Application Type	Original Application
STN	125518/0
CBER Received Date	July 28, 2014
PDUFA Goal Date	October 27, 2015
Division / Office	DCEPT /OCTGT
Committee Chair	Ramjay Vatsan
Clinical Reviewer(s)	Peter Bross; Robert Le; Maura C. O'Leary; Ke Liu; Marc R. Theoret
Project Manager	Mark L. Davidson
Priority Review	No
Reviewer Name(s)	Yuqun Abigail Luo
Review Completion Date / Stamped Date	August 24, 2015
Supervisory Concurrence	Boguang Zhen
	John Scott
Applicant	BioVex Inc., a subsidiary of Amgen, Inc.
Established Name	talimogene laherparepvec
(Proposed) Trade Name	IMLYGIC®
Pharmacologic Class	rHSV-1h^{GM-CSF} oncolytic immunotherapy
Formulation(s), including Adjuvants, etc	<No Formulations>
Dosage Form(s) and Route(s) of Administration	talimogene laherparepvec single-use vials (10^6 and 10^8 PFU/mL); intralesional injection into cutaneous, subcutaneous, and nodal lesions
Dosing Regimen	The initial dose of talimogene laherparepvec is up to 4 mL of 10^6 PFU/mL followed by 4 mL of 10^8 PFU/mL administered 3 weeks later; thereafter, subsequent doses of 4 mL of 10^8 PFU/mL are administered every 2 weeks.
Indication(s) and Intended Population(s)	Treatment of injectable regionally or distantly metastatic melanoma. [The indication is currently under discussion between the applicant and the FDA and as a result may be revised.]

Table of Contents

GLOSSARY

AE	Adverse Event
BD	Briefing Document for the Advisory Committee Meeting
BLA	Biologics License Application
CMC	Chemistry, Manufacturing, And Controls
CMV	Cytomegalovirus
CR	Complete Response
CSR	Clinical Study Report
CTGTAC	The Cellular, Tissue and Gene Therapies Advisory Committee
CTLA-4	Cytotoxic T Lymphocyte Antigen-4
DIS	Division of Inspections and Surveillance
DMC	Data Monitoring Committee
DR	Durable Response, Durable Responder
DRR	Durable Response Rate
DTIC	Dacarbazine
EAC	Endpoint Assessment Committee
ECOG	Eastern Cooperative Oncology Group
eCTD	electronic Common Technical Document
EIA	Efficacy Information Amendment
ES	Executive Summary
FACT-BRM	Functional Assessment of Cancer Therapy Biologic Response Modifier
GM-CSF	Granulocyte Macrophage Colony-Stimulating Factor
p	p-value
HR	Hazard Ratio
HSV-1	Herpes Simplex Virus Type-1
ICH	International Conference on Harmonisation (of Technical Requirements for Registration of Pharmaceuticals for Human Use)
IL-2	High-Dose Interleukin-2
ISE	Integrated Summary Of Efficacy

ITT	Intent-to-Treat
LDH	Lactate Dehydrogenase
mAb	Monoclonal Antibody
MAPK	Mitogen-Activated Protein Kinase
MedDRA	Medical Dictionary for Regulatory Activities
MOA	Mechanism of Action
NCI	National Cancer Institute
NDA	New Drug Application
NME	New Molecular Entity
OBE	Office of Biostatistics and Epidemiology
ODAC	Oncologic Drugs Advisory Committee
OCBQ	Office of Compliance and Biologics Quality
OR	Odds Ratio
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease
PD-1	Programmed Death 1 Molecule
(b) (4)	(b) (4)
(b) (4)	(b) (4)
PDcns	Central Nervous System Progressive Disease
PDn	Clinically Not Relevant Progressive Disease
PDr	Clinically Relevant Progressive Disease
PFU	Plaque Forming Unit
PFS	Progression Free Survival
PI	Package Insert
PK	Pharmacokinetics
PMC	Postmarketing Commitment
PMR	Postmarketing Requirement
PP	Per Protocol
PR	Partial Response

RECIST	Response Evaluation Criteria in Solid Tumors
RR	Relative Risk
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SC	Subcutaneous
SD	Stable Disease
SPA	Special Protocol Assessment
T-VEC	Talimogene Laherparepvec
WHO	World Health Organization

1. EXECUTIVE SUMMARY

BLA 125518/0 was submitted for talimogene laherparepvec (IMLYGIC®), for the proposed indication of treatment of regionally or distantly metastatic melanoma with injectable tumors.

IMLYGIC® is a first-in-class oncolytic virus immunotherapy. It is designed to genetically engineer Herpes Simplex Virus-1 to attenuate neuro-virulence, enhance preferential viral replication in tumor tissues, and to express the human Granulocyte Macrophage Colony-Stimulating Factor (GM-CSF) gene.

The primary efficacy and safety data come from study 005/05, an open-label study comparing intra-lesion administration of IMLYGIC® with subcutaneous administration of GM-CSF in treatment of melanoma patients with unresectable but injectable stage IIIB, IIIC, or IV disease. The intent-to-treat (ITT) population consists of 295 IMLYGIC® subjects and 141 GM-CSF subjects. The primary efficacy endpoint is durable response (DR), defined as a maintenance of response (>50% decrease of tumor burden) continuously for at least six months, which was also confirmed by an independent, blinded endpoint assessment committee. Overall survival (OS) is an important secondary endpoint. Power for both DR and OS were taken into account in the study design.

The primary analysis comparing the durable response rate (DRR) in the ITT population between the two arms is statistically highly significant ($p < .0001$), with a DRR of 16.3% (48/295) in the IMLYGIC® arm vs. 2.1% (3/141) in the GM-CSF arm. The GM-CSF subjects had on average a much shorter duration of study treatment and response assessment, compared to IMLYGIC® subjects. Because of this, the reported DRR in the GM-CSF arm may be an underestimate. However, I consider the statistical significance of the comparison of DRR between the two arms to be statistically robust. In addition, 19 of the 48 IMLYGIC® DRs are durable complete responders, with complete response maintained for at least six months, accounting for 6.4% of the ITT IMLYGIC® subjects.

The primary analysis comparing OS between the two arms in the ITT population, at the time of database lock, was just short of being statistically significant, at a p-value of 0.051, with a hazard ratio (HR) estimate of 0.79 and a 95% confidence interval (CI) of (0.62, 1.00). The median OS from the primary analysis is 23.3 months (95% CI: 19.6-29.7) in the IMLYGIC® arm and 18.9 months (95% CI: 16.2-24.0) in the GM-CSF arm. I identified a total of 10 subjects with potentially informatively censored event times. These 10 subjects were distributed disproportionately between the two arms, accounting for 5% (7/141) of the GM-CSF subjects and 1% (3/295) of the IMLYGIC® subjects, respectively. Additional retrospective information, on survival data up to the data cut-off date for the primary analysis, was subsequently obtained by the applicant for five of the 10 subjects. The updated analysis, incorporating this additional information from these five subjects, yields a p-value of 0.116, a hazard ratio estimate of 0.82 and a 95% CI of (0.65, 1.05). The median OS from the updated analysis is 22.9 months (95% CI: 19.6-29.7) in the IMLYGIC® arm and 19.0 months (95% CI: 16.2-24.3) in the GM-CSF arm. The updated survival curves are still visually separate, though to a lesser extent than the primary analysis, favoring the IMLYGIC® arm.

Subgroup analyses reveal that DRR in the IMLYGIC® arm is substantially higher in the subset of subjects with earlier stage disease, compared to later stage disease. The DRR is

33% in the 131 IIIB/IIIC subjects, 16.0% in the 118 IVM1a subjects, 3.8% in the 90 IVM1b subjects, and 3.4% in the 96 IVM1c subjects. Subgroup analyses of OS show a similar trend: in IIIB/IIIC, the median OS are 25.7 months vs "Not reached", for the GM-CSF and IMLYGIC® arms, respectively; the medians are 19.3 vs 29.9 in IVM1a, 12.9 vs 13.6 in IVM1b, and 16.2 vs 12.6 in IVM1c. Multiplicity control was not planned for subgroup analyses. Caution should be applied when considering the observed differences between subgroups in the comparison of the two arms. In particular, I recommend viewing the subgroup analyses of OS as supportive information for the subgroup analyses of DRR, rather than as definitive evidence of a survival benefit for IMLYGIC®.

A joint meeting of the Cellular, Tissue, and Gene Therapies Advisory Committee (AC) and Oncologic Drugs AC was held on April 29, 2015. The AC voted 22 "yes" and one "no" to the question of whether IMLYGIC® has a favorable benefit-risk profile to support traditional approval. A number of AC members qualified their votes by stating that they would want the approval to be limited to earlier stage disease or patients without visceral disease.

2. CLINICAL AND REGULATORY BACKGROUND

Note: Verbatim excerpts from the applicant's submission or identified literature are italicized when this reviewer would like to emphasize the source of the information.

BioVex, Inc., a subsidiary of Amgen, Inc., submitted this original biologics license application (BLA) for talimogene laherparepvec for the proposed indication of treatment of regionally or distantly metastatic melanoma with injectable tumors. Talimogene laherparepvec, during its clinical development, was also known as OncoVEX^{GM-CSF}, (b) (4) , and T-VEC. It will be referred to as IMLYGIC® in this review memo.

IMLYGIC® is designed to be an oncolytic virus immunotherapy, a replication-competent virus genetically engineered from an attenuated Herpes Simplex Virus-1 (HSV-1) isolate (newly isolated strain JS1; ECAAC Accession Number 01010209). HSV-1 is a non-integrating double stranded DNA virus. The genetic modification to the HSV-1 genome includes the following.

1. Deletion of ICP34.5 attenuates replication of the virus in normal tissues and thus allows selective replication of the virus in tumor tissues. This deletion also reduces neuro-virulence by 10,000 to 1,000,000 fold as compared to wild-type HSV-1.
2. Deletion of ICP47 prevents down-regulation of antigen presentation molecules and increases the expression of HSV US11 gene, which enhances viral replication in tumor cells.
3. Insertion of human Granulocyte Macrophage Colony-Stimulating Factor (GM-CSF) gene.

The therapeutic mechanism of action (MOA) of IMLYGIC® is postulated to be two-fold, after direct injection into a number of tumors in a patient: the virus selectively replicates inside and subsequently lyses a cancer cell, secreting the immune-stimulatory cytokine GM-CSF in the process and releasing an array of tumor-specific antigens to elicit a

systemic anti-tumor immune response. This MOA is postulated to result in the destruction of injected and non-injected tumors (including micrometastatic disease) and to reduce the development of new metastases. Clinically, the intended biologic effects are delay or prevention of disease progression and relapse, and the prolongation of overall survival (OS).

2.1 Disease or Health-Related Condition(s) Studied

Melanoma in adults is staged according to the American Joint Committee on Cancer (AJCC) melanoma tumor, node, metastasis (TNM) staging. Stages are based on the thickness and ulceration of the primary tumor, degree of lymph node involvement, and presence and location of metastases. Melanoma can spread by local extension (through lymphatics) or to distant sites (by hematologic routes) to any organ, most commonly lungs and liver. For the confirmatory trial supporting this BLA, eligibility required histologically confirmed diagnosis at study entry of stage IIIB, IIIC, or IV malignant melanoma that was not surgically resectable.

- Stage IIIB
    - T1-4b and N1-2a. The thickness of the primary tumor ranges from < 1mm to > 4mm, and "b" denotes "with ulceration". N1 and N2 denote 1 or 2-3 regional lymph node metastases, respectively, and "a" denotes "with micrometastasis".
    - T1-4a and (N1-2b, N2c). For the primary tumor, "a" denotes "without ulceration". For lymph node metastases, "b" denotes "with macrometastasis". N2c denotes "In transit met(s)/satellite(s) without metastatic lymph nodes".
- Stage IIIC
    - T1-4b and (N1-2b, N2c).
    - N3: ≥4 regional lymph node metastases; or matted nodes; or in transit met(s)/satellite(s) with metastatic lymph node(s).
- Stage IV
    - M1a: metastases to skin, subcutaneous, or distant lymph nodes and normal serum lactate dehydrogenase (LDH).
    - M1b: metastases to lung and normal serum LDH.
    - M1c: metastases to all other visceral sites and normal serum LDH; or distant metastases to any site and elevated serum LDH.

The applicant provided the following background in the CSR:

> *In adults, cutaneous melanoma is the fifth most common cancer in men and the seventh most common cancer in women in the United States. In the US (population: 317 million), an estimated 76,690 people (24 per 100,000) are diagnosed with melanoma and 9,480 people (3 per 100,000) die of melanoma annually. Of the people*

newly diagnosed with melanoma, approximately 10,000 people have regional or distant (metastatic) disease.

Melanoma that has spread to multiple regional nodal sites or presents with in transit/satellite lesions (Stage IIIB/C) is infrequently curable with standard therapy; 5-year survival rates range between 40% (for IIIC disease) and 59% (for IIIB disease). Melanoma that has spread to distant skin, nodes, or visceral organs (stage IV) is also infrequently curable with standard therapy. For patients with stage IV disease, 1-year survival rates are generally poor, ranging from 62% for M1a disease (skin or nodes only), to <53% for M1b disease (lung only), and 33% for M1c disease (other visceral lesions or high LDH).

2.2 Currently Available, Pharmacologically Unrelated Treatment(s)/Intervention(s) for the Proposed Indication(s)

Section 2.2 briefly summarizes available therapies and current development in the systemic treatment for unresectable Stage III, Stage IV, or recurrent melanoma, based on this statistical reviewer's layperson understanding. The purpose is to provide a context for the statistical evaluation of the submission. The summary is not intended to be comprehensive, nor up-to-date because of the ongoing rapid changes in the therapeutic landscape. Of note, the risk aspect of the benefit-risk profiles of these therapies is not summarized here. This summary is based heavily on *National Cancer Institute: PDQ® Melanoma Treatment. Bethesda, MD: National Cancer Institute. Date last modified <11/07/2014>. Available at: http://www.cancer.gov/cancertopics/pdq/treatment/melanoma/HealthProfessional. Accessed <12/07/2014>*. Many other references were consulted, but not identified individually here, to inform this summary.

The therapeutic landscape changed dramatically in 2011 with the approval of ipilimumab, ending a stretch from 1992 to 2011 without new approval by the FDA and an era where even first-line therapy was of questionable value as compared with supportive care. Approvals of treatment for advanced melanoma happened in rapid succession in the ensuing years, and clinical development of new agents and combinations of new agents is active.

