CONFIDENTIAL UNTIL PUBLISHED

Evidence Review Group Report commissioned by the NIHR HTA Programme on behalf of NICE

Cobimetinib in combination with vemurafenib for treating advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma

Produced by Southampton Health Technology Assessments Centre (SHTAC)

Authors Keith Cooper, Senior Research Fellow, SHTAC
Petra Harris, Research Fellow, SHTAC
Geoff Frampton, Senior Research Fellow, SHTAC
Christian E.H. Boehler, Independent Consultant
Jonathan Shepherd, Principal Research Fellow, SHTAC

Correspondence to Jonathan Shepherd
Southampton Health Technology Assessments Centre (SHTAC)
University of Southampton
First Floor, Epsilon House
Enterprise Road, Southampton Science park
Southampton SO16 7NS
www.southampton.ac.uk/shtac

Date completed 6th May 2016

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 15/06/08

Declared competing interests of the authors

None

Acknowledgements

We would like to thank Dr Paul Lorigan (Reader in Medical Oncology, Department of Medical Oncology, Christie Hospital NHS Trust, UK) and Professor Christian Ottensmeier (Professor of Experimental Medicine, University of Southampton) for providing clinical advice during the production of this report, and for commenting on a draft of the report.

We also thank: Karen Welch, Information Specialist, SHTAC, for appraising the literature search strategies in the company's submission, and running an update of the company's literature searches; and Professor Joanne Lord, Director, SHTAC, for acting as internal editor for the draft report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Cooper K, Harris P, Frampton G, Boehler C, Shepherd J. Cobimetinib in combination with vemurafenib for treating advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma: A Single Technology Appraisal. Southampton Health Technology Assessments Centre (SHTAC), 2016.

Contributions of authors

Keith Cooper (Senior Research Fellow) critically appraised the health economic systematic review, critically appraised the economic evaluation and drafted the report. Petra Harris (Research Fellow) critically appraised the clinical effectiveness systematic review and drafted the report. Geoff Frampton (Senior Research Fellow) critically appraised the clinical effectiveness systematic review and drafted the report. Christian Boehler (Independent

Consultant) critically appraised the economic evaluation and drafted the report. Jonathan Shepherd (Principal Research Fellow) critically appraised the clinical effectiveness systematic review, drafted the report, project managed the review and is the project guarantor.

Word count: 33,461

Commercial in confidence (CIC) information in blue
Academic in confidence (AIC) information in yellow

TABLE OF CONTENTS

The ERG conducted the following scenario analyses: 14
1 Introduction to ERG Report 16
2 BACKGROUND 16
2.1 Critique of the company's description of the underlying health problem 16
2.2 Critique of the company's overview of current service provision 17
2.3 Critique of company's definition of decision problem 18
3 CLINICAL EFFECTIVENESS 21
3.1 Critique of the company's approach to their systematic reviews 21
3.2 Summary statement of company's approach to evidence synthesis 45
3.3 Summary of submitted evidence 47
3.4 Overall summary of clinical effectiveness 64
4 COST EFFECTIVENESS 64
4.1 Overview of company's economic evaluation 64
4.2 Company's review of published economic evaluations 64
4.3 Critical appraisal of the company's submitted economic evaluation 65
4.4 Additional work undertaken by the ERG 93
4.5 Overall summary of cost effectiveness 100
5 END OF LIFE 101
6 INNOVATION 102
7 DISCUSSION 102
7.1 Summary of clinical effectiveness issues 102
7.2 Summary of cost effectiveness issues 103
8 REFERENCES 104

LIST OF TABLES

Table 1 NICE approved therapies for the treatment of stage IV melanoma 18
Table 2 Overview of baseline characteristics in the coBRIM trial 26
Table 3 Ongoing trials 27
Table 4: Company and ERG assessment of coBRIM trial quality 28
Table 5 Summary of the trials included in the company NMA (adapted from CS Table 20) 36
Table 6 ERG appraisal of NMA 36
Table 7 Quality assessment (CRD criteria) of CS review 45
Table 8 PFS: primary, secondary and updated analyses 47
Table 9 Results of sensitivity analyses on PFS 48
Table 10 OS: interim and final analyses 50
Table 11 Response rates outcomes: primary and updated analyses 52
Table 12 Most frequent adverse events (occurring in ≥20% of patients in either group) 58
Table 13 SAEs occurring in ≥1% of patients in either arm (safety evaluable population) (updated CSR) 59
Table 14 Grade ≥3 adverse events which differed ≥2% between arms 60
Table 15 Primary reasons patients discontinued from study treatment 61
Table 16 Adverse events leading to discontinuation or dose modification 62
Table 17 NICE reference case requirements 65

