STA ERRATUM

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Talimogene laherparepvec for treating metastatic melanoma [ID508]

Confidential until published

This report was commissioned by the NIHR HTA Programme as project number 14/206/04

Erratum completed 5th February 2016

CONTAINS ACADEMIC IN CONFIDENCE AND COMMERCIAL IN CONFIDENCE DATA

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

The company identified 5 overall issues in relation to factual inaccuracies in the original Evidence Review Group (ERG) report. All were considered by the ERG to require minor changes to the text. The pages of the report affected are presented here. Please note:

- New text added by the ERG is in italics and underlined.
- Text deleted completely (as opposed to being re-worded) is struck out.
- Unaltered text which is considered to be of relevant context to that added, amended or deleted (such as headings or sentences preceding or following the added, amended or deleted text) is presented in its original font.
- All other unaltered text is greyed out.

1 SUMMARY

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence has been submitted to NICE by Amgen Limited in support of the use of talimogene laherparepvec (Imlygic®) (hereafter referred to as T-VEC) to treat patients with non-visceral metastatic melanoma.

1.1 Critique of the decision problem in the company's submission

The intervention specified in the NICE scope is T-VEC. It has been recognised by the European Medicines Agency (EMA) as a novel, first-in-class oncolytic immunotherapy treatment. *A marketing authorisation was granted from the European Commission on 16 December 2015 for the treatment of adults with unresectable melanoma that is regionally or distantly metastatic (stage IIIB, stage IIIC and stage IV M1a) with no bone, brain, lung or other visceral disease.* The company estimates that, if recommended by NICE, 728 patients in England would be eligible for treatment with T-VEC in 2015.

The population specified in the NICE scope is adults with stage IIIB to stage IV melanoma. *This is the same patient population for which T-VEC is licensed.* However, as T-VEC is administered by intralesional injection, its use will be restricted to patients whose melanoma is considered injectable, i.e. there must be cutaneous, subcutaneous, and/or nodal lesions that are visible, palpable or detectable by ultrasound guidance.

The following comparators are specified in the NICE scope: ipilimumab, vemurafenib and dabrafenib. Unfortunately, none of these drugs has been studied in trials comprising only patients with non-visceral metastatic stage IIIB to stage IV M1a melanoma or in trials where these patients are a specified subgroup. Ipilimumab is considered by the company to be the primary comparator to T-VEC and vemurafenib and dabrafenib are not evaluated in the company submission (CS). However, with NICE's recent recommendation that pembrolizumab should be made available through the NHS as a treatment for some patients with metastatic melanoma, the ERG considers that, in future, all patients who are currently offered first- or second-line treatment with ipilimumab will now be offered pembrolizumab (if they have not already received it).

Clinical evidence is reported in the CS for all five outcomes specified in the NICE scope: overall survival (OS), progression-free survival (PFS), tumour response rate, adverse events (AEs) of treatment and health-related quality of life (HRQoL). These are all outcomes that

are commonly measured in metastatic melanoma drug trials. In addition, durable response rate (DRR) was also reported as the primary outcome in the OPTiM trial from which the majority of evidence for T-VEC is derived; DRR is a non-validated, albeit a clinically relevant, endpoint. The OPTiM trial reported time to treatment failure (TTF) instead of PFS since patients were permitted to continue to receive treatment despite showing evidence of disease progression with T-VEC.

1.2 Summary of clinical effectiveness evidence submitted by the company

Evidence for the relative efficacy of T-VEC was obtained from the OPTiM randomised controlled trial (RCT). Evidence from one Phase II non-RCT (Study 002/03) is also presented in the CS.

In the open label OPTiM trial patients with stage IIIB to stage IV M1c disease were randomised in a 2:1 ratio to receive either T-VEC (n=295) or granulocyte macrophage colony-stimulating factor (GM-CSF) (n=141). *The licence* for T-VEC is based on clinical data from a subgroup of these patients (n=249), namely patients with injectable non-visceral metastatic melanoma (i.e. stage IIIB to stage IV M1a disease); *although subgroup analyses by different stages of disease (stage IIIB to IIIC, stage IV M1a, stage IV M1b and stage IV M1c) were all pre-specified, stage IIIB to stage IV M1a disease was not a pre-specified subgroup but rather defined post-hoc.* Post-hoc analysis refers to those in which the hypotheses being tested are not specified before any examination of the data. *The results for analyses for this subgroup are:*

- DRR by Endpoint Assessment Committee (EAC) *(primary data cut)* assessment was higher in patients treated with T-VEC compared with GM-CSF (25.2% vs 1.2%; unadjusted odds ratio 28.6; [95% CI: 3.9 to 211.5]; p<0.0001)
- TTF by investigator assessment *(final data cut)* was longer in the T-VEC arm than in the GM-CSF arm (median 13.1 months vs 3.3 months; hazard ratio [HR]=0.28; [95% CI: 0.20 to 0.40]; p<0.0001)
- Overall tumour response rate by EAC assessment *(primary data cut)* was higher in the T-VEC arm than in the GM-CSF arm (40.5% vs 2.3%, p<0.0001)
- At the final OS analysis *(final data cut),* median OS gain was 25.3 months for patients in the T-VEC arm vs patients in the GM-CSF arm (median 46.8 months vs 21.5 months, unstratified HR=0.56; [95% CI: 0.40 to 0.79]; p=0.0008).

In patients with non-visceral metastatic disease treated with T-VEC, treatment-related Grade 3 to 5 AEs and treatment-related serious AEs (SAEs) were reported by 14% and 6% of patients respectively, and treatment emergent AEs leading to discontinuation were reported by 9% of patients. In the overall trial population, the most common AEs reported by patients receiving T-VEC were flu-like symptoms (90%) and injection-site reactions (42%)..

Overall, the ERG considers that patients with non-visceral metastatic disease in the OPTiM trial are generally similar to the patients with *stage IIIB* to stage IV M1a disease likely to be considered for treatment with T-VEC in clinical practice in England.