Agents in two novel classes have demonstrated improvement in progression-free survival (PFS) and overall survival (OS) in randomized trials, compared to chemotherapy dacarbazine (DTIC) or other comparators. The first class is immunotherapy using monoclonal antibody (mAb) as immune checkpoint blockade targeting T-cell inhibitory immune receptor, e.g., ipilimumab (anti-CTLA-4, cytotoxic T-lymphocyte antigen-4, approved in 2011) and pembrolizumab and nivolumab (anti-PD-1, programmed cell death-1, accelerated approvals in September and December 2014, respectively, breakthrough therapy designation). The second class consists of signal transduction inhibitors, targeted therapies using small molecules targeting proteins from activating mutations in oncogenes along the mitogen-activated protein kinase (MAPK: RAF–MEK–ERK) pathway, e.g, vemurafenib and dabrafenib (BRAF inhibitors, approved 2011 and 2013) and trametinib (MEK inhibitor, approved 2013). The latter three agents were

approved for treatment of patients with BRAF V600E or V600K mutations, which account for 40% to 60% of malignant melanomas.

Immune checkpoint blockades generate durable responses in relatively small proportions of patients, whereas BRAF and MEK inhibition in BRAFV600-mutated melanoma shows a high response rate with limited durability. Resistance develops in a majority of patients treated by single agents targeting the MAPK pathway, resulting in a median PFS of 6 to 7 months. In addition, BRAF-inhibitors can paradoxically activate the MAPK pathway, leading to secondary cancers, including cutaneous squamous-cell carcinoma, and may reactivate RAS-mutant tumors.

Clinical development is current focused on several fronts.

- Anti-PD-1 and Anti-PD-L1 (programmed death ligand 1) agents are viewed as more active and less toxic than ipilimumab due to their more tumor-specific mode of immune activation. Multiple agents are in active late phase development, some in multiple cancer types. There is also development in using tumor PD-L1 expression as a biomarker to predict response to PD-1 antibodies.
- Therapies that have improved OS in patients with recurrent or metastatic disease are now being tested as adjuvant therapy in clinical trials.
- Various combination therapies, including variation in sequencing and timing, are being actively developed. In 2014, the combination of dabrafenib and trametinib received accelerated approval from the FDA for patients with unresectable or metastatic melanomas that carry the BRAF V600E or V600K mutation. The combination demonstrated improved durable response rates over single-agent dabrafenib. On December 15, 2014, Genentech announced submission of a New Drug Application (NDA) for the combination of cobimetinib (a MEK inhibitor) with vemurafenib for a similar population, claiming demonstration of prolonged PFS compared to vemurafenib alone.

In addition to the recent novel classes of therapies summarized above, the following therapies are also available.

- Immunotherapy Interleukin-2 (IL-2) was approved by the FDA in 1998 on the basis of durable complete response (CR) in a minority of patients (6%–7%) with previously treated metastatic melanoma in eight phase I and II studies. A small subset of complete responders (6%) achieved long-term response or near cure.
- Chemotherapy includes DTIC and Temozolomide.
- Palliative local therapy. Melanoma metastatic to distant, lymph node-bearing areas may be palliated by regional lymphadenectomy. Isolated metastases to the lung, gastrointestinal tract, bone, or sometimes the brain may be palliated by resection, with occasional long-term survival.

2.5 Summary of Pre- and Post-submission Regulatory Activity Related to the Submission

- April 25, 2005. The original submission under IND 12412, the IND for all studies supporting this BLA, was received by the FDA.
- April 17, 2008. The protocol for the sole confirmatory study, study 005/05, submitted under IND serial #0041, received special protocol assessment (SPA) concurrence. The version date of the draft protocol was April 14, 2008. The protocol was later finalized (October 2, 2008), and Statistical Analysis Plan (SAP) was also finalized (September 9, 2008). The major amendments to the protocol and SAP are listed below, together with the corresponding SPA concurrence date.
    - Amendment #1. IND serial #0058, protocol version October 15, 2008, data monitoring committee (DMC) charter April 29, 2009. SPA concurred November 19, 2008.
    - Amendment #2. IND serial #0096, protocol version November 17, 2009, SAP November 30, 2009. SPA concurred December 1, 2009.
    - Amendment #3. IND serial #0119, protocol version July 18, 2010, SAP June 21, 2010. SPA concurred August 6, 2010.
    - Amendment #4. IND serial #0078, IND amendment #0246, protocol version November 30, 2011, SAP November 30, 2011, SPA concurred November 17, 2011.
    - Amendment #5. IND serial #0130, IND amendment #0297, protocol version January 4, 2013, SAP January 4, 2013, SPA concurred January 23, 2013.
- January 21, 2011. The FDA granted Fast Track designation to the clinical program of investigation of IMLYGIC® for treatment of unresectable Stages IIIb, IIIc and IV melanoma.
- March 14, 2011. The FDA granted IMLYGIC® Orphan Drug designation for treatment of stage IIb to stage IV melanoma, a broader indication than requested by the applicant in this BLA.
- December 06, 2013. The FDA granted rolling submission of the BLA.
- December 13, 2013. The applicant submitted the first component of the rolling BLA, a full Module 4 and corresponding sections in Module 2, along with some sections in Module 1.
- May 1, 2014. The applicant submitted the second component of the rolling BLA, a full Module 3 and corresponding sections in Module 2.
- July 28, 2014. The applicant submitted the third and final component of the rolling BLA, a full Module 5 and corresponding sections in Module 2. Electronic radiology and dermatology images were submitted separately on external drives and DVDs for archiving.

- December 11, 2014. The FDA notified the applicant that an amendment on facility received on November 28, 2014 was classified as a Major amendment, and therefore the review clock was extended for 3 more months. The action due date is now October 27, 2015.
- April 29, 2015. Joint meeting of two advisory committees (ACs): the Cellular, Tissue, and Gene Therapy AC (CTGTAC) and the Oncologic Drugs AC (ODAC).

3. SUBMISSION QUALITY AND GOOD CLINICAL PRACTICES

3.1 Submission Quality and Completeness

The submission was adequately organized for conducting a complete statistical review without unreasonable difficulty.

5. SOURCES OF CLINICAL DATA AND OTHER INFORMATION CONSIDERED IN THE REVIEW

5.1 Review Strategy

This statistical review will focus on the single confirmatory trial, Study 005/05.

5.2 BLA/IND Documents That Serve as the Basis for the Statistical Review

The basis of this statistical review includes documents in IND 12412 and the original BLA 125518. There have been multiple information requests (IRs) from the FDA and responses from the applicant. These interactions will be identified in the relevant review sections. The clinical study reports (CSRs) of Study 005/05, one on the primary analysis and the other on the supplemental overall survival (OS) and systemic effects, together with the relevant datasets submitted in the eCTD, form the primary basis of this review. The applicant's briefing document (BD) and presentation for the advisory committee meeting have also been reviewed:
http://www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/BloodVaccinesandOtherBiologics/CellularTissueandGeneTherapiesAdvisoryCommittee/ucm433808.htm

5.3 Table of Studies/Clinical Trials

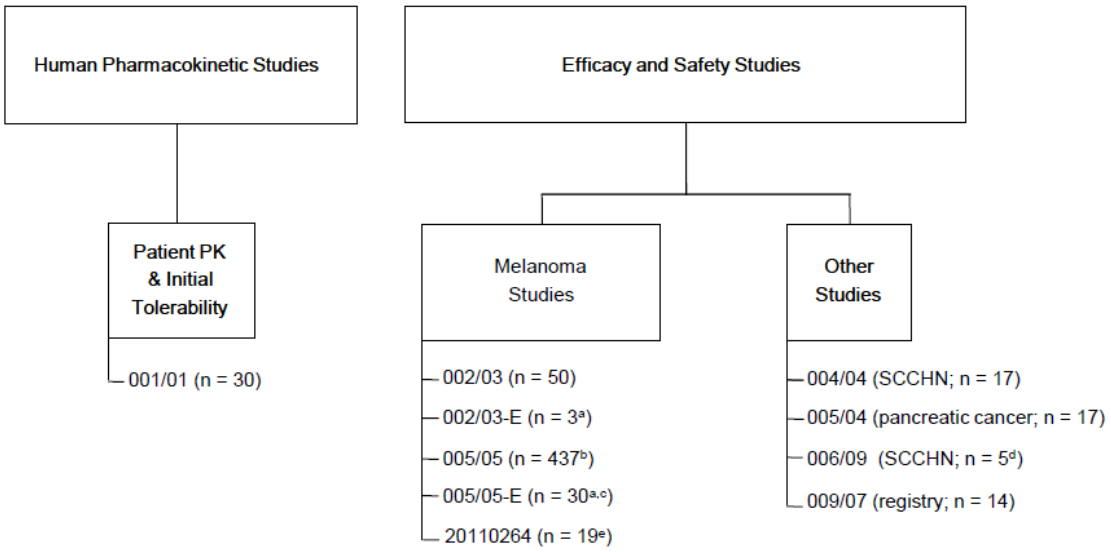
The clinical development of IMLYGIC® is actively ongoing. This section only summarizes the clinical trials included in this BLA. Figure 1 shows the clinical trials in the development program.

- Efficacy trials addressing the proposed indication in this BLA (Table 1) include
    - The single randomized, controlled, confirmatory trial, Study 005/05, and its extension, Study 005/05-E.
    - A supportive single-arm study, Study 002/03, and its extension, Study 002/03-E.

- Ongoing Phase 1/2 Combination trial of IMLYGIC® with ipilimumab in subjects with melanoma, Study 20110264.
- Trials for other tumor types include
  - A first-in-human study in subjects with solid tumors, Study 001/01
  - Efficacy/safety study in subjects with pancreatic cancer, Study 005/04
  - Efficacy/safety studies in subjects with squamous cell carcinoma of the head and neck (SCCHN), Studies 004/04 and 006/09.
- Ongoing observational registry study for subjects previously treated with IMLYGIC®, Study 009/07.

The registry study 009/07, the combination trial 20110264, and 005/05-E are ongoing. All other trials had either completed or were terminated early. Protocol 006/09 in SCCHN received a SPA concurrence on September 2, 2009. The first subject was treated on February 8, 2011. The study was terminated early on July 29, 2011 after enrolling five subjects.

Figure 1. Talimogene laherparepvec clinical program.



SCCHN = squamous cell carcinoma of the head and neck
[a] Reflects a subset of subjects enrolled in the principal study
[b] Talimogene laherparepvec: n = 296; granulocyte macrophage colony-stimulating factor: n = 141. Includes one subject who was randomized 3 times (at 3 different sites). The subject ultimately received talimogene laherparepvec and was included in the safety analyses, but excluded from the intent-to-treat analyses.
[c] Talimogene laherparepvec: n = 27; granulocyte macrophage colony-stimulating factor: n = 3.
[d] Talimogene laherparepvec + chemoradiation: n = 2; chemoradiation only: n = 3
[e] Reflects enrollment at data cutoff. Data from nine subjects in Ph 1b portion are summarized separately.

Source: Original BLA 125518, eCTD Section 2.7.3 Summary of Clinical Efficacy, p.9

Table 1. Efficacy studies related to the melanoma indication.

Study Number	Study Design	Study Population	Primary and Secondary Efficacy Endpoints	Region	Number of Subjects	Duration of Treatment
005/05	Phase 3, randomized, open-label, GM-CSF controlled	Unresectable stage IIIB, IIIC, or IV melanoma	Durable response rate[a] Overall survival Best overall response and disease burden Response onset Time to treatment failure Duration of response Response interval	US, Canada, South Africa, and UK	436 (ITT)[b]; 295 talimogene laherparepvec 141 GM-CSF	12 months (or 18 months if the subject was receiving clinical benefit)
005/05-E	Phase 3, open-label extension	Unresectable stage IIIB, IIIC, or IV melanoma	Overall response rate Durable response rate	US, Canada, South Africa, and UK	30; 27 talimogene laherparepvec 3 GM-CSF	12 additional months (or until disease progression if the subject was receiving clinical benefit)
002/03	Phase 2, open-label, single-arm	Unresectable stage IIIC or IV melanoma	Overall response rate[a] Time to tumor response Time to disease progression Overall survival	US and UK	50	Up to 47 weeks
002/03-E	Phase 2, open-label, single-arm extension	Unresectable stage IIIC or IV melanoma	Overall response rate Overall survival	US and UK	3	Up to 24 additional doses or 12 months of additional treatment, whichever was longer

[a]Primary endpoint
[b]A total of 437 subjects were randomized; 1 subject who was randomized 3 times at 3 different study centers was excluded from the ITT population.

Source: Original BLA 125518, eCTD Section 2.7.3 Summary of Clinical Efficacy, p.10

5.4 Consultations

5.4.1 Advisory Committee Meeting (if applicable)

A joint meeting of CTGTAC and ODAC was held on April 29, 2015. The single voting question was "Does talimogene laherparepvec have an overall favorable benefit-risk profile to support traditional approval for the treatment of injectable, regionally or distantly metastatic melanoma? ..." The vote was 22 "Yes" and 1 "no". A number of AC members qualified their votes by stating that they would want the approval to be limited to earlier stage disease or patients without visceral disease.

6. DISCUSSION OF INDIVIDUAL STUDIES/CLINICAL TRIALS

6.1 Trial #1

Study 005/05, also known as BVX00505, 20110263, or OPTiM, is the single confirmatory trial. The study protocol was titled "A Randomized Phase 3 Clinical Trial to Evaluate the Efficacy and Safety of Treatment with OncoVEX$^{GM\text{-}CSF}$ Compared to Subcutaneously Administered GM-CSF in Melanoma Patients with Unresectable Stage IIIB, IIIC and IV Disease."

Subsections 6.1.1 through 6.1.9 below are based on the final versions of the study protocol and the statistical analysis plan (SAP), and a supplemental statistical analysis plan (suppSAP).

While the final protocol and SAP were included in the BLA submission, this review used the corresponding documents, of the same version dates, submitted to IND 12412 amendment #0297 on January 23, 2013. Both the protocol and the SAP went through several amendments, with some substantial changes. The suppSAP was included only in the BLA with a version date of February 21, 2014. The suppSAP provided a plan for additional analyses of OS and systemic effects of IMLYGIC®.

Evaluation of the study outcome (Sections 6.1.10 – 6.1.12) is based primarily on two CSRs: the primary analysis CSR (paCSR) containing the primary analysis of the primary endpoint and an interim analysis (IA) of the secondary endpoint OS, and a supplemental CSR (suppCSR) containing the primary analysis of OS and analyses of systemic effect.

Additional documents were consulted during the review. These documents will be mentioned when their contents appear.

Reviewer Comment #1. The documents, including the protocol, SAP, data monitoring committee (DMC) charter, and endpoint assessment committee (EAC) charter, sometimes contain inconsistencies or ambiguities in their description of the same design element. This review is based on this reviewer's best understanding at such occurrences.

