Nivolumab for adjuvant treatment of resected stage III and IV melanoma

ERRATA

This report was commissioned by the NIHR HTA Programme as project number 17/156/02

BMJ Technology Assessment Group

This document contains errata in respect of the ERG report in response to the company's factual inaccuracy check.

The table below lists the page to be replaced in the original document and the nature of the change:

Page No.	Change
17	'CheckMate' amended to 'CheckMate 238', and 'CA8184-029' amended to 'CA184-029'
18	Spelling of the word 'nivolumab' amended
22	ICERs updated after corrections to the ERG model
24	Word 'underestimated' amended to 'overestimated
25-26	Table A results updated after correction to the ERG model
43	275 amended to 177
50	275 amended to 177
91	'CA8184-029' amended to 'CA184-029'
102	The word "predominantly" added to highlight that not only interferon trials were used in the surrogacy analysis.
111	"PF" corrected to "RF"
113	Text amended to clarify that local/regional recurrence data were used.
121-122	Minor edits to text and ICERs updated after corrections to the ERG model
129	Text on utility decrements per AE amended
132	Erroneous text about inclusion of AEs removed. Text about inpatient and outpatient data from trials amended.
151-153	Results updated after corrections to the ERG model
Abbreviations: AE, adverse event; ICER, incremental cost-effectiveness ratio; RF, recurrence-free;	

regional lymph node surgical removal. The ERG's clinical experts reported that the baseline characteristics for CA184-029 appear broadly consistent with those of CheckMate 238, although the ERG notes that median age of patients in CA184-029 is slightly younger compared to CheckMate 238 and patients in CheckMate 238 included patients with a more advanced disease stage (Stage IV) that were excluded from CA184-029. In addition, approximately 20% of the patients in CA184-029 had Stage IIIa disease and Stage IIIa patients were excluded from CheckMate 238. The ERG notes that the company applied co-variate adjustments in the ITC analyses to adjust for these differences in Stage and age. There were also comparability issues in terms of the RFS definitions used in CheckMate 238 and CA184-029 although the ERG considers the resulting analyses conducted by the company mean that the only difference was that RFS was assessed by the investigator in CheckMate 238 whereas in CA184-029 it was assessed by an independent panel. The ERG considers this difference in outcome assessment is likely result in a conservative estimate of the efficacy of ipilimumab versus placebo.

The ERG also noted that there was a difference in the maximum duration of ipilimumab treatment between CheckMate 238 and CA184-029; in CheckMate a maximum of one-years treatment was allowed whereas in CA184-029 ipilimumab treatment could continue for upto three-years and approximately [redacted] of patients continued ipilimumab treatment beyond one year. The company conducted additional analyses in their clarification response using data where ipilimumab patients who were treated beyond one-year were censored at one year which the ERG considers to be a more conservative estimate of the benefit of nivolumab compared to placebo as opposed to the over-optimistic results provided by the ITT ipilimumab population.

The ERG also noted that the subsequent therapies differed between CheckMate 238 and CA184-029 and the ERG's clinical experts reported that the types of subsequent therapies given in CheckMate 238 were likely to be generally more consistent with the types used in UK clinical practice. The ERG notes that part of the reason for the differences in subsequent therapy is likely to be related to advances in clinical practice since CA184-029. The ERG also notes that due to the outcome censoring selected for the ITC analyses, these differences in subsequent therapies will have the largest impact in the analysis of OS.

In terms of the ITC, the company conducted analyses for RFS using patient level data (PLD) meta-regression and parametric survival models; and also conducted a Bucher adjusted indirect comparison. For the PLD ITC the company included covariates for Gender, age, stage and trial, with the rationale for including a trial covariate being, "it will account for all unobserved differences between trials, thus

maintaining randomisation". The ERG is unclear as to exactly what these differences addressed by the trial level covariate are as the company has described them as "unobserved". The resulting parametric curves from the log-logistic meta-regression model (that was deemed to be the best-fitting model) suggest that for the matched population (CheckMate 238 and CA184-029), nivolumab is associated with the longest RFS compared to both ipilimumab and placebo.

The Bucher ITC analyses for RFS using the full ITT populations of CheckMate 238 and CA182-029 were conducted with and without covariate adjustments for age, gender and stage. [redacted] The CA184-029 data suggests that the ipilimumab censored at one-year population has a slightly shorter RFS compared to the ITT population beyond approximately 18 months; although this analysis is likely to be biased against ipilimumab as the censored patients are likely to be those who will have the best prognosis at 1 year. The HRs estimated by the Bucher ITC for nivolumab versus placebo were numerically [redacted] higher when the ipilimumab censored at one-year data were used rather than the full ipilimumab dataset, [redacted]

The ERG requested the company conduct a re-analysis of the clinical data from CheckMate 238 and CA184-029, re-staging patients into the new AJCC 8th edition disease stages for melanoma, and an analysis by baseline PD-L1 status, although the company reported that they were unable to conduct these analyses due to insufficient PLD being available from CA184-029. The ERG nevertheless considers them both to be potential subgroups of interest.