Table 18 Summary of utility values for the company's economic model (adapted from CS Table 36) 76
Table 19 Comparison of utility values from prior NICE technology appraisals relevant to scope population (CS Table 35) 79
Table 20 Weekly vemurafenib + cobimetinib drug costs at according to average dose taken in coBRIM study & weekly dabrafenib cost based on label doses. 83
Table 21 List of adverse events and summary of costs in the economic model 85
Table 22 Deterministic base-case results, (list prices), for direct treatment comparison (actual TOT used) (CS Table 46) 87
Table 23 Deterministic base-case results, (list prices) for indirect treatment comparison (PFS as surrogate for TOT) (CS Table 47) 87
Table 24 Mean results of PSA compared to base case (list prices) for vemurafenib + cobimetinib vs. vemurafenib (CS Table 57) 88
Table 25 Mean results of PSA compared to base case (list prices) for vemurafenib + cobimetinib vs. dabrafenib (CS Table 58) 88
Table 26 Deterministic sensitivity analyses (CS Table 59) (List prices) 90
Table 27 Scenario analysis results for vemurafenib + cobimetinib vs vemurafenib or dabrafenib (List prices) 91
Table 28 Comparison of results patients treated with vemurafenib in the CS with 93
Table 29 ERG deterministic analyses using 95% CIs/CrIs (using list prices) 93
Table 30 ERG scenario analyses (using list prices) 94
Table 31 Base case specification 98
Table 32 ERG analysis: Fully incremental analysis (using list prices) 100

LIST OF FIGURES

Figure 1 NICE Melanoma Pathway for the management of stage IV melanoma (2016)[3] 17
Figure 2 PRISMA diagram for systematic literature review of RCTs (search cut-off date: 8th or 9th September 2015; based on CS figure 2, page 40) 25
Figure 3 Illustration of the network meta-analysis presented in the company submission (reproduced from CS Figure 11). 35
Figure 4 Illustration of the network meta-analysis presented in the company submission showing closed loop (CS Figure 14) 41
Figure 5 Clinically meaningful improvement in HRQoL (QLQ-C30 score change ≥10 points relative to baseline) 54
Figure 6 Mean EQ-5D-3L utility estimates in the coBRIM trial ..... **Error! Bookmark not defined.**
Figure 7 Schematic of the company's economic model (CS Figure 15) 67
Figure 8 PFS observed data from coBRIM trial and the Log-logistic fitted curves used in the economic model (CS Figure 16) 70
Figure 9 OS for the economic model and Puzanov et al study for patients treated with vemurafenib 72
Figure 10 OS observed data from coBRIM trial and the Lognormal fitted curves used in the economic model (CS Figure 23) ..... **Error! Bookmark not defined.**

Figure 11 TOT for the economic model using the company base case parametric curves (CS Figure 26)..... 75
Figure 12 TOT for the economic model using the KM log-logistic parametric curve for the vemurafenib + cobimetinib treatment arm 75
Figure 13 Scatterplot of PSA results for cost effectiveness plane (CS Figure 32)..... 89

LIST OF ABBREVIATIONS

AE	Adverse events
AFT	Accelerated failure time
AIC	Academic in confidence
AkIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CI	Confidence interval
CIC	Commercial in confidence
Cob	Cobimetinib
CR	Complete response
CRL	Credible interval
CS	Company's submission
CSR	Clinical study report
DOR	Duration of response
DIC	Deviance Information Criterion
DTIC	Dacarbazine
ECOG	Eastern Cooperative Oncology Group
EORTC-QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30
EQ-5D	EuroQoL five dimension questionnaire
ERG	Evidence Review Group
HRQoL	Health-related quality of life
HR	Hazard ratio
HRG	Healthcare Resource Groups
ICER	Incremental cost-effectiveness ratio
ITT	Intention-to-treat
LDH	Lactate dehydrogenase
MAPK	Mitogen-activated protein kinase
MEK	MAP extracellular signal-regulated kinases
mITT	Modified Intention-to-Treat
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
ORR	Objective response rate
OS	Overall survival
PAS	Patient Access Scheme
PFS	Progression-free survival
PH	Proportional hazards
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RECIST	Response evaluation criteria in solid tumours
SAE	Serious adverse event
SD	Standard deviation
SE	Standard error
SG	Standard Gamble

STA	Single Technology Appraisal
SmPC	Summary of product characteristics
TOT	Time on Treatment
TTO	Time Trade Off
Vem	Vemurafenib

SUMMARY

Scope of the company submission

The company's submission (CS) generally reflects the scope of the appraisal issued by the National Institute for Health and Care Excellence (NICE). The scope was to consider adults with unresectable or metastatic BRAF V600 mutation-positive melanoma. The CS considers the same patient population, but due to evidence limitations was restricted to a treatment-naïve adult population with advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma. The CS therefore does not consider patients previously treated for advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma. Expert advice to the ERG is that many BRAF mutation positive patients (up to 70%) would be treated with immunotherapy first line before switching to BRAF inhibitor and MEK (MAP extracellular signal-regulated kinases) inhibitor treatment as necessary. However, there are no data available to suggest that outcomes would be worse for second line treatment.