The ERG has concerns that the population considered in this STA is one that has been constructed following the results of a post-hoc analysis of data collected during the OPTiM trial. The ERG is particularly concerned that the *differential survival trajectory* of patients with stage III disease is likely to differ from that of those with stage IV M1a disease. Furthermore, the ERG considers that the OPTiM trial may be subject to bias due to limited blinding, a higher proportion of dropouts in the GM-CSF arm (particularly in the first few months of the trial), and the use of DRR as the primary endpoint. It is noted in the draft European Public Assessment Report (EPAR) that DRR is a new, clinically relevant, endpoint that is non-validated and is potentially prone to bias. However, the ERG does not consider that the potential sources of bias fully explain the improvements in efficacy in the T-VEC arm compared with the GM-CSF arm. The ERG notes that a further uncertainty, raised by the US Food and Drug Administration (FDA), relates to the size of lesions. The results of an FDA post-hoc analysis of the overall intention-to-treat population (i.e. including those with stage IV M1b and stage IV M1c disease) suggest that patients who had very small lesions (<1 cm^2) were more likely to respond to T-VEC than were the overall population (10.1%). The ERG further notes that evidence for the effectiveness of T-VEC treatment is not presented by line of therapy in the subgroup of patients with non-visceral metastatic disease.

Results from the OPTiM trial suggest that T-VEC's safety profile compares favourably with those of the comparator treatments detailed in the NICE scope. The ERG, however, notes that there are limited data to support the long-term safety of treatment with T-VEC.

Although HRQoL data collected as part of the OPTiM trial show that, in general, quality of life for patients receiving T-VEC was better than for those receiving GM-CSF, a substantial proportion of patients in the GM-CSF arm did not complete HRQoL assessments, suggesting that the HRQoL findings should be interpreted with caution.

For reasons highlighted in Section 1.5, the ERG does not consider the ipilimumab survival estimates generated by the company, using either the modified Korn model or the two-step Korn model to be reliable. It is, therefore, impossible to determine the relative clinical effectiveness of T-VEC compared with any of the comparators listed in the NICE scope.

using data from patients with predominantly stage IV M1b and stage IV M1c disease, whilst it is patients with stage IV M1a disease who mostly feature in the OPTiM trial. Furthermore, in the OPTiM trial 54.7% of T-VEC patients had stage IIIB, stage IIIC or stage IV M1a disease compared with less than 20% in the ipilimumab trials

3. There is no information in the public domain relating to the way in which the original (published) Korn model has been modified or to the data used to calibrate the model. It is likely that the issues outlined in point 2 also hold for the modified Korn model. In addition, the modified Korn model includes an adjustment for elevated lactate dehydrogenase (LDH), which is not relevant for patients with stage IIIB, stage IIIC or stage IV M1a disease, but has the effect of reducing the size of the coefficients associated with other adjustment factors (and improving the relative efficacy of T-VEC)
4. The effectiveness of ipilimumab may vary significantly by stage of disease. The company has attempted to correct for this case-mix imbalance by using the two-step Korn model, which is a further application of the modified Korn model. This additional adjustment is likely to mean that the problems previously described are further compounded
5. The original Korn publication includes both PFS and OS models. The PFS model is quite different from the OS model. The ERG, therefore, concludes that the company's use of the same modified Korn model for both OS and PFS is inappropriate.

Within the company model, different methods are applied sequentially to estimate OS. A number of issues with this approach were identified by the ERG, including:

1. ~~OS data from the earlier, less mature, data cut of the OPTiM trial were used by the company~~
2. The exponential trend used by the company to project OS for patients treated with T-VEC deviates markedly from the final recorded OPTiM trial data
3. For patients with stage I, stage II and stage III disease, the American Joint Committee on Cancer (AJCC) survival trends provide results from the date of diagnosis, whilst for patients with stage IV disease trends are recorded from the time of identification of first distant metastases. The relevance of these mixed AJCC adjusted mortality estimates is highly questionable
4. The data on which the AJCC analyses were performed were gathered prior to the current era of novel immunological treatments and may be unrealistic as these newer treatments have significantly altered the prospects for many patients
5. A sudden increase in the mortality rate after 270 weeks (62.1 months) is observed in the company model. The ERG considers that this effect is arbitrary and without any clinical justification
6. After 10 years, UK life table mortality rates are applied within the company model without adjustment, other than for age and sex. This implies that the cohort of long-term survivors is suddenly cured at this time point.

Other model-related issues identified by the ERG include an error in the discounting calculation, poor choice of health state utility values, lack of use of a terminal state disutility, use of a half-cycle (rather than a mid-cycle) continuity correction and a PSA ICER calculation error.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

Clinical evidence

- Results from the OPTiM trial show that the effectiveness of T-VEC is markedly improved in the subgroup of patients with stage IIIB to stage IV M1a disease when compared with the overall trial population (which includes patients with stage IV M1b and M1c disease)
- Evidence from the OPTiM trial suggests that the safety profile of T-VEC compares favourably to the safety profile of the comparators listed in the NICE scope
- The company has made thorough attempts to identify studies that include both a relevant treatment comparator to T-VEC and a relevant patient population.

Cost effectiveness evidence

- The company supported the appraisal process by providing the additional analyses requested by the ERG in a timely manner.

1.6.2 Weaknesses and areas of uncertainty

Clinical evidence

- Following the very recent approval of pembrolizumab for the first- and second-line treatment of patients with metastatic malignant melanoma, clinicians' first choice of systemic treatment for this population is likely to shift away from ipilimumab towards pembrolizumab
- The efficacy data *for the T-VEC licensed population* (patients with non-visceral metastatic melanoma) has been extracted from a post-hoc subgroup data analysis from the OPTiM trial
- The OPTiM trial may be subject to bias due to limited blinding and a higher proportion of dropouts in the GM-CSF arm (particularly in the first few months of the trial)
- The use of DRR as the primary endpoint in the OPTiM trial raises concerns as DRR is a new, albeit clinically relevant, endpoint which is non-validated and is potentially prone to bias
- The results of an FDA post-hoc analysis suggest that patients who had very small lesions (<1 cm^2) were more likely to respond to T-VEC than the overall population
- Two areas where evidence relating to treatment with T-VEC is lacking are in relation to line of treatment and long-term safety
- The relative clinical effectiveness of T-VEC compared with any treatment currently used in clinical practice is unknown.

