STATISTICAL ANALYSIS PLAN

A Phase 1b Study of Talimogene Laherparepvec in Combination With Atezolizumab in Subjects With Triple Negative Breast Cancer and Colorectal Cancer With Liver Metastases

Protocol Number: **20140299**

Version: **2.0**

Date: **21 July 2016**

Amendment 1 Date: **25 November 2019**

Authors: [redacted]

NCT Number: NCT03256344
This NCT number has been applied to the document for purposes of posting on Clinicaltrials.gov

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	21Jun2016	Not applicable
Amendment 1 (v2.0)	25Nov2019	● Specify the censoring rule of efficacy endpoints. ● Update the definition of the maximum tumor burden decrease. ● Update the definition of durable response rate. ● Remove the definition of conventional durable response rate. ● Add response at the lesion-level definitions L-CR, L-PR and L-ORR. ● Add definitions baseline lesion and new lesion. ● Add lesion analysis sets, including the injected lesion analysis set, the uninjected lesion analysis set, the injected lesion evaluable and the uninjected lesion evaluable analysis sets. ● Add section 8 “Planned Analyses”. ● Add exploratory endpoints to be consistent with protocol amendment 5.0. ● Update “Table 1. Cumulative Increments of DLT-evaluable Subjects and Corresponding Acceptable Maximum Number and Percentage of DLTs at Each Interim Safety Analysis” to be consistent with protocol amendment 5.0. ● Update the modified irRC-RECIST criteria to be consistent with protocol amendment 5.0. ● Make administrative and editorial changes to be consistent with protocol amendment 5.0.

Table of Contents

List of Tables

List of Appendices

Table of Abbreviations

Abbreviation/Acronym	Definition
AE	Adverse Event
BC	Breast Adenocarcinoma
BOR	Best Overall Response
CI	Confidence Interval
CR	Complete Response
CRC	Colorectal Adenocarcinoma
CRF	Case Report Form
CT	Computerized Tomography
CTC	Common Toxicity Criteria
CTCAE	Common Toxicity Criteria for Adverse Events
DCR	Disease Control Rate
DLRT	Dose Level Review Team
DLT	Dose-Limiting Toxicity
DMP	Data Management Plan
DNA	Deoxyribonucleic Acid
DOR	Duration of Response
DRR	Durable Response Rate
DTP	Data Transfer Plan
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EOI	Event of Interest
HR	Hazard Ratio
HSV	Herpes Simplex Virus
HSV-1	Herpes Simplex Virus type 1
IP	Investigational Product
IPD	Important Protocol Deviation
irRC-RECIST	Immune-related Response Criteria (irRC) simulating RECIST version 1.1
IVRS	Interactive Voice Response System
KM	Kaplan-Meier
L-CR	Lesion Complete Response
L-ORR	Lesion Objective Response Rate
L-PR	Lesion Partial Response
MedDRA	Medical Dictionary for Regulatory Activities
MTC	Maximum Tolerated Concentration
MTD	Maximum Tolerated Dose
MTV	Maximum Tolerated Volume
ND	Not done
OR	Objective Response (CR or PR)

Abbreviation/Acronym	Definition
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PFS	Progression Free Survival
PFU	Plaque-Forming Unit
PR	Partial Response
PT	Preferred Term
Q21D	Every 21 days
qPCR	Real-time Polymerase Chain Reaction
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Stable Disease
SMQ	Standardized MedDRA Queries
SOC	System Organ Class
SOP	Standard Operating Procedure
TTF	Time to Treatment Failure
UE	Unevaluable for tumor response
ULN	Upper Limit of Normal
USA	United States of America
WHO	World Health Organization
WHODRUG	World Health Organization Drug Dictionary

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol amendment 5 for study 20140299 Talimogene Laherparepvec, dated 08 July 2019. The scope of this plan includes the primary analysis and the final analysis that is planned and will be executed by the Biostatistics department unless otherwise specified. [redacted]

2. Objectives

Primary Objective:

To evaluate the safety, as assessed by incidence of dose limiting toxicities (DLTs), of intrahepatic injection of talimogene laherparepvec into liver metastases in combination with intravenously administered atezolizumab separately in subjects with triple negative breast cancer and colorectal cancer.

Secondary Objectives:

To evaluate the efficacy of talimogene laherparepvec in combination with atezolizumab separately in subjects with triple negative breast cancer and colorectal cancer with liver metastases as assessed by:

- Objective response rate (ORR), best overall response (BOR), duration of response (DOR), lesion level responses in injected and uninjected tumor lesions (overall, hepatic, nonhepatic), disease control rate (DCR), durable response rate (DRR), progression-free survival (PFS), overall survival (OS) by cohort (triple negative breast cancer and colorectal cancer).

Safety Objective:

To evaluate the safety and tolerability of intrahepatic injection of talimogene laherparepvec into liver metastases in combination with intravenously administered atezolizumab separately in subjects with triple negative breast cancer and colorectal cancer.

Exploratory Objective:

The exploratory objectives of the study are as follows:

- To evaluate changes in tumor inflammation markers in tumor biopsies, such as programmed cell death ligand 1 (PD-L1) expression and cluster of differentiation 8 (CD8) density
- To explore blood and tissue biomarkers that may correlate with or predict treatment effect and/or clinical outcomes separately in subjects in triple negative breast cancer and metastatic colorectal cancer.

