysis for overall survival (OS) will be based on the ITT population.

Note that although this study is not powered to detect significant differences in Overall Survival at the time of the primary analyses of PFS, analyses of OS, with a statistical power = 63%, will be conducted after 50 deaths are observed, assuming an HR=0.63 and improvement in median OS from 8.5 months to 13.5 months. First type error will not be inflated as this statistical comparison will be conducted conditionally upon rejection of the null hypothesis for the PFS endpoint.

Duration of Objective Response (DoR) and Duration of Stable Disease (DoSD)

DMB02-GSOP-RF02 3.0 STATISTICAL ANALYSIS PLAN TEMPLATE 30-Jun-2015

DoR is defined as the time from the first confirmed documentation of objective tumor response to the first documentation of objective tumor progression or to death due to any cause, whichever occurs first. Censoring of the DOR data is the same as for PFS. DOR will only be calculated for the subgroup of patients with objective response. For Phase 2 data, a log-rank test will be used to compare the treatment groups with respect to DOR at a 2-sided alpha level of 0.20. In addition, Cox model will be used to estimate the hazard ratio and its 80% and 95% CIs for the treatment effect for Phase 2 data. The Kaplan-Meier estimates for the $\text{DoR} \geq 16$ weeks, and the 2-sided 80% and 95% confidence intervals of the rates using the Greenwood's formula will be reported. DOR will only be calculated for the subgroup of patients with response (CR or PR).

DoSD is defined as the time from the first documentation of stable disease to the first documentation of objective tumor progression or to death due to any cause, whichever occurs first. DoSD will be analyzed in the similar fashion as DoR. DoSD will only be calculated for the subgroup of patients with CR, PR or stable disease (SD).

Time to progression (TTP) with the double combination PF-04136309 and gemcitabine (maintenance therapy) after interruption of nab-paclitaxel

This type of TTP is defined as the time from interruption of nab-paclitaxel to the first documentation of objective tumor progression. Censoring of the TTP data is described in [Appendix 2](#). The Kaplan-Meier survival estimates, together with the number of subjects, percentage of subjects to experience the event, and the number and percentage of subjects censored will be summarized. This TTP is only calculated for the subgroup of patients who were treated with the maintenance therapy.

6.3.2. Pharmacokinetics Analyses

The concentration-time data of PF-04136309 will be summarized by descriptive statistics (n, mean, standard deviation, median, minimum, and maximum) according to dose level.

For patients enrolled in the Phase 1 dose finding cohorts, the concentration-time data of PF-04136309 within the 12-hour dose interval on Day 15 (with the pre-dose sample collected on Day 16 as the 12-hour post dose sample) will be analyzed for individual patients using non-compartmental methods. The noncompartmental analysis will estimate PK parameters including $C_{\text{max,ss}}$, T_{max} , $\text{AUC}_{\text{tau,ss}}$, $C_{\text{min,ss}}$, $\text{CL}_{\text{ss/F}}$, and as data permit, $t_{1/2}$ and $V_{\text{ss/F}}$. The PK parameters will be summarized using descriptive statistics according to dosing cohort.

The concentration-time data of PF-04136309 from all patients enrolled in the study will be using nonlinear mixed effect modeling. The modeling-based analysis will estimate the typical value and variability for PK parameters including oral clearance (CL/F) and apparent volume of distribution (V/F). Also, the influence of selected potential covariates on the PK parameters will be explored as appropriate; the potential covariates to be explored will include selected demographics (eg, body weight, age, sex) and selected patient characteristics (eg, baseline CCR2 expression, baseline CCL2 level, ECOG performance status, liver function markers).

Population PK assessment will be conducted with the concentration-time data of PF-04136309 from all patients, using the nonlinear mixed effect modeling approach in accordance with regulatory guidances. A structural PK model based on data from the Phase 1b portion will be used. The population PK analysis will estimate typical value and variability for PK parameters including oral clearance (CL/F) and apparent volume of distribution (Vd/F). Also, the influence of selected potential covariates on the PK parameters will be explored as appropriate; the potential covariates to be explored will include selected demographics (eg, body weight, age, sex) and selected patient characteristics (eg, baseline CCR2 expression, baseline CCL2 level, ECOG performance status, liver function markers).

The detailed procedures for the population PK analysis, including the model implementation and evaluation, will be described prospectively in the Population Modeling Analysis Plan (PMAP). The results of the analysis will be summarized in a Population Modeling and Analysis Report (PMAR), separate from the clinical study report.

6.3.3. Analysis of Pharmacodynamics

For FNA and core-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. For each pair of specimens, the mean change and/or mean percent change from baseline of these same parameters will also be calculated and the corresponding error bar plots over time be displayed. Other biomarker assay data will be analyzed in a similar fashion.