Cost effectiveness evidence

- The ERG does not consider that the synthesised ipilimumab comparator is sufficiently reliable to support a valid assessment of the cost effectiveness of treatment with T-VEC vs ipilimumab

3.1 Population

The population specified in the NICE scope is adults with stage IIIB to stage IV melanoma. *A marketing authorisation was granted from the European Commission on 16 December 2015 for the treatment of adults with unresectable melanoma that is regionally or distantly metastatic (stage IIIB, stage IIIC and stage IV M1a) with no bone, brain, lung or other visceral disease.* These patients with non-visceral metastatic melanoma are referenced in the company's description of the population in the decision problem. Therefore, the clinical evidence presented by the company is only applicable to a subgroup of the patients specified in the NICE scope.

Importantly, but not explicitly stated in either the NICE scope or company's decision problem or in *the licence,* as T-VEC is administered by intralesional injection, the patient population is further restricted to patient's whose melanoma is considered injectable, i.e. there must be cutaneous, subcutaneous, and/or nodal lesions that are visible, palpable or detectable by ultrasound guidance. Patient experience of injectable treatments is not discussed in the CS. The ERG is not confident that all patients with injectable melanoma will be accepting of this type of treatment every 2 weeks over a long period of time.

Just under three-quarters (73%) of patients with metastatic non-visceral disease are considered by the company to have injectable disease. The population in the OPTiM trial[4] is therefore not directly comparable with patients in other trials for two reasons: (i) as noted by the ERG in Section 2.2 (Table 1), no other trial has conducted a subgroup analysis of patients with stage IIIB to stage IV M1a disease and (ii) only the OPTiM trial[4] has included patients solely with injectable disease.

3.2 Intervention

The intervention specified in the CS and in the company's decision problem statement is an oncolytic virus, T-VEC, derived from the herpes simplex virus type-1 (HSV-1) that has been modified to efficiently replicate within tumours and to produce the immune stimulatory protein granulocyte-macrophage colony-stimulating factor (GM-CSF). The aim of treatment is to boost the body's immune system to protect itself from carcinogenesis and progression of cancer.[34,35]

T-VEC has two complementary mechanisms of action in/on cancerous cells:[36] (i) replication that causes cell rupture/lysis and death (intracellular or direct effect) (ii) post-lysis release of tumour-derived antigens and GM-CSF, stimulating a systemic immune response from antigen-presenting cells (APCs) upon distant tumour sites (extracellular or indirect effect).

A summary of the reasons for discontinuing treatment and the reasons for discontinuing to participate in the trial is presented in Table 9.

Table 9 Summary of the reasons for discontinuing treatment and the reasons for discontinuing to participate in the OPTiM trial (primary analysis)

Reason for discontinuing treatment and from study	**stage IIIB–stage IV M1a (T-VEC licensed population)**		**stage IIIB to stage IV M1c (ITT population)**	
	T-VEC (n=163)	**GM-CSF (n=86)**	**T-VEC (n=295)**	**GM-CSF (n=141)**
Not treated (%)	[redacted]	[redacted]	**1.4**	**9.9**
Discontinued from treatment (%)	[redacted]	[redacted]	*98.6*	*90.1*
• Maximum allowed dose without PR/CR	[redacted]	[redacted]	8.8	6.4
• PR or CR for at least 6 continuous months	[redacted]	[redacted]	14.2	0
• Progressive disease	[redacted]	[redacted]	64.7	67.4
• Adverse event	[redacted]	[redacted]	3.7	2.1
• Deaths	[redacted]	[redacted]	1.7	2.1
• Consent withdrawn	[redacted]	[redacted]	3.4	8.5
• Physician decision	[redacted]	[redacted]	2.0	3.5
Discontinued from trial after receiving treatment (%)	[redacted]	[redacted]	56.9	70.2
• Lost to follow up	[redacted]	[redacted]	[redacted]	[redacted]
• Deaths	[redacted]	[redacted]	[redacted]	[redacted]
• Consent withdrawn	[redacted]	[redacted]	[redacted]	[redacted]
• Physician decision	[redacted]	[redacted]	[redacted]	[redacted]
• Other	[redacted]	[redacted]	[redacted]	[redacted]

CR=complete response; PR=partial response

Source: CS, adapted from Figure 4-4 and Table 4-7, CSR (Primary Analysis), adapted from *Table 14-1.1* and company's response to clarification letter, adapted from Table A-13 and Figure A-6

Importantly, the EMA has noted that early treatment discontinuation in the GM-CSF arm could have potentially disproportionately affected the OS results in favour of T-VEC.5 However, the EMA also states that a sensitivity analysis submitted by the company clarified that the patients who discontinued early did not affect the observed treatment difference in the ITT population for OS (draft EPAR,5 Table 32) or DRR (draft EPAR,5 Table 37).

The EMA has also highlighted that there was a higher proportion of patients with major protocol deviations in the T-VEC arm (12.2%) than in the GM-CSF arm (3.5%).5 Missing confirmatory scans were reported to be the most common protocol deviation (6.1% vs 0.7%,

4.2.6 Results from OPTiM trial

All pre-specified primary, secondary and tertiary efficacy outcomes from the OPTiM trial[4] have been reported by the company. The key results are summarised in Table 10. In both the ITT population and subgroup of patients with non-visceral disease, T-VEC is significantly more efficacious than GM-CSF for all key outcomes.