Summary of submitted clinical effectiveness evidence

Overall, the literature searches conducted by the company were appropriate, comprehensive and well designed, although searches for the network meta-analysis (NMA) were approximately a year out of date. The ERG update of the searches for the NMA did not identify any potentially additional relevant studies. Similarly, updating searches for the clinical-effectiveness review for the last six months also revealed no additional relevant studies.

The CS presents evidence of the clinical effectiveness of the MEK inhibitor cobimetinib in combination with the BRAF inhibitor vemurafenib based on one multi-centre, phase III randomised controlled trial (RCT), the coBRIM trial. The RCT compared the combination of vemurafenib (960 mg orally twice daily for a 28-day cycle) plus cobimetinib (60 mg orally once daily for 21 days, followed by 7 days off) against vemurafenib (960 mg orally twice daily) plus placebo in patients with previously untreated unresectable locally advanced or metastatic BRAF V600 mutation–positive melanoma.

The coBRIM trial was considered by the ERG to be of reasonable methodological quality; however, there was a lack of clarity in the reporting of the randomisation and allocation concealment procedures which meant the risk of selection bias is uncertain. There were also

imbalances in discontinuations ‘for any reason’ between the randomised treatment groups (discontinued both vemurafenib + placebo: n=138; discontinued both vemurafenib + cobimetinib arm: n=102), indicating possible attrition bias. The company suggests that the imbalances in discontinuations between the groups were not unexpected, given that the cobimetinib was anticipated to extend survival.

Results of the coBRIM trial

Results were presented in the CS for various data cutoff time points. The primary end point of investigator-assessed progression-free survival (PFS) was defined as the time from randomisation to either the first occurrence of disease progression as assessed by the investigator according to response evaluation criteria in solid tumours (RECIST) 1.1 criteria, or death from any cause. The median PFS was 9.9 months in the vemurafenib + cobimetinib group and 6.2 months in the vemurafenib + placebo group, with a hazard ratio (HR) for death or disease progression of 0.51 (95% Confidence interval (CI), 0.39 to 0.68; primary investigator analysis, data cutoff 9th May 2014).

The median duration of overall survival (OS), defined as the time from randomisation to death from any cause, was 22.3 months in the vemurafenib + cobimetinib group and 17.4 months in the vemurafenib + placebo group, with a HR for death of 0.70 (95% CI, 0.55 to 0.90; final analysis, data cutoff 28th August 2015).

The proportion of patients with complete response (cobimetinib 10% vs 4% placebo) and partial response (cobimetinib 57% vs 40% placebo) favoured the vemurafenib + cobimetinib group (planned analysis, data cutoff 9th May 2014).

Limited results were reported for health-related quality of life (HRQoL) based on the EuroQoL Five Dimension Questionnaire (EQ-5D) and the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30) instrument. Overall, there appears to be some (non-statistically significant) pre-progression HRQoL benefit associated with vemurafenib + cobimetinib compared to vemurafenib + placebo. It is unclear whether the slightly higher HRQoL estimates in the vemurafenib + cobimetinib arm reflect HRQoL improvements resulting from less insomnia and/or other factors (e.g. the incidence of non-melanoma skin cancers which was lower in the vemurafenib + cobimetinib arm).

The most common adverse events (AE) for those treated with vemurafenib + cobimetinib (experienced by at least 20% of patients in either of the study group) included diarrhoea, photosensitivity, serous retinopathy, nausea and vomiting, elevated levels of creatine phosphokinase and aspartate aminotransferase. None of the observed differences between groups were tested statistically. Grade 4 AEs appeared to be more common in the vemurafenib + cobimetinib group (13% vs 9% placebo).

[redacted] An additional safety analysis will be conducted following the analysis for final OS, [redacted]

[redacted]

Results of the NMA

The NMA allowed a comparison between vemurafenib + cobimetinib combination therapy against dabrafenib monotherapy. The NMA used an accelerated failure time (AFT) model with outcomes for PFS and OS. The evidence network was sparse, with only one trial informing each comparison. Clinical heterogeneity between the trials in the network was not discussed. Results from the NMA were more favourable to the vemurafenib + cobimetinib combination therapy on measures of survival compared to treatment with dabrafenib monotherapy.