The correlations between FNA, core-biopsy samples and other biomarker results, and pharmacokinetic parameters and measures of anti-tumor/anti-cancer efficacy will be examined. And the modeling of pharmacokinetic and pharmacodynamics data as well as anti-tumor/anti-cancer efficacy data may include biomarker data as covariates. Pharmacodynamic biomarkers submitted to these analyses may include, but not be limited to:

- Ex vivo inhibition of CCL2-induced ERK kinase phosphorylation in the peripheral blood.
- Phenotype, relative proportions and CCR2 expression of IMs, TAMs, and other relevant immune cell subsets in the bone marrow, peripheral blood and tumor tissue.
- CCL2 levels in the peripheral blood.
- Relative expression of immune-related transcripts in tumor tissue and peripheral blood.
- Abundance and diversity of tumor-infiltrating and peripheral blood T cell clones.
- Prevalence and diversity of tumor antigenic epitopes.
- Presence or levels of expression of several tumor biomarkers and their relationship to clinical responsiveness (including but not limited to SPARC, tumor markers, CA 19.9).

6.3.4. Population Pharmacokinetic Analysis or Pharmacokinetic/Pharmacodynamic (PK/PD) Modeling (Phase 1b and Phase 2)

Pharmacokinetic and pharmacodynamic 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-04136309 exposure and biomarkers or significant safety endpoints. The results of these analyses, if performed, may be reported separately.

6.3.5. Analysis of Peripheral Neurological Adverse Events and EORTC QOL-C30 and QLQ-CIPN20 Questionnaire (Phase 2)

Summary statistics, including frequencies, means, medians, range will be reported by treatment for peripheral neurological adverse events according to NCI CTCAE v.4.03. Times to neurotoxic events may be analyzed by the Kaplan- Meier method.

Response to all questions of EORTC QOL-30 and QLQ-CIPN20 questionnaires will be shown in percentage. Logistic regression analyses may be conducted at each visit to detect potential differences regarding neuropathy problems at the corresponding visit (EORTC QLQ-CIPN20 or QOL-30 (except the last two questions on global health status); answer categories “quite a bit” or “very much” combined) between the two treatments in Phase 2.

Summary statistics will be provided by treatment for the raw scores of functional, Symptom and global health status raw subscales in QOL-C30 as well as for linearly converted scores on the scale of 0-100 for these subscales. Cronbach's alpha coefficient will be presented for each subscale in QOL-C30 and the correlations between the subscales will be also reported. Note The raw score is: $\text{RawScore} = \text{RS} = (I_1 + I_2 + \dots + I_n) / n$, where I_1 , I_2 , and I_n are the n component in the categories of functional, symptom or global health status scales.

The linearly converted score (LCS) for Functional scales is defined as:

$$\text{Linearly Converged Score} = \text{LCS} = (1 - (\text{RS}-1)/\text{Range})) \times 100,$$

And for Symptom scales or Global health status:

$$\text{Linearly Converged Score} = \text{LCS} = (\text{RS}-1)/\text{Range} \times 100$$

Summary statistics will also be provided for the raw scores of sensory, motor and autonomic scales raw subscales in QLQ-CIPN20 as well as for linearly converted scores on the scale of 0-100 for these subscales. Cronbach's alpha coefficient will be presented for each subscale in QLQ-CIPN20 and the correlations between the subscales will be also reported. For each of the scales in QLQ-CIPN20,

$$\text{Linearly Converged Score} = \text{LCS} = (\text{RS}-1)/\text{Range} \times 100$$

ANCOVA models which include covariates of treatment and baseline may be employed to compare the two treatments by visit in each of the three-category scales of QOL-C30 and each of the three-category scales of QLQ-CIPN20.

6.4. 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 [8].

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 will 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 3](#).

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

An interim Analysis (IA) will be conducted in the Phase 2 part of the study based on PFS after 34 patients have had an event, which is 60% of the total required number of events (ie, 57 events). The decision rules are detailed in [Section 5.2.2](#).

An external Data Monitoring Committee (DMC) will be established for the Phase 2 portion study. The external DMC will evaluate efficacy and safety data and make a recommendation to Pfizer according to the decision rules as specified in [Section 5.2.2](#). The logistical issues on the interim analysis will be detailed in DMC charter.

