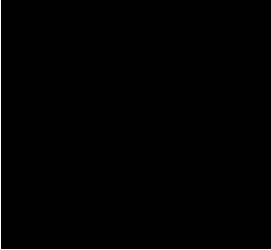


STATISTICAL ANALYSIS PLAN

A Phase 1, Multi-center, Open-Label, Dose De-escalation Study to Evaluate the Safety and Efficacy of Talimogene Laherparepvec in Japanese Subject With Unresectable Stage IIIB-IV Malignant Melanoma

Protocol Number: 20140270
Version: Version 3.0
Date: 04 November 2020
Authors: 

NCT Number: NCT03064763
This NCT number has been applied to the document for purposes of posting on Clinicaltrials.gov

Table of Contents

| | |
|---|----|
| Table of Abbreviations | 4 |
| 1. Introduction..... | 6 |
| 2. Objectives..... | 6 |
| 2.1 Primary | 6 |
| 2.2 Secondary..... | 6 |
| 2.3 Safety..... | 6 |
| 3. Study Overview | 6 |
| 3.1 Study Design..... | 6 |
| 3.2 Sample Size..... | 8 |
| 4. Study Endpoints and Covariates..... | 10 |
| 4.1 Study Endpoints | 10 |
| 4.1.1 Primary Endpoints | 10 |
| 4.1.2 Secondary Endpoints | 10 |
| 4.1.3 Safety Endpoints | 10 |
| 4.2 Planned Covariates..... | 10 |
| 5. Hypotheses and/or Estimations | 11 |
| 6. Definitions..... | 11 |
| 7. Analysis Subsets | 14 |
| 7.1 DLT Analysis Set..... | 14 |
| 7.2 Safety Analysis Set | 15 |
| 7.3 qPCR Analysis Set..... | 15 |
| 7.4 Subgroup Analyses | 15 |
| 8. Interim Analysis and Early Stopping Guidelines..... | 15 |
| 8.1 Interim Analysis..... | 15 |
| 8.2 Primary Analysis | 15 |
| 8.3 Final Analysis..... | 15 |
| 9. Data Screening and Acceptance..... | 15 |
| 9.1 General Principles..... | 15 |
| 9.2 Data Handling and Electronic Transfer of Data | 16 |
| 9.3 Handling of Missing and Incomplete Data | 16 |
| 9.4 Detection of Bias | 16 |
| 9.5 Outliers | 16 |
| 9.6 Distributional Characteristics..... | 16 |
| 9.7 Validation of Statistical Analyses..... | 16 |
| 10. Statistical Methods of Analysis..... | 17 |
| 10.1 General Principles..... | 17 |

| | | |
|---------|--|----|
| 10.2 | Subject Accountability | 17 |
| 10.3 | Important Protocol Deviations | 17 |
| 10.4 | Demographic and Baseline Characteristics | 18 |
| 10.5 | Efficacy Analyses | 18 |
| 10.6 | Safety Analyses | 18 |
| 10.6.1 | DLT Analysis | 18 |
| 10.6.2 | Adverse Events | 19 |
| 10.6.3 | Laboratory Test Results | 20 |
| 10.6.4 | Vital Signs | 20 |
| 10.6.5 | ECOG Performance Scores | 20 |
| 10.6.6 | Physical Measurements | 20 |
| 10.6.7 | Electrocardiogram (ECG) | 20 |
| 10.6.8 | Antibody Formation | 20 |
| 10.6.9 | Exposure to Investigational Product | 21 |
| 10.6.10 | Exposure to Concomitant Medication | 21 |
| 11. | Literature Citations / References | 22 |
| 12. | Appendices | 23 |

List of Tables

| | | |
|----------|--|----|
| Table 1. | Probability of Declaring a Cohort Safe or Unsafe | 9 |
| Table 2. | Potential Japanese Melanoma Phase 2 Single-Arm Designs to Test a Durable Response Rate Hypothesis Without Dose De-escalation | 10 |

List of Figures

| | | |
|-----------|---|----|
| Figure 1. | Probability of Declaring a Cohort Safe (Unsafe) | 9 |
| Figure 2. | The Definition of Analysis Sets | 14 |

List of Appendices

| | | |
|-------------|--|----|
| Appendix A. | Conventions for Clinical Data That Require Imputation for Partial or Missing Date | 24 |
| Appendix B. | Cutoff Values for qPCR Assay | 26 |

Table of Abbreviations

| Abbreviation or Term | Definition/Explanation |
|-------------------------------------|--|
| BRAF | serine/threonine protein kinase B-Raf V600E/K |
| CI | confidence interval |
| CR | complete response |
| CRF | case report form |
| CTCAE | Common Terminology Criteria for Adverse Events |
| DLRT | Dose Level Review Team |
| DLT | dose limiting toxicity |
| DNA | deoxyribonucleic acid |
| DOR | duration of response |
| DRE | Disease-related events |
| DRR | durable response rate |
| ECG | electrocardiogram |
| ECOG | Eastern Cooperative Oncology Group |
| End of Study for Individual Subject | The last day that protocol-specified procedures are conducted for an individual subject (ie, the date the subject withdraws full consent from the study, completes the safety follow-up visit or long-term follow-up [whichever is later] or death). |
| End of Study (primary completion) | The time when the last subject is assessed or receives an intervention for the purposes of final collection of data for the primary efficacy analysis. This will occur when a durable response outcome has been assessed for all subjects included in the primary efficacy analysis. |
| End of Study (end of trial) | The time when the last subject is assessed or receives an intervention for evaluation in the study. This is anticipated to occur when the last subject has had the opportunity to complete the safety follow-up visit or the last long term follow-up visit, whichever occurs later. |
| End of Treatment | The day of the last assessment for the protocol-specified treatment phase of the study for an individual subject |
| HCP | Health care providers |
| Heart rate | number of cardiac cycles per unit of time |
| HSV-1 | herpes simplex virus type-1 |
| ICH | International Conference on Harmonisation |
| IP | Investigational product |
| IPD | Important protocol deviations |
| LDH | lactate dehydrogenase |
| LTFU | Long term follow-up |
| ORR | overall response rate |
| OS | overall survival |
| PD | disease progression |
| PFS | progression-free survival |