3. Study Overview

3.1 Study Design

This is a phase 1b, multicenter, open-label study to evaluate the safety of intrahepatic injection of talimogene laherparepvec in combination with intravenously administered atezolizumab in subjects with triple negative breast cancer and colorectal cancer with liver metastases. Talimogene laherparepvec will be injected intrahepatically in combination with intravenous atezolizumab to approximately 36 DLT-evaluable subjects in 2 parallel cohorts. Cohort 1 will comprise subjects with triple negative breast cancer with liver metastases (n =18 DLT-evaluable subjects). Cohort 2 will comprise subjects with colorectal cancer with unresectable liver metastases (n =18 DLT-evaluable subjects). The DLT evaluation period for a given subject will consist of the period between the initial 10^6 PFU/mL dose of talimogene laherparepvec and atezolizumab and 3 weeks following the initial 10^8 PFU/mL dose of talimogene laherparepvec and atezolizumab or the start of cycle 3, whichever occurs first. DLTs will be evaluated based on the first 18 DLT-evaluable subjects in each cohort separately. A Dose Level Review Team (DLRT) will review the safety data to evaluate possible drug effects and DLT. To be evaluable for a DLT, subjects must have had the opportunity to be on treatment for at least 2 cycles from the initial dose of study treatment and have received at least 2 doses of talimogene laherparepvec and 2 doses of atezolizumab in combination, or have a DLT during the DLT evaluation period. Subjects may be replaced if they are not evaluable for DLT in order to obtain 18 DLT-evaluable subjects. There will be one safety interim analysis after the first 4 to 6 DLT-evaluable subjects have been enrolled in this study and a safety analysis after 18 DLT-evaluable subjects have been enrolled in a cohort. Enrollment will be suspended during the first safety interim analysis. At the discretion of the DLRT, additional safety analyses may be conducted as warranted. Treatment will continue until a subject experiences a DLT (during the DLT evaluation period), has complete response (CR), has need for an alternative anticancer therapy, or experiences an adverse event necessitating drug discontinuation. In addition, treatment will be discontinued for talimogene laherparepvec if the subject has no injectable lesions, upon confirmed progressive disease (PD) per modified immune related response criteria Response Evaluation Criteria in Solid Tumors (irRC-RECIST) or rapid clinical deterioration. Atezolizumab will be discontinued upon symptomatic disease progression. All subjects will complete a safety follow-up visit approximately 30 (+ 7) days after the last dose of study treatment. After the safety follow-up visit, all subjects will enter the long-term follow-up. Subjects will be followed for survival, subsequent anticancer

therapies and treatment-related adverse events every 12 weeks (± 28 days) for approximately 24 months after the last subject is enrolled.

3.2 Sample Size

Approximately 36 subjects will be enrolled (18 subjects in each cohort). The data will be analyzed by cohort. The null hypothesis (H_0) is that the combination of talimogene laherparepvec and atezolizumab has a DLT rate ≤ 10%. An unacceptable alternative hypothesis (H_a) is a true DLT rate ≥ 33%. The sample size goal is to have ≥ 80% power for a 1-sided ≤ 10% significance level test to reject H_0 when H_a is true (Goldman, 1987). Eighteen DLT-evaluable subjects in each cohort will be required to test H_0. Assuming the incidence of DLTs is evaluated as specified by Table 1, this design achieves a 7.7% 1-sided significance level and 81.6% power.

Table 1. Cumulative Increments of DLT-evaluable Subjects and Corresponding Acceptable Maximum Number and Percentage of DLTs at Each Interim Safety Analysis

Number of subjects	Acceptable Maximum Number (%) of DLTs	Stopping Number (%) of DLTs
2	1 (50)	2 (100)
3	1 (33)	2 (66)
4	2 (50)	3 (75)
5	2 (40)	3 (60)
6	2 (33)	3 (50)
7	2 (29)	3 (43)
8	2 (25)	3 (38)
9	3 (33)	4 (44)
10	3 (30)	4 (40)
11	3 (27)	4 (36)
12	3 (25)	4 (33)
13	3 (23)	4 (31)
14	4 (29)	5 (36)
15	4 (27)	5 (33)
16	4 (25)	5 (31)
17	4 (24)	5 (29)
18[a]	4 (22)	5 (28)

DLT = dose limiting toxicity
The design achieves a 7.7% 1-sided significance level and 81.6% power to test the null hypothesis of a DLT rate ≤ 10% versus the alternative hypothesis of a rate ≥ 33%.
[a] If > 18 subjects receive the combination they will contribute to the overall safety analysis, but only the first 18 DLT-evaluable will be considered in the decision to declare the combination safe.

4. Study Endpoints and Covariates

4.1 Study Endpoints

- Primary endpoints: Subject incidence of DLTs by cohort (triple negative breast cancer and colorectal cancer)
- Secondary endpoints: ORR, BOR, DOR, lesion level responses (≥ 30% and 100% decrease) in injected and uninjected lesions (overall, hepatic, nonhepatic), DRR, DCR, PFS, and OS by cohort (triple negative breast cancer and colorectal cancer)
- Safety endpoints: Subject incidence of adverse events and clinically relevant laboratory abnormalities by tumor type (triple negative breast cancer and colorectal cancer)
- Exploratory endpoints: The following endpoints are included as exploratory endpoints and their analysis plan may be described in a separate supplemental SAP if appropriate and they may be reported separately from the primary CSR:
    - Changes in tumor inflammation markers such as PD-L1 analysis and CD8 density
    - Identification of potential blood and tumor biomarkers which correlate with or predict clinical outcomes