6.1.1 Objectives (Primary, Secondary, etc)

The objective of the study was to evaluate the efficacy and safety of treatment with intra-lesion injection of IMLYGIC®, compared to subcutaneously administered GM-CSF, in melanoma patients with unresectable Stage IIIB, IIIC, or Stage IV disease.

6.1.2 Design Overview

The study was open-label. The study was planned to randomize 430 subjects 2:1 to the IMLYGIC® or the GM-CSF arm. Additionally, the local inflammatory response that can be seen with IMLYGIC® administration is potentially indicative of treatment arm.

6.1.3 Population

Eligible subjects were men and women ≥18 years of age with histologically confirmed diagnosis, at study entry, of stage IIIB, IIIC, or IV malignant melanoma that was not surgically resectable. Subjects were required to have measurable disease (at least 1 melanoma lesion ≥10 mm or multiple lesions totaling ≥10 mm) that was suitable for direct injection (or injection with ultrasound guidance). The injectable lesions could be cutaneous, subcutaneous, or nodal. Subjects were also required to have serum lactate dehydrogenase (LDH) levels ≤ 1.5 x ULN (the upper limit of normal), an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1, a life expectancy > 4 months from the date of randomization, and adequate organ function within 4 weeks before randomization. Stage IV M1c subjects were limited to no more than 40% in each study arm. Two exclusion criteria are listed below.

- Clinically active cerebral or any bone metastases. Patients with up to 3 (neurological performance status of 0) cerebral metastases may be enrolled, provided that all lesions have been adequately treated with stereotactic radiation therapy, craniotomy, gammaknife therapy, with no evidence of progression, and have not required steroids, for at least two (2) months prior to randomization.
- Greater than 3 visceral metastases (this does not include lung metastases or nodal metastases associated with visceral organs). For patients with ≤ 3 visceral metastases, no lesion may be >3 cm, and liver lesions must meet Response Evaluation Criteria in Solid Tumors (RECIST) for stable disease (SD) for at least 1 month prior to randomization.

There is no minimum size for a tumor mass to be eligible for injection. In order to enroll patients with lesions larger than 10 cm in longest diameter, or with a total cumulative tumor burden in excess of 20 cm based on the sum of the longest diameters of individual lesions, prior approval must be obtained from the Medical Monitor who will also discuss and approve the injection strategy to be employed.

Reviewer Comment #2. Study 005/05 excluded the most advanced patients, who the applicant postulated might not benefit from any delayed immune effect. Therefore subjects in Study 005/05 had less tumor burden on average than subjects enrolled in the confirmatory trials supporting the traditional or accelerated approvals of the agents described in Section 2.2. More details in the comparison of the composition of the treated populations in these trials can be found in the applicant's briefing document (p.19 and

p.35) for the AC meeting. Study 005/05 consists of 30% IIIB/C, 27% IVM1a, 21% IVM1b, and 22% IVM1c.

6.1.4 Study Treatments or Agents Mandated by the Protocol

Subjects in both the GM-CSF and the IMLYGIC® arms were treated in cycles, each of four (4) weeks (28 days). Treatment was required to begin within 24 hours after central randomization.

A GM-CSF treatment cycle consisted of daily subcutaneous (SC) injections (stomach or thigh) of GM-CSF for 14 days followed by a 14-day rest period. The daily dose was 125 $\mu g/m^2$. Except for the day 1 cycle 1 (D1C1) dose, which was administered in the study clinic to observe for any first-dose reactions, GM-CSF were to be administered at home by the patient or a caregiver. GM-CSF subjects were to return to the clinic on Day 1 (+/- 3 days) and Day 15 (+/- 3 days) of each 28-day cycle. On the Day 15 visits, patient diaries were to be reviewed by study staff.

A IMLYGIC® treatment cycle consisted of intra-lesion injections on Days 1 and 15 of each 28-day cycle, at up to 4mL of 10^8 pfu/mL per dose (nominal), with the following exceptions: Day 1 Cycle 1 dose was up to 4mL of 10^6 pfu/mL and the first cycle was 35 days with the second injection occurring on Day 22 (3 weeks). Investigators used their judgment to decide the order of lesions to be injected, with new lesions and larger injectable lesions given priority. The volume of IMLYGIC® delivered to individual tumor (s) followed a size-dependent algorithm, which scaled roughly linearly with the longest dimension of the lesion.

If any injected lesion progresses for a IMLYGIC® arm subject, injection frequency was to increase to once per week for four (4) weeks into the progressing lesion(s) only. Up to three (3) sets of four (4) accelerated injections might be given as long as after each set of four (4) accelerated injections clinically relevant progressive disease (PDr, see definition later) had not occurred and there was still residual tumor to inject. After completion of three (3) sets of accelerated injections, in the absence of PDr and where residual tumor remained, dosing should then return to once every two weeks. Accelerated dosing could only be used once per subject. Therefore a IMLYGIC® subject would be dosed for at most *45* injection days on this protocol: up to 18 months of bi-weekly dosing plus accelerated dosing of up to 12 weekly dosing. IMLYGIC® dosing might reinitiate if a subject without any injectable lesions developed a new injectable lesion within 12 months from the start day of treatment. If a new injectable lesion(s) appeared after 12 months from the start date of treatment, the patient would then be eligible for further dosing under an extension protocol. PDr is defined as progressive disease (PD) that is associated with a decline in performance status and/or in the opinion of the investigator the patient requires alternative therapy.

Treatment duration in both arms.

Patients were to be treated at least until Week 24 or until complete response (CR), even in the presence of PDr including the appearance of new lesions, to allow for delayed immune-based anti-tumor effects to occur, unless in the investigator's opinion other therapy for melanoma was required. After 24 weeks, patients would continue to be treated until PDr or CR. Patients would be treated for up to 12 months on this study.

Patients who were in response at 12 months (partial response (PR) or CR) were to continue to be treated until 18 months or progressive disease (PD), whichever was the earlier. The study duration for each subject, including the 28-day screening period, 12-month treatment period, and 3-year survival follow-up from randomization was up to approximately 37 months.

6.1.6 Sites and Centers

The study was conducted at 71 sites in the US, Canada, South Africa, and United Kingdom, of which 64 sites had at least one subject in the intent-to-treat (ITT) set. The minimum number of subjects at a site was one (12 sites) and the maximum was 28 subjects (one site). The majority of the subjects were from US sites, accounting for 87.6% of the ITT set subjects (382/436) and 73.4% of the sites (47/64).

6.1.8 Endpoints and Criteria for Study Success

The primary efficacy endpoint was durable response (DR), defined as an objective response (CR or PR) that initiated at any time within 12 months of commencing therapy and was maintained continuously for at least 6 months (183 days) from response initiation. DR was evaluated both by investigators and an independent EAC. The primary efficacy endpoint was DR based on EAC assessment.

The most important secondary efficacy endpoint was overall survival (OS), defined as the time from the date of randomization to the date of death from any cause.

Additional secondary efficacy endpoints included the following. All secondary endpoints were assessed by the investigators, as that provided a measurement for each subject in the ITT set, whereas EAC assessments of these endpoints were available for only a subset of the ITT subjects (see below).

- Best overall response and disease burden.
- Response onset (RO).
- Time to treatment failure (TTF), defined as time from baseline until the first PDr where there was no response achieved after the PDr.
- Duration of response, defined as the longest individual period from entering response to the first documented evidence of the patient no longer meeting the criteria for being in response or death, whichever is earlier. The duration of response was defined to be zero if no PR or CR was ever achieved.
- Response interval (RI), defined as the time from randomization to the date of the last documented evidence of response prior to any new anti-cancer therapy which may be given.
- Safety endpoints. Safety assessments will be based on adverse events, laboratory data, concomitant medications, the results of physical examinations and vital signs.

The exploratory endpoints included FACT-BRM (Functional Assessment of Cancer Therapy Biologic Response Modifier), a patient reported questionnaire on quality of life, impact of response on survival, and influence of BRAF mutation status.

Reviewer Comment #3. In what follows, I provide a high-level summary of the assessment procedures and schedule for DR and OS. The assessment of DR was quite complex, involving considerations of new lesions, previously unresectable lesions now becoming resectable, among other issues. The ultimate evaluation of matters related to DRR is deferred to the clinical review team.

Assessment of durable response.

The EAC charter went through eight (8) revisions, after the original version dated August 7, 2009, with the last version dated September 27, 2012, less than three (3) months before the data cut-off date for the primary analysis on December 21, 2012. The final versions of the EAC related documents, e.g. charters and manuals, were submitted to IND 12412 amendment 281 on October 11, 2012. Some of the study features were explained better in the EAC charter compared to the protocol. In what follows, this reviewer uses the information from the document that is felt to explain a particular feature better.

Response was defined according to the modified World Health Organization (WHO) criteria.

Baseline.

At baseline (the last assessment on or prior to the first dose of study drug being administered), a tumor lesion was categorized as measurable or non-measurable but evaluable (NMbE). Measurability is defined as the ability to measure a tumor bi-dimensionally with surface area determined by multiplying the longest diameter by the greatest perpendicular diameter.

Overall response at a time point.

Overall response was defined based on response in measurable disease and response in NMbE disease. At an assessment time point, status of measurable disease fell into one of four categories: CR, PR, stable disease (SD), or progressive disease (PD).

- CR was defined as the disappearance of all clinical evidence of tumor, including any new tumors which might have appeared. Any residual cutaneous or sub-cutaneous masses must be documented by representative biopsy to not contain viable tumor.
- PR was defined as *achieving a 50% or greater reduction in the sum of the products of the perpendicular diameters (SPPD) of all measurable tumors at the time of assessment as compared to the SPPD of all measurable tumors at baseline. If any new tumors have appeared, the SPPD of these must have reduced by 50% or more from when first documented. Any residual cutaneous or sub-cutaneous masses which must be tumor free for the patient to meet the criteria for PR must be documented as such by representative biopsy*. In the SPA concurrence letter dated April 17, 2008, "*…Thus, the definitions of CR and PR to be used in the study allow for progression before response, including the appearance of new lesions, as long as any new lesions have subsequently reduced in size by >50%*

from when first noted and that the overall disease burden, including any new lesions, is <50% of that at baseline (i.e. modified WHO criteria)."

- PD: A >25% increase in the SPPD of all measurable tumors since baseline, or the *unequivocal* appearance of a new tumor since the last response assessment time point.
- SD: Neither sufficient overall tumor shrinkage to qualify for response (PR or CR) nor sufficient tumor increase to qualify for PD.

At an assessment time point, status of NMbE fell into one of three categories: CR, incomplete response/SD, or progressive disease (PD).

- CR: Disappearance of all evaluable tumors
- Incomplete Response/SD: Persistence of one or more evaluable tumor(s)
- PD: *Unequivocal* appearance of one or more evaluable but non-measurable tumor.

The overall response at an assessment time point was given in the Table below.

Measurable Lesions including new lesions	Non-measurable Lesions	Overall Response
CR	CR	CR
PR	CR	PR
SD	CR	SD
CR or PR	SD	PR
SD	SD	SD
Any	PD	PDr
PDn	Not PD	PDn
PDr	Any	PDr
PDcns	Not PD	PDcns

Investigator assessment schedule.

As described above, patients were to be followed for at least 12 months following randomization for response duration, and if at 12 months a patient was in CR or PR, they should not have an end of treatment visit until 18 months or PD, whichever was earlier. Subjects in both arms were to be treated for up to 18 months, in the absence of meeting any discontinuation criteria. The End of Treatment (EOT) / Early Termination visit was to occur 30 (+/-7) days after the last injection, after which the subject entered the survival status follow-up period. Tumor and efficacy assessment occurred during the treatment period, prior to treatment on Day 1 of each 28-day treatment cycle and also on the EOT visit. The first response assessment occurred on Day 1 of Cycle 3, the Month 2 visit. Efficacy assessment was based on tumor measurement by clinical measurement, digital photographs of superficial tumors, ultrasonography of nodal tumors or in some cases other soft tissue tumors, representative biopsy of residual pigmented areas or other residual masses suspected to no longer contain tumor where necessary for the patient to be in the response observed (i.e. PR or CR), and imaging studies of measurable and evaluable disease. CT scans and ultrasound of nodal or other soft tissue masses were to be performed every 12 (+/-2) weeks from the start of therapy (or if a response, based on clinical assessment, is suspected to have initiated since the last visit, within 1 week), and

MRI of the brain every 16 weeks (or at any time when in the judgment of the investigator the patients displays signs or symptoms of CNS disease progression). For patients who reached 9 months of therapy but for whom a PR or CR had not been recorded, a whole body PET or PET/CT scan would be performed and representative biopsies taken from residual masses, as clinically warranted and feasible, to aid in determining their status. All objective responses were to be confirmed on 2 separate measurements no less than 1 week apart.

EAC assessment.

The EAC consisted of a team of board certified oncologists with experience treating melanoma patients, contracted by WorldCare Clinical (WCC). The EAC was blinded to the study treatment arm assignment and the response assessment of the study sites. Bi-dimensional tumor measurements were derived by a blinded radiologist and dermatologist and provided to the EAC. The EAC determined whether patients sent for EAC review were in objective response (CR or PR) at each response assessment time point, and the date of initiation and termination of response, with adjudication. Two EAC reviewers performed independent assessments of the same data, with adjudication by a third reviewer if there was disagreement between assessments by the two reviewers.

In general, it was intended that the patient's data would be sent for EAC review when the data for review were sufficiently complete, and one of the following criteria was met:

- A patient had been recorded as having achieved a response by the investigator and the EOT visit had occurred.
- A patient had reached 9 months on treatment, from the date of randomization, without a response having been declared by the investigator and the EOT visit had occurred.
- After all patients described above had been submitted for EAC review, any remaining patients who met EAC review criteria but had not yet ended treatment would be sent for EAC review in order to be included in the primary analysis.

Response initiation and date were determined based on the following.

- In general, initiation of response is declared to be the date on which the last measure is taken when all required data is available; the last value carry forward method for imputation of lesion measurements cannot be used to initiate response.
- For subjects with only clinically assessed disease or a majority of clinically assessed disease, the date the response is initiated is the first date when clinical and/or photography assessments show response if (1) the contribution of imageable-only disease measurement would not have altered the response assessment and (2) the next CT scan shows no new imageable-only disease.

Response termination and date were determined based on the following.

- Response is terminated when there is documented evidence of no longer meeting the criteria for response. Specifically, response is terminated at the time of the unequivocal appearance of a new lesion or when tumor burden has increased

(based on tumor burden actually assessed at the visit) such that the subject no longer meets the criteria for response.