| Abbreviation or Term | Definition/Explanation |
|----------------------|---|
| PFU | plaque-forming unit |
| PR | partial response |
| qPCR | real-time polymerase chain reaction |
| Source Data | Information from an original record or certified copy of the original record containing patient information for use in clinical research. The information may include, but is not limited to, clinical findings, observations, or other activities in the clinical trial necessary for the reconstruction and evaluation of the trial. Source data are contained in source documents (original records or certified copies). (ICH Guideline randomization identification, and stratification value. |
| TEAE | Treatment-emergent Adverse Events |
| TESAE | Treatment-emergent Serious Adverse Events |
| TTR | time to response |
| ULN | upper limit of normal |
| UE | unable to evaluate |
| vs | versus |
| WHO | World Health Organization |

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 for talimogene laherparepvec Study 20140270 (**Amendment 5 dated 01 September 2020**). The scope of this plan includes the primary analyses for safety and efficacy and the final analysis that are planned and will be executed by the Biostatistics department unless otherwise specified.

2. Objectives

2.1 Primary

- To evaluate the safety and tolerability of talimogene laherparepvec, as assessed by incidence of dose limiting toxicities (DLTs), in subjects with unresectable stage IIIB-IV malignant melanoma
- These subjects will also be evaluated for the anti-tumor activity of talimogene laherparepvec as assessed by durable response rate (DRR) using modified World Health Organization (WHO) response criteria. DRR is defined as the rate of objective response (Complete response [CR] or partial response [PR]) lasting continuously for 6 months or greater and starting any time within 12 months of initiating therapy.

2.2 Secondary

- To evaluate the anti-tumor activity of talimogene laherparepvec, as assessed by overall response rate (ORR), time to response (TTR), duration of response (DOR), progression free survival (PFS) using modified World Health Organization (WHO) response criteria by investigators
- To evaluate overall survival (OS)

2.3 Safety

To evaluate the safety of talimogene laherparepvec as determined by subject incidence of adverse events and clinically relevant laboratory abnormalities where not defined as DLTs.

3. Study Overview

3.1 Study Design

This is a phase 1, multicenter, open-label study of talimogene laherparepvec in Japanese subjects with unresectable stage IIIB-IVM1c malignant melanoma that are candidates for intralesional therapy.

Approximately 18 subjects will be enrolled in the study. The DLT evaluation period is 35 days from the initial administration of talimogene laherparepvec. Initially, 6 DLT-evaluable subjects will be enrolled and treated at 100% of the dose regimen of talimogene laherparepvec (Dose 1) as described below. The first dose administered will be up to 4.0 mL of 10^6 PFU/mL (day 1, week 0) followed by dose of up to 4.0 mL of

10⁸ PFU/mL 3 weeks (+ 5 days) later. Subsequent doses of up to 4.0 mL of 10⁸ PFU/mL will be administered every 2 weeks (\pm 3 days) thereafter (see Protocol Section 6.2.1.1). A Dose Level Review Team (DLRT) meeting will be convened after 6 DLT-evaluable subjects are enrolled. Further enrollment will be temporarily paused until the DLRT meeting is convened and a decision is made on the dose level. Upon demonstration of safety based on DLT incidence < 33% in the first 6 subjects, an additional 12 subjects will be enrolled and treated at Dose 1 to obtain additional safety data.

However, if at any time there are 2 or more subjects with a DLT prior to enrollment of the first 6 DLT-evaluable subjects, an ad hoc DLRT meeting will be convened to determine if dose de-escalation is required.

The safety data during the enrollment of the additional 12 subjects will be monitored on an ongoing basis and an ad hoc DLRT meeting may be called, if needed, to determine the further course of the study.

Dose De-escalation

Following a decision to dose de-escalate, an additional 6 subjects will be treated at the lower dose regimen of talimogene laherparepvec (Dose -1). The initial dose administered will be up to 4.0 mL of 10⁶ PFU/mL (day 1, week 0) followed by dose of up to 4.0 mL of 10⁷ PFU/mL 3 weeks (+ 5 days) later. Subsequent doses of up to 4.0 mL of 10⁷ PFU/mL will be administered every 2 weeks (\pm 3 days) thereafter (see Protocol Section 6.2.1.2). The DLRT will review the safety data after the first 6 DLT-evaluable subjects at Dose -1 have been enrolled. Subject treatment will continue, however, further enrollment will be temporarily paused during the DLT review period to evaluate the 6 DLT-evaluable subjects at Dose -1. Upon demonstration of safety, additional subjects will be enrolled (up to a total of 18 subjects) at Dose -1 to obtain additional safety data.