4.2 Planned Covariates

For each tumor type the following covariates may be used to examine efficacy and safety in subgroups or in multivariate analyses as appropriate:

- Region, if applicable (USA vs non-USA)
- Age at baseline: < 50, ≥ 50; < 65, ≥ 65; < 75, ≥ 75 years
- Sex (female vs male)
- Prior lines of cancer therapy in metastatic setting (0, 1, 2, >2)
- HSV-1 serostatus (positive versus negative)
- Baseline lactate dehydrogenase (LDH) ≤ ULN vs > ULN
- ECOG performance status (0 versus 1)
- Brain metastases for triple negative breast cancer (yes or no)
- Baseline overall tumor burden (sum of length of baseline measurable lesion diameters)
- Baseline liver tumor burden (sum of length of baseline liver measurable lesion diameters)
- Prior exposure to chemotherapy (yes or no)
- Extrahepatic visceral metastases at baseline (yes or no)
- PD-L1 status (positive versus negative)
- MSI phenotype (yes, no, or unknown) (colorectal cancer cohort only)
- Receipt of non-hepatic talimogene laherparepvec (yes or no)
- Other covariates reported in the literature or from other Amgen/Genentech studies may be considered as appropriate at the time of analysis.

5. Hypotheses and/or Estimations

It is hypothesized that intrahepatic injection of talimogene laherparepvec in combination with intravenously administered atezolizumab in subjects with triple negative breast cancer and colorectal cancer with liver metastases will be safe and well tolerated with a DLT rate ≤ 10%.

6. Definitions

The definition of DLT is provided in Protocol Section 6.2.1.2.3. Additional definitions are as follows:

DLT evaluation period

The DLT evaluation period for a given subject will consist of the period between the initial 10^6 PFU/mL dose of talimogene laherparepvec and atezolizumab and 3 weeks following the initial 10^8 PFU/mL dose of talimogene laherparepvec and atezolizumab or the start of cycle 3, whichever occurs first.

Enrollment date

The date subject is enrolled to the study.

A subject is considered enrolled when the investigator decides that the subject has met all eligibility criteria. The investigator is to document this decision and date, in the subject's medical record and in the enrollment case report form (CRF).

Investigational Product

Investigational product (IP) refers to Amgen investigational product talimogene laherparepvec and non-Amgen investigational product atezolizumab in this study.

Last IP Dose Date

Last IP Dose Date for each subject is defined as the latest date IP is administered in this study.

Screening Phase

The screening phase is the time period after subject signing off the informed consent form and before enrollment when study-specific laboratory tests and procedures are performed, and medical history is reviewed to confirm subject eligibility for the study.

Study Day

Study day is calculated from the first day IP is administered to the subject.
Study day = visit date – first dose date +1 if visit date is on or after the first dose date.
Study day = visit date – first dose date, if visit date is before the first dose date.

Study Day 1

Study day 1 is the first day that protocol-specified investigational products are administered to the subject. The day before study day 1 is study day '-1'.

Baseline

Baseline in general refers to study day 1. The baseline value of a parameter (eg, vital signs, laboratory tests, and tumor measurement) is considered to be the latest value prior to receiving the study drugs (ie, on or prior to the first date of dosing).

Best overall response (BOR) per Modified irRC-RECIST

Best overall response (BOR) of complete response (CR), partial response (PR), stable disease (SD), progressive disease (PD) or un-evaluable (UE) will be derived based on investigator assessment using modified irRC-RECIST as defined in protocol Appendix D. Any baseline tumor markers above the ULN must normalize for an overall visit response of CR.

Overall visit response assessments occurring on or after the start of the first subsequent anticancer therapy, as well as those occurring on or after complete or partial removal/reduction of any target lesion which contained cancer on pathology evaluation or pathology results were unknown, will not be included. Consecutive confirmation of CR, PR, and PD is required where the second overall response must be at least 28 days from the first, the only exception being when the investigator reports that an initial PD cannot be confirmed due to rapid clinical deterioration. The following will be considered a confirmed PR: CR followed by PR or PR followed by CR.

BOR is defined as the best visit response in the following order: CR, PR, SD, PD, or UE (see Table 2). BOR is defined as SD if the best overall visit response is an unconfirmed CR or PR, and UE if it is either SD earlier than 63 days after the date of first dosing or an unconfirmed PD when confirmation of PD is required (ie, initial PD without rapid clinical deterioration).

Table 2. Matrix of Determining BOR per Modified irRC-RECIST

Visit Overall Response Sequence	Examples	Best Overall Response	Confirmation specifications
, CR, CR,	PR, CR, CR CR, CR, PD	CR	The confirmatory CR must be at least 4 weeks (28 days) later; a subsequent CR within 28 days will not be valid for confirmation and will be ignored; the CR will also not be confirmed if there is a subsequent PR/SD/PD at any time prior to the next CR.
*, PR, PR, * *, PR, CR/PR, non-CR, * *,CR, PR,*	PR, PR, PD PR, CR, PD CR, PR, PD	PR	Criteria for BOR=CR not met. The confirmatory PR/CR must be no less than 4 weeks (28 days) later; a subsequent PR/CR within 28 days will not be valid for confirmation and will be ignored; the PR will also not be confirmed if there is a subsequent SD/PD at any time prior to the next PR/CR .
*, SD, * *,CR, non-PR/CR, * *,PR, non-PR/CR, *	CR PD, CR PR PD, PR, SD SD PD,SD,PD	SD	Criteria for BOR=CR or PR not met. SD must be ≥ 63 days from first dose; however, this is not required for an unconfirmed CR/PR.
*,PD, PD, * *, PDr, * PDr = PD with concurrent or subsequent rapid clinical deterioration as the reason for ending radiographic follow-up.	PD, PD PD, SD, PD, PD PDr	PD	Criteria for BOR= CR, PR, or SD not met. The confirmatory PD must be no less than 4 weeks (28 days) later unless there is rapid clinical deterioration; PDr does not require confirmation; a subsequent PD within 28 days will not be valid for confirmation and will be ignored.
*, SD, * PD	SD UE, SD PD UE, PD	UE	Criteria for BOR=CR, PR, SD, or PD not met. SD must be < 63 days from first dose.

