STATISTICAL ANALYSIS PLAN

A Phase 1, Multi center, Open label, Dose De-escalation Study to Evaluate the Safety and Efficacy of Talimogene Laherparepvec in Pediatric Subjects With Advanced Non Central Nervous System Tumors That are Amenable to Direct Injection

Protocol Number: **(Talimogene Laherparepvec) 20110261**

EudraCT number 2015-003645-25

Version: **Version 3.0**

Date: **August 19, 2020**

Authors: [redacted]

NCT Number: NCT02756845
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)	19JAN2016	Not applicable
Amendment 1 (v2.0)	18APR2017	The summary of changes document is saved in RIM with file name "AMG 678 20110261 Statistical Analysis Plan version 2 changes Memo"
Amendment 2 (v3.0)	19AUG2020	**Rationale** Changes (as per protocol amendment 3) in this version are primarily focus on sample size change (due to enrollment challenge). Minimum enrollment number for younger cohorts is removed and changed total sample size to 18 dosed subjects from 18 DLT evaluable subjects. In current version, sample size 18-24 was used in protocol and SAP description which includes all possible cohorts (current cohorts at 2 different age ranges + de-escalated cohorts at 2 different age ranges). All changes are noted in this version of the statistical analysis plan by the use of bold letters. **Description of Changes** **Section 2.2. Secondary Objectives** **Rewrite:** • To evaluate the anti-tumor activity of talimogene laherparepvec, as assessed by overall response rate (ORR), duration of response (DOR), time to response (TTR), time to progression (TTP), progression-free survival (PFS) using immune related Response Criteria Simulating Response Evaluation Criteria in Solid Tumors(RECIST 1.1) [Modified irRC RECIST], and overall survival (OS). [redacted]

		[redacted] **As:** • To evaluate the anti-tumor activity of talimogene laherparepvec, as assessed by overall response rate (ORR), duration of response (DOR), time to response (TTR), time to progression (TTP), progression-free survival (PFS) using immune related Response Criteria Simulating Response Evaluation Criteria in Solid Tumors(RECIST 1.1) [Modified irRC RECIST], and overall survival (OS). **Section 2.4. Exploratory** **Rewrite:** [redacted] **As:** [redacted]

		[redacted] **Section 3.1. Study Design** **Rewrite:** This is a phase 1, multicenter, open label study of talimogene laherparepvec in pediatric subjects with advanced non-CNS tumors that are amenable to direct injection in the clinical setting. Approximately 18 treated pediatric subjects are expected to be enrolled into – 2 cohorts stratified by age (permissible based on the incidence of DLTs, a minimum of 6 subjects/cohort and a minimum of 18 subjects total) as follows: • Cohort A1 (12 to ≤ 21 years of age) • Cohort B1 (2 to < 12 years of age) Initially, 3 subjects 12 to ≤ 21 years of age are to be enrolled and treated at 100% of the recommended adult dose regimen of talimogene laherparepvec (Cohort A1). The first dose administered will be up to 4.0 mL of 10^6 PFU/mL followed by a dose of up to 4.0 mL of 10^8 PFU/mL 21 days (+3 days) later. Subsequent doses of up to 4.0 mL of 10^8 PFU/mL will be administered approximately every 14 days (± 3 days) thereafter (see protocol Section 6.2.1.1). The DLT evaluation period is 35 days from the initial administration of talimogene laherparepvec. Rules for DLT evaluation are described in Section 6.2.1.2.1.of the protocol.The dose level review team (DLRT) will review the safety data of the first 3 subjects in the older age cohort A1 to decide if the younger age cohort B1 can be

		opened for enrollment. If none of the first 3 DLT-evaluable subjects in cohort A1 experiences a DLT then cohort B1 will be opened for enrollment and treatment at the same dose level per the section Study Design and Treatment Schema of the protocol. If a DLT occurs in the first 3 DLT-evaluable subjects in the older age cohort (A1 or A2), the younger age cohort will not open until a DLT rate $< 33\%$ is observed with at least 6 DLT-evaluable subjects in the older age cohort (A1 or A2). If any of the first 3 subjects in cohort A1 experiences a DLT, the DLRT will make a recommendation on whether an additional 3 subjects should be enrolled in cohort A1. Cohort A1 or B1 can enroll up to a maximum of 12 DLT-evaluable subjects. If permissible based on the incidence of DLTs, the minimum number of DLT-evaluable subjects per cohort will be 6 and the minimum number of DLT-evaluable subjects across cohorts A1 and B1 for the study will be 18. In the case of dose de-escalation, the minimum number of DLT-evaluable subjects at the de-escalated dose will be 9 (permissible based on the incidence of DLTs). After an age cohort is closed for further enrollment, if $< 33\%$ of all DLT-evaluable subjects in the cohort experiences a DLT at a dose level, the dose will be declared safe for the cohort. If dose de-escalation is needed and if permissible based on the incidence of DLTs, a minimum of 9 additional DLT evaluable subjects will be enrolled and treated at a lower dose level of talimogene laherparepvec (Dose Level -1). The initial dose administered will be up to 4.0 mL of 10^6 PFU/mL followed by dose of up to 4.0 mL of 10^6 PFU/mL 21 days ($+3$ days) later. Subsequent doses of up to 4.0 mL of 10^6 PFU/mL will be administered approximately every 14 days ($\pm$ 3 days) thereafter (see Table 1). Dose de-escalation cohorts will obey the following naming convention, based on age at baseline, and the same DLT rules will be applied to the de-escalated cohorts: • Cohort A2 (12 to $\leq$ 21 years of age)