- Since CTs are available only every 12 weeks, tumor assessments performed at clinical visits in between the CT scans may not have the concurrent assessment of all lesions. A last value carry forward imputation method will be used to impute the lesion sizes from the previous CT scan. In addition, if other lesion assessments at clinical visits are missing, last value carry forward imputation will also be used. These imputations are allowed only for the purpose to determine if response is maintained or terminated.
- Response will also terminate if new systemic anti-cancer therapy is initiated, not including radiation or surgery, which are allowed per protocol. The date of any subsequent anti-cancer therapies will be given to the EAC if available at the time patients are triggered for EAC review and used in the analyses performed by the Sponsor as termination of response when appropriate.
- Response will be considered continuously maintained in the presence of missing/incomplete assessment(s) as long as there is no other information (including clinical, radiographic and photographic assessments) that indicates progression between two assessments showing response (PR or CR). When a response period continues to the end of study without a documented evidence of no longer meeting the criteria for response, the response is considered to be maintained until the last tumor assessment prior to the end of study.

Additional considerations in EAC assessment are given below.

- Data (such as imaging and photography) may be acquired at times different from the clinical visits (e.g. in between two clinical visits). In assessing response, the data is associated to the appropriate cycle within which it occurred (Subject cycle schedules will be provided by the Sponsor).
- Surgery is allowed on study where previously unresectable tumors become resectable. If the response of other tumors is at least PR, the patient should be designated PR with the date of surgery as the date of PR. If no residual disease remains following surgery, the response definition remained PR.
- In the case of missing data (i.e. scans) after querying the study site, if it involves a baseline assessment, this scan would be considered not readable. For a missed follow-up assessment, in the event that a response is documented on multiple scan visits and intervening scan visits or individual scans are missing, then the patient will be deemed to have remained in response unless there is any information to the contrary (e.g. clinical data).
- It was intended that an entire patient's data package would be reviewed in one batch wherever possible.
- Where residual tumors have been biopsied, it is at the discretion of the EAC if these and all tumors for which they are representative (i.e. which appear similar by radiology, photography or clinical assessment) are to be included.

- If tumors are resected and concluded not to contain viable tumor tissue (by histology or judged by the EAC), the resected tumor measurements will be measured as (0.00 x 0.00) on the resected visit but remain in the sum of bi-dimensional measurements at baseline or when they first appeared. On the other hand, if resected tumors are judged to contain viable tumor tissue or histology is not performed to indicate resected lesion is tumor free, the lesion will be removed from the sum of bi-dimensional measurements at baseline or when they first appeared, at the resected visit and at all subsequent visits as well.

Follow-up for survival status.

All patients were to be followed for survival status at 3-month intervals after concluding the treatment period of the trial. Patients were to be followed for survival status until End of Study (EOS). EOS was defined as 36 months from the date the last patient was randomized, or until the last patient had died, whichever was the earlier. The follow-up plan included patients who discontinued after randomization but prior to receiving the first dose of study treatment. After three years, patients would then be followed for survival on the Registry Protocol which was in place for all patients treated with IMLYGIC®. OS time will be censored at the last date the patient is known to be alive when the confirmation of death is absent or unknown. Patients are censored at the date of randomization if no additional follow-up data are obtained.

In response to an information request from the FDA, the applicant communicated that "Prior to both the primary and 36-month (last planned) survival analyses, sites were instructed to conduct a search of the US Death Index to determine if any subject had died and, if confirmed, to report the date of death on the CRF."

6.1.9 Statistical Considerations & Statistical Analysis Plan

Randomization.

Subjects were randomized 2:1 to IMLYGIC® or GM-CSF, using a central interactive voice response system (IVRS), stratified by the following known prognostic factors.

1. Site of first recurrence (3 levels): in transit or distant skin, lymph node, visceral.
2. Presence of liver metastases (2 levels): no, yes.
3. Stage of disease (3 levels): IIIB/C, IVM1a or IVM1b, IVM1c.
4. Prior treatment and outcome (3 levels):
    a. No prior nonsurgical melanoma treatment other than adjuvant therapy,
    b. Prior nonsurgical melanoma treatment other than adjuvant therapy and recurrence <1 year from primary diagnosis,
    c. Prior nonsurgical melanoma treatment other than adjuvant therapy and recurrence >1 year from primary diagnosis.

Note that the protocol and SAP both listed 4 levels for the stratification factor "stage of disease" by separating IVM1a and IVM1b into two levels, inconsistent with the actual randomization scheme that used the 3 listed levels instead.

Sample size consideration.

The trial was designed to randomize 430 subjects at a 2:1 ratio, to yield 360 subjects evaluable for DR. The trial would have 90% power, at a size 0.05 2-sided test, to detect a DRR difference of 13% (IMLYGIC®) versus 3% (GM-CSF), or 21% versus 8%.

Analysis population.

The analysis populations or analysis sets are defined below.

- The screened population is defined as all patients who had signed an informed consent and participated in screening procedures at the investigative site to assess eligibility.
- The Intent-to-Treat (ITT) population is defined to include all patients who had been randomized to receive study treatment.
- The safety population is defined as all randomized and treated patients.
- The Per Protocol (PP) population is defined as all patients who were randomized, eligible and treated, received at least 2 cycles of therapy and completed the assessment after 8 weeks (and at other time of termination for those who stay on study past 8 weeks), unless taken off therapy due to progression or due to safety issues before two cycles had been received. Patients with major protocol violations were excluded from this population. All major protocol violations were determined following Sponsor standard operating procedures prior to the data base lock.

Unless noted otherwise, the primary analysis of endpoints would use the ITT population and the analysis using the PP population would be supportive.

Analysis of DRR, the primary endpoint.

The primary analysis of DRR and all response based endpoints occurred when no further patients had the possibility of meeting the criteria for durable response, or all patients reached 18 months from first dose (whichever is the earlier). The primary analysis of DRR was a two-sided unadjusted Fisher's exact test. Study success was defined as the test being statistically significant at the 0.0488 level. A significance level of 0.0488 was used because of plans for interim analysis (IA). See below for discussion on the interim analysis plan.

Analysis of OS.

OS was to be tested for superiority in the IMLYGIC® arm compared to the GM-CSF arm at the following occasions.

- Interim analysis (IA) of OS would occur at each IA of DRR and at the time of the primary analysis of DRR, but only in the event of a statistically significant difference on DRR.
- The primary analysis of OS would occur at the time of 290 deaths if that was later than the time of the primary analysis of DRR.

- A descriptive OS analysis would occur when all subjects had been followed for 3 years after randomization (EOS).

The primary analysis of OS was the un-adjusted log-rank test. The Cox proportional hazard model was used to estimate the hazard ratio for the treatment effect. With respect to Type 1 error control, the applicant stated that "*a nominal 0.0001 one-sided alpha spending will be used to account for the possibility of an unexpected survival outcome prior to the primary OS analysis (including the analyses at each interim and at the primary DRR analysis if applicable). Given the minimal alpha spending on OS prior to the primary analysis, the primary OS analysis will have one-sided significance level of 0.025.*"

The decision to set the primary analysis of the OS at 290 deaths was made in one of the revisions to the protocol and SAP. The applicant stated that with 290 events, there would be 90% power to detect a hazard ratio of 2/3.

Interim analysis.

Two formal interim analyses (IA) with respect to efficacy were planned. The first IA was to occur after the first 75 subjects had been on study for 9 months. This IA was intended to assess futility, align hypothesis (per the applicant), and to determine timing of the second IA, based on response rate (PR+CR) and DRR (in the GM-CSF arm). The significance level for this IA was set to be one-sided 0.0001 for the DRR endpoint.

The second IA was to occur once all planned subjects had been randomized and on study for 9 months, at a time to be determined by the DMC after performing the first IA. After the first IA, the DMC recommended performing the second IA once there had been 42 EAC-confirmed DRs. The significance level for this IA was set to be one-sided 0.0005 for the DRR endpoint.

The second IA was eventually cancelled. The applicant stated that the timing of the second IA would have occurred within one month of the primary (final) DRR analysis, which was to occur after the last randomized subject reached 18 months on study. The reason was that the EAC did not start response assessment until October 2012, only 2 months before the data cut-off date for the primary analysis of DRR. Alpha spend for both IAs, however, was accounted for in the primary analysis of DRR. That is, the primary analysis of DRR used a nominal significance level of one-sided 0.0244 (=0.025-0.0001-0.0005), or 2-sided 0.0488.

6.1.10 Study Population and Disposition

6.1.10.1 Populations Enrolled/Analyzed

There were 439 randomizations. One subject was randomized three times. See subsection 6.1.10.1.3 "Subject Disposition" for a brief narrative on this subject. The applicant defined the ITT population to include all subjects that were randomized once, excluding the subject who was randomized three times. The ITT population consists of 436 subjects, 295 assigned to the IMLYGIC® arm and 141 assigned to the GM-CSF arm.

The Per Protocol (PP) population consists of 372 subjects, 262 in the IMLYGIC® arm and 110 in the control arm. The PP subjects are 89% and 79% of the ITT subjects for the IMLYGIC® and GM-CSF groups, respectively.

The Safety population consists of 419 subjects, 292 subjects randomized and received at least one dose of IMLYGIC®, and 127 received as least one dose of GM-CSF.

6.1.10.1.1 Demographics

Baseline demographics for the ITT population are summarized in Table 2 below. Baseline demographics were generally balanced between the two study groups. Overall, 57.3% were male and 97.9% were white. The mean (range) age was 63 (22 to 94) years.

6.1.10.1.2 Medical/Behavioral Characterization of the Enrolled Population

Baseline disease and subject characteristics for the ITT population are summarized in Table 3 below. These characteristics were generally balanced between the two study groups. Most subjects (70%) had an ECOG performance status of 0. Fifty-seven percent of subjects had earlier stages of disease (ie, stage IIIB, IIIC, and IVM1a) and 43% of subjects had more advanced disease (ie, stage IVM1b and IVM1c). The baseline LDH level was above the ULN in 4.6% of subjects. The 3 most common prior therapies were biologic therapy (33%), chemotherapy (29.1%), and investigational treatment (17%). Fifty-eight percent of subjects were known seropositive for HSV-1 at baseline. Per the IVRS, 46.6% of subjects were first-line (excluding surgery, adjuvant, or radiation) and 53.4% of subjects had received prior therapy other than or in addition to surgery, adjuvant, or radiation. The applicant reported that 40 subjects in the ITT population (9%) had discordant disease stage reported between IVRS at randomization and the CRF.

Table 2. Baseline demographics for the ITT population.

	GM-CSF (N = 141)	Talimogene Laherparepvec (N = 295)	Total (N = 436)
Sex - n (%)			
Male	77 (54.6)	173 (58.6)	250 (57.3)
Female	64 (45.4)	122 (41.4)	186 (42.7)
Ethnicity - n (%)			
Hispanic or Latino	1 (0.7)	9 (3.1)	10 (2.3)
Not Hispanic or Latino	139 (98.6)	285 (96.6)	424 (97.2)
Missing	1 (0.7)	1 (0.3)	2 (0.5)
Race - n (%)			
White	138 (97.9)	289 (98.0)	427 (97.9)
Black	2 (1.4)	1 (0.3)	3 (0.7)
Asian	0 (0.0)	1 (0.3)	1 (0.2)
Native Hawaiian or Other Pacific Islander	0 (0.0)	1 (0.3)	1 (0.2)
Other	1 (0.7)	3 (1.0)	4 (0.9)
Age (years)			
n	141	295	436
Mean	62.92	63.14	63.07
SD	14.13	13.67	13.80
Median	64.00	63.00	63.00
Q1, Q3	54.00, 74.00	54.00, 74.00	54.00, 74.00
Min, Max	26.0, 91.0	22.0, 94.0	22.0, 94.0
Age Group - n (%)			
< 55 years	36 (25.5)	80 (27.1)	116 (26.6)
≥ 55 years	105 (74.5)	215 (72.9)	320 (73.4)
< 65 years	72 (51.1)	152 (51.5)	224 (51.4)
≥ 65 years	69 (48.9)	143 (48.5)	212 (48.6)
< 75 years	109 (77.3)	229 (77.6)	338 (77.5)
≥ 75 years	32 (22.7)	66 (22.4)	98 (22.5)

N = Number of subjects in the analysis set; SD = sample standard deviation; Q1 = first quartile; Q3 = third quartile .
Intent to treatment population includes all subjects that have been randomized. Subjects will be analyzed using the randomized treatment.

Source: Original BLA 125518, eCTD Section 5.3.5.1, Study 005/05 paCSR, p.68, Table 9-2.

Table 3. Key baseline disease and subject characteristics.

	GM-CSF (N = 141)	Talimogene Laherparepvec (N = 295)	Total (N = 436)
ECOG performance status - n(%)			
0	97 (68.8)	209 (70.8)	306 (70.2)
1	32 (22.7)	82 (27.8)	114 (26.1)
Missing	12 (8.5)	4 (1.4)	16 (3.7)
Histogenetic classification at orginal diagnosis			
Histogenetic subtype - n(%)			
Superficial spreading melanoma	39 (27.7)	72 (24.4)	111 (25.5)
Lentigo maligna melanoma	3 (2.1)	11 (3.7)	14 (3.2)
Acral lentiginous melanoma	6 (4.3)	17 (5.8)	23 (5.3)
Nodular melanoma	39 (27.7)	76 (25.8)	115 (26.4)
Desmoplastic	1 (0.7)	10 (3.4)	11 (2.5)
Unclassifiable	38 (27.0)	99 (33.6)	137 (31.4)
Missing	15 (10.6)	10 (3.4)	25 (5.7)
Disease stage from CRF - n(%)			
Stage IIIb	12 (8.5)	22 (7.5)	34 (7.8)
Stage IIIc	31 (22.0)	66 (22.4)	97 (22.2)
Stage IV M1a	43 (30.5)	75 (25.4)	118 (27.1)
Stage IV M1b	26 (18.4)	64 (21.7)	90 (20.6)
Stage IV M1c	29 (20.6)	67 (22.7)	96 (22.0)
Missing	0 (0.0)	1 (0.3)	1 (0.2)
LDH			
≤ ULN	124 (87.9)	266 (90.2)	390 (89.4)
>ULN	5 (3.5)	15 (5.1)	20 (4.6)
BRAF status			
Mutation	23 (16.3)	46 (15.6)	69 (15.8)
Wild-type	23 (16.3)	45 (15.3)	68 (15.6)
Unknown	5 (3.5)	9 (3.1)	14 (3.2)
Missing	90 (63.8)	195 (66.1)	285 (65.4)
HSV-1 status			
Negative	45 (31.9)	97 (32.9)	142 (32.6)
Positive	78 (55.3)	175 (59.3)	253 (58.0)
Unknown	18 (12.8)	23 (7.8)	41 (9.4)
Prior non-surgical procedures (CRF)			
Yes	89 (63.1)	202 (68.5)	291 (66.7)
No	36 (25.5)	80 (27.1)	116 (26.6)
Unknown	16 (11.3)	13 (4.4)	29 (6.7)
Line of therapy per IVRS			
First line	65 (46.1)	138 (46.8)	203 (46.6)
≥ Second line	76 (53.9)	157 (53.2)	233 (53.4)

CRF = case report form; ECOG = Eastern Cooperative Oncology Group; GM-CSF = granulocyte macrophage colony-stimulating factor; HSV-1 = herpes simplex virus, type 1; ITT = intent-to-treat; IVRS = interactive voice response system; LDH = lactate dehydrogenase; ULN = upper limit of the normal range

Source: Original BLA 125518, eCTD Section 5.3.5.1, Study 005/05 paCSR, p.70, Table 9-3.