OS for nivolumab versus routine surveillance for use in the economic model was estimated using a surrogacy analysis. The ERG requested [redacted]

A further issue relating to OS is that the subsequent treatments received in the CA184-029 trial have generally been superseded by more effective drugs such as immunotherapies in current UK clinical practice. Therefore, the more effective immunotherapies are likely to improve OS for patients after disease recurrence for those who receive routine surveillance and the relative benefit of adjuvant nivolumab may not be a great as the company's analyses suggest.

If the use of the OS data identified in the metastatic setting is robust and reliable, this approach potentially resolves the issue of subsequent treatments, or at least allows for the exploration of alternative estimates of PRS by varying the proportions of subsequent treatments in the model. The ERG conducted a scenario that used nivolumab as the subsequent therapy for all distant recurrence after routine surveillance and this increased the incremental cost-effectiveness ratio (ICER) substantially from £18,685 to £96,443 per QALY. Another scenario was conducted that also applied ipilimumab to all distant recurrence patients after adjuvant nivolumab. This ICER was also much greater than the company's base case at £31,663 per QALY; above the upper £30k per QALY threshold.

There were also some minor issues with excess resources use for imaging, whereby the majority of patients in the model were assumed to receive regular CT and PET scans of the chest and abdomen. Clinical experts suggested that it is unlikely for both to be given in UK clinical practice. The ERG assessed the impact of removing the PET scan costs from the model and this had a negligible effect on the results. A similar issue was noted with the use of both CT and MRI for the head. The ERG found the impact of this to be minimal also.

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

1.5.1 Strengths

Clinical

- The data for nivolumab is based on evidence from an international phase III, double-blind, high-quality RCT (CheckMate 238), which is closely aligned with the NICE final scope requested population, intervention and outcomes.

- The company's statistical approach was generally appropriate and well described.

The company conducted a comprehensive SLR to identify clinical effectiveness evidence of relevance to the decision problem in the NICE final scope.

- The ERG considers the validity and generalisability of the results of the ITC to be questionable based on differences in the ipilimumab treatment duration between the CheckMate 238 and CA184-029 studies (upto 1 year in CheckMate 238 and upto 3 years in CA184-029). The ERG, therefore, considers the analysis provided by the company where ipilimumab patients continuing treatment beyond one-year are censored in CA184-029 to be a more robust analysis and results in a more conservative estimate of the benefit of nivolumab compared to placebo as opposed to the over-optimistic results provided by the use of the ITT ipilimumab CA184-029 population.

- The ERG considers no suitable clinical effectiveness data were presented in the CS for the comparison of nivolumab versus routine surveillance for the outcomes of DMFS, HRQoL or AEs of treatment although data were provided for these outcomes for nivolumab versus ipilimumab.

- The ERG is concerned about the use of non-standard methods for the surrogacy analysis which was reported to be a first-step 'statistical exercise' to estimate OS for nivolumab versus routine surveillance. In addition, the surrogacy relationship was based on predominantly interferon studies which is potentially unreliable when applied to data for an immunotherapy, and used trial-level data rather than the PLD which was used in the methods which the surrogacy analysis is based on ([redacted].

- Data for nivolumab, and in particular for the outcome of OS [redacted] and ongoing nature of the CheckMate 238 study. [redacted]

- The ERG is also concerned that the subsequent therapies in CA184-029 do not reflect clinical practice in the UK. Following routine surveillance, a larger proportion of patients are likely to receive more effective subsequent immunotherapies than in the CA184-029 trial, meaning that the OS estimate for nivolumab versus routine surveillance generated from the ITC is potentially overestimated.

Economic

- The key weakness in the company's economic analysis is the lack of mature PLD OS data to inform nivolumab. This prevented the use of an indirect comparison between the CheckMate 238 and CA184-029 trials, which could have been adjusted for, including adjustments for subsequent treatments, to provide a potentially more reliable analysis.

- The changing pathway in recent years has also made the OS data from the CA184-029 trial less applicable, given that more effective treatments are now available for use as subsequent treatments for advanced melanoma. Adding to this, the CA184-029 trial also used a different treatment duration for the ipilimumab group, making the ITC between the trials potentially unreliable. However, this was explored for RFS with the use of censored data.

- The use of a surrogate relationship to estimate OS from RFS, which the company describe as being derived in an exploratory analysis, is not reliable enough to be considered for the base case analysis. The alternative structures with alternative OS data sources for subsequent treatments demonstrate the extent of the potential uncertainty in OS estimates with ICER increasing substantially, meaning that the company's base case analysis cannot be confidently relied upon for decision making.

- The company apply multiple estimates of effect using HRs, where the assessment of PH has not been appropriately assessed, such as the estimation of PRS from OS and the long-term prediction of RFS from OS. The company also apply HRs to survival models that do not support the use of PH, which is methodologically flawed.

1.6 Summary of exploratory and sensitivity analyses undertaken by the ERG

Economic

The ERG's preferred base case results are given in Table A, showing the impact of using the censored RFS ITC and changing the subsequent treatments. This base case used the company's alternative Markov Option 2 model as its foundation.