Table 10 Summary of key efficacy results in the OPTiM trial ~~(final data cut)~~

Outcome	~~Patients with each type of AE (%)~~			
	Patients with non-visceral metastatic disease		ITT population	
	T-VEC (n=163)	GM-CSF (n=86)	T-VEC (n=295)	GM-CSF (n=141)
DRR by EAC assessment (%)*	25.2	1.2	16.3	2.1
Unadjusted odds ratio (95% confidence interval)	28.6 (3.9 to 211.5)		8.9 (2.7 to 29.2)	
P-value	<0.0001		<0.0001	
ORR by EAC assessment (%)*	40.5	2.3	26.4	5.7
P-value	<0.0001		<0.0001	
Median TTF by investigator assessment (months)†	13.1	3.3	8.1	2.9
Hazard ratio (95% confidence interval)	0.28 (0.20 to 0.40)		0.43 (0.33 to 0.56)	
P-value	<0.0001		<0.0001	
Median OS (months)†	46.8	21.5	23.3	18.9
Hazard ratio (95% confidence interval)	0.56 (0.40 to 0.79)		0.79 (0.62 to 1.00)	
P-value	0.0008		0.0494	

DRR=duration of response rate; ITT=intention to treat; OS=overall survival; TTF=time to treatment failure
**Primary data-cut*
†Final data-cut
Source: CS, adapted from Table 4-13, Table 4-16, Table 4-14 and clarification response, Table A-12 (patients with non-visceral metastatic disease) and appendices to CS, adapted from Table 1-13, Table 1-15, Table 1-17 and Table 1-14 (ITT population)

Subgroup analyses of ITT population

Subgroup analyses for DRR and OS suggested that the treatment effect of T-VEC may differ according to disease stage, prior non-surgical melanoma treatment, line of therapy, presence of visceral disease, and (for DRR only) by geographic region. The p-values for the tests for interaction for these subgroup analyses are provided in appendices to this ERG report (Section 11.1).

In an exploratory post-hoc analysis of data for patients in the ITT population which was presented in the FDA briefing document,[37] a larger proportion (30.4%) of patients with a DR had only very small lesions (<1cm^2) compared to the overall population (10.1%). The FDA interpreted this to suggest that patients who had larger lesions were less likely to respond to T-VEC, although it also cautioned that the clinical meaningfulness of a response (and therefore DRR) is questioned for patients with already relatively small baseline lesions.

Subgroup of patients with non-visceral metastatic disease

In the subgroup of patients with non-visceral metastatic disease, it was noticeable that the CR rate was higher in the T-VEC arm than in the GM-CSF arm (16.6% vs 0.0%; $p < 0.001$; primary data cut). Furthermore, results of an analysis presented in the draft EPAR5 show that in patients with non-visceral metastatic disease, patients receiving ≥ second-line T-VEC also had improved DRR (17% vs 2%) and objective response (28% vs 2%) relative to GM-CSF. However the p-values for the tests for interaction for these subgroup analyses were not provided.

After treatment failure, a greater proportion of patients in the GM-CSF arm received subsequent ipilimumab, vemurafenib, dabrafenib, trametinib or an anti-PD1 antibody (such as pembrolizumab) than patients in the T-VEC arm (50% and 41% respectively in T-VEC licensed population). Ipilimumab was the most common subsequent treatment (37% of patients in both arms). Vemurafenib and anti-PD1 antibodies (such as pembrolizumab) were both more commonly given to patients who failed treatment with GM-CSF than T-VEC: 15% vs 9% (vemurafenib) and 5% vs 1% (anti-PD1 antibodies) respectively.

The annual survival rates for patients in the T-VEC licensed population were consistently higher in the T-VEC treatment group compared with the GM-CSF arm. After 3 years, the survival rate for patients in the T-VEC treatment group was 54.9% compared with a survival rate of 34.6% for patients in the GM-CSF treatment group. Moreover, the survival rate in the T-VEC arm appeared to be stable over 4 and 5 years, and the difference in long-term survival rates at 4-years between T-VEC patients and GM-CSF patients was more than 20% (48.9% vs 27.5%).

Summary of findings and ERG comment

The company states that the results from patients with non-visceral metastatic disease are in line with the results from the ITT population. The ERG notes that the magnitude of difference between arms for all endpoints is much greater in patients with non-visceral metastatic disease than in the ITT population. Given the potential risks of bias identified in Section 4.2.5, the ERG cautions that it is difficult to argue that there is a demonstrable OS benefit for T-VEC over GM-CSF in the ITT population. On the other hand, in patients with non-visceral metastatic disease, there does seem to be a demonstrable benefit; the difference in efficacy endpoints between arms is large and is unlikely to be explained by methodological bias.

Although subgroup analyses by different stages of disease (stage IIIB to IIIC, stage IV M1a, stage IV M1b and stage IV M1c) were all pre-specified, the definition of non-visceral disease used by the company (stage IIIB to stage IV M1a disease) was not a pre-specified subgroup

but rather defined post-hoc. Hence the findings for patients with non-visceral metastatic disease are derived solely from an analysis of an exploratory post-hoc subgroup. Carrying out such analyses risks identifying subgroups in which superior drug efficacy occurs only by chance. However, the ERG's primary concern is that the subgroup comprises a mixture of patients with stage III and patients with stage IV disease. This is an issue as the *differential survival trajectory* for patients with stage III disease is likely to differ from that of patients with stage IV disease.

The company did not attempt to employ the modified Korn model or the two-step Korn model to adjust the survival curves of patients receiving BRAF inhibitors. The reason given for this was that the trials included in the meta-analysis which forms the basis for the original Korn[51] model did not differentiate patients by BRAF status. The ERG concurs with the company.

The results of the two-step Korn model are more conservative than the results from implementing the modified Korn model as the two-step approach assumes that ipilimumab is more effective in patients with non-visceral metastatic melanoma than in the wider population of patients with metastatic melanoma (predominantly later stage disease). Hence, the latter is considered to generate "best case" findings and the former "worst case" findings. More information about the Korn models is presented in the appendices to this ERG report (Section 11.2).