Summary of submitted cost effectiveness evidence

A systematic search of the literature was conducted by the company to identify economic evaluations of cobimetinib in combination with vemurafenib compared to any other BRAF inhibitor for advanced BRAF V600 mutation-positive melanoma. The review did not identify any relevant studies.

The company's de novo cost effectiveness analysis used a partitioned survival model to estimate the cost-effectiveness of cobimetinib in combination with vemurafenib compared against vemurafenib and against dabrafenib. The model adopted a time horizon of 30 years and

a cycle length of one week. The model consisted of three health states: progression-free survival, progressed disease and death. As recommended by NICE, a discount rate of 3.5% was used for both costs and health outcomes.

The economic evaluation used data from the coBRIM trial for the comparison between vemurafenib + cobimetinib against vemurafenib, and the NMA for the indirect comparison of vemurafenib + cobimetinib against dabrafenib. Health related quality of life utility values were calculated from data collected from the coBRIM trial, but these data were predominately for patients whose disease had not yet progressed. Utility values for progressed disease was used from a separate published study.

Results of the economic model were presented as the incremental cost per quality adjusted life year (QALY) and the incremental cost per life years gained. Two of the comparators (vemurafenib and dabrafenib) have a confidential patient access scheme (PAS) in place. Results were presented in the CS at the drug list price and at estimated discounted PAS prices. The results of the cost effectiveness analyses at list prices showed an incremental cost effectiveness ratio (ICER) for vemurafenib + cobimetinib of £150,514 per QALY compared to vemurafenib, and £209,942 compared to dabrafenib.

The company performed a range of deterministic and probabilistic sensitivity analyses to assess model uncertainty. The ICER remained above £50,000 per QALY in all sensitivity and scenario analyses, including when the cost of cobimetinib was reduced to zero. The probabilistic sensitivity analysis (PSA) estimated a 0% probability that vemurafenib + cobimetinib is cost effective at a willingness to pay threshold of £50,000 per QALY gained.

Commentary on the robustness of submitted evidence

Strengths

The company's systematic review of clinical effectiveness followed standard procedures and is of good quality. The ERG is not aware of any additional relevant published trials that could be included.

The key RCT, coBRIM, is generally well-designed and provides an appropriate evidence base to inform the assessment of clinical and cost-effectiveness in this appraisal.

In the absence of head-to-head direct evidence the NMA enabled indirect comparisons to be made between vemurafenib + cobimetinib against dabrafenib monotherapy.

The structure of the economic model was appropriate, comprehensive and reflected the clinical pathway for patients with advanced melanoma. The model was well-structured and consistent with the outcomes from the coBRIM trial.

The methods chosen for the analysis were generally appropriate and conformed to NICE methodological guidelines.

The company performed a wide range of sensitivity analyses including one-way, probabilistic and scenario analyses to assess model uncertainty.

Weaknesses and areas of uncertainty

The quality assessment of the coBRIM trial pointed to some areas of uncertainties based on randomisation and allocation concealment procedures, and a potential risk of attrition bias favouring the vemurafenib + cobimetinib treatment group.

The evidence network of the NMA was sparse, with only one trial informing each comparison, and there was no discussion of clinical heterogeneity between the trials in the network.

The two comparator treatments have not been compared in a fully incremental analysis. Rather, two separate analyses with different assumptions have been conducted that compare vemurafenib + cobimetinib against vemurafenib and against dabrafenib. These analyses use different assumptions and so it is not possible to integrate the two analyses into a fully incremental analysis.

The model results are sensitive to the parametric curves chosen to extrapolate beyond the coBRIM trial data for PFS, OS and time on treatment. Other parametric curves may also provide plausible extrapolation and these result in less favourable ICERs for vemurafenib + cobimetinib.

Furthermore, the ERG considered that a more reasonable approach to estimating the time on treatment was to use the Kaplan-Meier data with a loglogistic tail.

There is inconsistency in the dosing assumptions used in the comparison between vemurafenib + cobimetinib against dabrafenib. For vemurafenib + cobimetinib the actual dose is used and for dabrafenib the planned dose is used. However the actual dose of vemurafenib + cobimetinib, and hence the estimated cost of the intervention was lower than the planned dose in the coBRIM trial, which in turn made the intervention appear more favourable when compared to dabrafenib.