Duration of Response (DOR)

Duration of response is defined as the time from the date of an initial response that is subsequently confirmed to the earlier of PD (see definition of BOR per modified irRC-RECIST) or death. Subjects who have not ended their response at the time of analysis will be censored at the last evaluable tumor assessment.

Durable Response Rate (DRR)

Durable response rate is defined as the percent of subjects with a CR or PR per modified irRC-RECIST with a DOR ≥ 6 months. One month will be calculated based on 365.25 days per year.

Disease Control Rate (**DCR** also called the **Clinical benefit response rate**)

Disease control rate is the proportion of subjects that have a best overall response of confirmed CR/ PR or SD.

Objective Response Rate (ORR)

The subject incidence rate of a best overall response of either confirmed CR or PR.

Baseline Lesion

Baseline lesion is defined as the lesions identified on or before study day 1 with a positive tumor length. If a baseline lesion has a positive tumor length reported at more than one visit, then the length from the most recent visit will be selected as the baseline length.

New Lesion

New lesion is defined as the lesions that first appear after study day 1 with a positive tumor length. “Baseline” (ie, initial) tumor length of a new lesion will be obtained from the earliest date with a positive tumor length. Lesions that split from other lesions will not be considered as new lesions.

Lesion complete response (L-CR)

Lesion complete response (L-CR) is a response at the lesion level in injected vs uninjected lesions and applies only to target lesions and new measurable lesions and is defined as complete disappearance of a lesion with confirmation by a repeat consecutive assessment no less than 4 weeks (28 days) from the date first documented.

Lesion partial response (L-PR)

Lesion partial response (L-PR) is a response at the lesion level in injected vs uninjected lesions and applies only to target lesions and new measurable lesions and is defined as ≥ 30% decrease in lesion length relative to baseline with confirmation by a repeat consecutive assessment no less than 4 weeks (28 days) from the date first documented.

Lesion objective response rate (L-ORR)

Lesion objective response rate (L-ORR) is a response rate at the lesion level in injected vs uninjected lesions and is defined as the incidence among a set of lesions analyzed of L-CR or L-PR.

Response at the lesion-level will be censored on or after the earliest event of start of non-study anticancer therapy, merged with another lesion, or complete or partial removal/reduction of any target lesion which contained cancer on pathology evaluation or pathology results were unknown. If a lesion splits, the sum of the length of the split lesions will be added together prior to analysis.

Maximum Tumor Burden Decrease

Tumor burden is defined as the sum of the diameters (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions and up to 10 (maximum 5 per organ) new measurable lesions. Tumor burden ratio is defined as the ratio of the tumor burden at each assessment divided by the baseline tumor burden. The Maximum tumor burden decrease is defined as smallest tumor burden ratio among all assessments -1.

The maximum tumor burden decrease at the subject-level will be censored on or after the earliest event of start of non-study anticancer therapy, or complete or partial removal/reduction of any target lesion which contained cancer on pathology evaluation or pathology results were unknown.

The maximum tumor burden decrease at the lesion-level will be censored on or after the earliest event of start of non-study anticancer therapy, merged with another lesion, or complete or partial removal/reduction of any target lesion which contained cancer on pathology evaluation or pathology results were unknown. If a lesion splits, the sum of the length of the split lesions will be added together prior to analysis.

Overall Survival (OS)

Overall survival (OS) is defined as the time from the date of first dose to the date of death from any cause. Death is the event of interest. OS time will be censored at the

last date the subject is known to be alive when the confirmation of death is absent or unknown.

Evaluable tumor assessment

An overall visit response other than unevaluable (UE) or not done (ND).

Evaluable lesion

A target lesion or a new measurable lesion with a baseline and at least subsequent non-missing dimension.

Progression-free survival (PFS)

PFS per modified irRC-RECIST is defined as the interval from first dose date to the earlier event of PD per the response criteria, or death from any cause. Subjects without an event will be censored at the latter of their last evaluable tumor assessment or first dose date.

Treatment Period

Treatment period is defined as the period between the first date of study therapy administration in the study and 30 days after the last IP administration in the study.

Treatment-emergent Adverse Events (TEAE)

Treatment-emergent adverse events are defined as any adverse event occurring in the treatment period. Adverse events that occur on the same day as the first dose of study therapy will be treated as treatment-emergent events unless indicated otherwise. (For example, if an event occurs on the same date as the first administration of talimogene laherparepvec and the check box indicating prior to the first dose of IP is checked, then the event will not be counted as a treatment-emergent AE. Additionally, if an event is identified as disease-related on the eCRF, it will not be counted as a TEAE).