		• Cohort B2 (2 to < 12 years of age) **As:** This is a phase 1, multicenter, open label study of talimogene laherparepvec in pediatric subjects with advanced non-CNS tumors that are amenable to direct injection in the clinical setting. Approximately 18 **to 24** pediatric subjects are expected to be enrolled and **treated with at least 1 dose of talimogene laherparepvec** into 2 cohorts stratified by age (permissible based on the incidence of DLTs, a minimum of 18 **dosed** subjects total **for the primary analysis**) as follows: • Cohort A1 (12 to ≤ 21 years of age) • Cohort B1 (2 to < 12 years of age) Initially, 3 subjects 12 to ≤ 21 years of age are to be enrolled and treated at 100% of the recommended adult dose regimen of talimogene laherparepvec (Cohort A1). The first dose administered will be up to 4.0 mL of 10^6 PFU/mL followed by a dose of up to 4.0 mL of 10^8 PFU/mL 21 days (+3 days) later. Subsequent doses of up to 4.0 mL of 10^8 PFU/mL will be administered approximately every 14 days (± 3 days) thereafter (see protocol Section 6.2.1.1). The DLT evaluation period is 35 days from the initial administration of talimogene laherparepvec. Rules for DLT evaluation are described in Section 6.2.1.2.1.of the protocol.The dose level review team (DLRT) will review the safety data of the first 3 subjects in the older age cohort A1 to decide if the younger age cohort B1 can be opened for enrollment. If none of the first 3 DLT-evaluable subjects in cohort A1 experiences a DLT then cohort B1 will be opened for enrollment and treatment at the same dose level per the section Study Design and Treatment Schema of the protocol. If a DLT occurs in the first 3 DLT-evaluable subjects in the older age cohort (A1 or A2), the younger age cohort will not open until a DLT rate < 33% is observed with at least 6 DLT-evaluable subjects in the older age cohort (A1 or A2). If any of the first 3 subjects in cohort A1

		experiences a DLT, the DLRT will make a recommendation on whether an additional 3 subjects should be enrolled in cohort A1. Cohort A1 or B1 can enroll up to a maximum of **18 subjects treated with at least 1 dose of talimogene laherparepvec with at least 9** DLT-evaluable subjects **in cohort A1.** In the case of dose de-escalation, the minimum number of DLT-evaluable subjects at the de-escalated dose will be **6** (permissible based on the incidence of DLTs). After an age cohort is closed for further enrollment, if $< 33\%$ of all DLT-evaluable subjects in the cohort experiences a DLT at a dose level, the dose will be declared safe for the cohort. If dose de-escalation is needed and if permissible based on the incidence of DLTs, a minimum of **6** additional DLT evaluable subjects will be enrolled and treated at a lower dose level of talimogene laherparepvec (Dose Level -1). The initial dose administered will be up to 4.0 mL of 10^6 PFU/mL followed by dose of up to 4.0 mL of 10^6 PFU/mL 21 days ($+3$ days) later. Subsequent doses of up to 4.0 mL of 10^6 PFU/mL will be administered approximately every 14 days ($\pm$ 3 days) thereafter (see Table 1). Dose de-escalation cohorts will obey the following naming convention, based on age at baseline, and the same DLT rules will be applied to the de-escalated cohorts: • Cohort A2 (12 to $\leq$ 21 years of age) • Cohort B2 (2 to $<$ 12 years of age) **Section.3.1.1.2 Long-term Follow-up** **Rewrite:** After the safety follow up visit, all subjects will enter the long term follow up. Subjects will be followed for survival and use of subsequent anticancer therapies every 12 weeks ($\pm$ 28 days) from the safety follow-up until death, subject or legally acceptable representatives withdraws full consent/ assent, or up to approximately 24 months after the last subject is enrolled

		in the study. In addition, talimogene laherparepvec related adverse events that occur through the end of the long term follow up will be reported. **As:** After the safety follow up visit, all subjects will enter the long term follow up. Subjects will be followed for survival and use of subsequent anticancer therapies every 12 weeks (± 28 days) from the safety follow-up visit until death, subject or legally acceptable representative withdraws full consent/ assent, or up to approximately 24 months after the last subject is enrolled in the study **whichever occurs first.** In addition, talimogene laherparepvec related adverse events that occur through the end of the long term follow up will be reported. **Section 3.2. Sample Size** **Rewrite:** The sample size of 6 to 12 DLT-evaluable subjects for each age cohort is determined empirically and is consistent with those used in 3+3 phase 1 designs assuming a true DLT incidence rate < 33%. A dose level will be considered safe for a cohort if < 33% of all DLT-evaluable subjects in a given cohort experiences a DLT. Table 2. presents the probability of declaring a dose level safe (unsafe) for a range of true DLT rates for the protocol therapy in the "3+3" phase. For example, the probability of declaring a dose level safe (unsafe) is 89% (11%), 42% (58%), and 11% (89%) if the true DLT rate is 10%, 30%, or 50%, respectively. **As:** The sample size of **18** to **24 subjects will be enrolled and treated with at least 1 dose of talimogene laherparepvec with at least 9** DLT-evaluable subjects **in cohort A1. For age cohort opening and dose de-escalation, criteria from** 3+3 phase 1 designs assuming a true DLT

		incidence rate < 33% **is used. Minimum sample size with age between 2 and 12 years cohort is not required.** A dose level will be considered safe for a cohort if < 33% of all DLT evaluable subjects in a given cohort experiences a DLT **(minimum of 6 DLT evaluable subjects)**. **Table 2**. **and Figure 1.** presents the probability of declaring a dose level safe (unsafe) for a range of true DLT rates for the protocol therapy based **on 6 DLT evaluable subjects (see the triangle symbols)**. For example, the probability of declaring a dose level safe (unsafe) is 89% (11%), 42% (58%), and 11% (89%) if the true DLT rate is 10%, 30%, or 50%, respectively. **Section 4.2. Planned Covariates Delition:** • GM-CSF receptors / subunits in archival tumor tissue **Section 6. Definitions** **Treatment-emergent Adverse Events (TEAE)** **Rewrite:** Treatment-emergent adverse events (TEAE) are defined as any adverse event with an onset date 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 laherparepvec and the check box indicating prior to the first dose of IP is checked on eCRF, 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 treatment-emergent AE). **As:**