Table A. ERG base case ICER

Results per patient	Nivolumab	Routine surveillance	Incremental value
Company's alternative model (Markov Option 2)			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	14.08	[redacted]
ICER			**£18,685**
RFS using censoring at one-year of treatment continuation			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	14.19	[redacted]
ICER (compared with company ICER)			**£18,960**
ICER with all changes incorporated			**£18,960**

Nivolumab as subsequent therapy for distant recurrence after routine surveillance			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	17.05	[redacted]
ICER (compared with company ICER)			**£96,443**
ICER with all changes incorporated			**£107,787**
Ipilimumab as subsequent therapy for distant recurrence after adjuvant nivolumab			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	17.05	[redacted]
ICER (compared with company ICER)			**£10,202**
ICER with all changes incorporated			**£32,758**
Abbreviation used in the table: ICER, incremental cost-effectiveness ratio; LY, life-year; QALYs, quality-adjusted life years; RFS, recurrence-free survival.			

Other outcomes used in the economic model/specified in the scope	DMFS, determined based on the first date of distant metastasis provided by the investigator and was defined as the time between the date of randomisation and the date of first distant metastasis or death, whatever the cause.[a] AEs according to the CTCAE v4.0. Immune-mediated AEs were determined on the basis of a prespecified list of terms from the MedDRA. HRQL according to the EORTC QLQ-C30 and the EQ-5D. HRQL was assessed at baseline, Weeks 5, 7, 11, 17, 25, 37 and 49, and then at two follow-up visits.
Abbreviations: AEs, adverse events; AJCC, American Joint Committee on Cancer; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CNS, central nervous system; CT, computed tomography; CTCAE, Common Terminology Criteria for Adverse Events; DMC, data monitoring committee; DMFS, distant metastasis-free survival; ECOG, Eastern Cooperative Oncology Group; EORTC QLQ C30 European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C-30; HRQL, health-related quality of life; IVRS, interactive voice response system; MRI, magnetic resonance imaging; RFS, recurrence-free survival; VEGF, vascular endothelial growth factor; WBC, white blood cell.	

4.2.1 Trial conduct

CheckMate 238 was an international randomised, double-blind, phase III trial of nivolumab compared with ipilimumab as adjuvant therapy in patients with completely resected Stage IIIB, IIIC or IV melanoma.[1] The company reported that CheckMate 238 was conducted at 130 centres in 25 countries, including 9 sites in England and Wales.

Patients were enrolled in CheckMate 238 between 30 March 2015 and 30 November 2015 and were required to be a minimum of 15 years old and to have Stage IIIB, IIIC or IV melanoma, according to the 2009 classification of the American Joint Committee on Cancer (AJCC) 7th edition.[1] However, as discussed in Section **Error! Reference source not found.**, only patients aged 18 or over were actually enrolled into the trial and the AJCC classification has recently been updated to the 8th edition[12] which means some Stage IIIA patients were enrolled.

Patients in CheckMate 238 were randomised 1:1 to nivolumab or ipilimumab study arms using an interactive voice response system (IVRS). The randomisation was stratified according to disease stage and programmed death receptor ligand-1 (PD-L1) status. The company focused their report in the CS on the 18-month follow-up data-cut of May 2017, although some results using [redacted]-month follow-up data from [redacted] were also provided. The ERG reports only the latter data-cut, although in some instances the [redacted] data were not provided in the CS and so the May 2017 data-cut is used by the ERG where necessary. The Consolidated Standards of Reporting Trials (CONSORT) diagram provided in Appendix D of the CS indicates that, at the interim data-cut in May 2017, a total of 906 patients had undergone randomisation resulting in 453 patients in each of the nivolumab and ipilimumab groups. As discussed in Section **Error! Reference source not found.**, there was an imbalance in the reasons for discontinuation from study drug between the two study arms. Based on the May 2017 data-cut, 177 patients in the nivolumab arm had discontinued study drug with the most common reasons for discontinuation being disease recurrence (121 patients) and study drug toxicity (41 patients). In the ipilimumab arm, 331 patients discontinued with the most common reason being study drug toxicity (208 patients) and the second most common reason was disease recurrence (101 patients).

4.2.4 Summary statement

In summary, CheckMate 238, an international randomised, double-blind, phase III trial of nivolumab compared with ipilimumab as adjuvant therapy in patients with completely resected Stage IIIB, IIIC or IV melanoma provided the clinical effectiveness evidence for adjuvant nivolumab in the CS.[1] CheckMate 238 was conducted at 130 centres in 25 countries, including 9 sites in England and Wales. The ERG considers CheckMate 238 addressed the population, intervention and outcomes requested in the NICE final scope although it was for the comparison of nivolumab versus ipilimumab and not the required comparison of nivolumab versus routine surveillance. However, the company also conducted an ITC to enable estimates of nivolumab versus routine surveillance to be presented in the CS (discussed in Section **Error! Reference source not found.**).