In summary, the trial results for T-VEC are: median OS: 46.8 months; mean OS: 36.9 months; median TTF: 13.1 months; ***mean TTF: 20.6 months***; TTF is considered by the company to be a proxy for PFS. For ipilimumab, the adjusted results, as presented in the company's response to the ERG's clarification letter, are:

- Modified Korn model results for ipilimumab:
  - median OS increases from 10.9 months to 21.3 months (95% prediction interval: 14.6 months to upper interval not reached)
  - mean OS increases from 19.5 to 29.2 months (95% prediction interval: 23.8 months to 34.6 months)
  - median PFS increases from 2.8 months to 5.3 months
  - mean PFS increases from 8.0 to 15.2 months.
- Two-step Korn model results for ipilimumab:
  - median OS increases from 10.9 months to median not reached (95% prediction interval: 27.0 months to upper interval not reached)
  - mean OS increases from 18.0 to 32.3 months (95% prediction interval: 28.1 months to 35.8 months)
  - median PFS increases from 2.8 months to 17.6 months
  - mean PFS increases from 7.4 to 18.6 months.

Given the lack of clinical effectiveness evidence available, the ERG considers that the company was correct to attempt to apply alternative approaches for the comparison of T-VEC with ipilimumab. However, for reasons described in Section 5.5.1 the ERG does not consider that the use of either of the Korn models was appropriate. Therefore, the ERG does not consider the findings reported by the company when utilising the modified Korn model or the two-step Korn model to be either reliable or robust.

4.4 Safety

AE data are available for patients treated with T-VEC; these data have been previously reported for the OPTiM trial overall safety population (patients with stage IIB to stage IV M1c disease) in the published paper[4] and in the draft EPAR.[5] In the CS, the company reports only AEs for patients with non-visceral metastatic disease. Data for both populations are summarised by the ERG in Table 13 and a summary of the specific types of AEs and serious AEs (SAEs) is presented in the appendices of the ERG report (Section 11.3, Table 50).

Table 13 Summary of safety profiles of T-VEC and GM-CSF in the OPTiM trial

Type of safety concern	**Patients with each type of AE (%)**			
	Patients with non-visceral metastatic disease		**Overall safety population***	
	T-VEC (n=163)	**GM-CSF (n=76)**	**T-VEC (n=292)**	**GM-CSF (n=127)**
All cause and any Grade treatment emergent AE	99	93	99	95
All cause treatment emergent Grade 3 to 5 AEs	33	23	[redacted]†	[redacted]†
All cause and any Grade treatment emergent SAE	20	13	26	13
All cause treatment emergent Grade 3 to 5 SAEs	*18*	*10*	[redacted]†	[redacted]†
Treatment-related AEs	93	79	93	80
Treatment-related Grade 3 to 5 AEs	14	5	[redacted]†	[redacted]†
Treatment-related SAE	6	0	7	0
Treatment emergent AE leading to discontinuation	9	7	10	6
Fatal AEs on study	1	0	3	2

AE=adverse event; NR=not reported; SAE=serious adverse event
Source: CS, adapted from Table 4-32 and *draft EPAR,[5] Table 46 except † taken from CSR, Table 12-2

The ERG concurs with the company that treatment emergent AEs, SAEs and treatment-related AEs were higher in the T-VEC arm than in the GM-CSF arm. In patients with non-visceral metastatic disease, AEs leading to treatment discontinuation were reported to be similar between arms and there was only one fatal AE, in the T-VEC arm, but this was not related to treatment. The ERG notes that treatment discontinuation rates due to AEs were marginally higher in the T-VEC arm than in the GM-CSF arm in the overall safety population.

Adverse events of special interest (AEOSIs) have also been identified by the company, and feature in the risk management plan (RMP), agreed with the EMA,[5] as being important safety concerns. These AEOSIs were not fully reported in the CS. The ERG has summarised the AEOSI data in Table 14; these events include flu-like symptoms, injection site reactions and cellulitis. The draft EPAR[5] states that the majority (70% to 90%) of the flu-like symptoms were reported to resolve within 72 hours. These events were also reported more frequently within the period of the first six treatments, particularly in patients who were HSV-1 negative at baseline, due to the intratumoral injection route of administration of T-

Table 15 Adverse events reported during pivotal trials with ipilimumab, vemurafenib, dabrafenib, pembrolizumab and T-VEC

Trial/ treatment	Frequency of any treatment emergent and/or treatment-related AEs, dose discontinuations and/or modifications due to AEs (%)	Common AEs
MDX010-20[19]/ Ipilimumab (Previously treated)	Grade 3 or 4 treatment-related AEs 23 Treatment-related AEs leading to discontinuation 10	AEs were mostly immune-related which may involve the gastrointestinal, liver, skin, nervous, endocrine, ocular, or other organ systems
BRIM-3[17]/ Vemurafenib (First-line)	Grade 3 to 5 AEs 50 SAEs 33 AEs leading to treatment discontinuation 7 AEs leading to dose modification/ interruption 38	Most frequently occurring Grade 3 or 4 AEs (%): Cutaneous SCC 19 Increase in LFT 11 Keratoacanthoma 10 Rash 9 Arthralgia 6
BREAK-3[18]/ DTIC (First-line)	Grade 3 to 5 AEs 42 SAEs 23 Treatment-related SAEs 15 AEs leading to treatment discontinuation 3 AEs leading to dose reduction 18 AEs leading to dose interruption 27	Most frequently occurring Grade 3 to 5 AEs (%): Back pain 4 Hyperglycaemia 3 Pyrexia 3 GGT increased 3
KEYNOTE-006[21]/ Pembrolizumab (First-line)	Grade 3 to 5 AEs 35 Grade 3 to 5 treatment-related AEs 12 SAEs 25 Treatment-related SAEs 9 Treatment-related AEs leading to discontinuation 9	Most frequently occurring Grade 3 to 5 AEOSIs (%): Colitis 3 Hepatitis 2 Diarrhoea 1
KEYNOTE-006[21]/ Ipiliumumab (First-line)	Grade 3 to 5 AEs 37 Grade 3 to 5 treatment-related AEs 20 SAEs 30 Treatment-related SAEs 18 Treatment-related AEs leading to discontinuation 9	Most frequently occurring Grade 3 to 5 AEOSIs (%): Colitis 7 Diarrhoea 4 Hypophysitis 2
OPTiM trial[4]*/ T-VEC (Previously treated and untreated)	Grade 3 to 5 AEs 33 Grade 3 to 5 treatment-related AEs 14 SAEs 20 Treatment-related SAEs 6 *AEs leading to discontinuation* 9	Most frequently occurring Grade 3 to 5 AEs (%): Fatigue 2 Injection-site pain 1