		Treatment-emergent adverse events (TEAE) are defined as any adverse event **occurring after first dose through 30 days after last dose of talimogene laherparepvec.** 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 laherparepvec and the check box indicating prior to the first dose of IP is checked on eCRF, 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 treatment-emergent AE). **Addition:** **Treatment-emergent Serious Adverse Events (TESAE)** **Treatment-emergent serious adverse events are defined as any serious adverse event occurring after initiation of the first dose of study therapy through 30 days after the last administration of study therapy.** **Treatment-related emergent Adverse Events** **Treatment-related adverse events are defined as treatment-emergent adverse events which are considered to be related to the treatment of talimogene laherparepvec, as determined by the investigator.** **Treatment-related Adverse Events in Long Term Follow-up** **Treatment -related adverse events in long term follow-up are collected after the safety follow-up visit through the end of the long-term follow-up.** **Section 7.1. DLT Analysis Set Rewrite:** 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 and received at least two treatments

		of talimogene laherparepvec (except subjects who had a DLT after the first dose) **.** 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). **As:** **The DLT analysis set will include DLT evaluable subjects defined as subjects who had the opportunity to be followed for at least 35 days from the initial dosing of talimogene laherparepvec** and received at least two treatments of talimogene laherparepvec (except subjects who had a DLT after the first dose). 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). **Delition:** Section 7.3. GM-CSF Receptors/Subunits Analysis Set The GM-CSF receptors/subunits analysis set will include all treated subjects with available baseline result of GM-CSF receptors/subunits. **Addition:** **8.1 Primary Analysis** **The primary analysis will occur 35 days after the last subject has enrolled and received at least 1 dose of talimogene laherparepvec.** **8.2 Final Analysis** **The final analysis will occur when all the subjects have discontinued the study treatment and have had the opportunity to complete the long-term follow-up visit.**

		Deleted the below text at Section 10.4 : • GM-CSF receptors / subunits in archival tumor tissue **Deleted the below text at Section 10.5.2 :** Analysis of GM-CSF receptors/subunits will be based on GM-CSF Receptors/Subunits Analysis Set. If adequate sasmples are received, the association between GM-CSF receptors/subunits in archival tumor tissue and efficacy endpoints (such as ORR, DOR, TTR, TTP, PFS, and OS) may be examined by multiple functional relationships, for example: (a) The correlation of baseline value as a continuous variable with the odds of OR, or the hazard of ending response/death/progression; (b) A test that the ORR, or hazard of ending response/death/progression is greater in subgroups of subjects with a baseline value at or above a cutoff versus below. Association between GM-CSF receptors/subunits in archival tumor tissue and safety endpoints may be examined by multiple functional relationships as described in above. **Section 10.5.3. Analyses of Exploratory Endpoints** [redacted]

		[redacted] **Section 10.6.2. Adverse Events** MedDRA version is updated to 23.0 **Addition:** **Section 11. Changes From Protocol-specified Analyses**

		In order to define DLT evaluable subjects consistently with protocol section 6.2.1.2.1, definition of DLT analysis set is updated in statistical analysis plan section 7.1 from protocol section 10.1.2.1. Subjects who had the opportunity to be followed for at least 35 days from the initial administration of talimogene laherparepvec and received at least two treatment of talimogene laherparepvec, with the exception when a DLT occurs after the first dose, will be considered DLT evaluable.

Table of Contents

List of Tables

List of Figures

List of Appendices

[redacted]

Table of Abbreviations or Term

Abbreviation or Term	Definition/Explanation
AE	adverse event
BDSG	Biomedical Data Stewardship Governance
BOR	best overall response
[redacted]	
CI	confidence interval
CRF	case report form
CNS	central nervous system
CR	complete response
CSR	clinical study report
CT	computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DILI	drug-induced liver injury
DLRT	Dose Level Review Team
DLT	dose limiting toxicity
DOR	duration of response
DRE	Disease-related events
ECG	Electrocardiogram
EDC	electronic data capture
End of Study for Individual Subject	Defined as the last day that protocol-specified procedures are conducted for an individual subject (ie, the date the subject withdraws full consent/assent 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 primary completion date is defined as the date when the last subject is assessed or receive**d** an intervention for the final collection of data for the primary endpoint(s), whether the study concluded as planned in the protocol or was terminated early. The primary completion date is anticipated to occur **35 days after** the last subject has enrolled and **received at least 1 dose of talimogene laherparepvec**.
End of Study	The end of study date is defined as the date when the last subject across all sites is assessed or receives an intervention for evaluation in the study (ie, last subject last visit), following any additional parts in the study (eg, long-term follow-up), as applicable.
End of Treatment	defined as the last assessment for the protocol-specified treatment phase of the study for an individual subject
FDA	Food and Drug Administration
GM-CSF	granulocyte macrophage colony-stimulating factor
HSV, HSV-1	herpes simplex virus, herpes simplex virus type 1