The safety data of the subjects enrolled at the Dose-1 level will be monitored on an ongoing basis and an ad hoc DLRT meeting may be called, if needed, to determine the further course of the study.

All subjects will complete a safety follow-up visit approximately 30 (+7) days after the last dose of study treatment. Adverse events and disease related events will be collected as described in Protocol Section 9.2. After the safety follow-up visit, all subjects will enter the long-term follow-up. **If subjects receives therapy beyond 24 months after the last subject was enrolled in the study, when the LTFU period ends, they will not enter LTFU and their last vist will be safety follow-up visit. Otherwise,** subjects will be

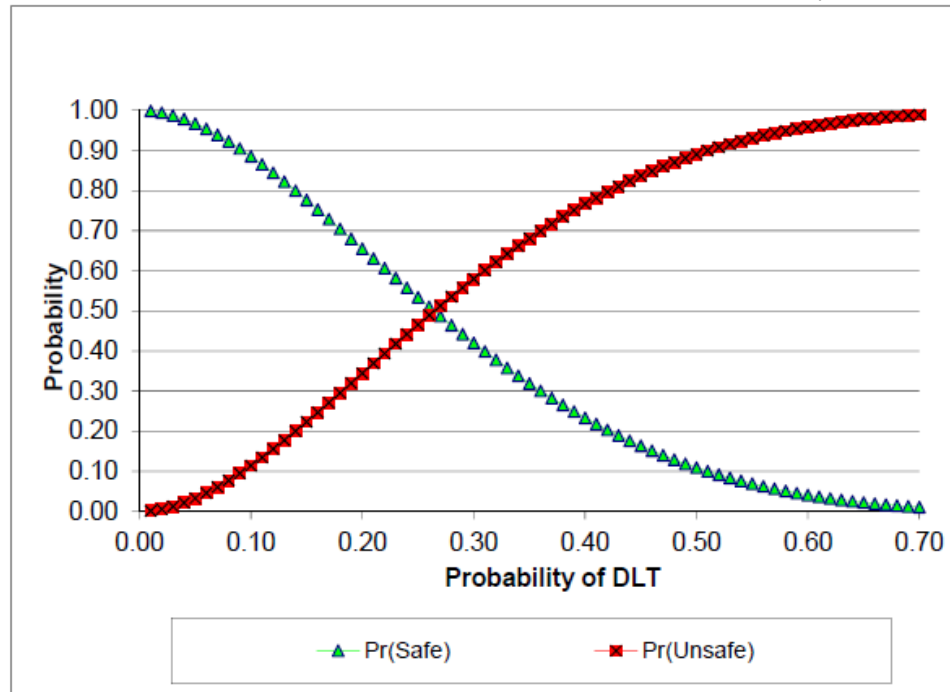
followed for survival and subsequent anticancer therapies every 12 weeks (\pm 28 days) for approximately 24 months after the last subject is enrolled. In addition, talimogene laherparepvec related adverse events that occur **after the safety follow-up visit** through the end of the long-term follow-up will be reported. Maximum treatment duration is **48 months or until drug is commercially available in Japan**, total study duration of an individual subject can be up to around 4 years plus enrollment duration. The hypothesis test for DRR will include all subjects enrolled after dose de-escalation who received at least 1 dose of talimogene laherparepvec. If no dose de-escalation is needed, the DRR hypothesis test will include the first 18 subjects dosed.

3.2 Sample Size

Approximately 18 subjects will be enrolled. See protocol Section 10.2. The total number of subjects that will participate in the study and be used in the DLT analysis is approximately 18. The first 6 DLT-evaluable subjects will be used for the decision of dose de-escalation. If there is no dose de-escalation, the first 18 subjects in the safety analysis set will be used for the efficacy analysis of DRR, and the whole safety analysis set will be used for other efficacy analyses including ORR, TTR, DOR, PFS and OS. If dose de-escalation occurs, the efficacy analysis will be performed using the safety analysis set at Dose -1. In the current study design, a dose level will be considered safe if there are 0 to 1 DLTs observed in the initial 6 DLT evaluable subjects.

[Table 1](#) and [Figure 1](#) presents the probability of declaring a dose level safe (unsafe) for a range of true DLT rates for the protocol therapy for the 6 DLT-evaluable subjects (see triangle symbols). As shown below in [Table 1](#), the probability of declaring a dose level safe (unsafe) in the current study (based on only 6 DLT evaluable subjects) is 89% (11%), 42% (58%), and 11% (89%) if the true DLT rate is 10%, 30%, or 50%, respectively. Hypothetically, if the study were to enroll an additional 6 DLT-evaluable subjects after the initial 6 subjects (total of 12 DLT evaluable subjects), the probability of observing a DLT rate \geq 33% in the 12 subjects following the observation of 0-1 DLTs in the first 6 DLT evaluable subjects would be 0.6%, 8.6%, and 6.7% if the true DLT rate is 10%, 30%, or 50%, respectively.

Figure 1. Probability of Declaring a Cohort Safe (Unsafe)



DLT= dose-limiting toxicity; Pr = probability

Table 1. Probability of Declaring a Cohort Safe or Unsafe

| True Cohort DLT Probability | Probability Declare Cohort Safe | Probability Declare Cohort Unsafe |
|-----------------------------|---------------------------------|-----------------------------------|
| 10% | 89% | 11% |
| 20% | 66% | 34% |
| 30% | 42% | 58% |
| 40% | 23% | 77% |
| 50% | 11% | 89% |

DLT= Dose Limiting Toxicity

The total number of subjects that will participate in the study and be used in DLT analyses is approximately 18.