6.1.10.1.3 Subject Disposition

Milestone dates of Study 005/05 are listed below. Note that for each study report multiple cut-off dates may be involved. No attempt is made to comprehensively list all dates.

- The first subject was enrolled on April 29, 2009.
- The data cut-off date for the primary analysis of the primary efficacy endpoint, DRR, was December 21, 2012. The primary analysis CSR (paCSR) was dated April 14, 2014.
- The data cut-off date for the primary analysis of the OS endpoint was March 31, 2014. The supplemental CSR (suppCSR) was dated June 9, 2014.
- *The 120-day safety update includes updated data from Study 005/05 (data cutoff 05 September 2014) and Study 005/05-E (data cutoff 02 June 2014).*

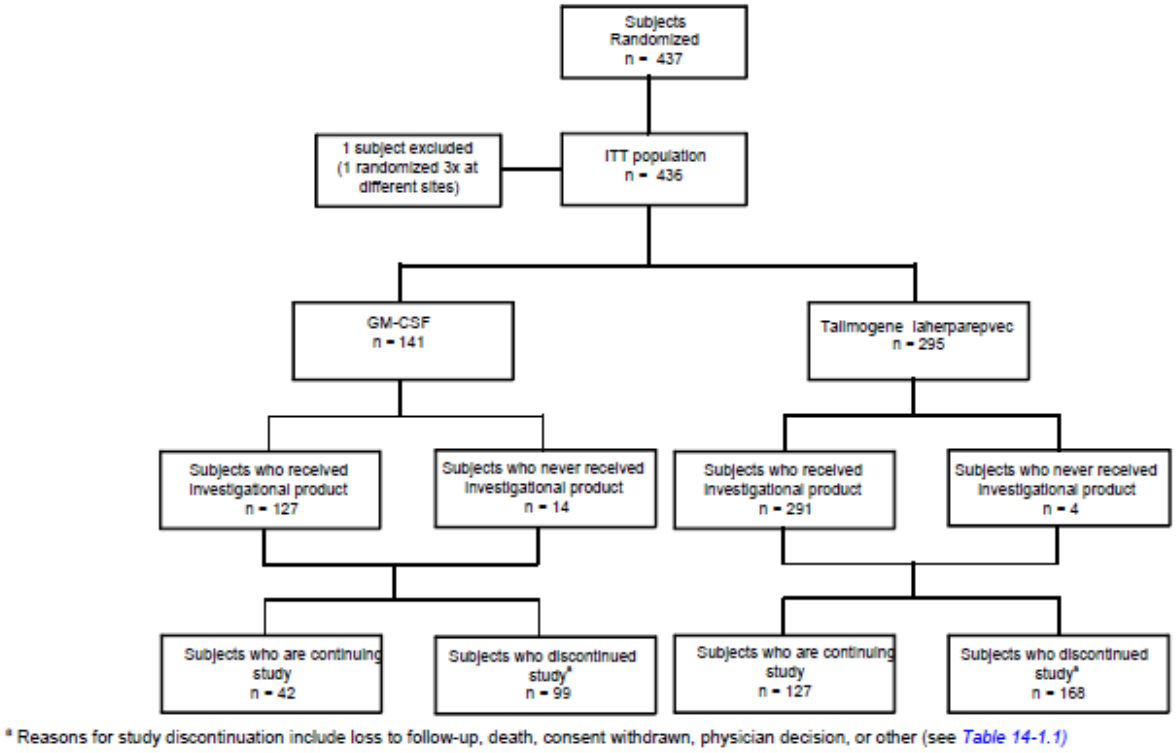
Subject disposition as of the December 21, 2012 is summarized in Figure 2 below. The ITT population consists of 436 subjects, 295 in the IMLYGIC® arm and 141 in the GM-CSF arm. Of these subjects, 18 did not receive any study treatment, 4 in the IMLYGIC® arm and 14 in the GM-CSF arm. All subjects had discontinued treatment, predominantly due to progressive disease: 65.6% in the IMLYGIC® arm and 74.8% in the GM-CSF arm. As of the primary analysis cutoff date, 56.9% of subjects in the IMLYGIC® arm and 70.2% of subjects in the GM-GSF arm had discontinued from the study, predominantly due to death: 97.6% (164/168) in the IMLYGIC® arm and 86.9% (86/99) in the GM-CSF arm.

The subject incidence of sponsor-defined *important* protocol deviations was 12.2% (36/295) in the IMLYGIC® group and 3.5% (5/141) in the GM-CSF group. The most common important protocol deviation was due to subjects missing confirmatory scans (19 subjects, 4.4%), which was defined as not having a scan performed prior to the next scheduled radiologic assessment after a response (CR or PR) was determined by clinical assessment. Overall, 33 subjects (7.6%) experienced inclusion and exclusion criteria violations, 8.8% (26/295) in the IMLYGIC® group and 5.0% (7/141) in the GM-CSF group.

Narrative on the subject that was randomized three times.

A subject was randomized three times in this study. The first two randomizations were to the GM-CSF arm, at site #69 on June 18, 2010 and site #74 on July 6, 2010, respectively. Each time the subject withdrew consent without receiving any study treatment. The third randomization was to the IMLYGIC® arm at site #67 on July 15, 2010. The subject received IMLYGIC® treatment from July 16, 2010 to October 1, 2010. The subject died on [redacted] from the last randomization date. This subject was not among the durable responders confirmed by EAC.

Figure 2. Subject Disposition as of December 21, 2012



Source: Original BLA 125518, eCTD Section 5.3.5.1, Study 005/05 paCSR, p.65, Figure 9-1.

6.1.11 Efficacy Analyses

6.1.11.1 Analyses of Primary Endpoint(s)

The EAC reported 48 DRs in the IMLYGIC® arm and 3 DRs in the GM-CSF arm among the ITT population, resulting in a 16.3% DRR in the IMLYGIC® arm and a 2.1% DRR in the GM-CSF arm. The primary analysis, using an unadjusted Fisher's exact test in the ITT population, returned a p-value of <0.0001. The applicant reported an unadjusted odds ratio (OR) of 8.9 with a 95% confidence interval (CI) of (2.7, 29.2). I also calculated the relative risk (RR) and the corresponding 95% exact CI, 7.6 and (2.6, 25.5), respectively. There were no important differences between the results from the per-protocol analysis population and the ITT population.

Durable responders reported by the FDA and the investigators.

The FDA clinical review team considered 3 of the 51 DRs reported by the EAC to not qualify for DR, for the following reasons.

- Subject #003023 (IMLYGIC®): *PD not a DR for PR (Lesion 2 progressed by CT)*
- Subject #053004 (GM-CSF): *PD not a DR for PR (per EAC)*

- Subject #066017 (IMLYGIC®): *not evaluable, too many missed visits*.

The investigators reported 56 DRs in the IMLYGIC® arm and 2 DRs in the GM-CSF arm. The DRRs reported by the FDA and the investigators both result in very small p-values in the comparison between the two arms, confirming the statistical robustness of the EAC-confirmed DRR.

The EAC received and reviewed information on only a subset of 143 (33%) subjects of the ITT population to determine whether a subject had a DR. Per protocol, the investigators were to send subject information for EAC review, if they determined that a subject had a response (PR or CR), or that a subject had reached 9 months of treatment without a response having been recorded. Of these 143 EAC-reviewed subjects, 124 were from the IMLYGIC® arm, and 19 were from the GM-CSF arm, representing 42% and 13% of the ITT population of the respective arm, respectively. Table 4 below summarizes the agreement on DR status between the EAC and the investigators, among these 143 EAC-reviewed subjects. The EAC agreed with investigators on the DR status determination in 85% of these subjects. The EAC was more likely to reclassify an investigator-determined DR to be a non-DR (14/58, 24%), than to reclassify an investigator-determined non-DR to be a DR (7/85, 8%). Note that although the EAC reviewed information submitted by the investigators, the actual information reviewed by the EAC may be different from that by the investigators, e.g., for some lesions, the EAC reviewed photos while investigators measured the lesions directly. In addition, the EAC might have categorized lesions as measurable or non-measurable differently from the investigators.

Table 4. Agreement on durable response status between investigators and EAC among the 143 EAC evaluated subjects.

	Durable Response per Investigator		
Durable Response per EAC	Durable Responder	Non-Durable Responder	Total
Durable Responder - n (%)	44(30.8)	7(4.9)	51(35.7)
Non-Durable Responder - n (%)	14(9.8)	78(54.5)	92(64.3)
Total - n (%)	58(40.6)	85(59.4)	143(100.0)

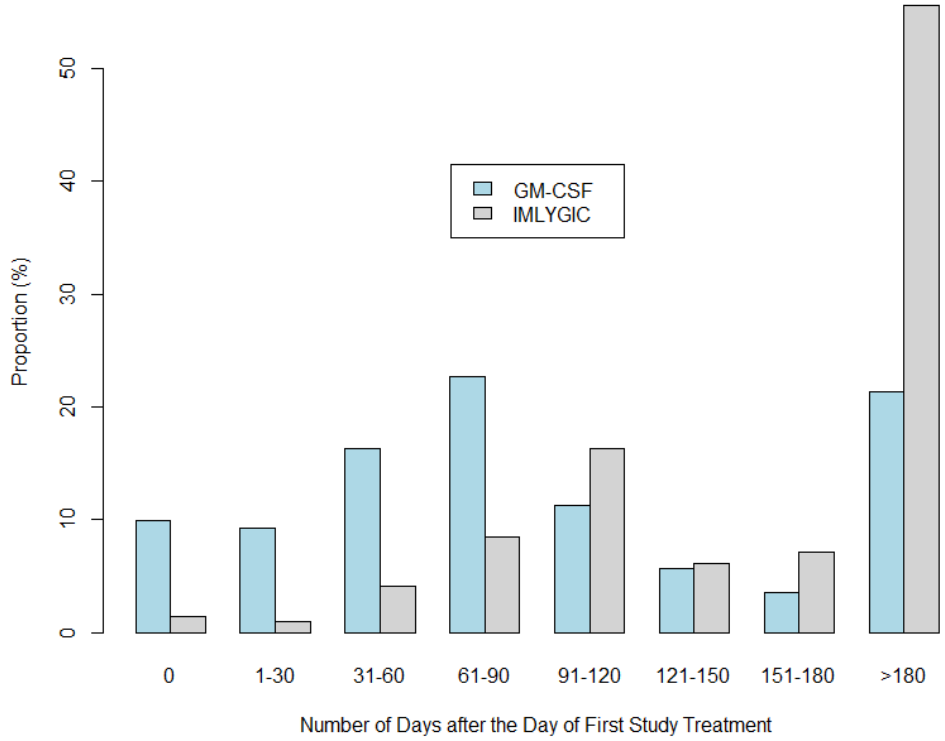
Source: Original BLA 125518, eCTD Section 5.3.5.1, Study 005/05 paCSR, p.213, Table 14-4.2.24.

Actual duration of response assessment.

Several sources of information from the BLA submission, such as the number of subjects randomized but not treated, indicate that there might be a difference between the two study arms in the duration of response assessment. To better understand this difference, we considered the number of days from initiation of study treatment to the "End of Treatment / Early Termination Visit (EOT)". Response assessment ended on the EOT visit, if not earlier. Figure 3 compares the distribution of the elapsed time when EOT occurred between the study arms. Per protocol, subjects should be treated through 24 weeks (approximately 6 months or 180 days) even in the face of PD. However, by Day 180, 78.7% of the GM-CSF subjects had already had the EOT visit and therefore the last chance for response assessment, compared to 45.4% in the IMLYGIC® arm. The

difference was more pronounced for earlier time points, e.g., the proportions were 58.1% versus 15% for Day 90. The median time to EOT was 85 days for the GM-CSF arm and 197 days for the IMLYGIC® arm. The GM-CSF subjects had a much shorter duration of response assessment, compared to IMLYGIC® subjects.

Figure 3. Distribution of End of Treatment or Early Termination (EOS) Visit. The number of days from the day of first study treatment to the EOS visit is plotted. The number of days is 0 for subjects who did not receive any study treatment.