Abbreviation or Term	Definition/Explanation
HCP	Health Care Professional
IP	Investigational Product
IPD	Important Protocol Deviations
irRC-RECIST	Immune-related Response Criteria (irRC) simulating Response Evaluation Criteria in Solid Tumors (RECIST)
L-CR	Lesion complete response
L-DOR	Lesion duration of response
LLOQ	Lower Limit of Quantification
L-ORR	Lesions objective response rate
L-PR	Lesion partial response
L-TTR	Lesion time to response
MedDRA	Medical Dictionary for Regulatory Activities
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression-free survival
PFU	plaque-forming unit
PR	partial response
[redacted]	
QT interval	QT interval is a measure of the time between the start of the Q wave and the end of the T wave in the heart's electrical cycle as measured by ECG.
QTc	QT interval corrected for heart rate using accepted methodology
RECIST	Response Evaluation Criteria in Solid Tumor
SAE	serious adverse event
SD	stable disease
SOC	System organ class
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 a 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 [E6]). Examples of source data include subject identification, randomization identification, and stratification value.
Study Day 1	defined as the first day that Talimogene laherparepvec administered to the subject
TEAE	Treatment-Emergent Adverse Event
TESAE	Treatment-Emergent Serious Adverse Event

Abbreviation or Term	Definition/Explanation
ULN	Upper Limit of Normal
[redacted]	

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 3** for talimogene laherparepvec Study 20110261 dated **12 June 2020.** The scope of this plan includes the primary analysis and the final analysis that are planned and will be executed by the Biostatistics department unless otherwise specified.

The clinical study report (CSR) will be written based on the results of the primary analysis. Data collected and analyzed in Amgen-owned databases and systems will adhere to approved Data Element Standards and International Case Report Form (CRF) Standards established by Biomedical Data Stewardship Governance (BDSG).

2. Objectives

2.1 Primary

To evaluate the safety and tolerability of talimogene laherparepvec, as assessed by incidence of dose-limiting toxicities (DLT), in pediatric subjects with advanced non-central nervous system (CNS) tumors that are amenable to direct injection.

2.2 Secondary

- To evaluate the anti-tumor activity of talimogene laherparepvec, as assessed by overall response rate (ORR), duration of response (DOR), time to response (TTR), time to progression (TTP), progression-free survival (PFS) using immune related Response Criteria Simulating Response Evaluation Criteria in Solid Tumors(RECIST 1.1) [Modified irRC RECIST], and overall survival (OS).

2.3 Safety

To evaluate the safety and tolerability of talimogene laherparepvec

2.4 Exploratory

3. Study Overview

3.1 Study Design

This is a phase 1, multicenter, open label study of talimogene laherparepvec in pediatric subjects with advanced non-CNS tumors that are amenable to direct injection in the clinical setting.

Approximately 18 **to 24** pediatric subjects are expected to be enrolled and **treated with at least 1 dose of talimogene laherparepvec** into 2 cohorts stratified by age (permissible based on the incidence of DLTs, a minimum of 18 **dosed** subjects total **for the primary analysis**) as follows:

- Cohort A1 (12 to $\leq$ 21 years of age)
- Cohort B1 (2 to $<$ 12 years of age)

Initially, 3 subjects 12 to $\leq$ 21 years of age are to be enrolled and treated at 100% of the recommended adult dose regimen of talimogene laherparepvec (Cohort A1). The first dose administered will be up to 4.0 mL of 10^6 PFU/mL followed by a dose of up to 4.0 mL of 10^8 PFU/mL 21 days (+3 days) later. Subsequent doses of up to 4.0 mL of 10^8 PFU/mL will be administered approximately every 14 days ($\pm$ 3 days) thereafter (see protocol Section 6.2.1.1). The DLT evaluation period is 35 days from the initial administration of talimogene laherparepvec. Rules for DLT evaluation are described in Section 6.2.1.2.1.of the protocol.The dose level review team (DLRT) will review the safety data of the first 3 subjects in the older age cohort A1 to decide if the younger age cohort B1 can be opened for enrollment. If none of the first 3 DLT-evaluable subjects in cohort A1 experiences a DLT then cohort B1 will be opened for enrollment and treatment at the same dose level per the section Study Design and Treatment Schema of the protocol. If a DLT occurs in the first 3 DLT-evaluable subjects in the older age cohort (A1 or A2), the younger age cohort will not open until a DLT rate $<$ 33% is observed with at least 6 DLT-evaluable subjects in the older age cohort (A1 or A2). If any of the first 3 subjects in cohort A1 experiences a DLT, the DLRT will make a recommendation on whether an additional 3 subjects should be enrolled in cohort A1.

Cohort A1 or B1 can enroll up to a maximum of **18 subjects treated with at least 1 dose of talimogene laherparepvec with at least 9** DLT-evaluable subjects **in cohort A1.** In the case of dose de-escalation, the minimum number of DLT-evaluable subjects at the de-escalated dose will be **6** (permissible based on the incidence of DLTs). After an age cohort is closed for further enrollment, if $<$ 33% of all DLT-evaluable subjects in

the cohort experiences a DLT at a dose level, the dose will be declared safe for the cohort.