As shown in [Table 2](#), without dose de-escalation, 18 subjects provides maximum power to test the DRR null hypothesis in the range of 16 to 20 subjects, therefore only the first 18 subjects in the safety analysis set will be used in the analysis.

Table 2. Potential Japanese Melanoma Phase 2 Single-Arm Designs to Test a Durable Response Rate Hypothesis Without Dose De-escalation

| Null Hypothesis Durable Response Rate (DRR) | Alternative Hypothesis DRR | Target Significance Level | Target Power | Number of Evaluable Subjects | Minimum Number of Durable Responders (DRR) to Reject Null | Achieved Significance Level (Power) |
|---|----------------------------|---------------------------|--------------|------------------------------|---|-------------------------------------|
| 2% | 16.3% | 1-sided 5% | 80% | 16 | 2 (12.5) | 1-sided 4.0% (76.1) |
| | | | | 17 | 2 (11.8) | 1-sided 4.5% (79.1) |
| | | | | 18 | 2 (11.1) | 1-sided 5.0% (81.7) |
| | | | | 19 | 3 (15.8) | 1-sided 0.6% (61.9) |
| | | | | 20 | 3 (15.0) | 1-sided 0.7% (65.5) |

4. Study Endpoints and Covariates

4.1 Study Endpoints

4.1.1 Primary Endpoints

- Subject incidence of DLTs
- DRR using modified WHO response criteria by investigators

4.1.2 Secondary Endpoints

- ORR, TTR, DOR, and PFS using modified WHO response criteria by investigators
- OS

4.1.3 Safety Endpoints

Subject incidence of all the following, where not defined as DLT:

- adverse events
- grade \geq 3 adverse events
- serious adverse events
- clinically significant laboratory changes
- changes in vital signs

4.2 Planned Covariates

The following covariates will be used to examine efficacy and safety in subgroups or in multivariate analyses as appropriate:

- Grouped disease stage at baseline: IIIB-IVM1a vs IVM1b-c
- Line of therapy for current disease: 1st line vs 2nd line or later
- Dose level (Dose 1 vs Dose -1, if applicable)

5. Hypotheses and/or Estimations

No formal statistical hypothesis will be tested for safety endpoints in this trial. Based on the well tolerated safety profile from a global phase 3 study, it is hypothesized that talimogene laherparepvec will be safe and well tolerated in Japanese subjects with unresectable stage IIIB-IVM1c malignant melanoma.

A DRR is hypothesized to be consistent with results from the global phase 3 study ([Andtbacka et al., 2015](#)).

6. Definitions

1-year, 2-year survival

The Kaplan-Meier (K-M) estimate of the proportion of subjects alive at 1 year and 2 years from the first dose, respectively.

Actual Follow-up Time

Actual follow-up time for a subject is calculated from the first dose to the last on-study date (ie, death date, or date last known to be alive for patients still alive).

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 any study drug. If a subject did not receive study drug, then the latest value on or prior to the enrollment date is to be used.

Duration of Response (DOR)

DOR is defined as the time from the date of an initial response (CR or PR) to the earlier of PD or death. Subjects who have not ended their response at the time of analysis will be censored at their last evaluable tumor assessment.

Durable Reponse Rate (DRR)

DRR is the incidence of a durable response defined as an OR of CR or PR per modified WHO response criteria lasting continuously for ≥ 6 months and beginning at any point within 12 months of initiating therapy among a set of subjects analyzed. Subjects with an OR who do not have any follow-up tumor assessments will not be regarded as having a durable response.

Evaluable tumor assessment

An overall visit response other than UE.

Event of Interest (EOI)

MedDRA dictionary preferred terms for each EOI search strategy are defined and maintained by Amgen Safety Medical Coding.

Investigational product (IP)

Investigational product refers to talimogene laherparepvec in this study.

Objective Response (OR) per modified WHO criteria

Objective response will be assessed based on the response of the index lesions and nonindex lesions, and presence or absence of new lesions. Confirmation of complete or partial response is not required. The overall response is derived from time point response assessments as described in Protocol Tables 5, 6, 7.

Overall Response Rate (ORR)

ORR is defined as the incidence of an OR of CR or PR per modified WHO response criteria among a set of subjects analyzed. Subjects who do not have any follow-up tumor assessments will be regarded as non-responders.

Overall Survival (OS)

OS is defined as the interval from the first dose to the event of death from any cause; otherwise, OS is censored at the date the subject was last known to be alive.

Positive qPCR

By performing qPCR analysis, the talimogene laheparepvec DNA is \geq cutoff in the sample. Details are provided in [Appendix B, Section 12](#).

Potential Follow-Up Time

Potential follow-up time for a subject is calculated from the first dose to the analysis data cutoff date.

Progression Free Survival (PFS)

Progression-free survival (PFS) per modified WHO response criteria is defined as the interval from the first dose to the earlier of PD per modified WHO response criteria or death from any cause; otherwise, PFS is censored at the last evaluable tumor assessment.

Quantifiable qPCR

By performing qPCR analysis, the talimogene laheparepvec DNA is \geq LLOQ in the sample. Details are provided in [Appendix B, Section 12](#).

Study day

Study day is calculated from the first day when investigational product is administered to the subject; mathematically stated as Study day = (visit date – first dose date) + 1.