8. REFERENCES

1. A9421018 Protocol Amendment 1, Nov. 25, 2015.
2. Ji Y, Wang SJ., Modified toxicity probability interval design: a safer and more reliable method than the 3 + 3 design for practical phase I trials, *J Clin Oncol.*, 2013; 10;31(14):1785-91.
3. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958; 53:457-481.
4. Greenwood, M. (1926) The natural duration of cancer. Reports on Public Health and Medical Subjects 33, 1–26. Her Majesty's Stationery Office, London.
5. Blyth, C. R. and Still, H. A. (1983). Binomial Confidence Intervals. *Journal of the American Statistical Association* **78**, 108–116.
6. Chan, I. S. F. and Zhang, Z. (1999), "Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions," *Biometrics*, 55, 1202–1209.
7. 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
8. ICH E14 - The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. CHMP/ICH/2/04.
9. E.A. Eisenhauer, P. Therasse, J. Bogaerts, L.H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbu, 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
10. 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.
11. 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.
12. Guha M: The new era of immune checkpoint inhibitors; *Pharm J* Volume 293, 18 NOV 2014, No 7837/8 online URL 200671127

9. APPENDICES

Appendix 1. Details of Definitions of Endpoints

DLT Definitions

DLTs will be defined as any of the following events described below that occur in the first cycle of treatment (Day 1 through 28) and are attributed (ie, judged to be at least possibly related) to the combination of PF-04136309 with nab-paclitaxel and gemcitabine where relationship with the combination cannot be ruled out. DLTs will be classified according to CTCAE version 4.03.

Hematologic:

- Grade 4 neutropenia lasting >5 days;
- Febrile neutropenia [defined as absolute neutrophils count (ANC) <1,000/mm³ with a single temperature of >38.3°C (101°F) or a sustained temperature of ≥38°C (100.4°F) for more than one hour];
- Grade ≥3 neutropenic infection;
- Grade ≥3 thrombocytopenia with Grade ≥2 bleeding;
- Grade 4 thrombocytopenia.

Non-Hematologic:

- Grade 3 toxicities, except:
 - Nausea and vomiting responding to prophylaxis and/or treatment and lasting less than 7 days from each chemotherapy infusion period;
 - Diarrhea responding to treatment and lasting less than 7 days;
 - Grade 3 QTc prolongation (QTc >500 msec) will be considered a DLT only if persisting after correction of any reversible causes. So, Grade 3 QTc prolongation will first require repeat testing, re-evaluation by a qualified person, and correction of reversible causes such as electrolyte abnormalities or hypoxia for confirmation;
 - Grade 3 aspartate aminotransferase (AST) and/or alanine aminotransferase (ALT) increase lasting ≤7 days.
- All Grade 4 toxicities;

- Delay of more than 2 weeks in receiving the next scheduled cycle due to persisting treatment-related toxicities.

A subject is classified as DLT-evaluable if he/she experiences a DLT or if he/she otherwise in the absence of a DLT receives at least 85% of the planned doses of each study drug in the first 28 day cycle. If a subject is withdrawn from study for any reason other than a DLT prior to completion of the 28-day safety observation period, a replacement subject will be assigned the same dose as the replaced subject.

Adverse Effects of Interest are defined as those adverse events that fulfill the same grade and terms that would qualify as a DLT but occur beyond the protocol defined 28-day DLT window. Any Adverse Event of Interest that occurs should be reported to Pfizer as they occur and entered into the electronic case report form (eCRF).

Late Onset Toxicity and Toxicities

Adverse events that meet the same grading criteria as the DLT criteria listed above occurring after the DLT observation period will lead Pfizer to immediately schedule a meeting with the investigators to review the details of the potential late onset toxicity and determine if the enrollment has to be held for this dose level or if a dose reduction should be implemented for all ongoing patients. Late onset toxicities meeting the definition of a DLT will be used in the evaluation of the MTD.

Maximum tolerated dose (MTD) and Recommended Phase 2 Dose (RP2D)

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%]. However, in this study, MTD and RP2D may be determined after a full evaluation of safety and efficacy data from patients treated in the escalation phase of the study. See details in [Section 5.2.1](#).

Appendix 2. Time to Event Data Analysis Censoring Rules**Table 2. Progression Free Survival and Duration of Response**

Situation	Date of Progression/Censoring¹	Outcome
Inadequate baseline assessment	Randomization date	Censored
No on-study assessments	Randomization date	Censored
Alive, on treatment ² and no Progression	Date of last objective tumor assessment	Censored
Progression Documented on or between scheduled tumor assessments prior to treatment discontinuation ²	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 3. Time to Progression with the double combination PF-04136309 and gemcitabine (maintenance therapy) after interruption of nab-paclitaxel