If dose de-escalation is needed and if permissible based on the incidence of DLTs, a minimum of **6** additional DLT evaluable subjects will be enrolled and treated at a lower dose level of talimogene laherparepvec (Dose Level -1). The initial dose administered will be up to 4.0 mL of 10^6 PFU/mL followed by dose of up to 4.0 mL of 10^6 PFU/mL 21 days (+ 3 days) later. Subsequent doses of up to 4.0 mL of 10^6 PFU/mL will be administered approximately every 14 days (± 3 days) thereafter (see Table 1). Dose de-escalation cohorts will obey the following naming convention, based on age at baseline, and the same DLT rules will be applied to the de-escalated cohorts:

- Cohort A2 (12 to ≤ 21 years of age)
- Cohort B2 (2 to < 12 years of age)

Table 1. Dose De-escalation

Dose Level	Initial Dose	Subsequent Doses
1	**10^6 PFU/mL**	**10^8 PFU/mL**
-1	**10^6 PFU/mL**	**10^6 PFU/mL**

3.1.1 Tumor Response Assessment

Assessment of tumor response will be performed according to modified irRC-RECIST (Nishino et al, 2013 and Nishino et al, 2014). Details can be found in protocol Appendix D.

3.1.1.1 Safety Follow-up

Adverse events will be collected as described in protocol Section 9.2. All subjects will complete a safety follow up visit approximately 30 (+ 7) days after the last dose of study treatment. Adverse events and any concomitant medications associated with adverse events that occur 30 (+ 7) days following cessation of treatment will be reported, followed, and recorded in the case report form (CRF).

3.1.1.2 Long-term Follow-up

After the safety follow up visit, all subjects will enter the long term follow up. Subjects will be followed for survival and use of subsequent anticancer therapies every 12 weeks (± 28 days) from the safety follow-up **visit** until death, subject or legally acceptable representative withdraws full consent/ assent, or up to approximately 24 months after the last subject is enrolled in the study **whichever occurs first.** In addition, talimogene

laherparepvec related adverse events that occur through the end of the long term follow up will be reported.

3.2 Sample Size

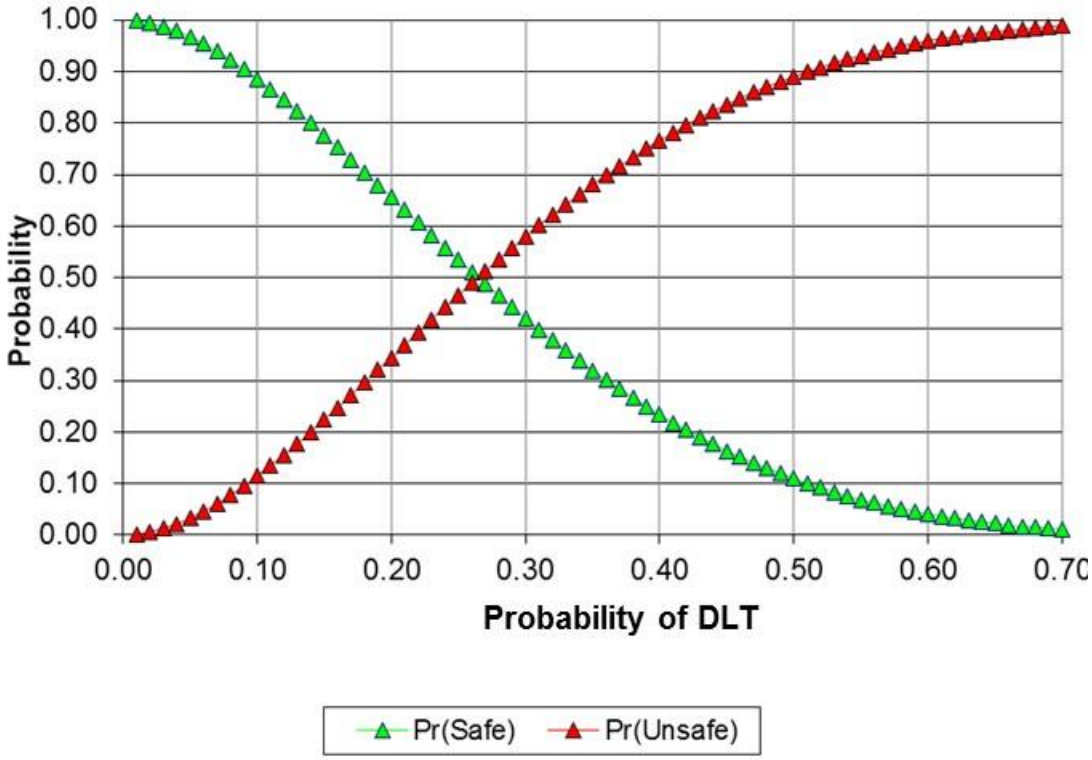
The sample size of **18** to **24 subjects will be enrolled and treated with at least 1 dose of talimogene laherparepvec with at least 9** DLT-evaluable subjects **in cohort A1. For age cohort opening and dose de-escalation, criteria from** 3+3 phase 1 designs assuming a true DLT incidence rate < 33% **is used. Minimum sample size with age between 2 and 12 years cohort is not required.**

A dose level will be considered safe for a cohort if < 33% of all DLT evaluable subjects in a given cohort experiences a DLT **(minimum of 6 DLT evaluable subjects).** Table 2. and Figure 1. presents the probability of declaring a dose level safe (unsafe) for a range of true DLT rates for the protocol therapy **based on 6 DLT evaluable subjects (see the triangle symbols).** For example, the probability of declaring a dose level safe (unsafe) is 89% (11%), 42% (58%), and 11% (89%) if the true DLT rate is 10%, 30%, or 50%, respectively.

Table 2. 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%

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



4. Study Endpoints and Covariates

4.1 Study Endpoints

4.1.1 Primary Endpoints

- Subject incidence of DLTs

4.1.2 Secondary Endpoints

- ORR, DOR, TTR, TTP, and PFS using modified irRC-RECIST
- OS

4.1.3 Safety Endpoints

- Subject incidence of adverse events and clinical significant laboratory abnormalities

4.1.4 Exploratory Endpoints

[redacted]

4.2 Planned Covariates

The following covariates may be used to examine clinical activity in subgroups or in multivariate analysis.