AE=adverse event; AEOSI=adverse event of special interest; CS=company submission; GGT= Gamma-glutamyl transferase; LFT=liver function tests; SCC=squamous-cell carcinoma
*T-VEC licensed population only
Source: CS, adapted from Table 4-38 and text of pages 198 to 109 with additional data reported for BRIM-3[17] and BREAK-3[18] taken from ERG report submitted during the dabrafenib STA[53] and from the company's submission (Merck) for pembrolizumab for previously untreated ipilimumab naïve patients[54]

4.7 Conclusions of the clinical effectiveness section

The majority of evidence for the clinical effectiveness of T-VEC is derived from the OPTiM trial,[4] a relatively large (*n=436*), open-label, multi-centre, international Phase III trial which included patients from the UK (n=33 [8%]). ITT population (patients with stage IIIB to stage IV M1c disease) results show statistically significant improvements in favour of T-VEC vs GM-CSF for DRR, TTF (a proxy for PFS in this trial) and ORR but not for OS (although the OS gain was close to being statistically significant).

Findings from the OPTiM trial[4] were reported for patients with non-visceral metastatic melanoma (patients with stage IIIB to stage IV M1a disease); these patients are the focus of this appraisal as these are the patients for whom T-VEC will be licensed. Statistically significant improvements in DRR, TTF, ORR and OS were reported for patients treated with T-VEC compared with those treated with GM-CSF. The magnitude of the effect in the licensed population is much greater for all outcomes than in the ITT population. These findings were derived from an exploratory post-hoc analysis of *249* patients.

The ERG has concerns that the population considered in this STA comprises a mixture of patients with stage III and stage IV M1a disease as it is likely that the *differential survival trajectory* of patients with stage III disease differs from that of patients with stage IV disease. The ERG also considers that there are a number of potentially important sources of bias in the OPTiM trial[4] due to limited blinding, a higher proportion of drop-outs in the GM-CSF arm (particularly in the first few months of the trial), and the use of DRR as the primary endpoint. However, the ERG does not consider that the potential sources of bias explain the improvements in efficacy in the T-VEC arm compared with the GM-CSF arm reported for patients with non-visceral disease.

An area of uncertainty that has been raised by the FDA[37] relates to the size of lesions. The results of an FDA post-hoc analysis suggest that patients who had very small lesions (<1cm^2) were more likely to respond to T-VEC than the overall population (30.4% vs 10.1% respectively).[37]

In both the overall trial population and the subgroup of patients with non-visceral metastatic melanoma, there were more treatment emergent AEs, SAEs and treatment-related AEs in the T-VEC arm of the OPTiM trial[4] than in the GM-CSF arm. The types of AEs included flu-like symptoms (very common), injection site reactions (very common) and cellulitis (common and potentially serious). Careful wound care is important to minimise risk of infection, particularly if tissue necrosis results in open wounds. In terms of the types of AEs observed, T-VEC compares favourably in terms of safety to other recommended treatments

5.4 ERG's critique of the submitted economic evaluation

5.4.1 NICE reference case checklist

Table 33 NICE Reference case checklist completed by ERG

Attribute	Reference case	Does the de novo economic evaluation match the reference case?
Decision problem	The scope developed by NICE	Partial. The economic evaluation considers a subgroup of that issued in the final NICE scope in line with *the marketing authorisation for T-VEC.* The decision problem addressed in the submission is adults with unresectable melanoma that is regionally or distantly metastatic with no bone, brain, lung or other visceral disease (disease stage IIIB–stage IV M1a)
Comparator(s)	As listed in the scope developed by NICE	Partial. The company considers that BRAF inhibitors are unlikely to be treatment options for the stage IIIB to stage IV M1a population and that ipilimumab is the primary comparator
Perspective on costs	NHS and PSS	Partial. The model only includes NHS costs. Personal Social Service costs have not been considered
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Patient related direct health effects are considered. No impact on carers has been considered in the model
Form of economic evaluation	Cost–utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes – 30 year time horizon
Synthesis of evidence on outcomes	Based on systematic review	No – no connected evidence network is possible. A synthesised comparator was developed from three arms of two ipilimumab trials with adjustments to match baseline patient characteristics
Outcome measure	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults	Yes – health effects are expressed in QALYs, using utility estimates from other NICE appraisals which used the EQ-5D instrument
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	Yes, HRQoL data were collected as part of the OPTiM trial[4] but these were not suitable for utility estimation
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	Yes
Discount rate	The same annual rate for both costs and effects (currently 3.5%)	Benefits and costs have been discounted at the 3.5% rate
Equity	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	All QALYs estimated by the economic model have the same weight
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes, partially - NHS costs, valued at relevant prices, have been used. PSS costs are not included in the model

EQ-5D=EuroQol-5 dimension; PPS=personal social services; QALY=quality adjusted life year

It is generally appropriate to use K-M analysis results directly in a model prior to use of projection methods. ~~However, in this case, it appears that the final analysis of the trial data (CS, Figure 4-6) has not been used in the model, which includes only OS data from the earlier, less mature, data cut.~~ The ERG has carried out a curve-fitting exercise to a re-analysis of the final data cut which was requested during the clarification process. The ERG has found that a 2-part exponential model (Figure 9) closely follows the trial OS data from 9 months until the last recorded death (47 months).