A total of 906 patients were randomised in CheckMate 238, resulting in 453 patients in each of the nivolumab and ipilimumab groups; study drug treatment in both groups was continued up to a maximum of one year. The ERG notes that the median treatment duration was [redacted] for the nivolumab group compared to [redacted] for the ipilimumab group.[22] Based on a May 2017 data-cut, 177 patients in the nivolumab arm had discontinued study drug with the most common reason for discontinuation being disease recurrence (121 patients). In contrast, in the ipilimumab arm, 331 patients had discontinued by the May 2017 data-cut with the most common reason being study drug toxicity (208 patients). The ERG's clinical experts reported that these differences in reasons for study drug discontinuations were not unexpected given the known toxicity profile of ipilimumab.

The baseline characteristics for each trial arm in CheckMate 238 appear to be well balanced and the company response to clarification suggests that approximately 50% of randomised patients in CheckMate 238 were from Western Europe sites although less than 10% of the randomised patients were from UK sites. The ERG's clinical experts reported that the baseline characteristics of patients in CheckMate 238 were generally in keeping with those expected of the equivalent patients in the UK although the ERG notes that only Stage IIIB, IIIC or IV melanoma patients were enrolled in CheckMate 238 when defined using the AJCC 7th edition.[1] However, the AJCC classification has recently been updated to the 8th edition[12] which means some Stage IIIA patients were also enrolled.

At the [redacted] data-cut, a total of [redacted] patients ([redacted]) from the nivolumab group and [redacted] patients ([redacted]) from the ipilimumab group had received subsequent melanoma anti-cancer therapies. The ERG's clinical experts reported that they were unsure as to what treatments would routinely be used following disease progression in patients who have received adjuvant nivolumab, as adjuvant therapy is not currently given in the UK. However, they considered that the subsequent therapies used in CheckMate 238 were reasonable. For the outcomes of RFS and DMFS, patients who received

approximately [redacted] of patients continued ipilimumab treatment beyond one year). The company conducted additional analyses in their clarification response using data where ipilimumab patients who were treated beyond one-year were censored at one year, which the ERG considers to be a more conservative estimate of the benefit of nivolumab compared to placebo, as opposed to the over-optimistic results provided by using the ITT ipilimumab population of CA184-029.

- The ERG also noted that the subsequent therapies differed between CheckMate 238 and CA184-029 and the ERG's clinical experts reported that the types of subsequent therapies given in CheckMate 238 were likely to be generally more consistent with the types used in UK clinical practice. The ERG notes that part of the reason for the differences in subsequent therapy is likely to be related to advances in clinical practice since CA184-029. The ERG also notes that due to the outcome censoring selected for the ITC analyses, these differences in subsequent therapies will have the largest impact in the analysis of OS.

- The company conducted an ITC analysis for RFS using PLD meta-regression and parametric survival models. For the PLD ITC the company included covariates for gender, age, stage and trial, with the rationale for including a trial covariate being, "it will account for all unobserved differences between trials, thus maintaining randomisation". The ERG is unclear as to exactly what these differences addressed by the trial level covariate are. The resulting parametric curves from the log-logistic meta-regression model (that was deemed to be the best-fitting model) suggest that for the matched population (CheckMate 238 and CA184-029), nivolumab is associated with the longest RFS compared to both ipilimumab and placebo.

- The company also conducted a Bucher adjusted indirect comparison for RFS using the full ITT populations of CheckMate 238 and CA182-029. This was conducted with and without covariate adjustments for age, gender and stage. [redacted] The HRs estimated by the Bucher RFS ITC for nivolumab versus placebo were numerically [redacted] higher when the ipilimumab censored at one-year data were used rather than the full ITT ipilimumab dataset, [redacted].

- The ERG requested the company conduct a re-analysis of the clinical data from CheckMate 238 and CA184-029, re-staging patients into the new AJCC 8th edition disease stages for melanoma, and an analysis by baseline PD-L1 status although the company reported that they were unable to conduct these analyses due to insufficient PLD. The ERG nevertheless considers them both to be potential subgroups of interest.

between the two is the proportion in the PR state. Appropriate costs and utility values are applied in each health state, which are described in Section 5.4.8 and Section 5.4.7, respectively.

This approach is a simple application of the key outcome data relating to disease-free survival and mortality that is often collected in cancer drug trials. For this reason, it is a common approach taken to model the cost effectiveness of cancer drugs for NICE technology appraisals, and generally considered appropriate.

For this appraisal, RFS was informed by an indirect treatment comparison (ITC) between the CA184-029 trial, which compared ipilimumab with placebo, and the CheckMate 238 trial, which compared nivolumab with ipilimumab. The ipilimumab groups of the two trials provides the link to indirectly form the desired comparison of nivolumab and placebo.[1, 21] This is discussed further in Section 4.4.

The proportion of patients in the death state at any given cycle was informed by a surrogate relationship between RFS and OS that had previously been estimated using predominantly interferon trials in the adjuvant setting. The implementation of this and the estimation of long term treatment effects is discussed further in Section 5.4.5.

5.4.4.2 Markov model (alternative scenario analyses)

The company's economic model also includes an alternative Markov structure, which was used to test structural uncertainty and to provide a range of scenario analyses that enabled the modelling of alternate assumptions for long term treatment effectiveness. The Markov structure has the same health states and the same time horizon as the PSM but the key differences lie in application of the effectiveness data.