- Age at baseline: (2 to $<$ 12 years, 12 to $\leq$ 21 years)
- Sex (female versus male)
- HSV-1 baseline serostatus (positive versus negative)
- Baseline sum of largest diameter of target lesions
- Extent of disease (localised versus metastatic)
- Performance status (Karnofsky/Lansky play scale (70% versus $>$ 70%))
- Prior lines of therapy in the recurrent/metastatic setting (0,1, 2, $>$ 2)
- Prior surgery for current malignancy (yes versus no)
- Prior radiotherapy for current malignancy (yes versus no)
- Complete surgical resection of all gross diseases (yes versus no)
- 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.

5. Hypotheses and/or Estimations

Talimogene laherparepvec injected in pediatric subjects with advanced non-CNS tumors that are amenable to direct injection will be safe as assessed by subject incidence of DLTs and the safety profile.

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, 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.

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 after the start of the first subsequent anticancer therapy, including complete or partial removal/reduction of any target lesion, 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 3). 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 49 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 3. 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 ≥ 49 days from first dose; however, this is not required for an unconfirmed CR/PR.

Table 3. Matrix of Determining BOR per Modified irRC-RECIST

Visit Overall Response Sequence	Examples		Best Overall Response	Confirmation specifications
*,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 < 49 days from first dose (or randomization if not treated).

Page 2 of 2

Duration of response (DOR) per Modified irRC-RECIST

Duration of response (DOR) 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 their last evaluable tumor assessment.

Evaluable tumor assessment

An overall visit response other than UE.

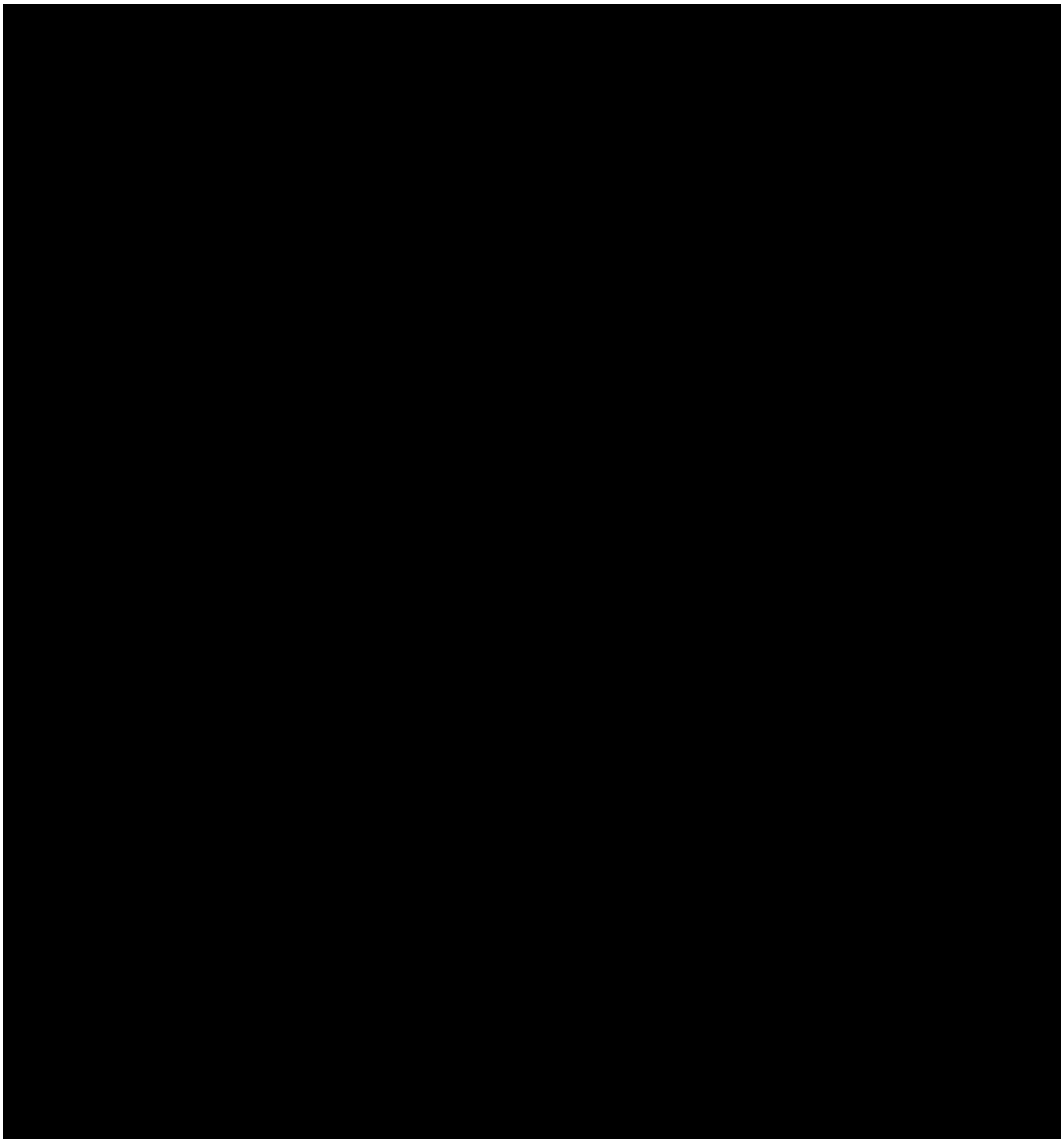
Event of Interest (EOI)

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

Investigational product (IP)

Investigational product refers to talimogene laherparepvec in this study.