Study Day 1

Study day 1 is defined as the first day of administration of the investigational product after enrollment. The day prior to Study Day 1 is considered Study Day -1.

Study Week 0

The start of investigational product administration to the subject is study week 0. Study day 1 corresponds to the first day of study week 0.

Time to Response (TTR)

This endpoint will only be calculated for those subjects in the safety analysis set with CR or PR. It is defined as the time from first dose to the date of the first CR or PR.

Treatment period

Study day 1 through 30 days after the last administration of investigational product.

Treatment-emergent Adverse Events (TEAE)

Treatment-emergent adverse events are defined as any adverse event during the treatment period. Adverse events that occur on the same day as the first dose date of IP 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 laheparepvec and the check box indicating prior to the first dose of IP is checked on CRF, then the event will not be counted as a treatment-emergent emergent AE). Additionally, if an event is identified as disease-related on the eCRF, it will not be counted as a treatment-emergent AE).

Treatment-emergent Disease-related Events (TEDRE)

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 IP 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).

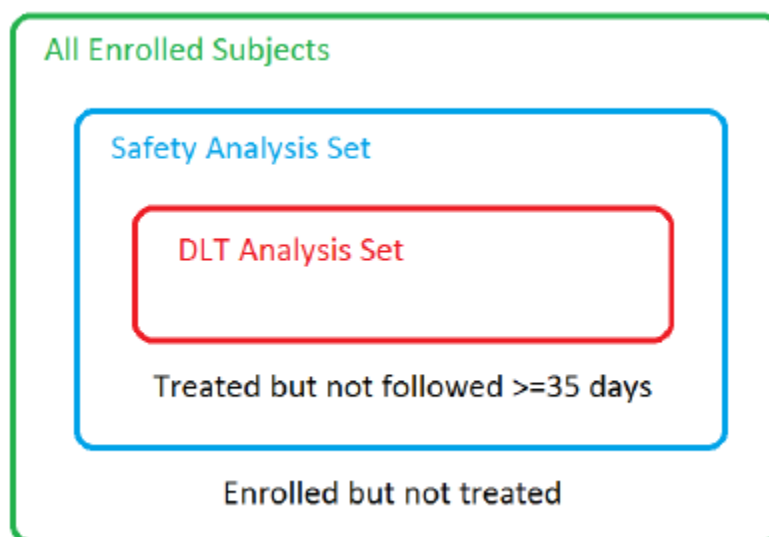
Treatment-emergent Serious Adverse Events (TESAE)

Treatment-emergent serious adverse events are defined as any serious adverse event with an onset date from first dose date to 90 days after the last dose of talimogene laherparepvec or 30 days following cessation of treatment if the subject initiates new anti-cancer therapy, whichever is earlier.

7. Analysis Subsets

The DLT analysis set will be used to summarize the subject incidence of DLTs for the study and the safety analysis set will be used for all other analyses of safety endpoints. The efficacy analysis will be conducted using the safety analysis set unless otherwise specified.

Figure 2. The Definition of Analysis Sets



7.1 DLT Analysis Set

The DLT analysis set will include DLT-evaluable subjects defined as subjects who had the opportunity to be followed for at least 35 days on treatment from the initial dosing (unless discontinued due to DLT) and received at least one dose of talimogene laherparepvec. Subjects may be replaced in a cohort if they are not evaluable for DLT (eg, a subject did not receive study treatment, or ended the study treatment before completion of DLT evaluation period for a reason other than experiencing a DLT).

7.2 Safety Analysis Set

The safety analysis set will include all subjects who received at least 1 dose of talimogene laherparepvec. The safety analysis set will be used for efficacy endpoints with the exception that only the first 18 subjects will be used to test the DRR null hypothesis for Dose 1.

7.3 qPCR Analysis Set

The qPCR Analysis Set includes subjects in the Safety Analysis Set with a swab sample obtained for qPCR testing of talimogene laherparepvec DNA from any lesion suspected to be of herpetic origin.

7.4 Subgroup Analyses

See [Section 4.2](#).

8. Interim Analysis and Early Stopping Guidelines

8.1 Interim Analysis

No formal interim efficacy analysis is planned for this study. Interim safety analyses will be performed to support the evaluation of safety by the DLRT (see protocol section 10.3.2).

8.2 Primary Analysis

The primary analysis for safety will occur 3 months after the last subject is enrolled. The primary analysis for efficacy will occur when a durable response outcome has been assessed for all subjects included in the primary efficacy analysis. Safety will be re-analyzed at the primary efficacy analysis.

8.3 Final Analysis

The final analysis will occur 24 months after the last subject is enrolled.

If there are subjects who receive therapy beyond 24 months, a post final analysis of these subjects will occur after the last subject completes safety follow-up.

Adverse events and tumor assessments evaluated after final analysis will be listed for these subjects.

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

The Amgen Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. This study will use the RAVE database.

9.3 Handling of Missing and Incomplete Data

Every effort will be made to obtain complete data in the clinical study. Partial or missing dates of adverse events, disease-related events, and concomitant medications and incomplete death dates will be imputed. Details of the imputation algorithm are provided in [Appendix A, Section 12](#). Adverse events with missing IP relatedness, seriousness, or CTCAE severity grades will be included in all analyses of TEAEs as long as the events qualify for the reporting period. Events with missing relatedness, seriousness, and severity grades will be excluded from corresponding analyses of TEAEs that are treatment-related, serious, and had a specific CTCAE grade.