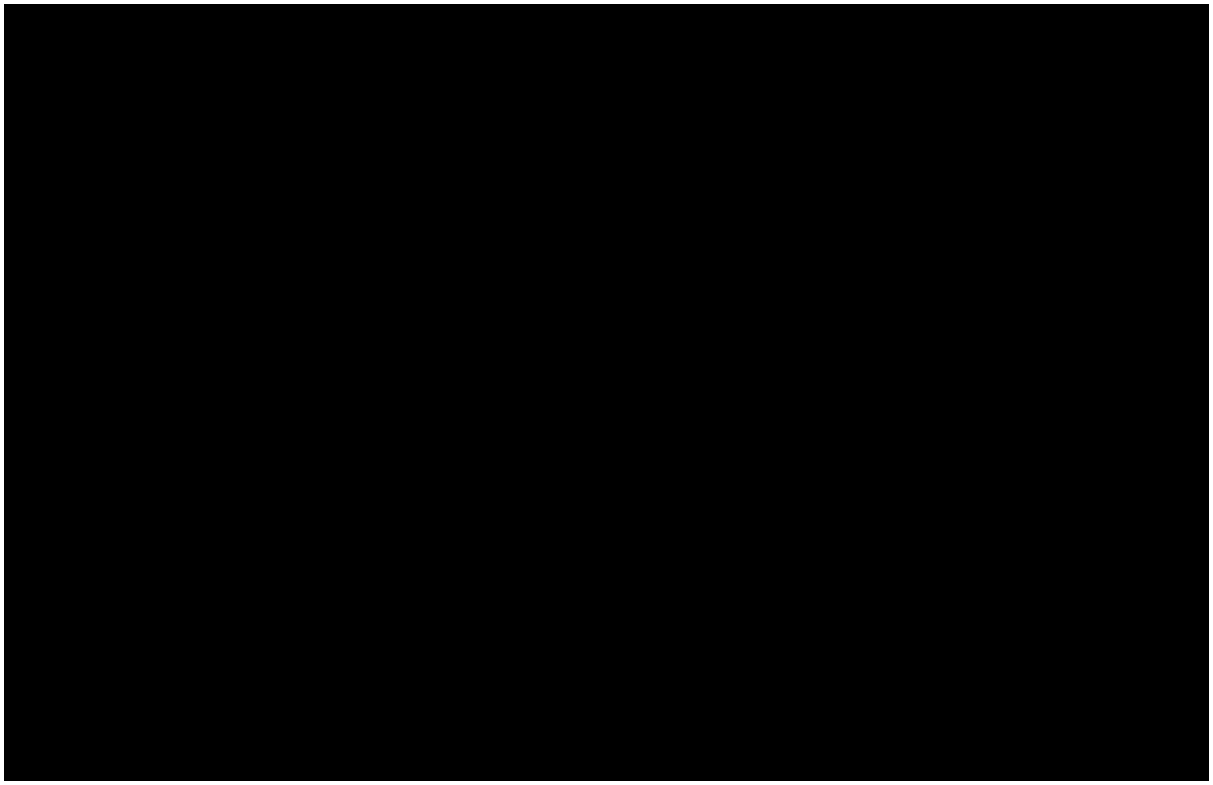
In contrast to the PSM, the Markov model relies on transition probabilities between each of the states, which are applied at each cycle to determine the proportion of patients in a particular health state at a particular time. This structure allows for alternative data sources to inform post-progression survival, which, in a PSM, would be inherently determined for the time horizon of the model by survival models used to inform the analysis.

5.4.4.3 ERG critique

The ERG considers the company's PSM and Markov structures used in the base case analysis and scenarios to be suitable structures for the decision problem. The model is generally well constructed and the ERG did not identify any errors in the functioning of the model.

The model contains relevant health states to capture the key changes in the natural history of the disease; namely, recurrence-free, disease recurrence and death. The time horizon is long at 60 years but the ERG considers this to be appropriate given that patients as young as 18 are included in the population. The

[redacted]



[redacted]5.4.5.2 Markov model (Option 1)

The Markov model uses the same RFS modelling from the ITC as per the PSM. However, to estimate the probability of remaining in the RF state and the probability of transitioning to death from the RF state, the composite RFS measure needed to be "split" to separate out the rates of recurrence and the rates of death. This was approximated by weighting the RFS curves by the proportion of patients from the CheckMate 238 trial who had disease recurrence and who died, from those patients who had experienced an event. This split was used up to 10 years after which the proportion of patients used to determine the weights was taken from the Agarwala *et al*. 2017 trial; a longer-term trial comparing interferon with routine surveillance in the adjuvant setting.[46]

Post-recurrence survival (PRS) uses the same OS data as in the PSM. To estimate PRS transition probabilities from the OS data, Cox PH models were fitted to the data in the CA184-029 trial to estimate treatment specific HRs for PRS compared with OS. The resulting HRs, which were applied to the OS modelling used in the PSM, were [redacted] and [redacted] for ipilimumab and placebo, respectively, which were used to estimate PRS for nivolumab and routine surveillance, respectively.