Summary of additional work undertaken by the ERG

The ERG conducted the following scenario analyses:

i) Cure rate fraction removed
ii) Time to treatment (TOT) extrapolation curve changed to Kaplan-Meier (KM) with log-logistic tail
iii) Changes to utility values
iv) Consistency in dosing between vemurafenib + cobimetinib and dabrafenib
v) Shorter treatment duration
vi) Inclusion of subsequent treatment costs
vii) Assuming equal efficacy between vemurafenib and dabrafenib for OS
viii) Combination analysis (scenario ii, iii and iv)

The results shown in this report are based on the drug list prices. The analyses have also been repeated in a separate confidential appendix for the NICE Appraisal Committee using the PAS drug discount prices for vemurafenib and for dabrafenib. Of these scenarios, the two with the largest impact on the model results were changing the parametric curve used for TOT (scenario ii) which increased the ICER to £204,340 per QALY for vemurafenib + cobimetinib compared against vemurafenib; and reducing the dosage for dabrafenib (scenario iv) which increased the ICER to £223,277 per QALY for vemurafenib + cobimetinib compared against dabrafenib. The ERG's preferred base case compared vemurafenib + cobimetinib to vemurafenib and dabrafenib with results presented as an incremental analysis. The preferred base case included changes to the TOT extrapolation curve (scenario ii), changes to the utility values for the progressed disease health state (scenario iii) and changes to the dosing for dabrafenib

(scenario iv). With these changes the ICER for vemurafenib + cobimetinib compared to vemurafenib is £223,738 per QALY gained. In summary, all of the additional sensitivity and scenario analyses conducted by the ERG resulted in ICERs above £100,000 per QALY gained.

1 Introduction to ERG Report

This report is a critique of the company's submission (CS) to NICE from Roche on the clinical effectiveness and cost effectiveness of cobimetinib (brand name: Cotellic) in combination with vemurafenib for treating advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma. It identifies the strengths and weakness of the CS. Clinical experts were consulted to advise the ERG and to help inform this review.

Clarifications on some aspects of the CS were requested from the manufacturer by the ERG via NICE on 1st March 2016. A response from the company via NICE was received by the ERG on 16th March 2016 and this can be seen in the NICE committee papers for this appraisal.

2 BACKGROUND

2.1 Critique of the company's description of the underlying health problem

The description of melanoma appears to be appropriate and the CS outlined the different forms and its natural history. Early stages of melanoma (stage I or II), in which the cancer has not spread, are generally asymptomatic and can often be cured by surgery (resection). Tumour spread can occur, either to nearby lymph nodes (stage III) or other parts of the body (stage IV) and a mutated form of the BRAF gene (called BRAF V600) is found in about half of melanomas. The mutated gene means that the cells produce too much BRAF protein, leading to uncontrolled cell division and growth of the tumour.

The NICE guidance (NG14)[1] for assessment and management of melanoma published in 2015 report 13,348 new cases of melanoma in 2011, with 2209 related deaths. At diagnosis, around 1% of melanomas are stage IV. The guidance states melanoma is the second most common cancer in adults aged between 25 and 49, with more than 900 adults under the age of 35 diagnosed each year in the UK. This means that melanoma leads to more years of life lost than many more common cancers,[1] and the incidence of malignant melanoma is reported to be increasing every year.[2]

2.2 Critique of the company's overview of current service provision

The CS provides a detailed list of existing NICE guidelines, pathways and technology appraisals relevant to this appraisal (CS Section 3.5, page 35). Figure 1 (CS page 33) illustrates the NICE stage IV melanoma treatment pathway.