Treatment-emergent Disease-Related Events (DREs)

Treatment-emergent Disease-Related Events (DREs) are defined as adverse events, determined by the investigator to be disease-related, with an onset during the treatment period. DREs that occur on the same day as the first dose of study therapy will be treated as treatment emergent events unless indicated otherwise (for example, if an event occurs on the same date as the first administration of talimogene laherparepvec and the check box indicating prior to the first dose of IP is checked on the eCRF, then the event will not be counted as a treatment-emergent DRE).

Event of Interest (EOI) for talimogene laherparepvec

An event of interest (EOI) is a noteworthy treatment-emergent adverse event for a particular product or class of products that may warrant careful monitoring. It could be serious or non-serious, and could include events that might be potential precursors or prodromes for more serious medical conditions in susceptible individuals (Council for International Organizations of Medical Sciences (CIOMS) VI, 2005). MedDRA dictionary preferred terms for each EOI search strategy is defined and maintained by Amgen Safety Medical Coding. EOIs for talimogene laherparepvec are defined in the version-controlled EOI search strategies maintained by Global Regulatory Affairs and Safety (GRAAS) at Amgen for talimogene laherparepvec. The MedDRA version used for the EOI analysis will be documented in the clinical study report.

On-study Death

Deaths of all causes that occur any time during the study are defined as on-study deaths. This may include deaths that occur beyond the end of the treatment period. Note that this study has a long-term survival follow-up after the safety follow-up, deaths that occur after the study long-term survival follow-up will be reported for subjects that enroll in study 20120139, which is the registry study for long-term follow up of clinical trial subjects who have completed talimogene laherparepvec treatment.

7. Analysis Subsets

7.1 DLT Analysis Set

The DLT analysis set will include DLT-evaluable subjects who have had the opportunity to be on treatment for at least 2 cycles from the initial dose of study treatments and have received at least 2 doses of talimogene laherparepvec and 2 doses of atezolizumab in combination, or have a DLT during the DLT evaluation period as described in Section 3.1 of the protocol.

7.2 Safety Analysis Set

The safety analysis set will include all subjects who have received at least 1 dose of talimogene laherparepvec or atezolizumab.

7.3 qPCR Analysis Set

The qPCR analysis set includes subjects in the safety analysis set with a sample obtained for qPCR testing of talimogene laherparepvec DNA from a swab of cold sore, vesicles and other lesions suspected to be herpetic in origin (if any).

7.4 Subgroup Analyses

The covariates in Section 4.2 may be used to examine efficacy and safety in subgroups or in multivariate analyses.

7.5 Lesion Analysis Sets

The Injected Lesion Analysis Set includes any target lesion and new measurable lesion that was ever injected and the Uninjected Lesion Analysis Set includes any target lesion and new measurable lesion that was never injected. The Injected Lesion Evaluable and Uninjected Lesion Evaluable analysis sets will include the corresponding subset of target lesions and new measurable lesions with at least one post-baseline measurement prior to a censoring event, ie, non-study anticancer therapy, merged with another lesion, or resection (except if it is partial or complete resection and pathology indicates absence of malignant cells). If a lesion splits, the sum of the length of the split lesions will be added together prior to analysis.

8. Planned Analyses

8.1 Interim Analysis and Early Stopping Guidelines

No formal interim efficacy analysis is planned for this study. Interim safety analyses will be performed to support the evaluation of safety of the combination therapy by the Dose Level Review Team (DLRT). Enrollment will be suspended at the first safety interim analysis when the first 4 to 6 DLT-evaluable subjects have been enrolled in the study. At the discretion of the DLRT, additional safety analyses may be conducted as warranted. If the DLRT does not recommend prematurely ending enrollment, safety of the combination therapy will be conducted on the first 18 DLT-evaluable subjects in each cohort.

8.1.1 Dose Level Review Team (DLRT)

DLRT meetings will be held to review data, monitor safety, and make dose recommendations.The DLRT will be composed of the investigator(s), Amgen Medical Monitor, Genentech representative, Amgen Global Safety Officer or designated safety scientist, Amgen Global Clinical Trial Manager, and Amgen Biostatistics representative. Additional members may be added as needed (eg, Global Development Leader). A quorum, defined as > 50% of the participating investigators who have enrolled subjects in the study or their qualified designee [ie, sub-PI or research nurse or study coordinator possessing hard copy documentation (eg, email) of the PI's vote regarding the dose level review], must be in attendance for each DLRT meeting. The DLRT will be rescheduled if a quorum is not reached. The DLRT members are responsible for dosing

recommendations, which may include continuation, delay, or termination of dosing. All available study data, including demographics, investigational product administration, medical history, concomitant medications, adverse events, ECGs, vital signs, laboratory results will be reviewed. In addition to DLTs, all ≥ grade 3 treatment emergent toxicities not meeting DLT criteria will be reviewed by the team and can be considered in the DLRT's recommendations. Data to be reviewed will be queried.

8.2 Primary Analysis

The primary analyses will be performed separately for each cohort if the expected data cutoff dates for the two cohorts exceed 3 months. The primary safety analyses will occur when the last DLT-evaluable subject in the cohort is enrolled and has had the opportunity to complete the DLT evaluation period. The database will be cleaned, and a locked database will be used in the analysis.

The primary efficacy analyses will occur 19 weeks after the last DLT-evaluable subject in the cohort is enrolled. The database will be cleaned, and a locked database will be used in the analysis.