The resulting health-state partitions for the company's Markov scenario (Option 1), demonstrating the proportions of patients in each health-state across the time horizon of the model, are depicted in Figure 27 and Figure 28, for nivolumab and routine surveillance, respectively.

[redacted]

[redacted]

5.4.5.3 Markov model (Option 2)

The company also provided a second Markov option in the model, which used the same approach as the first option in terms of RFS, but had differences to the approach for estimating PRS.

The OS estimates for this modelling approach were derived from numerous data sources. For patients with a local/regional recurrence, the survival curves that were fitted to the local/regional recurrence data from the CA184-029 trial were used. For patients with distant recurrence, a range of data sources, including Kaplan–Meier (KM) data from drug trials for advanced and/or metastatic melanoma, and registry data, were used to fit survival curves, which were then weighted to produce estimates expected to be reflective of the relevant population.

These curves were weighted by the subsequent treatments as per the treatments assessed in each of the data sources, and were assumed to apply up to 10 years. Beyond 10 years, registry data were used to estimate the proportion of patients alive at each cycle, and again, general population mortality data were used to impose a minimum mortality rate.

For patients in either the CheckMate 238 or CA184-029 trial who received nivolumab, ipilimumab or a combination of the two as a subsequent therapy, PRS was determined by the OS in the relevant group of the CheckMate 067 trial.[49]

5.4.5.4.2 Markov structure (ERG critique)

The first modelling option within the company's Markov structure (Option 1) provides an alternative to the surrogacy relationship to estimate OS; however, this alternative brings with it, different issues that do not necessarily mitigate the uncertainty in the results.

This Markov model structure avoids the requirement for the surrogate relationship to predict the OS benefits based on the RFS benefits, by using alternative data sources to estimate PRS transition probabilities. However, a key concern that the ERG has is with the application data used to inform PRS and the assumptions made in the application.

In particular, the ERG is uncertain that the assumption of a constant relative effect on the OS hazard, regardless of the applicability in the Agarwala *et al.* 2017 study, may not be a reliable measure of estimating PRS.[46] However, the company also provide a scenario analysis using the CA184-029 trial, which may be preferable given that it is a trial assessing an immunotherapy in the adjuvant setting rather than interferon. This modelling approach does not, however, resolve the issue of inappropriate subsequent therapies influencing OS outcomes.

The alternative modelling approach provided by the company using the Markov structure (Option 2) allows for the issues of inappropriate subsequent treatments to be explored. The data sources used are from various trials of different drugs used in the post-recurrence setting, meaning that both of these issues can be explored by applying a suitable proportion of patients to the subsequent treatments as used in UK clinical practice. The validity of applying the OS data from these alternative sources also needs to be considered, which is discussed later in this section.

The ERG sought clinical expert opinion to inform the expected proportions of subsequent therapies in order to assess the potential impact on OS and on the ICER. The experts suggested that a greater use of immunotherapies such as pembrolizumab and ipilimumab would be used in clinical practice. A key point that the experts raised is that nivolumab would be used in a metastatic setting following routine surveillance, and may in fact have the same efficacy as using nivolumab in the adjuvant setting. The ERG conducted a scenario using the Markov Option 2 model, which set all patients in the routine surveillance group who had a distant recurrence to receive subsequent nivolumab. All other subsequent treatments were kept the same. This increased the ICER from £18,685 per QALY to £96,443 per QALY.

The ERG emphasises the fact that this analysis uses a range of potentially disparate sources of evidence to inform PRS, so it is unlikely that the estimates of PRS and the applicability to the population on which the ITC was formed is robust and reliable. However, even if the analysis was considered reliable, the range of ICERs resulting from plausible scenarios demonstrates the potentially serious uncertainty

that currently exists within the results. The ERG considers it difficult to fully account for this uncertainty without an OS ITC using the CheckMate 238 trial data, with appropriate adjustments for subsequent treatments.

A scenario assuming all patients with a distant recurrence in the routine surveillance group receive subsequent nivolumab using data from the metastatic NMA with fractional polynomial models, resulted in an ICER of £324,108 per QALY. This merely acts to reiterate the uncertainty, as the applicability of these data is not certain, nor has it fully been justified by the company. Further to this, the company applied the underlying time-varying HRs to an exponential model for dacarbazine. This is methodologically flawed as it breaks the transitivity between survival models in the NMA and, hence, breaks the applicability of the relative treatment effect. The company should have instead fitted a fractional polynomial with parameters P1=0 and P2=1 to the dacarbazine data in order to retain the transitivity of the treatment effect across the network.

5.4.5.4.3 Summary of critique

Overall, the ERG considers there to be a large degree of uncertainty remaining in the analysis, although it appears difficult to mitigate this uncertainty without the availability of PLD from the CheckMate 238 trial to form an adjusted ITC for OS. Based on the data available, the ERG considers it preferable to assume OS for nivolumab is equivalent to ipilimumab for the PSM, and for the Markov Option 2 structure, to apply subsequent therapy costs and survival outcomes reflective of treatments expected to be received in the UK. The latter will form the ERG's preferred base case. The ERG acknowledges that this is still a very uncertainty analysis and only partially mitigates the uncertainty in the company's analysis. However, the ERG considers the company's analyses to be no less certain than the ERG's scenario that resulted in an ICER of £324,108 per QALY, hence, emphasising the potential impact of the uncertainty.