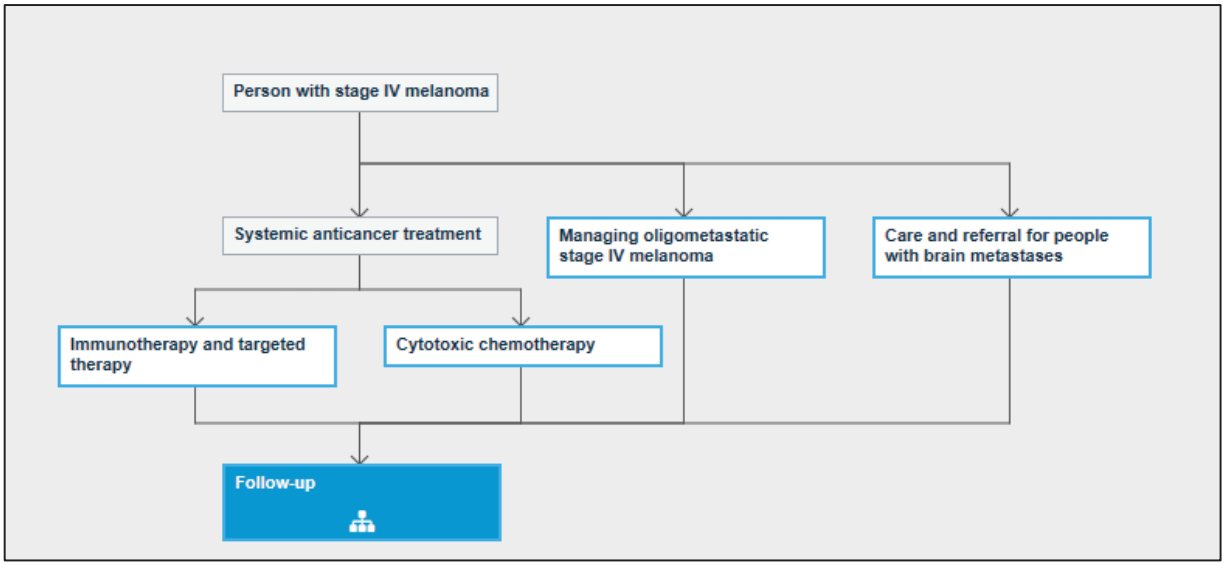


Figure 1 NICE Melanoma Pathway for the management of stage IV melanoma (2016)[3]

NICE has recommended immunotherapy and targeted therapy, and chemotherapy treatment for the treatment of stage IV melanoma (line 3 in Figure 1). These include nivolumab,[4] ipilimumab,[5, 6] dabrafenib,[7] vemurafenib,[8] and pembrolizumab[9,10] (Table 1) each varying in their marketing authorisation and some dependant on a patient access scheme (PAS). NICE clinical guidelines recommend dacarbazine for people with stage IV metastatic melanoma if immunotherapy or targeted therapy are not suitable.[1] Nivolumab or pembrolizumab are said to be used in about 60 - 70% of BRAF mutation positive patients according to clinical advice to the ERG, with a preference for pembrolizumab to nivolumab due to a shorter treatment cycle.

Certain drugs selectively target the BRAF enzyme, inhibiting its action and the mitogen-activated protein kinase (MAPK) pathway, which is a signalling pathway important for cell growth, proliferation, and survival. Cobimetinib is an antineoplastic agent targeting the MEK enzyme in the MAPK pathway. It inhibits the action of the abnormal BRAF protein, with the aim of slowing the growth and spread of the cancer.

Table 1 NICE approved therapies for the treatment of stage IV melanoma

Drugs	NICE guidance – drug recommended for:
Immunotherapy	
Ipilimumab	Previously treated advanced (unresectable or metastatic) melanoma (NICE TA 268); Previously untreated advanced (unresectable or metastatic) melanoma (NICE TA319)
Nivolumab	Advanced (unresectable or metastatic) melanoma (NICE TA384)
Pembrolizumab	Disease has progressed following ipilimumab and, if BRAF V600 mutation positive, also a BRAF inhibitor (NICE TA357); Advanced melanoma not previously treated with ipilimumab (NICE TA366)
Targeted therapy	
Dabrafenib	Unresectable or metastatic BRAF V600 mutation‑positive melanoma (NICE TA321)
Vemurafenib	Locally advanced or metastatic BRAF V600 mutation‑positive malignant melanoma (NICE TA269)
Cytotoxic chemotherapy drugs	
Dacarbazine	Stage IV metastatic melanoma if immunotherapy or targeted therapy are not suitable (NICE NG14)

Cobimetinib is licensed to be used in combination with the BRAF inhibitor vemurafenib. The rationale for combination therapy is stated in the CS to be to reduce the disease progression, following a period of tumour response, which is common with BRAF inhibition monotherapy. According to expert clinical advice provided to the ERG, 30 - 40% of BRAF mutation positive patients in clinical practice would potentially start with combination therapy.

2.3 Critique of company's definition of decision problem

Population

The patient population addressed in the CS is broadly similar to that specified in the NICE scope and for whom cobimetinib is licensed. The patient population specified in the scope is 'Adults with unresectable or metastatic BRAF V600 mutation-positive melanoma. The CS includes the same patient population, but currently relevant comparative evidence exists only for treatment-naïve patients and hence the company presents clinical and cost effectiveness evidence to support the use of cobimetinib in combination with vemurafenib for treatment-naïve patients with advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma. In contrast to the description of current practice in the CS (which states that a BRAF inhibitor is the usual first-line treatment in this patient group), expert clinical advice to the ERG suggested that some patients (estimated up to 70%) with advanced (unresectable or metastatic) BRAF mutation positive melanoma would receive immunotherapy as first line treatment (e.g. ipilimumab, pembrolizumab, or nivolumab), before potentially switching to targeted BRAF mutation inhibitor therapy as necessary. A BRAF inhibitor might be a more commonly used first line treatment for BRAF mutation positive patients with a higher burden of disease or faster disease progression, given that the onset of action with certain immunotherapies may be relatively slow (e.g. around three months for ipilimumab). The submission, therefore, does not specifically cover patients with previous treatment experience. However, expert clinical opinion is that efficacy and safety of vemurafenib + cobimetinib in this group would be similar to that seen in treatment naïve patients.