[redacted]



Objective response rate (ORR) per Modified irRC-RECIST

ORR is defined as the incidence of a BOR of CR or PR per Modified irRC-RECIST 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 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.

Potential Follow-Up Time

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

[redacted]

Progression-free survival (PFS) per Modified irRC-RECIST

Progression-free survival (PFS) per Modified irRC-RECIST is defined as the interval from first dose to the earlier of PD per Modified irRC-RECIST or death from any cause; otherwise, PFS is censored at the last evaluable tumor assessment. The initial date of PD will be the PFS date when it is consecutively confirmed.

[redacted]

Study day

Study day is calculated from the first day when investigational product is administered to the subject.

That is, Study day = visit date – first dose date + 1, if visit date is after first dose date

Study Day = first dose date – visit date, if visit date is prior to first dose date

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 1

The start of Investigational product administration to the subject is study week 1. Study day 1 is corresponding to the first day of study week 1.

Time to progression (TTP): defined as the interval from first dose to the date of the first PD per Modified irRC-RECIST; otherwise, TTP is censored at the last evaluable tumor assessment.

Time to response (TTR): (calculated only for those subjects in the safety analysis set with confirmed CR or PR) time from first dose to the date of the first confirmed 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 (TEAE) are defined as any adverse event **occurring after first dose through 30 days after last dose of talimogene laherparepvec.** 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 laherparepvec and the check box indicating prior to the first dose of IP is checked on eCRF, 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 treatment-emergent AE).

Treatment-emergent Serious Adverse Events (TESAE)

Treatment-emergent serious adverse events are defined as any serious adverse event occurring after initiation of the first dose of study therapy through 30 days after the last administration of study therapy.

Treatment-related emergent Adverse Events

Treatment-related adverse events are defined as treatment-emergent adverse events which are considered to be related to the treatment of talimogene laherparepvec, as determined by the investigator.

Treatment-related Adverse Events in Long Term Follow-up

Treatment -related adverse events in long term follow-up are collected after the safety follow-up visit through the end of the long-term follow-up.

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).

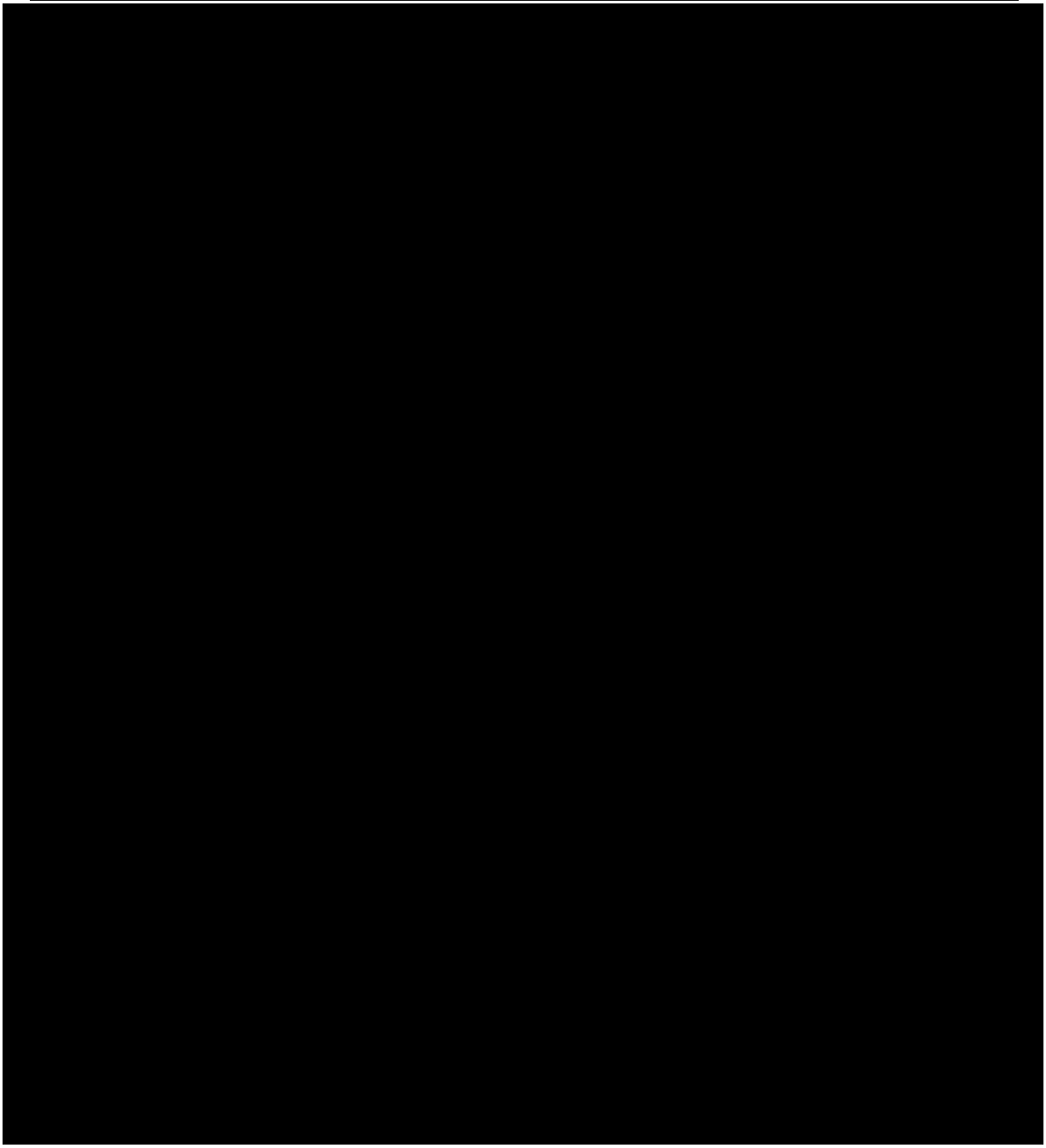
7. Analysis Subsets

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 from the initial dosing of talimogene laherparepvec and received at least two treatments of talimogene laherparepvec (except subjects who had a DLT after the first dose). 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.