9.4 Detection of Bias

Lack of protocol compliance may introduce potential bias in the estimation of protocol endpoints. All important protocol deviations (IPDs) will be reported, documented and stored in eClinical (a clinical trial management system). IPD reports will be produced using Cognos by the study manager and will be regularly reviewed in the study team's IPD review meetings as well as before analysis.

Protocol compliance will be examined by tabulating subjects with IPDs.

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

All binary endpoints will be assumed to follow a binomial distribution. The Kaplan-Meier estimates for the probability of time-to-event endpoints are non-parametric.

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 9.2 or later.

10. Statistical Methods of Analysis

10.1 General Principles

The DLT Analysis Set will be used to summarize the subject incidence of DLT and the Safety Analysis Set will be used for all other analyses of safety and efficacy endpoints.

Descriptive statistics will be provided for efficacy and safety endpoints for all subjects. Summary statistics including mean, standard deviation, median, first and third quartiles, minimum and maximum will be provided for continuous variables. Frequency and percentage will be summarized for binary and categorical variables. Proportions and the corresponding exact 95% confidence intervals using F distributions will be calculated. Exact tests will be considered for subgroup analyses. Time to-event endpoints will be estimated using the K-M method per [Kalbfleisch and Prentice \(1980\)](#) and confidence intervals for quartiles will be estimated per [Brookmeyer & Crowley \(1982\)](#).

Separate reporting at a program-level may occur of close contact and HCP events with or without a known unintended exposure.

10.2 Subject Accountability

The number of subjects enrolled will be tabulated by investigator sites. Subject disposition (including the number screened, enrolled, treated, ended treatment, that completed the safety follow-up visit, and that completed the study) will be summarized separately for all enrolled subjects. Reasons for not receiving treatment, for ending treatment, not completing the 30-day safety follow-up visit, and ending the study will be provided.

10.3 Important Protocol Deviations

Important Protocol Deviations (IPDs) categories are defined by the study team before the first subject's visit and updated during the IPD reviews throughout the study 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 and IPD and Eligibility Deviation listings.

10.4 Demographic and Baseline Characteristics

Summary statistics of the following demographic and baseline characteristics will be tabulated using the Safety Analysis Set.

- Age at baseline: < 50, ≥ 50; < 65, ≥ 65; < 75, ≥ 75 years
- Sex (female vs male)
- Disease stage at baseline: IIIB and IIIC vs IVM1a vs IVM1b vs IVM1c
- Line of therapy for current disease
- Type of prior therapy
- BRAF status
- Baseline HSV-1 serostatus
- Baseline LDH ≤ ULN vs > ULN
- ECOG (0 vs 1)
- Other covariates reported in the literature or from other Amgen studies maybe considered in the analysis as appropriate and if feasible at the time of analysis.

10.5 Efficacy Analyses

In general, the analysis of efficacy endpoints will be based on the Safety Analysis Set with the exception that only the first 18 subjects will be used to test the DRR null hypothesis for Dose 1.

- DRR and ORR (CR+PR) will be summarized with associated exact 95% CIs for binomial proportions (Clopper & Pearson 1934). A 1-sided 5% significance level exact binomial test will be performed for a DRR > 2%. If there are more than 18 subjects in the safety set for Dose Level 1, the test will be based on the first 18 enrolled subjects; otherwise, it will be based on all subjects in the safety set for either Level 1 or -1). If there is no dose de-escalation, the DRR/ORR for Dose 1 will be estimated using both the first 18 subjects and all treated.
- DRR/ORR will be summarized by planned covariates (grouped disease stage and lines of therapy) as in Section 4.2.
- DOR among responders, PFS and OS will be estimated using the Kaplan-Meier method and estimates of event time quartiles, event-free rates at selected times and the corresponding 95% CIs will be provided.
- Summary statistics will be estimated for TTR among responders.

10.6 Safety Analyses

10.6.1 DLT Analysis

The subject incidence of DLT based on CRF will be summarized as a binary variable using the DLT Analysis Set by calculating the frequency, percentage, and exact 95% CI.

10.6.2 Adverse Events

The Medical Dictionary for Regulatory Activities (MedDRA) version 18.0 or later will be used to code all adverse events (AE) to a system organ class and a preferred term. The CTCAE version 4.0 will be used to grade severity of adverse events. In general, events with missing IP relatedness, seriousness, or CTCAE severity grades are included in the analysis of treatment emergent AE as long as the event meets the criteria for a TEAE. However, analyses of treatment-related, SAE, or grade 3 or higher, or combination thereof will exclude events with missing relatedness, seriousness, and severity grades, respectively.

The analyses for AEs will include TEAEs (ie, occurring in the treatment period) unless otherwise specified. The subject incidence of TEAEs will be summarized for all AEs, serious AEs, AEs leading to withdrawal of investigational product, worst grade 3 or 4 AEs, and fatal AEs. The subject incidence of all treatment-related AEs, serious AEs, AEs leading to withdrawal of investigational product, worst grade 3 or 4 AEs, and fatal AEs will be tabulated by system organ class (SOC) and preferred term in descending order of frequency.

A listing of fatal AEs will be provided. A listing of all SAEs reported in the clinical database with an event onset from consent will be provided with onset reported relative to the first and most recent dose of study therapy and identification as a TEAE.

Subject incidence of DREs (disease-related events) and fatal disease related events will be tabulated by system organ class and preferred term. Selected safety analyses may also be performed in which DREs are considered to be a TEAE.