Intervention

The intervention addressed in the CS reflects the NICE scope and the marketing authorisation. The recommended dose of cobimetinib is 60 mg (3 tablets of 20 mg each), taken orally once daily for 21 days, followed by a 7 day break (Days 22 to 28). The dose for vemurafenib is 960 mg taken orally twice daily (4 tablets of 240 mg, equivalent to a total daily dose of 1,920 mg of 8 tablets in total) taken without a break (days 1-28 of each cycle). For both drugs, down-dosing in response to toxicity is possible as deemed clinically appropriate and both are continued until disease progression. The CS suggests that the combination treatment should only be initiated and supervised by a qualified physician experienced in the use of anticancer medicinal products in specialist secondary, or tertiary care centres. As both drugs in the combination treatment are taken orally, it is stated that there is no impact on NHS staff in terms of administering the drugs. However, monitoring and dose adjustments due to AEs may be required and although it is suggested that this would not require additional NHS resources as this is established clinical practice in England and Wales for this type of treatment, this has not been established as yet.

Comparators

Two comparators are specified in the NICE scope: vemurafenib and dabrafenib, which are both used in treatment-naïve patients and those with prior treatment experience for malignant melanoma. Direct evidence was only available for the comparison with vemurafenib, from the pivotal coBRIM trial. However, comparative evidence for dabrafenib was provided in the form of an indirect comparison in the CS (see Section 3.1.7 of this report for a description and critique of the indirect comparison).

Outcomes

Clinical evidence in the CS is provided for all five outcomes specified in the NICE scope: progression-free survival (PFS), overall survival (OS), response rates (reported as objective response rate (ORR), best overall response (BORR) and duration of response), adverse effects (AEs) of treatment and health-related quality of life (HRQoL). These outcomes are widely used and accepted endpoints in oncology trials. Pharmacokinetic measures were also employed, but not included in the CS.

Economic analysis

As specified in the final NICE scope, the cost effectiveness of treatments were expressed in terms of the incremental cost per quality adjusted life year (QALY) gained. Outcomes were assessed over a 30-year time horizon (deemed equivalent to a lifetime horizon for this patient population) and costs were considered from the perspective of NHS England.

Other relevant factors

Subgroups

No subgroups are specified in the NICE scope or the decision problem. As noted above, evidence in the CS was only available from treatment-naïve patients. The CS presents clinical effectiveness results from the coBRIM trial for 12 pre-specified subgroups (NB. cost effectiveness evidence is not presented for these sub-groups):

- Disease stage (IIIc, M1a, M1b, M1c)
- Disease stage (IIIc/M1a/M1b, M1c)
- Age (≤ 65 years, > 65 years) at randomisation

• Race (non-White, White)
• Sex (female, male)
• Geographic region (North America, Europe, Australia/New Zealand/others)
• Eastern Cooperative Oncology Group (ECOG) performance status at randomisation (0, 1)
• LDH (lactate dehydrogenase) (normal, elevated)
• Presence of brain metastases (yes, no)
• Time since metastatic disease diagnosis (< 6 months, ≥ 6 months)
• Prior adjuvant therapy (Yes, No)
• BRAF V600 mutation status (V600E, V600K)

Expert clinical advice to the ERG is that these subgroups are appropriate, with no further subgroups of clinical importance suggested.

Equity or equality issues
The CS states that it is not believed that the use of cobimetinib (in combination with vemurafenib) will be associated with any equality issues (CS Section 3.8, page 36) and the ERG concurs with this assessment.

Other valid issues
The CS points out that BRAF mutation testing is part of routine management for patients with advanced melanoma in the UK, therefore is not considered an additional cost or resource burden to the system (CS page 128). However, as per the prior vemurafenib technology appraisal (TA269),[8] a cost of £95 per test is incorporated into the model, which has no incremental effect as this cost is applied to both the intervention and all comparators. Clinical opinion provided to the ERG concurs with this approach.

There is currently no PAS in place for cobimetinib, though there is one for vemurafenib and for dabrafenib.