8.3 Final Analysis

The final analyses will be performed separately for each cohort and will occur when the last subject in the cohort has discontinued study treatment and has had the opportunity to complete the long-term survival follow-up visit. The final analyses will be performed separately for each cohort if the expected data cutoff dates for the two cohorts exceeds 3 months. The database will be cleaned, and a locked database will be used in the analysis.

9. Data Screening and Acceptance

9.1 General Principles

The objective of the data screening is to assess the quantity, quality and statistical characteristics of the data relative to the requirements of the planned analyses.

9.2 Data Handling and Electronic Transfer of Data

Amgen's Clinical Data Management (CDM) department will provide all data to be used in the planned analyses. This protocol will use the RAVE database. In addition, [redacted] HSV antibody serostatus samples, and swab samples collected from lesions suspected to be herpetic in origin, will be processed and analyzed by a vendor of Amgen. Analysis results of these samples will

be transferred to Amgen CDM on a regular basis as per a Data Transfer Plan (DTP) and will be stored on the CDM server.

9.3 Handling of Missing and Incomplete Data

Adverse events with missing IP relatedness, seriousness, or CTCAE severity grades are included in TEAE as long as the events qualify for the reporting period. Events with missing relatedness, seriousness, and CTCAE severity grades will be excluded from treatment-related, serious, and with a CTCAE grade of 3 or higher AE analysis, respectively.

Partial or missing dates of adverse events and concomitant medications will be imputed.

[redacted]

9.4 Detection of Bias

Not Applicable.

9.5 Outliers

Descriptive statistics will be used to identify outliers in any key variables. Suspected outliers will be included in all analyses unless there is sufficient scientific justification to exclude them.

9.6 Distributional Characteristics

Not Applicable

9.7 Validation of Statistical Analyses

Programs will be developed and maintained, and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software; for example, the SAS System version 9.4 or later.

10. Statistical Methods of Analysis

10.1 General Principles

The data will be analyzed by cohort. Besides a summary of the incidence of DLTs, descriptive statistics will be provided for demographic, safety, efficacy, and biomarkers

as appropriate. The DLT analysis set will be used to summarize the subject incidence of DLT for the study and the safety analysis set will be used for all other analyses of safety endpoints (including but not limited to all adverse events, grade ≥ 3 adverse events, serious adverse events, fatal adverse events, adverse events requiring discontinuation of study drug, and AEs defined as events of interest). The efficacy analysis will be conducted using the safety analysis set unless otherwise specified.

In principle, mean, standard deviation, median, first and third quartiles, minimum and maximum will be calculated for continuous variables; frequency count and percent will be calculated for binary and categorical variables. The analysis of response variant endpoints will be based on investigator assessments per modified irRC-RECIST as described in the protocol unless otherwise specified.

10.2 Subject Accountability

The number of subjects enrolled will be tabulated by investigator sites. Subject disposition including screened, enrolled, treated and ended treatment, and completed safety follow-up visit will be summarized for all enrolled subjects. Reasons for discontinuation of study drug, reason for not completing the study, and the reason for not completing safety follow-up will also be summarized overall or by cohort if appropriate. Key study dates for the first subject enrolled, last subject enrolled, last subject's end of study, last subject's end of investigational product, and data cut-off date will be presented.

10.3 Important Protocol Deviations

Important Protocol Deviation (IPD) categories are defined by the protocol team before the first subject's visit and updated during the IPD reviews throughout the protocol prior to database lock. These definitions of IPD categories, sub-category codes and descriptions will be used during the course of the study.

Eligibility deviations are defined in the protocol. Eligibility deviations that are defined as IPDs will be summarized in both the IPD and Eligibility Deviation table. An IPD and an Eligibility Deviation listing will be provided. The number of subjects with at least one IPD during study will be tabulated.

10.4 Demographic and Baseline Characteristics

Summary statistics of the demographic, baseline characteristics, and disease characteristics described in Section 4.2 will be tabulated. They will also be used

whenever applicable to examine efficacy and safety in subgroups or in multivariate analyses.

10.5 Efficacy Analyses

ORR, DRR, and DCR will be summarized with the associated 95% CI (Clopper and Pearson, 1934). The proportion of all evaluable injected and uninjected lesions with a response (≥ 30% and 100% decrease) will be summarized by lesion type (overall, hepatic, non-hepatic). Wilson's score method with continuity correction (Newcombe, 1998) will be used to calculate an approximate exact CI for between-group differences in binary rates if appropriate. Waterfall plots for maximum tumor burden decrease in the lesion diameter at subject-level and lesion-level will also be provided.

Kaplan-Meier (K-M) estimates of landmarks (eg, 1-, 2-, and 3-year rates) and quartiles for DOR/PFS/OS will be provided (Kaplan and Meier, 1958). DOR will be analyzed only for subjects with a best overall response of CR or PR. Greenwood's formula (Kalbfleisch and Prentice, 1980) for standard error will be used to calculate CIs for landmark K-M rates. CIs for K-M quartiles will be estimated (Brookmeyer and Crowley, 1982).

10.6 Safety Analyses

10.6.1 Analysis of Primary Safety Endpoint

The subject incidence of DLT will be summarized as a binary variable using the DLT analysis set.

10.6.2 Adverse Events and Disease Related Events

Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 17.0 or later will be used to code all events categorized as adverse events (AEs), disease-related events (DREs) to a system organ class and a preferred term. The CTCAE version 4.0 will be used to grade severity of AEs.

A table will be provided with a high-level summary of the subject incidence of any treatment-emergent adverse event and any treatment-emergent, treatment-related adverse event within the following categories: worst grade 3, worst grade 4, worst grades 3 or 4, fatal adverse events; serious AEs; and AEs leading to permanent discontinuation of study therapy.