Situation	Date of Progression/Censoring¹	Outcome
Inadequate assessment at the start of maintenance therapy	the start of maintenance therapy	Censored
No on-study assessments after the start of maintenance therapy	the start of maintenance therapy	Censored
Alive, on treatment ² and no Progression	Date of last objective tumor assessment in the maintenance therapy period	Censored
Progression Documented on or between scheduled tumor assessments prior to treatment	Date of first objective tumor assessment showing objective progression in the maintenance	Progressed (Event)

DMB02-GSOP-RF02 3.0 STATISTICAL ANALYSIS PLAN TEMPLATE 30-Jun-2015

Table 3. Time to Progression with the double combination PF-04136309 and gemcitabine (maintenance therapy) after interruption of nab-paclitaxel

Situation	Date of Progression/Censoring ¹	Outcome
discontinuation ²	therapy period	
Treatment discontinuation for undocumented progression in the maintenance therapy period	Date of last objective tumor assessment prior to discontinuation ² in the maintenance therapy period	Censored
Treatment discontinuation due to toxicity or other reason in the maintenance therapy period	Date of last objective tumor assessment prior to discontinuation ² in the maintenance therapy period	Censored
New anticancer treatment <28 days after discontinuation of treatment without progression in the maintenance therapy period	Date of last objective tumor assessment prior to new anticancer treatment in the maintenance therapy period	Censored
Death prior to first planned tumor assessment in the maintenance therapy period	Start date of the maintenance therapy period	Censored
Death without objective progression prior to treatment discontinuation ² in the maintenance therapy period	Date of last objective tumor assessment prior to death in the maintenance therapy period	Censored
Progression after 2 or more missed tumor assessments in the maintenance therapy period	Date of last objective tumor assessment prior to the event in the maintenance therapy period	Censored

1: 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.

2: 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 \leq max. < 60	max. ≥ 60	

Categories for PR and QRS

PR (ms)	max ≥ 300	
PR (ms) increase from baseline	Baseline > 200 and max. $\geq 25\%$ increase	Baseline ≤ 200 and max. $\geq 50\%$ increase
QRS (ms)	max ≥ 200	
QRS (ms) increase from baseline	Baseline > 100 and max. $\geq 25\%$ increase	Baseline ≤ 100 and max. $\geq 50\%$ 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. Therasse, 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 [9].

CATEGORIZING LESIONS AT BASELINE

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 disease

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 subjected 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 as close 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 be indeterminate.

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

Target lesions

All measurable lesions up to a maximum of 2 lesions per organ, 5 lesions in total, representative of all involved organs, should be identified as target lesions at baseline. Target lesions should be selected on the basis of size (longest lesions) and suitability for accurate repeated measurements. Record the longest diameter for each lesion, except in the case of pathological lymph nodes for which the short axis should be recorded. The sum of the diameters (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions at baseline will be the basis for comparison to assessments performed on study.

- If two target lesions coalesce the measurement of the coalesced mass is used. If a large target lesion splits, the sum of the parts is used.
- Measurements for target lesions that become small should continue to be recorded. If a target lesion becomes too small to measure, 0 mm should be recorded if the lesion is considered to have disappeared; otherwise a default value of 5 mm should be recorded.

NOTE: When nodal lesions decrease to <10 mm (normal), the actual measurement should still be recorded.

Non-target disease

All non-measurable disease is non-target. All measurable lesions not identified as target lesions are also included as non-target disease. Measurements are not required but rather assessments will be expressed as ABSENT, INDETERMINATE, PRESENT/NOT INCREASED, INCREASED. Multiple non-target lesions in one organ may be recorded as a single item on the case report form (eg, 'multiple enlarged pelvic lymph nodes' or 'multiple liver metastases').

OBJECTIVE RESPONSE STATUS AT EACH EVALUATION

Disease sites must be assessed using the same technique as baseline, including consistent administration of contrast and timing of scanning. If a change needs to be made the case must be discussed with the radiologist to determine if substitution is possible. If not, subsequent objective statuses are indeterminate.

Target disease

- Complete Response (CR): Complete disappearance of all target lesions with the exception 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: 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 disease

- CR: Disappearance of all non-target lesions and normalization of tumor marker levels. All lymph nodes must be 'normal' 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 is equivocal, 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 to necrosis, 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 after discontinuation of treatment.

Table 4. Objective Response Status at each Evaluation			
Target Lesions	Non-target Disease	New Lesions	Objective status
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	Stable
Indeterminate or Missing	Non-PD	No	Indeterminate

Table 4. Objective Response Status at each Evaluation			
Target Lesions	Non-target Disease	New Lesions	Objective status
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

If the protocol allows enrollment of patients with only non-target disease, the following table will be used:

Table 5. 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