It is noteworthy that the company model exponential trend (Phase 1b in Figure 8) deviates markedly from the final recorded trial data and leads to a clear separation from the exponential trend identified by the ERG. This results in a much more advantageous OS estimate for T-VEC compared to the long-term projection resulting from the fitted ERG curve (Figure 9).

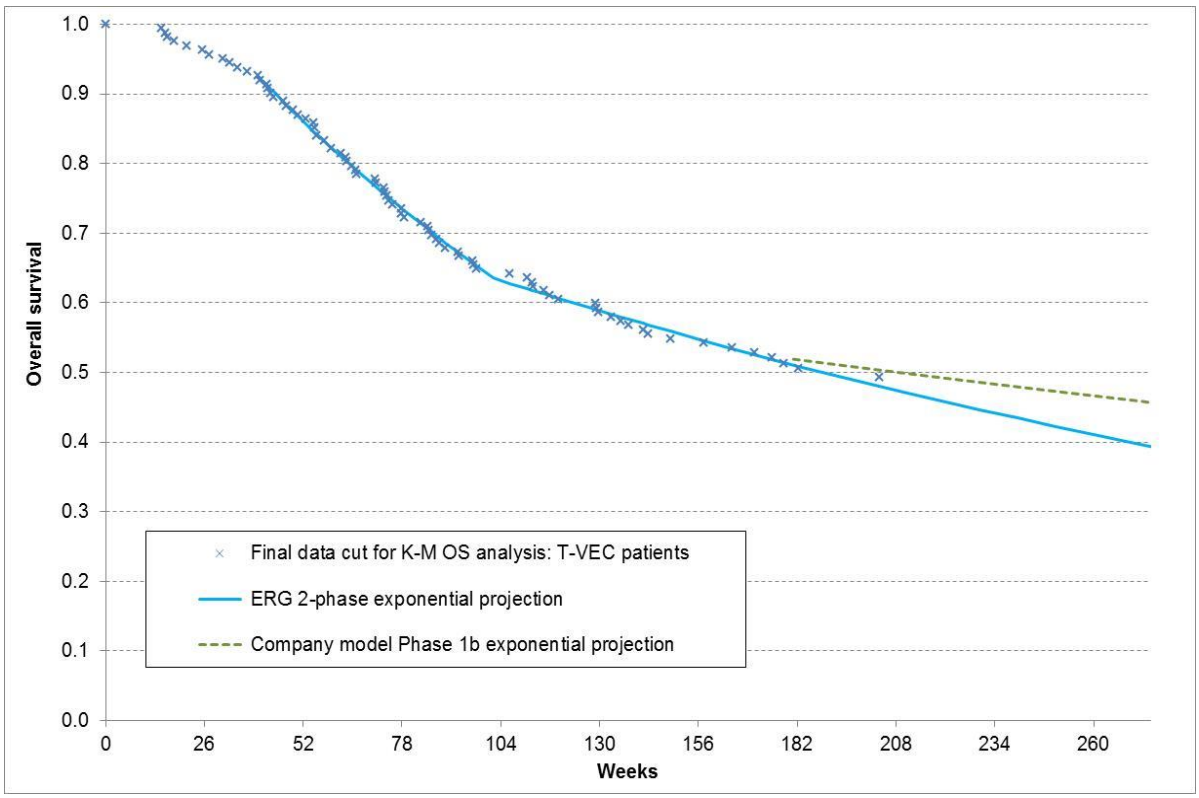

Figure 9 Comparison of company Phase 1b OS projection and ERG exploratory projection

5.7 Conclusions of the cost effectiveness section

In the absence of direct trial evidence for a clinically appropriate comparator, estimation of the relative cost effectiveness of T-VEC vs current clinical practice is rendered extremely difficult. The company's proposal for a constructed comparator, based on the pooling of data from two ipilimumab trials[19,22] adjusted for baseline characteristics and using a proportional hazard model derived from a patient population that is very different from *the T-VEC licensed population*, is considered by the ERG to be ill-conceived and unreliable as the basis for determining cost effectiveness. Moreover, due to the high degree of volatility exhibited in model-generated quantitative estimates of cost effectiveness when ERG amendments are implemented, the ERG does not consider that it is appropriate to present detailed alternative ICERs for this questionable comparison.

The ERG has also identified serious problems relating to the long-term projection of survival. These relate to the selective use of registry data and life table estimates. The company appeals to precedents from previous appraisals in melanoma to justify their approach to projecting survival. However, the ERG considers that the populations studied previously differ substantially from the target population proposed for T-VEC and from the population on which the Korn model[51] was based, so that the appeal to such precedents is not appropriate.

Had the OPTiM trial[4] included an alternative treatment arm involving a recognised alternative treatment (e.g. DTIC), then indirect evidence synthesis may have been appropriate. Unfortunately, it is not possible to determine whether GM-CSF constitutes an active or inactive comparator for T-VEC, so the data from the comparator arm of the OPTiM trial[4] can play no part in assessing the extent to which T-VEC benefits patients with non-visceral metastatic disease compared to current practice.

8 DISCUSSION

Evidence from the OPTiM trial

T-VEC *is licensed* for the treatment of patients with unresectable melanoma that is regionally or distantly metastatic (stage IIIB, stage IIIC and stage IV M1a) with no bone, brain, lung or other visceral disease. In practice, the melanoma must also be injectable. Evidence for the efficacy of T-VEC treatment in this population has been obtained from a post-hoc analysis of data from patients with non-visceral metastatic disease who took part in the OPTiM trial[4].