Subject incidence of all treatment-emergent AEs, worst grade ≥3 adverse events, serious AEs, AEs leading to withdrawal of investigational product (investigational product

refers to talimogene laherparepvec and atezolizumab), and fatal AEs will be tabulated by system organ class, preferred term in descending order of frequency and grade.

Summaries of treatment-emergent and serious AEs occurring in at least 5% of the subjects by preferred term in any cohort will be provided in descending order of frequency.

Suspicious Herpetic Lesion

The proportion of subjects with positive qPCR analysis result of talimogene laherparepvec DNA detection in swab samples taken from cold sore, vesicles, and other lesions suspected to be herpetic in origin (if any) may be summarized. qPCR positive subjects may be subgrouped by quantifiable vs too low to quantify (≥LLOQ vs <LLOQ).

Summary statistics may be presented for rate of talimogene laherparepvec DNA detection (in samples with positive qPCR analysis results) in swab samples taken from cold sore, vesicles, and other lesions suspected to be herpetic in origin (if any). qPCR positive samples may be subgrouped by quantifiable vs too low to quantify (≥LLOQ vs <LLOQ).

Disease Related Events

Subject incidence of disease-related events and fatal disease-related events will be tabulated by system organ class and preferred term. A sensitivity analysis of treatment-emergent adverse events will be conducted that considers any disease-related event as an adverse event if the disease-related event was reported in the study for any subject as a treatment-emergent adverse event.

Adverse Events of Interest

The subject incidence of treatment-emergent and treatment-emergent, treatment-related events of interest (including all events of interest, serious events of interest, non-serious events of interest) according to the EOI search strategy categories will be summarized.

Subject incidence of events of interest will also be summarized by EOI category. In addition, events of interest may be tabulated by worst grade within EOI category by preferred term.

10.6.3 Laboratory Test Results

Lab-defined lower and upper limits of normal ranges will be used for chemistry and hematology laboratory tests. NCI Common Toxicity Criteria (CTC) version 4.0 grading will be used. Laboratory results will be summarized with descriptive statistics at baseline

and selected time points. Grade shifts in important laboratory results from baseline to worst on-protocol value will be presented. Subject incidence of potential hepatoxicity as identified by the Hy's Law criteria (FDA guidance for Industry Drug Induced Liver Injury: pre-marketing evaluation, July 2009) as well as confirmed DILI events as reported by investigators will be presented.

10.6.4 Vital Signs

Descriptive analyses of temperature, blood pressure, heart rate, and respiratory rate will be conducted at baseline and selected time points. Subject incidence of abnormal vital signs will be provided.

10.6.5 Physical Measurements

Height and weight will be summarized at each assessed time point. The change in weight from baseline to each assessed time point will also be summarized as a difference in original units of measurement and as percentage difference.

10.6.6 Electrocardiogram (ECG)

The ECG measurements from this clinical study will be performed as per standard of care for routine safety monitoring, rather than for purposes of assessment of potential QTc effect. Because these evaluations may not necessarily be performed under the rigorous conditions expected to lead to meaningful evaluation of QTc data; neither summaries nor statistical analyses will be provided, and these data would not be expected to be useful for meta-analysis with data from other trials.

10.6.7 Antibody Formation

Not Applicable.

10.6.8 Exposure to Investigational Product

Descriptive statistics will be produced to describe the exposure to Amgen investigational product talimogene laherparepvec and non-Amgen investigational product atezolizumab by cohort. Summary statistics for exposure to talimogene laherparepvec and atezolizumab, including total doses administered, total volume administered, duration from the first to the last administration of talimogene laherparepvec and atezolizumab, and the average volume received by subject per visit will be provided and will be separated by the first (concentration of 10^6 PFU/ml) and subsequent doses for talimogene laherparepvec. The subject incidence rate and reasons for IP change/withdrawl and missed treatment will be tabulated.

10.6.9 Exposure to Other Protocol-specified Treatment

Not applicable.

10.6.10 Exposure to Concomitant Medication

The number and proportion of subjects receiving therapies of interest during the treatment period will be summarized by preferred term or category by cohort as coded by the World Health Organization Drug (WHO DRUG) dictionary.

10.7 Pharmacokinetic or Pharmacokinetic/Pharmacodynamic Analysis

Not Applicable.

11. Changes from Protocol-specified Analyses

There are no changes to the protocol-specified analyses.

12. Literature Citations / References

Clopper, C.; Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934:26:404-413.

Kaplan and Meier. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association. 1958:53 457-481

Brookmeyer, R. and Crowley, J. A Confidence Interval for the Median Survival Time. Biometrics. 1982:38:29-41.

Kalbfleisch, J.D. and Prentice, R.L. The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons. 1980

Leemis, L.M. and Trivedi, K.S. A comparison of approximate interval estimates for the Bernoulli parameter. The American Statistician. 1996:50:63-68.

Newcombe, R.G. Interval estimation for the difference between independent proportions: comparison of eleven methods. Stat Med. 1998:17:873-890.