5.4.6 Adverse events

Adverse events (AEs) were included in the economic analysis to account for resources required to management those that occur in patients who receive nivolumab. The company included immune-related AEs of any grade, diarrhoea of grade 2 or above, and any other AE of grade 3 or above.

The company noted that previous metastatic melanoma NICE submissions of immunotherapies had applied a similar approach, using endocrine disorders instead of immune-related AEs. The company considered immune-related AEs to be more appropriate given that the comparator is routine surveillance, making the broader inclusion of AEs a potentially more important factor.

For nivolumab, data were taken directly from the all-cause data in the CheckMate 238 trial. For the proportions of immune-related AEs and diarrhoea in the routine surveillance group, the risk of AEs

trial.[60] They are applied as a one-off impact at the start of the model, with immune-related AEs and other AEs weighted by inpatient visits to apply the appropriate toxicity decrement.

The company acknowledge that there may be some double counting with adverse events experienced in CheckMate 238, but they consider the impact likely to be low given that the utilities are assumed to be equal across treatment groups. A scenario analysis was also conducted to assess the impact of potential double counting and this is provided in Section 5.5.

5.4.7.4 ERG critique

The ERG considers the company's approach to utility estimation within the economic analysis to be generally sound. The company appropriately applied EQ-5D-3L values from the trial informing the treatment effectiveness where data were available, i.e. CheckMate 238 trial for nivolumab. Where evidence was not available, i.e. for routine surveillance, the company's approach to estimating utilities by applying a mapping algorithm was suitable and they provided a clear justification for the selection of the mapping study used. The company also considered the use of published utilities from other studies to assess the potential uncertainty in this data. The ERG considers the company's approach to be thorough and appropriate.

However, for the impact of AEs, the ERG considers inclusion of AE decrements using an external source to be unnecessary. The company apply the same utilities per treatment arm and same utility decrements per AE to each treatment group so there are small differences between treatments. Further to this, the impact of AEs would already be captured within the trial data itself, so the purpose of applying these utility decrements is unclear. The ERG considers a more plausible approach would be to attempt to remove the impact of AEs from the health-state utility values (HSUVs) by adding the decrements to the utility values for the routine surveillance group. However, the ERG considers this to be unlikely to affect the results of the cost-effectiveness analysis.

5.4.8 Resources and costs

The company conducted a SLR to identify evidence relating to resource use and costs to be unnecessary, given that they had access to PLD for both the CheckMate 238 and CA184-029 trials that inform the treatment effectiveness within the economic analysis. The company also considered that the treatment pathways in the metastatic setting have changed considerably in recent years making it unlikely that relevant and up-to-date evidence would be publicly available.

5.4.8.1 Drug costs

The unit cost of nivolumab per vial based on the list price and the current agreed patient access scheme (PAS) are given in Table 38. There are two vial sizes available; 4ml and 10ml, both with a concentration of 10mg/ml, providing a dose per vial of 40mg and 100mg, respectively.

5.4.8.3 Resource use for adverse reactions

As discussed in Section 5.4.6, the company included costs associated with any grade of immune-related disorders, diarrhoea grade ≥2, and other AEs grade ≥3. For nivolumab, this data was taken from the CheckMate 238 trial, which also collected data for the number of hospitalisations required for treatment of AEs. For routine surveillance, AEs were based on the relative difference between the ipilimumab and placebo groups of the CA184-029 trial and the ipilimumab and nivolumab groups of the CheckMate 238 trial, with the exception of "other" AEs, which were assumed to be the same as nivolumab in the base case analysis because of an implausible difference making AEs higher for the placebo group.

The model captures the different resources expected to be incurred by inpatient and outpatients. Costs for inpatients were calculated using the weighted non-elective excess bed day unit costs for endocrine disorders, and for other AEs, the total HRG excess bed day cost was used; both from NHS reference costs 2016-17. These costs are summarised in Table 41.

Table 41. Adverse event inpatient costs (CS, page 157, Table 45)

Treatment	Hospital cost (£)	Type of stay	Reference
Hospital bed day (immune-related)	£297.41	Non-elective excess bed days	Weighted average between KA08A, DZ29H and FD01C- NHS reference costs 2016/17[64]
Hospital bed day (other AEs)	£305.85	Total HRGs - Non-elective inpatients	Excess bed days - NHS reference costs 2016/17[64]
Abbreviations in table: AE, adverse events; HRG, healthcare resource group; NHS, National Health Service. **Note:** Endocrine disorders used as costs for immune-related disorders.			