The ERG considers that the efficacy results for the OPTiM trial[4] ITT population (a broader patient population that also includes patients with stage IV M1b and stage IV M1c disease), all of which favour T-VEC, may be subject to bias. This is because the trial lacked blinding, employed limited central assessment and the proportion of patients dropping out of the GM-CSF arm was higher than that associated with the T-VEC arm. All of these limitations also apply to the analyses carried out on data collected from the subgroup of patients with non-visceral metastatic melanoma, with the additional concern that these analyses were not pre-specified. The ERG notes that, for this non-visceral metastatic disease subgroup, the differences in treatment effect between the two trial arms, for all efficacy outcomes, were large. This suggests that, despite the identified limitations, for these patients, the conclusion that T-VEC is a more efficacious treatment option than GM-CSF may be credible. However, the ERG has concerns relating to the validity of this subgroup as it comprises both patients with stage III and stage IV disease. This is of concern as it is likely that the *differential survival trajectory* of patients differs by stage of disease which means this is not a relatively homogeneous patient group.

In summary, results from the OPTiM trial[4] show that T-VEC is clinically superior to GM-CSF. However, GM-CSF is not used in the NHS to treat patients with melanoma and, therefore, for the purposes of this STA, is not considered to be a relevant comparator

Applicability of the OPTiM trial results to clinical practice

The ERG considers that the characteristics of patients included in the OPTiM tria,[4] with non-visceral metastatic disease, are generally similar to the patient population likely to be considered for treatment with T-VEC in clinical practice in England. In this respect, the results from the OPTiM trial[4] are generalisable to patients seen in clinical practice in England.

Results from the OPTiM trial[4] show that, for patients treated with T-VEC, measures of ORR, DRR and TTF were better in the subgroup of patients with non-visceral metastatic disease than in the whole trial arm: 40.5% vs 26.4%, 25.2% vs 16.3% and 13.1 months vs 8.1

setting. In the FDA briefing document[37] it is suggested that the overall risk-benefit profile of T-VEC shows most benefit to patients receiving first-line treatment. Furthermore, within the draft SmPC,[36] there is a caution that the efficacy data supporting the use of T-VEC in second or later line treatment settings are limited.

The lack of confidence in the efficacy of T-VEC as a second (or later) line of treatment is largely due to the fact that, during the period when the OPTiM trial[4] was conducted, first-line treatment options for patients were different from those available to such patients today. This means that the patients in the OPTiM trial[4] who received T-VEC as a second- (or later) line of treatment will be different from the patients receiving T-VEC as a second- (or later) line of treatment in clinical practice today. In addition, it is reported in the draft EPAR[5] that there is a strong correlation between line of therapy and disease stage; line of therapy was not retained as an independent predictor for DR in a multivariate analysis considering disease stage.

Treatment with T-VEC can be continued even if there is some evidence of disease progression, with a minimum of 6 months of treatment being recommended. The EMA[5] raised concern that, for some patients, next-line treatment may commence later than if an alternative to T-VEC had been administered at the time of disease progression. The ERG considers that, because injectable melanoma entails lesions that can be clearly seen by the treating clinician, unnecessary treatment delays are unlikely since, if there is evidence of rapid progression, clinicians would not delay next-line treatment in clinical practice.

Company's cost effectiveness estimates

The ERG does not consider that the cost effectiveness results presented by the company are reliable. The reasons that support this conclusion relate primarily to the clinical evidence employed within the model and the methods used in the company model to project survival.

There are four main clinical issues that cast doubt on the reliability of the company's cost effectiveness results. The first issue is whether ipilimumab is the most appropriate comparator to include in the company's baseline cost effectiveness analysis. The second and third issues relate to factors that affect patients' *differential survival trajectory*, namely (a) that the subgroup of patients with injectable non-visceral metastatic disease includes both patients with stage III and those with stage IV disease, and (b) that this subgroup includes both patients receiving T-VEC as a first-line treatment and those receiving it as a later line of treatment. The fourth issue is that the relative clinical effectiveness of T-VEC compared with any treatment currently used in clinical practice is unknown.

Table 36 List of studies included in the evidence base for the modified Korn and two-step Korn models

Study	Treatments	Patient population
OPTiM trial[4]*	T-VEC GM-CSF	Previously treated and untreated patients with *stage IIIB* to stage IV M1c disease
MDX010-20[19]	Ipilimumab monotherapy ipilimumab in combination with gp100 gp100	Previously treated patients with stage III or stage IV disease
CA184-024[22]	Ipilimumab + DTIC DTIC monotherapy	Previously untreated with stage III or stage IV disease

Bristol-Myers Squibb=Bristol Myers Squibb; DTIC=dacarbazine; GM-CSF= granulocyte-macrophage colony-stimulating factor
*The company cites the primary reference for the OPTiM trial to be a 2014 conference abstract by Kaufman et al[48]
Source: CS, adapted from Table 4-22

11.2.1 The modified Korn model

The model originally reported by Korn[51] can be used to predict OS for melanoma patients using four prognostic characteristics; gender, ECOG PS, presence of visceral metastases, and presence of brain metastases. The coefficients for the effects of these variables on relative risk were obtained using prediction models based on individual patient data from 42 Phase II studies, in 2100 patients with metastatic melanoma, and are provided in Equation 1.

Equation 1)

$$\log(\widehat{HR}) = 0.248X_{Gender=Male} + 0.436X_{ECOG=1} + 0.948X_{ECOG\geq 2} \mp 0.421X_{Visceral=YES} + 0.304X_{Brain=YES}$$

The proportion of patients with each specified characteristic are inputted into the equation in order to give the log(HR) for each treatment group.

However, the company decided that a modified Korn model, which would take elevated LDH levels into consideration as a prognostic factor, was more appropriate to adjust the data as elevated LDH levels has been found to be an important independent prognostic factor in metastatic melanoma.[52] Bristol-Myers Squibb developed such a model in their recent submission to NICE for the use of ipilimumab in previously untreated metastatic malignant melanoma.[11] The modified Korn equation with the estimated coefficients is:

Equation 2)

$$\log(\widehat{HR}) = -0.154X_{Gender=Female} - 0.400X_{ECOG=0} - 0.285X_{Visceral=NO} - 0.306X_{Brain=NO} - 0.782X_{LDH=Normal}$$