13. Prioritization of Analyses

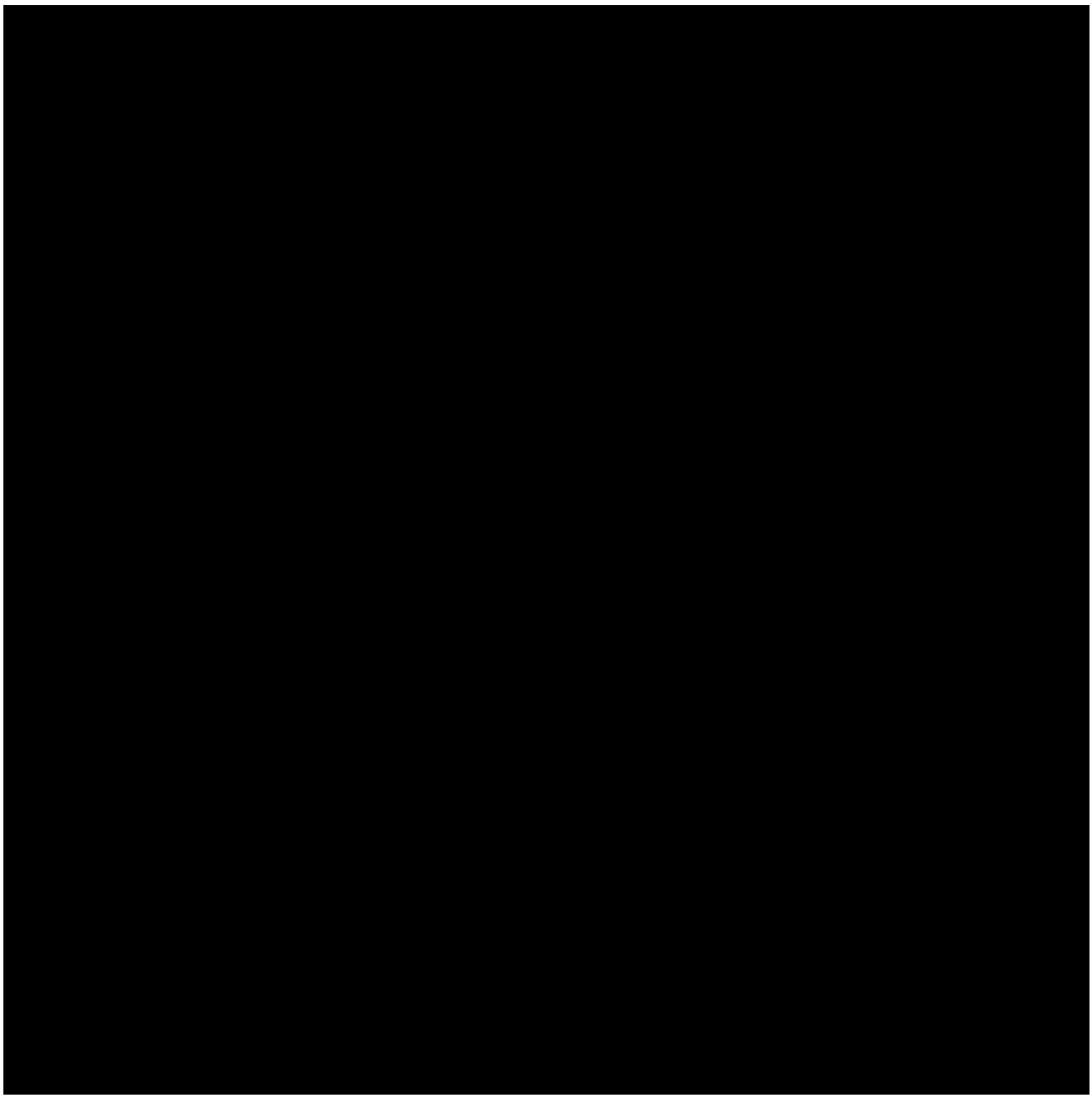
Not Applicable.

14. Data Not Covered by This Plan

Analyses for the exploratory objectives may be documented in a supplemental SAP.

15. Appendices

Appendix A. Conventions for Clinical Data That Require Imputation for Partial or Missing Date



Appendix B. Wilson's Score Method With Continuity Correction in Calculating the 95% Confidence Intervals for Difference in 2 Independent Proportions

The following formula of interval estimation is extracted from Newcombe 1998 paper:

Assuming

$$X_i \overset{i.i.d.}{\sim} Bernoulli\ (\pi_1), where\ \ i = 1, 2, ..., m.$$

$$Y_j \overset{i.i.d.}{\sim} Bernoulli\ (\pi_2), where\ \ j = 1, 2, ..., n.$$

$$\theta = \pi_1 - \pi_2$$

$$\hat{\pi}_1 = \sum_{i=1}^{m} x_i / m = a / m\ .$$

$$\hat{\pi}_2 = \sum_{j=1}^{n} y_j / n = b / n\ .$$

Then the 100(1- α)% confidence interval for $\hat{\theta} = \hat{\pi}_1 - \hat{\pi}_2 = a/m - b/n$, with continuity correction is

$$\delta = \sqrt{\{(a/m - l_1)^2 + (u_2 - b/n)^2\}}$$

$$\varepsilon = \sqrt{\{(u_1 - a/m)^2 + (b/n - l_2)^2\}}$$

l_1 and u_1 delimit the interval $\{\pi_1 : |\pi_1 - a/m| - 1/(2m) \le z_{1-\alpha/2}\sqrt{\pi_1(1-\pi_1)/m}\}$ and

l_2 and u_2 delimit the interval $\{\pi_2 : |\pi_2 - b/n| - 1/(2n) \le z_{1-\alpha/2}\sqrt{\pi_2(1-\pi_2)/n}\}$.