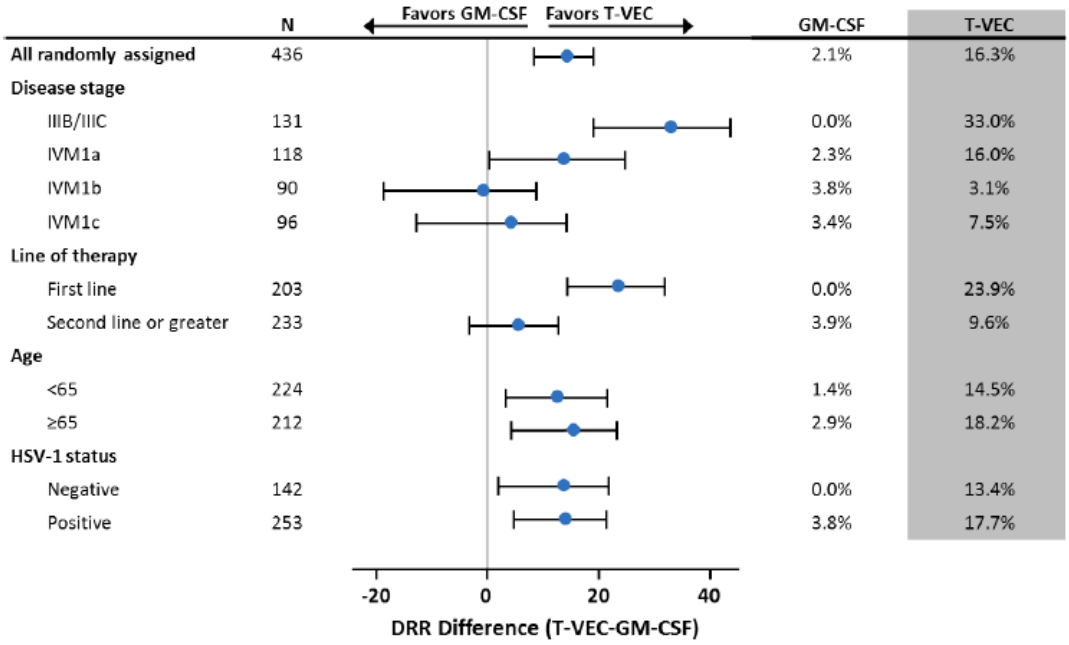
Reviewer Comment #4. As seen above, a substantially greater proportion of GM-CSF subjects terminated treatment and follow-up for response assessment early, compared to IMLYGIC® subjects. This difference might be due mostly to the open-label nature of the trial. While employing EAC to confirm DR instilled a degree of objectivity into the adjudication of response, EAC only reviewed about 1/3 of the ITT subjects sent by the investigators and therefore would not be able to rectify any bias that might have been introduced earlier on. However, to negate the statistical significance comparing DRR between the two arms, the GM-CSF arm would need a DRR of 9.2% (13/141). Some clinicians consider subcutaneous GM-CSF as equivalent to a placebo control, in which case it is implausible to expect a 9.2% DRR even when the GM-CSF subject had been treated and followed for response for durations comparable to the IMLYGIC® subjects. On the other hand, some clinicians consider GM-CSF to be somewhat effective, in which case the comparator was an active control and therefore comparison between the treatment arms would have been viewed differently from when the comparator was a placebo. Regardless, the DRR of 16.3% in the IMLYGIC® arm is an accurate estimate of the effect of IMLYGIC® because of the EAC adjudication and further verification by the FDA clinical reviewers.

The FDA clinical reviewers are also interested in the durable complete response (DCR), defined as a complete response maintained for at least six months. The FDA clinical reviewers determine that there are 19 DCRs, all in the IMLYGIC® arm, accounting for 6.4% (19/295) of the IMLYGIC® subjects, which is of a similar magnitude to the DCR rate in DTIC. Of these 19 DCRs, one was stage IVM1c, four were stage IVM1a/IVM1b, and the rest were stage IIIB/IIIC.

Subgroup analysis on DRR.

Because the DRR comparison between the two arms in the ITT population is statistically significant, the applicant and FDA have explored consistency of effects in subgroups. Figure 4 plots the DRR comparison in subgroups formed by key prognostic and potentially predictive factors. Subgroups formed by age, HSV-1 status, or sex (not reported in the Figure here) demonstrate treatment effects of similar magnitude. Numerically larger treatment effect is seen for subjects where the study treatment was 1st line than when the study treatment was 2nd line or greater. Note that the protocol definition of "line of therapy" excluded surgery, adjuvant, or radiation. Most notably, treatment effect decreases with later stage of disease. Stage IIIB/IIIC subjects have 33% DRR in the IMLYGIC® arm versus 0% in the GM-CSF arm. Stage IVM1a show a DRR of 16% vs 2.3%. Stage IVM1b and IVM1c both have low DRR that are of similar magnitude in both study arms. The applicant also tested qualitative interactions of treatment-by-line of therapy and treatment-by-stage, using the Gail and Simon likelihood ratio tests. The applicant concluded that there was not enough evidence to show that the direction of the treatment effect on DRR is different across the different stages or different line of therapy (p=0.5 and 0.9, respectively).

Figure 4. Absolute difference in EAC-confirmed durable response rates between study arms, in subgroups formed by key covariates based on case report form.



Source: Applicant's briefing document for the April 29, 2015 AC meeting, p.52, Figure 10.

Baseline lesion size of measurable lesions at baseline.

The clinical review discipline requested assistance to understand the baseline characteristics of the subject population and the responders. I performed several analyses to examine the distribution of baseline size of measurable lesions, among both the ITT population and the durable responders. Because the investigators at the study sites and the EAC selected baseline lesions for assessment of responses independently from each other, and because they might have reported different sizes for the same lesions, data from the investigators and the EAC were used in FDA analyses. Note that in this document, the size was determined by "*multiplying the longest diameter by the greatest diameter perpendicular to the longest diameter*." Thus, the size may or may not have matched the actual surface area of a lesion, depending on the actual shape of the lesion. The analyses results were provided to the clinical review discipline. One analysis is presented here in Table 5.

Table 5 lists the number and percentage of subjects whose largest baseline lesion fell within one of four size categories: < 0.5cm^2, 0.5 to 1, 1 to 2, or 2 to 1164 (the largest lesion among all subjects), based on measurements recorded by the investigators. The distributions in these size categories are comparable between the two treatment arms, as expected. However, among the durable responders, a larger proportion (30.4%) of subjects had only very small lesions (< 1cm^2) compared to the overall subject population (10.1%). This suggests that subjects who had smaller lesions were more likely to respond to IMLYGIC®. On the other hand, 45.7% of the DRs in the IMLYGIC® arm had at least one lesion that was greater than 2cm^2.

Table 5. Distribution of subjects according to the baseline size of the largest baseline measurable lesions, recorded by the investigators, in the ITT population*, by treatment arm and status of being a durable responder.

	IMLYGIC®			GM-CSF		
Baseline Size of Largest Lesion at Baseline (cm^2)	All (N=289)	Durable Responder (N=46)	Not Durable Responder (N=243)	All (N=127)	Durable Responder (N=2)	Not Durable Responder (N=125)
<0.5	12 (4.2%)	7 (15.2%)	5 (2.1%)	7 (5.5%)	0	7 (5.6%)
0.5 to (<1)	17 (5.9%)	7 (15.2%)	10 (4.1%)	6 (4.7%)	0	6 (4.8%)
1 to (<2)	34 (11.8%)	11 (23.9%)	23 (9.5%)	16 (12.6%)	0	16 (12.8%)
2 to 1164	226 (78.2%)	21 (45.7%)	205 (84.4%)	98 (77.2%)	2 (100%)	96 (76.8%)

*Data source: Original BLA 125518, amendment 24, dated 3/11/2015, submitted in response to an information request from the FDA. * The source data contained 5209 records for 419 subjects. The table is generated using 3442 records on 416 subjects, which were marked with "Y" for baseline flag and which did not miss the product value. Some records were taken at the "screening" visit while others were taken at the "cycle 1 day 1" visit.*

6.1.11.2 Analyses of Secondary Endpoints

I will discuss only the analyses of the most important secondary endpoint, overall survival (OS), in this section. Additional secondary endpoints included best overall response and disease burden, response onset, time to treatment failure, duration of response, and response interval. Quality of life (QOL) was assessed by the FACT-BRM questionnaire as an exploratory endpoint. The overall response rate was 26.4% among IMLYGIC® subjects (78/295, 32 CR, 46 PR) and 5.7% among the GM-CSF subjects (8/141, 1 CR, 7 PR). There was no provision in the applicant's statistical analysis plan to control the type 1 error rate in testing these additional secondary endpoints. In addition, OS, as the first among the list of the secondary endpoints, did not reach statistical significance in its pre-specified primary analysis. Furthermore, the results of these additional secondary endpoints are susceptible to the same potential biases identified previously in the consideration of DR, the primary endpoint. The FACT-BRM result is not readily interpretable; the data are limited by the low rate of completion of questionnaires in the GM-CSF group compared with the IMLYGIC® group during the study. Therefore, we do not further evaluate any of these other secondary and exploratory endpoints.

An interim analysis (IA) of OS occurred as planned at the time of the primary analysis of DRR, following a significant difference in DRR. At that time, 250 deaths had been recorded. This IA of OS yielded a p-value of 0.075. The primary analysis of OS was to occur after accumulation of 290 deaths. No other IA of OS occurred. The descriptive analysis of OS at the end of study (EOS) identified one additional death, from subject #035020 of the IMLYGIC® arm, during the additional follow-up period between the time of primary analysis of OS and EOS. The applicant reported a p-value of 0.0494 for this descriptive final analysis.

The data cut-off date (DCO) for the OS primary analysis was set to March 31, 2014. As of the DCO, there were 189/295 (64%) confirmed deaths in the IMLYGIC® arm and 101/141 (72%) confirmed deaths in the GM-CSF arm. The primary analysis using the un-adjusted log-rank test yielded a p-value of 0.051, just short of statistical significance. The estimates of median OS (in months) and the 95% CIs were 23.3 (19.6, 29.7) for the IMLYGIC® arm and 18.9 (16.2, 24.0) for the GM-CSF arm. The estimate of the hazard ratio was 0.79 with a 95% CI of (0.62, 1.00). See Figure 5 for the survival curves from this primary analysis.

Imbalance in potentially informative censoring and sensitivity analyses.

Per protocol, subjects were to be followed for survival status at 3-month intervals. To evaluate the impact of potential follow-up difference on OS results, we examined the reasons for subject censoring. There were two categories of reasons for censoring. The first category included all subjects who were followed for survival status until the DCO (March 31, 2014). These censorings were administrative and therefore were non-informative. The second category included all subjects whose follow-up stopped before the DCO. These censorings were considered potentially informative.

We identified 12 subjects whose end-of-study reason was not administrative censoring or death (Table 6). Upon close inspection, the two subjects from site #35 both had a last observation date that was within 3 months prior to the DCO, indicating that they had survival status follow-up until DCO. Therefore, only the remaining 10 subjects had non-administrative and therefore potentially informative censoring. These 10 subjects were distributed disproportionally, accounting for 5% (7/141) of the GM-CSF subjects but only 1% (3/295) of IMLYGIC® subjects. Two sites, site #3 and site #15, each accounted for three of the 10 subjects (Table 6).

The FDA had several interactions with the applicant in an attempt to ascertain the survival status of these 10 subjects at the time of the DCO. The applicant informed FDA that "*All ten subjects listed are being followed using the US Death Index and no deaths were found prior to study 005/05 final OS analysis*." The data cut-off date for the final, descriptive OS analysis was a few months after the DCO for the primary analysis of OS.

To understand the robustness of the OS result in the face of the imbalance in potentially informative censoring, we performed several sensitivity analyses. These sensitivity analyses included several approaches of imputing censoring times, and excluding site #15, or site #3, or both sites. All sensitivity analyses resulted in an increase above the reported .051 in the p-value of a log-rank test comparing OS between the two arms. For example, one post-hoc sensitivity analysis used the DCO of March 31, 2014 as the censoring time for all 10 subjects (Table 6, last column). This sensitivity analysis results in a p-value of 0.155, and a hazard ratio of 0.84 with a 95% CI of (0.66, 1.07). The presence and imbalance of potentially informative censoring increase the uncertainty about the presence and magnitude of the comparative effect on OS in the study. Nonetheless, the two survival curves in the sensitivity analysis continue to visually suggest some difference in time to death favoring the IMLYGIC® arm.

Updated survival status for the 10 subjects and updated analysis.

The applicant enlisted the service of a company to assist study sites to locate subjects lost to follow-up and to obtain vital status information (amendment #19, sequence #19, received February 18, 2015, Efficacy Information Amendment (EIA), pp.20-23 of 372).

- The applicant updated the survival status of three subjects, as of March 11, 2015, in amendment #27 (sequence #27, received March 26, 2015, EIA, 16 pages).
- The applicant updated the survival status of five of the 10 subjects as of April 3, 2015 (amendment #31, sequence #30, received April 14, 2015, EIA, p.12 of 16). In particular, one subject from the IMLYGIC® arm died in late 2011. This subject is identified with "§" in Table 6. Four GM-CSF subjects were confirmed alive after the DCO. These four subjects are identified with "*" in Table 6.

We performed an "updated" analysis that used the DCO as the censoring times for the four GM-CSF subjects confirmed alive after the DCO and treated the additional IMLYGIC® subject who died before DCO as a known event. The remaining five subjects without an update are treated the same as in the primary analysis. The updated analysis returns a p-value of 0.116, a hazard ratio of 0.82 with 95% of (0.65, 1.05). The median OS is 22.9 months (19.6, 29.7) in the IMLYGIC® arm and 19.0 months (16.2, 24.3) in the GM-CSF arm. See Figure 5 for the survival curves of the primary and updated analyses. The trend demonstrated by the survival curves in the two arms is similar between the updated analysis and the primary analysis, albeit with increase uncertainty in the latter, as we expected from sensitivity analyses performed prior to receiving the updated information. Although there is no further sensitivity analysis performed for the five remaining subjects currently without an update, the sensitivity analyses discussed previously demonstrate that no qualitative changes are expected should more precise information on these five subjects become available.

Subjects may have received various treatments following discontinuation of study therapy, which may affect the estimate of treatment effect on overall survival. Because subjects in the GM-CSF arm were more likely to discontinue therapy and did so at a faster rate than subjects in the IMLYGIC® arm, they may have had a greater exposure to various other therapies. However, in the absence of a pre-planned systematic collection of information on exposure to other therapies and an understanding of the activity of other therapies in the context of this trial, the effect of this potentially confounding factor on the OS result remains unknown.

Subgroup analysis on OS using the updated survival status of the five subjects.

The primary endpoint of DR shows greater effect in earlier stage disease and first-line therapy. We performed subgroup analyses of OS to further explore the subgroup trend observed in DR. For the subgroup analyses, we updated the data on OS with the updated survival status of the five subjects as described above, while keeping the rest of the OS data as in the primary analysis in the ITT population. That is, there are now only five subjects with potentially informative censoring; three in the GM-CSF and two in the IMLYGIC® arm. Further sensitivity analyses considering these remaining subjects do not impact the results qualitatively.

Figure 6 reports the subgroup analyses on OS using the updated information. We have the following observation.

1. Overall, the direction and relative magnitude of IMLYGIC® comparative effects on OS in subgroups resemble the trend observed in the DRR subgroup result.
2. OS effects in subgroups formed by age, sex, and HSV-1 status are consistent.
3. The OS effect is more pronounced in earlier stage disease (IIIB/IIIC/IVM1a), with a median OS of 23.7 months in the GM-CSF arm vs 41.1 months in the IMLYGIC® arm, a 17.4 months improvement. The nominal p-value comparing the two study arms is 0.004. There appears be no effect in either direction in later stage disease (IVM1b/IVM1b). There is also a monotone trend that OS in GM-CSF subjects is longer as stage becomes lower, consistent with the expectation that stage is a prognostic factor.
4. The OS effect is pronounced in subjects where the study treatments were first line therapy, with a median OS of 17.3 months in the GM-CSF arm versus 33.1 months in the IMLYGIC® arm, a 15.8 months improvement. The nominal p-value comparing the two study arms is 0.0008. IMLYGIC® performs slightly worse than GM-CSF numerically when the study treatments were 2nd line or greater.
5. The applicant reported covariate-by-treatment interaction tests, and concluded that "Results of the interaction tests for line of therapy and disease stage showed that the magnitude of the treatment effect was statistically different for line of therapy and disease stage although there was no statistically significant difference in the direction of the treatment effect between the subgroups."