The proportion of patients treated for AEs in an inpatient setting was recorded in CheckMate 238; however, the number of hospitalisation days was not recorded. The proportion of patients treated for AEs in an outpatient setting was not recorded in either the CA184-029 or CheckMate 238 trials. To inform this, the company used a study by Oxford Outcomes,[66] which was designed to estimate resource use associated with advanced melanoma in the UK, Italy, Sweden, Spain and Portugal. The UK resource estimates were used to inform the model and these were inflated to 2016-17 prices using the PSSRU inflation indices.[65]

Table 42. Adverse event outpatient costs (CS, page 158, Table 46)

Outpatients and unit costs	Value	Reference
% Treated as outpatient (immune-related disorders)	24.2%	Oxford Outcomes. Table 91
Unit outpatient cost (immune-related disorders)	£428.08	Oxford Outcomes. Table 17
% Treated as outpatient (diarrhoea)	19.2%	Oxford Outcomes. Table 91
Unit outpatient cost (diarrhoea)	£649.85	Oxford Outcomes. [66 27] Table 17
% Treated as outpatient (other Grade 3+ AEs)	21.7%	Oxford Outcomes. Table 91
Unit outpatient cost (other Grade 3+ AEs)	£403.68	Oxford Outcomes. Table 16/17
Abbreviations in table: AE, adverse events.		

The results of these scenario analyses are given in Table 66 along with the company's results of their preferred analysis for the Markov Option 2 model.

Table 66. Scenario analyses for Markov Option 2 model

	Results per patient	NIvolumab (1)	Routine surveillance (2)	Incremental value (2-1)
0	**Company's Markov Option 2**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	14.08	[redacted]
	ICER			**£18,685**
1	**RFS using censoring ipilimumab at one-year**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	14.19	[redacted]
	ICER (compared with base case)			**£18,960**
2	**Nivolumab as subsequent therapy for distant recurrence after routine surveillance**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	16.89	[redacted]
	ICER (compared with base case)			**£96,443**
3	**Nivolumab after routine surveillance; ipilimumab after adjuvant nivolumab (distant recurrence only)**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	16.89	[redacted]
	ICER (compared with base case)			**£31,663**
	Abbreviation used in the table: ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALYs, quality-adjusted life years.			

6.3 ERG base case ICER

The ERG's preferred base case is based on the company's alternative model; the Markov Option 2. The ERG made three key changes to the company's preferred assumptions for this model, outlined in the following bullets:

- RFS based on the ITC analysis that used censoring for patients who received treatment beyond one year in the ipilimumab group of the CA184-029 trial;
- nivolumab applied as subsequent therapy for patients with a distant recurrence after routine surveillance;
- ipilimumab applied as subsequent therapy for patients with a distant recurrence after adjuvant nivolumab.

Although subsequent therapies can vary depending on the patient and the provider, the ERG considered a simplistic approach to assume that all patients within a particular treatment group have the same subsequent treatments following a distant recurrence.

The chosen therapies are both recommended in the metastatic setting by NICE and, therefore, represent a cost-effective use of resource. Clinical experts suggested that nivolumab is the most effective and should be the first choice; however, they were uncertain as to whether it should be used following adjuvant nivolumab. The clinicians considered ipilimumab to be an appropriate immunotherapy that could be used after adjuvant nivolumab as it is a different class of drug.

The results of the ERG's preferred base case ICER are given in Table 67.

Table 67. ERG base case ICER

Results per patient	Nivolumab	Routine surveillance	Incremental value
Company's alternative model (Markov Option 2)			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	14.08	[redacted]
ICER			**£18,685**
RFS using censoring at one-year of treatment continuation			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	14.19	[redacted]
ICER (compared with company ICER)			**£18,960**
ICER with all changes incorporated			**£18,960**

Nivolumab as subsequent therapy for distant recurrence after routine surveillance			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	17.05	[redacted]
ICER (compared with company ICER)			**£96,443**
ICER with all changes incorporated			**£107,787**
Ipilimumab as subsequent therapy for distant recurrence after adjuvant nivolumab			
Total costs (£)	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]
Lys	[redacted]	17.05	[redacted]
ICER (compared with company ICER)			**£10,202**
ICER with all changes incorporated			**£32,758**
Abbreviation used in the table: ICER, incremental cost-effectiveness ratio; LY, life-year; QALYs, quality-adjusted life years; RFS, recurrence-free survival.			

6.3.1 Scenarios using ERG's preferred base case

The ERG conducted a range of scenario analyses using the ERG's preferred base case as a basis. The results are presented in Table 68.

Table 68. Scenario analyses using ERG's preferred base case

	Results per patient	Nivolumab (1)	Routine surveillance (2)	Incremental value (2-1)
0	**ERG's preferred base case**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	17.05	[redacted]
	ICER			**£32,758**
1	**50% of patients with distant recurrence in each group receive dabrafenib+trametinib**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	14.21	[redacted]
	ICER (compared with base case)			**£15,245**
2	**All patients with distant recurrence in the nivolumab group receive dabrafenib+trametinib**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	17.05	[redacted]
	ICER (compared with base case)			**£238,154**
3	**Using metastatic fractional polynomial-based NMA to inform PRS**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	Lys	[redacted]	15.16	[redacted]
	ICER (compared with base case)			**£34,354**
	Abbreviation used in the table: ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; PRS, post-recurrence survival; QALYs, quality-adjusted life years.			