Treatment-related AEs reported in long term follow-up will be summarized in this study at the final analysis and also separately at a program-level.

Potential or known unintended exposure to talimogene laherparepvec, related suspected herpetic signs or symptoms, and detection of talimogene laherparepvec in a subject's household member, caregiver, or healthcare provider will be reported to Amgen.

Exposures, signs or symptoms will be summarized and also analyzed separately at a program-level.

10.6.3 Laboratory Test Results

Shifts in grades of safety laboratory values between the baseline and the worst on-study value will be tabulated overall and by dose level group(if applicable). See Protocol Table 3 for list of laboratory analytes.

A listing will be provided for subjects in the qPCR Analysis Set by assessed time point of qPCR analysis results (positive, and numeric value if quantifiable) for talimogene laherparepvec DNA in swab samples taken from cold sore, vesicles, and other lesions suspected to be herpetic in origin.

10.6.4 Vital Signs

Descriptive statistics will be presented for systolic/diastolic blood pressure, heart rate, respiratory rate, temperature and temperature for baseline, each post-baseline visit, change from baseline, and percent change from baseline.

10.6.5 ECOG Performance Scores

Shifts in scores for ECOG performance status scores between baseline and safety follow-up will be tabulated.

10.6.6 Physical Measurements

Height and weight will be summarized at each assessed time point. The change in weight and percent change in weight from baseline to each assessed time point will also be summarized.

10.6.7 Electrocardiogram (ECG)

The ECG measurements from this clinical study were 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.

10.6.8 Antibody Formation

Baseline HSV-1 serostatus will be summarized.

The incidence of HSV-1 seroconversion will be summarized for subjects that are baseline seronegative.

10.6.9 Exposure to Investigational Product

Descriptive statistics will be produced to describe the exposure to investigational product. Summary statistics for exposure to talimogene laherparepvec, including total doses administered, total volume administered, duration from the first to the last administration of talimogene laherparepvec, 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 (concentration of 10^8 PFU/mL or 10^6 PFU/mL). Subject incidence rate and reasons for IP delay, dose change and withdrawal will be tabulated.

10.6.10 Exposure to Concomitant Medication

The number and proportion of subjects receiving therapies of interest will be summarized by preferred term or category overall and for each dose level as coded by the World Health Organization Drug (WHO DRUG) dictionary.

11. Literature Citations / References

Clopper, C., *et al* (1934) "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial" *Biometrika*, **26**, 404-413

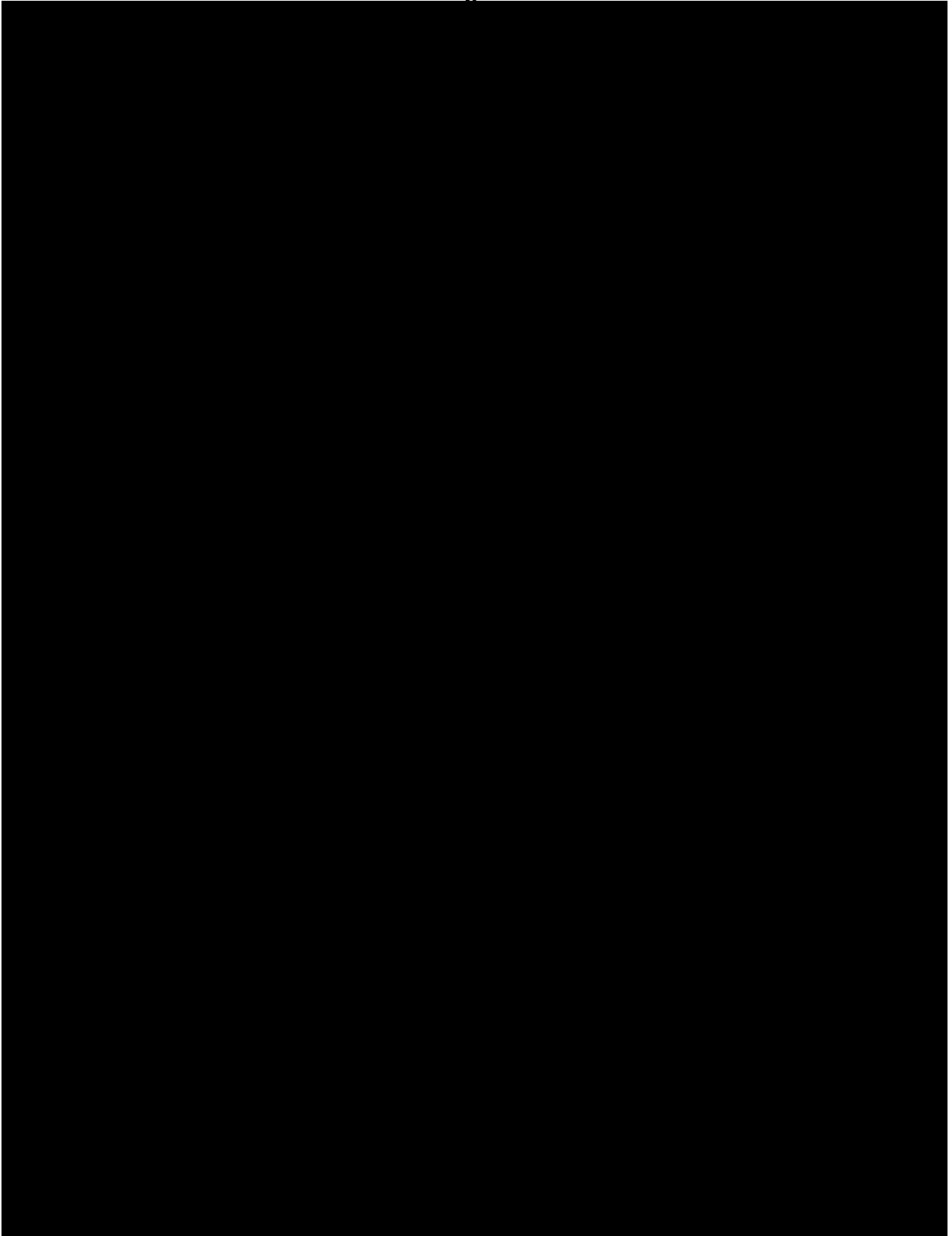
Kalbfleisch, J.D. and Prentice, R.L. (1980) *The Statistical Analysis of Failure Time Data*, New York: John Wiley & Sons