Figure 7 shows the post-hoc subgroup analysis in earlier stage disease (IIIB/IIIC/IVM1a) versus later stage disease (IVM1b/IVM1c). Note that there is no multiplicity correction for these post-hoc analyses, and the p-values and confidence levels are nominal and should be interpreted with caution.

Table 6. Observations with potentially informative censoring.

Arm	#	Site ID	Age	Sex	F1	F2	F3	F4	End-of-Study Reason	Randomization Date	Last Observation Date	Censoring Time in Days (Primary Analysis)	Censoring Time in Days (Sensitivity Analysis)
GM-CSF	1*	15	59	F	No	LN	No Prior	3b/c	CW	4/13/2010	4/13/2010	1	1449
	2*	15	67	M	No	IT-DS	No Prior	3b/c	CW	6/22/2011	6/22/2011	1	1014
	3	3	55	F	No	LN	No Prior	3b/c	CW	12/10/2009	12/11/2009	2	1573
	4	15	79	F	No	Visceral	Prior, <1 year	4M1a/M1b	LFU	2/25/2010	3/12/2010	16	1496
	5*	9	60	M	No	IT-DS	Prior, >1 year	4M1a/M1b	CW	12/8/2009	3/3/2010	86	1575
	6	3	58	M	No	LN	Prior, >1 year	3b/c	CW	6/22/2009	1/27/2011	585	1744
	7*	69	38	F	No	LN	No Prior	4M1a/M1b	CW	2/28/2011	2/26/2013	730	1128
	8	35	71	F	No	LN	Prior, >1 year	4M1a/M1b	CW	12/15/2009	2/6/2014	1515	-
IMLYGIC®	1§	79	54	M	No	Visceral	Prior, <1 year	4M1a/M1b	O	11/22/2010	11/22/2010	1	1226
	2	3	60	F	No	IT-DS	Prior, >1 year	4M1a/M1b	CW	4/25/2011	5/2/2011	8	1072
	3	66	54	F	No	IT-DS	No Prior	4M1a/M1b	CW	7/30/2010	11/9/2012	834	1341
	4	35	85	F	Yes	IT-DS	No Prior	4M1c	CW	5/4/2010	2/6/2014	1375	-

F1 through F4 stand for levels of the stratification factors recorded in IVRS. (1) F1: Presence of liver metastasis. (2) F2: Site of First Recurrence. LN: Lymph Nodes. IT-DS: In Transit or Distant Skin. (3) F3: Whether received prior non-surgical melanoma treatment and time to recurrence. (4) F4: Stage of disease.

CW: Consent withdrawn. LFU: Lost to follow-up. O: Other, "*subject randomized in error; subject was ineligible [for enrollment] due to brain mets*".

Last Column: Censoring time imputed using the primary analysis data cut-off date (DCO), 3/31/2014, as the last follow-up time for the 10 observations with potentially information censorings in the primary analysis. The FDA sensitivity analysis using this imputation is summarized in the text above.

Rows highlighted as "*" are subjects who were later updated to be verified alive as of a date after the DCO. For these subjects, the last column is actual, updated data, not imputed data.

Row highlighted as "§" is the subject who was verified to have died prior to the DCO. For this subject, the last column imputation in the FDA sensitivity analysis, therefore, assumes the subject was alive longer than he actually was. See text for more information

Figure 5. Overall survival in the ITT population: primary analysis vs updated analysis.

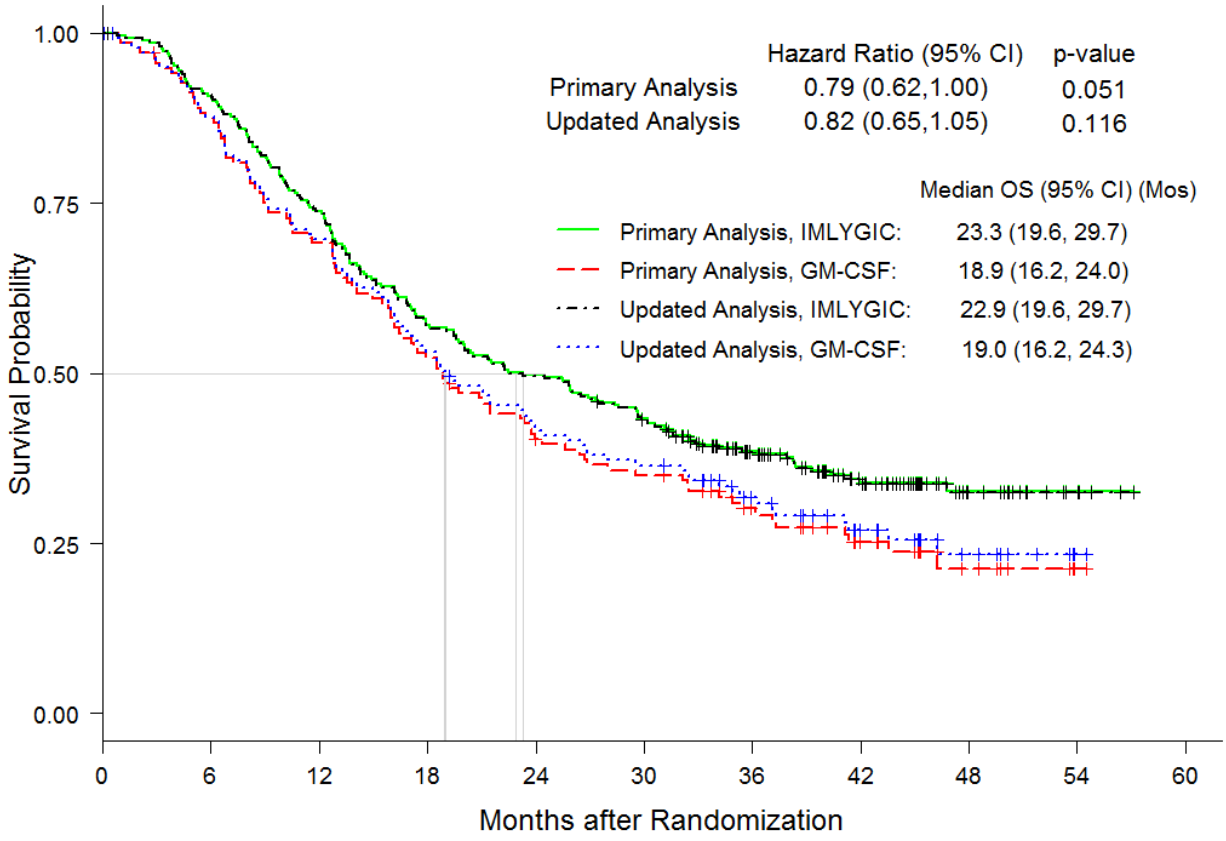
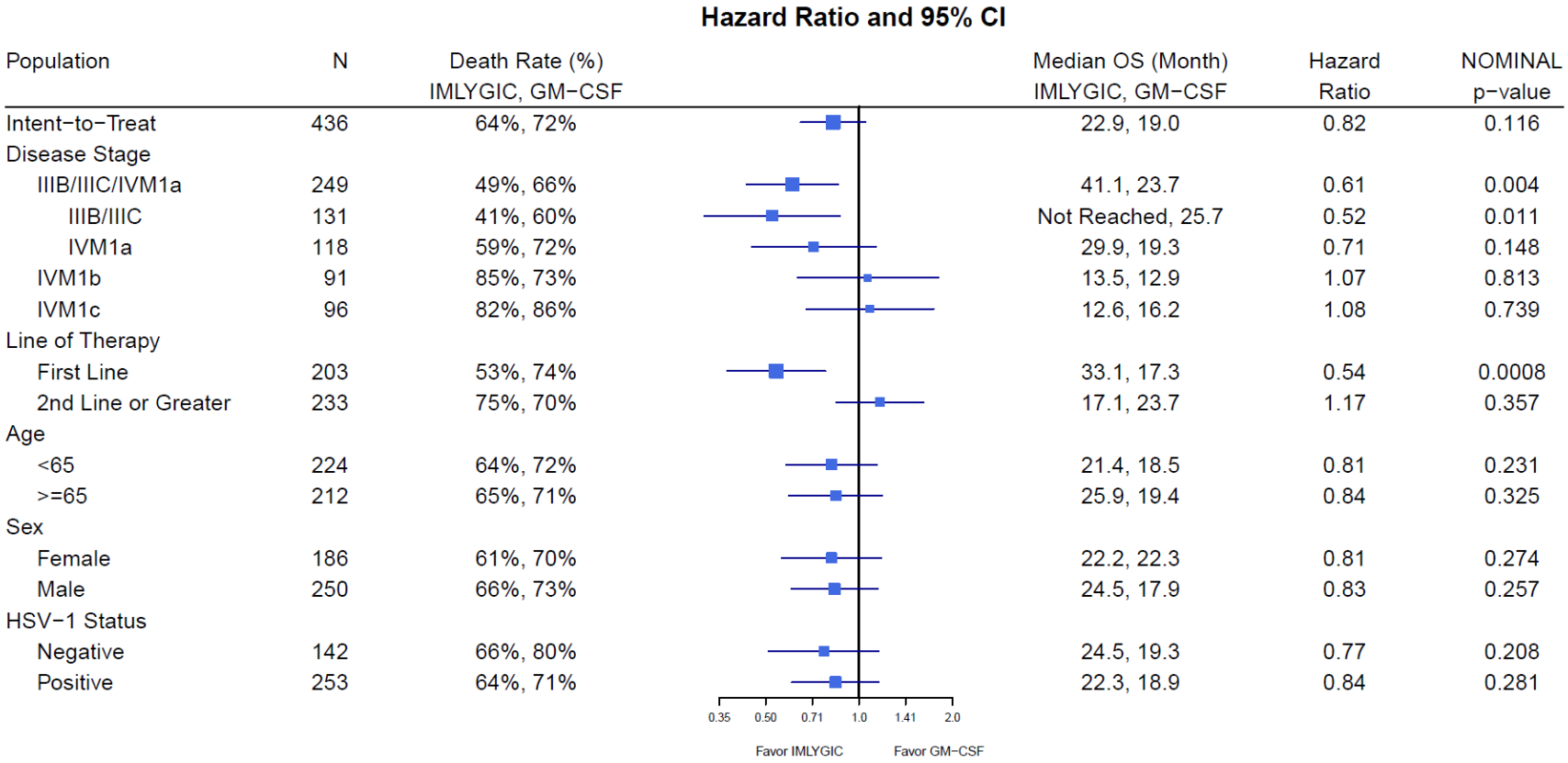
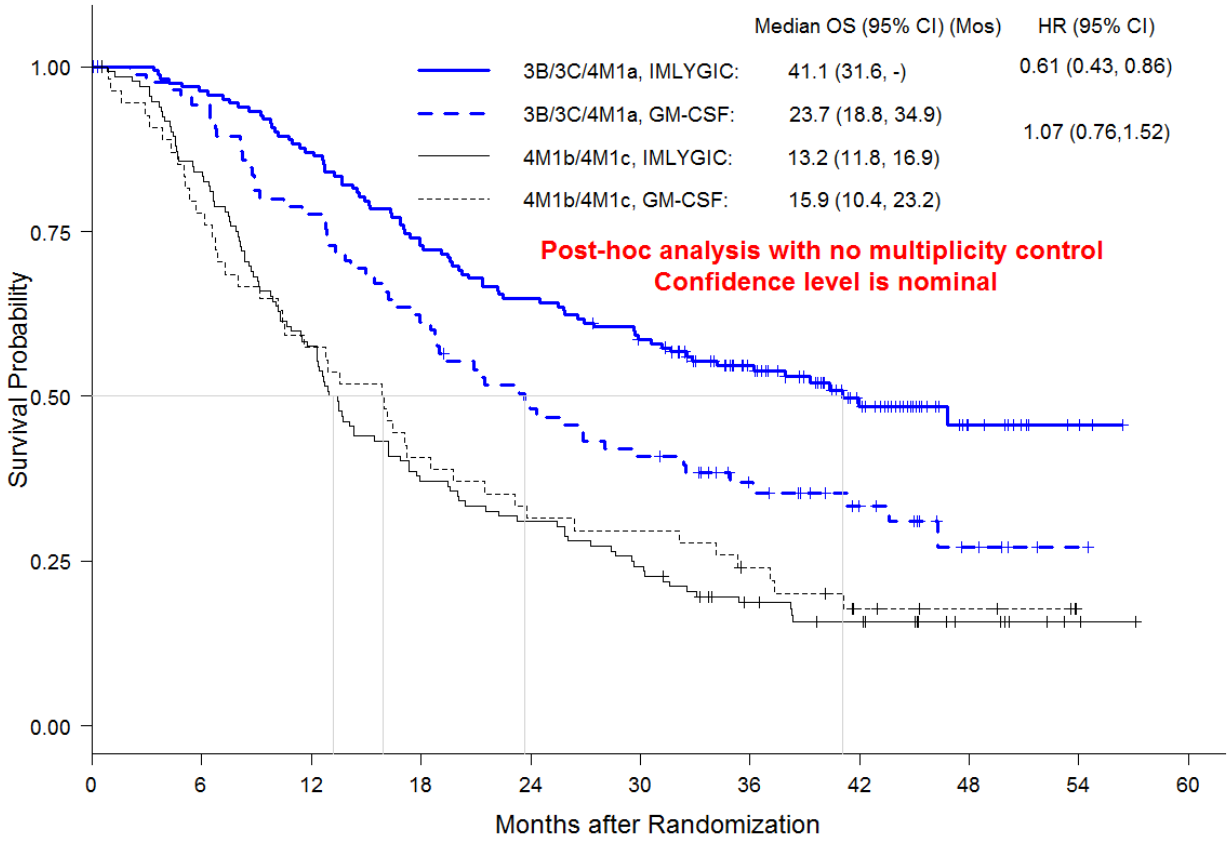


Figure 6. Updated overall survival comparison between study arms, in subgroups formed by key covariates reported by case report forms. Updated information on five of the 10 subjects with potentially informative censoring in the primary analysis is used. For the other five subjects the same data as in the primary analysis are used. Note that p-values are nominal, due to absence of pre-specified multiplicity control.