7.7 Subgroup Analyses

See Section 4.2.

8. 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 by the DLRT (see protocol section 10.3.2).

8.1 Primary Analysis

The primary analysis will occur 35 days after the last subject has enrolled and received at least 1 dose of talimogene laherparepvec.

8.2 Final Analysis

The final analysis will occur when all the subjects have discontinued the study treatment and have had the opportunity to complete the long-term follow-up visit.

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

[redacted]

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 and concomitant medications and incomplete death dates will be imputed. Details of the imputation algorithm are provided in Appendix A, section 13. 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). IPDs 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, safety, talimogene laherparepvec DNA, and other exploratory endpoints for all subjects and by cohort. 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 (Clopper & Pearson 1934). Exact tests will be considered for subgroup analyses. Time to-event endpoints will be estimated using the K-M method.

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 countries and investigator sites overall and by cohort (based on age). 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: (2 to $<$ 12 years, 12 to $\leq$ 21 years)
- Sex (female versus male)
- HSV-1 baseline serostatus (positive versus negative)
- Baseline sum of largest diameter of target lesions
- Extent of disease (localized versus metastatic)
- Performance status (Karnofsky/Lansky play scale (70% versus $>$ 70%)
- Prior lines of therapy in the recurrent/metastatic setting (0,1, 2, $>$ 2)
- Prior surgery for current malignancy (yes versus no)
- Prior radiotherapy for current malignancy (yes versus no)
- Complete surgical resection of all gross diseases (yes versus no)
- 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

10.5.1 Analyses of Secondary Endpoints - ORR, DOR, TTR, TTP, PFS, OS

Analysis of secondary endpoints such as ORR, DOR, TTR, TTP, PFS, OS will be based on Safety Analysis Set and excluding subjects without baseline measurable disease, if applicable.

- ORR (CR+PR) will be summarized with associated exact 95% CIs for binomial proportions (Clopper & Pearson 1934).

- DOR among responders, TTP, 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.

10.5.2 Analyses of Exploratory Endpoints

[redacted]

10.6 Safety Analyses

10.6.1 Analyses of Primary Endpoints

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 **23.0** 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, 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, 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), fatal disease related events, serious disease-related events, and disease-related events leading to withdrawal from study will be tabulated by system organ class and preferred term.

Treatment-related AEs reported in long term follow-up will be summarized in this study and analyzed at a program-level.

Potential or known unintended exposure to talimogene laherparepvec, related suspected signs or symptoms, and detection of exposure to talimogene laherparepvec in a subject's household member, caregiver, or healthcare provider will be reported to Amgen. Exposures, signs or symptoms will be summarized in this study and analyzed 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. For a list of laboratory measurements see protocol Section 7.4 Table 3. 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 statistics will be presented for systolic/diastolic blood pressure, heart rate, respiratory rate, and temperature for baseline, each post-baseline visit, and change from baseline.

10.6.5 Lansky Play Scale and Karnofsky Performance Status

Performance status scores will be summarized at each assessed time point. The change in scores from baseline to each assessed time point will also be summarized.

10.6.6 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.

10.6.7 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. Since 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 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, withdrawal and/or injections with $\leq$ 4 mL will be tabulated.

10.6.9 Exposure to Concomitant Medication/On-study Surgery or Procedure/Subsequent Anti-cancer Therapy

The concomitant medications include all medications during the treatment period except for planned study drugs. The number and proportion of subjects receiving therapies of interest will be summarized by preferred term or cohort as coded by the World Health Organization Drug (WHO DRUG) dictionary.

11. Changes From Protocol-specified Analyses

In order to define DLT evaluable subjects consistently with protocol section 6.2.1.2.1, definition of DLT analysis set is updated in statistical analysis plan section 7.1 from protocol section 10.1.2.1. Subjects who had the opportunity to be followed for at least 35 days from the initial administration of talimogene laherparepvec and received at least two treatment of talimogene laherparepvec, with the exception when a DLT occurs after the first dose, will be considered DLT evaluable.

12. Literature Citations / References

Agresti, A. (2002) *Categorical Data Analysis*, 2nd Edition, New York: John Wiley & Sons.

Brookmeyer, R., *et al* (1982) “A Confidence Interval for the Median Survival Time.” *Biometrics*, **38**, 29-41 Clopper, C., *et al* (1934) “The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial,” *Biometrika*, **26**, 404-413

Eisenhauer, A., *et al* (2009) “New Response Evaluation in Solid Tumors: Revised RECIST Guideline (version 1.1)”, *EJC*, **45**, 228-247.

Kalbfleisch, J.D., *et al* (1980) *The Statistical Analysis of Failure Time Data*, New York: John Wiley & Sons.

Nishino, M., *et al* (2014) “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.

Nishino, M., *et al* (2013) “Developing a common language for tumor response to immunotherapy: Immune-related response criteria using unidimensional measurements”. *Clin Cancer Res*; **19**(14); 3936-43.

Wolchok, J.D., *et al*. (2009) “Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria.” *Clin Cancer Res,* **15**(23), 7412-7420.

13. Appendices