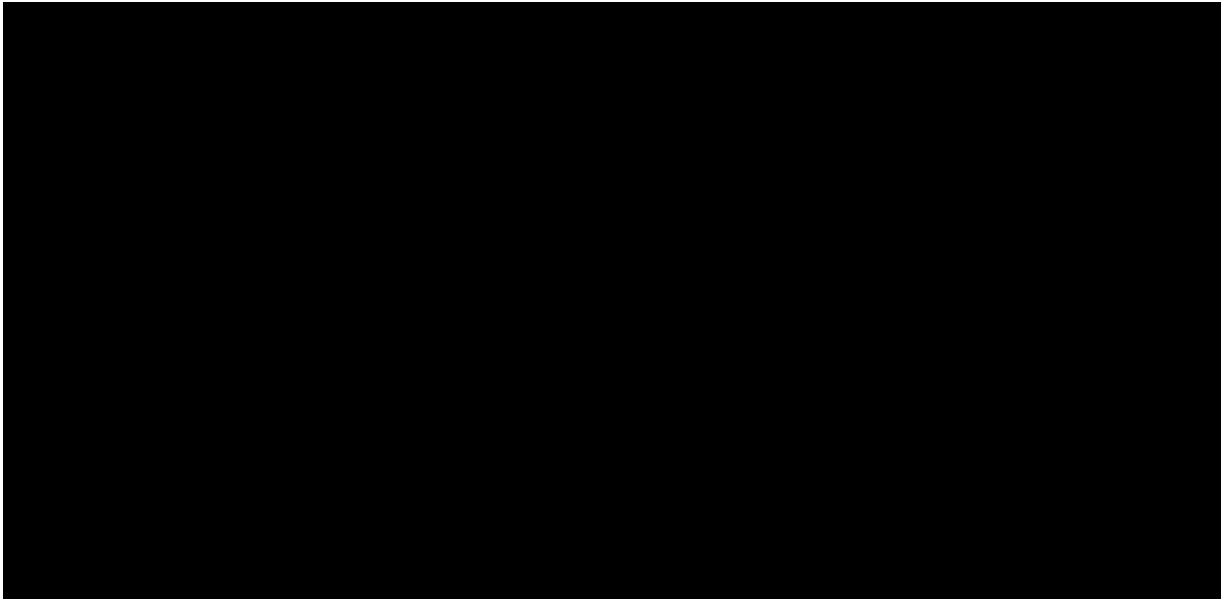
Brookmeyer R. and Crowley J. (1982) "A Confidence Interval for the Median Survival Time" *Biometrics*; **38**, 29-41

Andtbacka R., *et al* (2015) "Talimogene Laherparepvec Improves Durable Response Rate in Patients With Advanced Melanoma." *Journal of Clinical Oncology*; 33(25), 2780-2788

12. Appendices

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





Appendix B. Cutoff Values for qPCR Assay

The current assay cutoff values are as follows:

For qPCR, the cutoff (LLOQ) is 0.6 (1.76) copies/ μ g for blood, 5.8 (24) copies/ μ g for urine, and 7.5 (18) copies/ μ g for swabs.

For the most recent cutoff and LLOQ values and other details regarding their determination, refer to the following validation report:

Validation Report to Establish the Performance Characteristics of the Talimogene Laherparepvec (T-VEC) Quantitative and Qualitative Real Time PCR (qPCR) assay in Whole Blood, Swabs and Urine qPCR (see table on page 6 in the validation report for qPCR LLOD and LLOQ values)

Table B1. Definition of qPCR Results

| | Viracor Result to be reported | Result Unit | Definition of Results | Amgen T-VEC qPCR result Definition | |
|---------------------------|-------------------------------|---------------------|----------------------------|------------------------------------|------------------|
| Blood/ Urine/ Swabs | Number | Copies/ μ g DNA | \geq LLOQ | Positive | Quantifiable |
| | Detected BQL | NA | \geq cutoff and $<$ LLOQ | | Non-quantifiable |
| | Not detected | NA | $<$ cutoff | Negative | |

Table B2. Cutoff Values for qPCR Results Definition*

| | LLOQ (Copies/ μ g DNA) | cutoff (Copies/ μ g DNA) |
|-------|----------------------------|------------------------------|
| Blood | 1.76 | 0.6 |
| Urine | 24 | 5.8 |
| Swabs | 18 | 7.5 |

* The analytical method has no upper limit of quantitation.