Population	N	Death Rate (%) IMLYGIC, GM−CSF	Median OS (Month) IMLYGIC, GM−CSF	Hazard Ratio	NOMINAL p−value
Intent−to−Treat	436	64%, 72%	22.9, 19.0	0.82	0.116
Disease Stage					
IIIB/IIIC/IVM1a	249	49%, 66%	41.1, 23.7	0.61	0.004
IIIB/IIIC	131	41%, 60%	Not Reached, 25.7	0.52	0.011
IVM1a	118	59%, 72%	29.9, 19.3	0.71	0.148
IVM1b	91	85%, 73%	13.5, 12.9	1.07	0.813
IVM1c	96	82%, 86%	12.6, 16.2	1.08	0.739
Line of Therapy					
First Line	203	53%, 74%	33.1, 17.3	0.54	0.0008
2nd Line or Greater	233	75%, 70%	17.1, 23.7	1.17	0.357
Age					
<65	224	64%, 72%	21.4, 18.5	0.81	0.231
>=65	212	65%, 71%	25.9, 19.4	0.84	0.325
Sex					
Female	186	61%, 70%	22.2, 22.3	0.81	0.274
Male	250	66%, 73%	24.5, 17.9	0.83	0.257
HSV−1 Status					
Negative	142	66%, 80%	24.5, 19.3	0.77	0.208
Positive	253	64%, 71%	22.3, 18.9	0.84	0.281

Figure 7. Overall survival: post-hoc subgroup analysis by disease stage.



6.1.11.3 Subpopulation Analyses

The ITT population is 97.9% white. Sex and age dichotomized by 65 years both show consistent effect in both the DR (Table 7) and OS endpoint. Refer to sections on the analysis of these two endpoints for details on other subpopulations, e.g., disease stage.

Table 7. Durable response rates in subpopulations formed by age and sex.

Subpopulation	N	IMLYGIC® DRR	GM-CSF DRR
Age ≥ 65	212	14.5%	1.4%
Age < 65	224	18.2%	2.9%
Female	186	15.6%	1.6%
Male	250	16.8%	2.6%

6.1.12 Safety Analyses

The safety summary below is based on the AC meeting briefing documents from both the FDA and the applicant. At this time clinical review of the safety data is still ongoing. Please refer to the final clinical review memo for detailed and critical evaluation of the safety data.

The applicant's safety analyses include three safety analysis sets. Note that the summary below is on Study 005/05.

- Primary Melanoma Analysis Set includes data from Study 005/05 submitted in the primary analysis clinical study report.
- Supportive Melanoma Analysis Set includes data from Studies 002/03 and 005/05 and their respective extensions.
- Program-Wide Analysis Set includes data from the Supportive Melanoma Analysis Set and from several smaller studies in various tumor types, including melanoma.

Exposure.

Across the clinical program, 408 subjects were exposed to at least one dose of IMLYGIC®; 269 subjects were exposed for less than months and 20 subjects were exposed for ≥18 months.

In Study 005/05, the safety population, the Primary Melanoma Analysis Set, consisted of 419 subjects: 292 IMLYGIC® and127 GM-CSF. The median duration of treatment was 23 weeks (range: 0.1 to 78.9) in the IMLYGIC® arm and 10 weeks (range: 0.6 to 72) in the GM-CSF arm. For subjects who continued into the extension study, the maximum duration of treatment with IMLYGIC® was 30.8 months. The following summary is based on the Primary Melanoma Analysis Set.

The most common treatment-emergent adverse events with IMLYGIC® were fatigue, chills, pyrexia, nausea, influenza-like illness and injection site pain.

6.1.12.3 Deaths

Fatal adverse events were reported for 10 (3.4%) of the IMLYGIC® subjects and two of the GM-CSF subjects. Eight of the IMLYGIC® deaths and both of the GM-CSF deaths were due to disease progression. The remaining two IMLYGIC® deaths were due to myocardial infarction and sepsis, which were considered by the applicant to be due to other underlying disease processes. No fatal events were reported as treatment-related.

6.1.12.4 Nonfatal Serious Adverse Events

The incidence of serious adverse events was 25.7% in the IMLYGIC® arm and 13.4% in the GM-CSF arm. See Table 8 for a list of treatment-emergent serious adverse events with a ≥ 1% incidence in either treatment group. Cellulitis at the injection site was the most commonly reported treatment-related serious adverse event, occurring to 1.7% of the IMLYGIC® subjects and none of the GM-CSF subjects.

After IMLYGIC® administration, a wound became resistant to medical therapy, and required a below-the-knee amputation.

6.1.12.5 Adverse Events of Special Interest (AESI)

Adverse events of special interest (Table 9) included immune-mediated adverse events, cellulitis, flu-like symptoms, HSV-1 infections, hypersensitivity reactions, injection site reactions, vitiligo, impaired wound healing at the injection site, plasmacytoma at the injection site, and other neoplastic events.

Table 8. Treatment-Emergent Serious Adverse Events by Preferred Term With ≥1 % Subject Incidence in Either Treatment Group (Primary Melanoma Analysis Set).

Preferred Term	GM-CSF (N = 127) n (%)	Talimogene Laherparepvec (N = 292) n (%)	Total (N = 419) n (%)
Number of subjects reporting serious treatment-emergent adverse events	17 (13.4)	75 (25.7)	92 (22.0)
Disease progression	2 (1.6)	9 (3.1)	11 (2.6)
Cellulitis	1 (0.8)	7 (2.4)	8 (1.9)
Pyrexia	0 (0.0)	5 (1.7)	5 (1.2)
Tumour pain	0 (0.0)	4 (1.4)	4 (1.0)
Cerebral haemorrhage	0 (0.0)	3 (1.0)	3 (0.7)
Deep vein thrombosis	0 (0.0)	3 (1.0)	3 (0.7)
Gastrointestinal haemorrhage	0 (0.0)	3 (1.0)	3 (0.7)
Infected neoplasm	0 (0.0)	3 (1.0)	3 (0.7)
Metastases to central nervous system	1 (0.8)	3 (1.0)	4 (1.0)
Metastatic malignant melanoma	0 (0.0)	3 (1.0)	3 (0.7)
Pleural effusion	0 (0.0)	3 (1.0)	3 (0.7)

Source: Applicant's briefing document for the April 29, 2015 AC meeting, p.67, Table 18.

Table 9. Subject Incidence of Adverse Events of Interest by Category (Primary Melanoma Analysis Set)

Event of Interest Category	GM-CSF (N = 127) n (%)	Talimogene Laherparepvec (N = 292) n (%)	Total (N = 419) n (%)
IMMUNE-MEDIATED EVENTS (AUTOIMMUNE DISORDERS)[a]			
Adverse event	2 (1.6)	5 (1.7)	7 (1.7)
Serious adverse event	0 (0)	1 (0.8)	1 (0.2)
CELLULITIS AT THE INJECTION SITE (BACTERIAL CELLULITIS)			
Adverse event	2 (1.6)	18 (6.2)	20 (4.8)
Serious adverse event	1 (0.8)	7 (2.4)	8 (1.9)
FLU LIKE SYMPTOMS			
Adverse event	83 (65.4)	264 (90.4)	347 (82.8)
Serious adverse event	0 (0.0)	9 (3.1)	9 (2.1)
HERPES SIMPLEX VIRUS (HSV) INFECTIONS			
Adverse event	2 (1.6)	16 (5.5)	18 (4.3)
Serious adverse event	0 (0.0)	0 (0.0)	0 (0.0)
HYPERSENSITIVITY			
Adverse event	25 (19.7)	53 (18.2)	78 (18.6)
Serious adverse event	0 (0.0)	0 (0.0)	0 (0.0)
INJECTION SITE REACTIONS			
Adverse event	64 (50.4)	122 (41.8)	186 (44.4)
Serious adverse event	0 (0.0)	0 (0.0)	0 (0.0)
VITILIGO			
Adverse event	2 (1.6)	15 (5.1)	17 (4.1)
Serious adverse event	0 (0.0)	0 (0.0)	0 (0.0)
IMPAIRED WOUND HEALING AT THE INJECTION SITE[a]			
Adverse event	3 (2.4)	16 (5.5)	19 (4.5)
Serious adverse event	1 (0.8)	0 (0.0)	1 (0.2)
OTHER NEOPLASTIC EVENTS (MALIGNANT OR UNSPECIFIED TUMORS)			
Adverse event	3 (2.4)	16 (5.5)	19 (4.5)
Serious adverse event	1 (0.8)	9 (3.1)	10 (2.4)
PLASMACYTOMA			
Adverse event	0 (0.0)	1 (0.3)	1 (0.2)
Serious adverse event	0 (0.0)	1 (0.3)	1 (0.2)

Source: Applicant's briefing document for the April 29, 2015 AC meeting, pp.70-71, Table 20.

9. ADDITIONAL STATISTICAL ISSUES

9.2 Aspect(s) of the Statistical Evaluation Not Previously Covered

Study site issue.

The Office of Compliance and Biologics Quality (OCBQ) identified an issue regarding past inspection of Site #66, On December 23, 2014, I provided information on data from site #66 to the clinical review discipline, with the following observation.

- There are 25 subjects from site #66 in the ITT population, 8 in the GM-CSF arm and 17 in the IMLYGIC® arm. Seven GM-CSF subjects and 14 IMLYGIC® subjects died by DCO. There is a greater mortality rate in each arm at site #66 compared to the ITT population as a whole. This may be due to random variation. All subjects at this site were randomized in 2010 and 2011.
- One subject in the IMLYGIC® arm at Site #66 is one of the 10 subjects with potentially informative censoring (Table 6). This subject had a censoring time at 834 days. No updated survival status information was obtained by the applicant.
- Preliminary sensitivity analyses at the time did not raise concerns from a statistical perspective with including the site #66 data in the ITT population. For example, excluding all subjects from site #66 in the OS analysis resulted in a p-value of 0.056, versus 0.051 in the primary analysis on OS. There is only one subject from this site reported as a durable responder. The clinical reviewers determined that this subject should not qualify as DR due to too many missed visits.

10. CONCLUSIONS

10.1 Statistical Issues and Collective Evidence

Study 005/05 was an open-label study comparing the efficacy and safety of intra-lesional injection of IMLYGIC® to subcutaneous GM-CSF in treatment of melanoma patients with unresectable but injectable stage IIIB, IIIC, or IV disease. The ITT population consists of 295 IMLYGIC® subjects and 141 GM-CSF subjects. The primary efficacy endpoint is EAC-confirmed durable response, defined as maintenance of response (PR or CR) continuously for six months. Overall survival (OS) is an important secondary endpoint. Power for both DR and OS were taken into account in the study design.

The primary analysis comparing the durable response rate (DRR) in the ITT population between the two arms is statistically highly significant ($p < .0001$), with a DRR of 16.3% (48/295) in the IMLYGIC® arm vs. 2.1% (3/141) in the GM-CSF arm. The GM-CSF subjects had on average a much shorter duration of study treatment and response assessment, compared to IMLYGIC® subjects. Because of this, the reported DRR in the GM-CSF arm may be an underestimate. However, I consider the statistical significance of the comparison of DRR between the study arms to be statistically robust. In addition, 19 of the 48 IMLYGIC® DRs are durable complete responders, with complete response maintained for at least six months, accounting for 6.4% of the ITT IMLYGIC® subjects.

The primary analysis comparing OS in the ITT population at the time of database lock between the two arms was just short of being statistically significant, at a p-value of 0.051, with a hazard ratio (HR) estimate of 0.79 and a 95% confidence interval (CI) of (0.62, 1.00). The median OS for the primary analysis is 23.3 months (95% CI: 19.6-29.7) in the IMLYGIC® arm and 18.9 months (95% CI: 16.2-24.0) in the GM-CSF arm. I identified a total of 10 subjects with potentially informatively censored event times. These 10 subjects were distributed disproportionately between the two arms, accounting for 5% (7/141) of the GM-CSF subjects and 1% (3/295) of the IMLYGIC® subjects, respectively. Additional retrospective information, on survival data up to the data cut-off date for the primary analysis, was subsequently obtained by the applicant for five of the 10 subjects. The updated analysis, incorporating this additional information from these five subjects, yields a p-value of 0.116, a hazard ratio estimate of 0.82 and a 95% CI of (0.65, 1.05). The median OS for the updated analysis is 22.9 months (95% CI: 19.6-29.7) in the IMLYGIC® arm and 19.0 months (95% CI: 16.2-24.3) in the GM-CSF arm. The updated survival curves are still visually separate, though to a lesser extent than the primary analysis, favoring the IMLYGIC® arm.

Subgroup analyses reveal that DRR in the IMLYGIC® arm is substantially higher in the subset of subjects with earlier stage disease, compared to later stage disease. The DRR is 33% in the 131 IIIB/IIIC subjects, 16.0% in the 118 IVM1a subjects, 3.8% in the 90 IVM1b subjects, and 3.4% in the 96 IVM1c subjects. Subgroup analyses of OS show a similar trend: in IIIB/IIIC, the median OS are 25.7 months vs "Not reached", for the GM-CSF and IMLYGIC® arms, respectively; the medians are 19.3 vs 29.9 in IVM1a, 12.9 vs 13.6 in IVM1b, and 16.2 vs 12.6 in IVM1c.

A joint meeting of the Cellular, Tissue, and Gene Therapies AC and Oncologic Drugs AC was held on April 29, 2015. The AC voted 22 "yes" and one "no" to the question of whether IMLYGIC® has a favorable benefit-risk profile to support traditional approval. A number of AC members qualified their votes by stating that they would want the approval to be limited to earlier stage disease or patients without visceral disease.

10.2 Conclusions and Recommendations

Study 005/05 demonstrated a highly statistically significant improvement in durable response rate in the ITT population, from 2.1% in the GM-CSF arm to 16.3% in the IMLYGIC® arm. In addition, 6.4% of the IMLYGIC® subjects maintained a complete response for six months or longer. The comparison of overall survival in the ITT population was not statistically significant, but showed a favorable trend towards IMLYGIC®. For DRR, IMLYGIC® had a substantially greater effect in earlier stage disease (IIIB and IIIC, and possibly IVM1a) than later stage disease. Subgroup analyses of OS showed a similar pattern, i.e., point estimates suggest that the product has greater activity in earlier stage disease than in later stage disease.

There was no planned multiplicity control for subgroup analyses. Caution should be applied when considering the observed differences between subgroups in the comparison of the two arms. In particular, I recommend viewing the subgroup analyses of OS as supportive information for the subgroup analyses of DRR, rather than as definitive evidence of a survival benefit for IMLYGIC®.