3 CLINICAL EFFECTIVENESS

3.1 Critique of the company's approach to their systematic reviews

3.1.1 Description of the company's search strategies

The CS reports separate literature searches for the following systematic reviews:

- Clinical-effectiveness (searched to September 2015)
- Cost-effectiveness (searched to December 2015)
- Indirect treatment comparison (searched to April 2015)
- Measurement and valuation of health effects (searched to December 2015)
- Cost and resource use (searched to December 2015)

The search sources and strategies for each of these searches are reported in appendices to the CS. The ERG regards the searches to be comprehensive, well designed, explicitly documented and fit for purpose. The search terms have been documented line-by-line as applied to the databases, though for the clinical-effectiveness search the number of hits per line is not documented, which lessens transparency.

An appropriate range of databases were included in each search, including the core databases of MEDLINE (and MEDLINE In-Process), Embase and the Cochrane Library. Additional specialist databases were included as appropriate to certain reviews (e.g. Econlit was searched for the cost-effectiveness and the costs and resources search). Conference proceedings were searched for all the reviews, with the exception of the indirect treatment comparison. The conference proceedings were searched from 2013-2015 for all reviews except the clinical effectiveness review which was searched from 2014-2015. The conferences searched were appropriate to the scope of the appraisal, including key oncology conferences (American Society of Clinical Oncology; European Society for Medical Oncology), and melanoma-specific conferences (Society for Melanoma Research). The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conference was searched for the cost-effectiveness review, the measurement and valuation of health effects review and the cost and resource use review.

Some hand-searching took place, with reference lists of included studies reported for all searches. Additional sources were handsearched as appropriate to the review in question (e.g. the Research Papers in Economics (RePEc) resource was searched for the cost effectiveness review and the costs and resources review).

The searches were generally up to date, with the exception of the indirect treatment comparison search which was current only to April 2015. The ERG updated this search on MEDLINE and MEDLINE In-Process and Embase but did not identify any potentially relevant new studies (see

section 3.1.7). The clinical-effectiveness review was current to September 2015, and the ERG also updated the searches for this review covering the period 2015 to 2nd March 2015 on MEDLINE, MEDLINE In-Process and Embase. The results were screened by an ERG systematic reviewer and no additional relevant studies were identified.

The CS does not mention if searches were conducted to identify on-going studies. The company responded to a clarification question from the ERG (clarification response A10) that the following databases were searched on April 7th 2015: ClinicalTrials.gov; European Union Clinical Trials Register; World Health Organisation Trials Registry; British Association of Dermatologists guidelines; Scottish Intercollegiate Guidelines Network (SIGN): Mapi Institute; and the International Network of Agencies for Health Technology Assessment (INAHTA). In terms of the results of this search all that is stated is that it did not identify any studies expected to report results within the following 12 months. In advance of the company's response to the clarification question, the ERG conducted a separate search of the following trials databases Sources: UK Clinical Research Network Study Portfolio (UKCRN), ISRCTN, WHO International Clinical Trials Registry Platform; clinicaltrials.gov; and rochetrials.com. One potentially relevant trial was identified, discussed in Section 3.1.3.1.

The ERG also checked the ScHARRHUD (Health Utilities Database) database for studies reporting health utility papers appertaining to melanoma, however the only relevant result was already cited in the CS.

In summary, we consider that the searches conducted by the company to support the systematic reviews in the submission are generally comprehensive and are reported transparently. We updated two of the database searches but did not identify any additional relevant published studies.

3.1.2 Statement of the inclusion/exclusion criteria used in the study selection.

The CS provides a clear overview of the inclusion and exclusion criteria (CS Table 6, page 38). The criteria appear to be in line with the marketing authorisation, the NICE scope, and the company's decision problem.

No limits were placed on inclusion relating to the quality of the RCTs, but only phase II, III or IV RCTs, or systematic reviews or meta-analyses of RCTs were eligible for inclusion. Setting was not an inclusion criterion. A PRISMA diagram, illustrating the numbers of references included and excluded at each stage of the systematic literature review conducted for the CS, is provided in Figure 2. The total number of records identified through database searching seems low (n=57), even after duplicate removal. However, as we are not aware of any relevant studies that were not identified in the company's systematic review, this does not necessarily indicate that their search strategy was flawed.

The company did not address any potential bias that may have arisen in relation to their searches or inclusion/exclusion criteria, but processes appear to have been robust.

3.1.3 Identified studies

The CS identified one relevant RCT (coBRIM).[12] The majority of the details of the RCT in the CS were summarised in tables such as trial design, intervention, population, patient numbers and statistical analysis. Trial outcome measures and subgroups were described in text.

References including the coBRIM trial[13-15] were provided electronically, as was the clinical study report (CSR). The coBRIM trial was sponsored by f. Hoffmann – La Roche/Genentech.

The company did not identify any non-randomised or non-controlled evidence relevant to the decision problem (CS Section 4.11). However, it appears from the CS that non-randomised studies were not searched for (an RCT filter was applied to the clinical effectiveness search strategies).

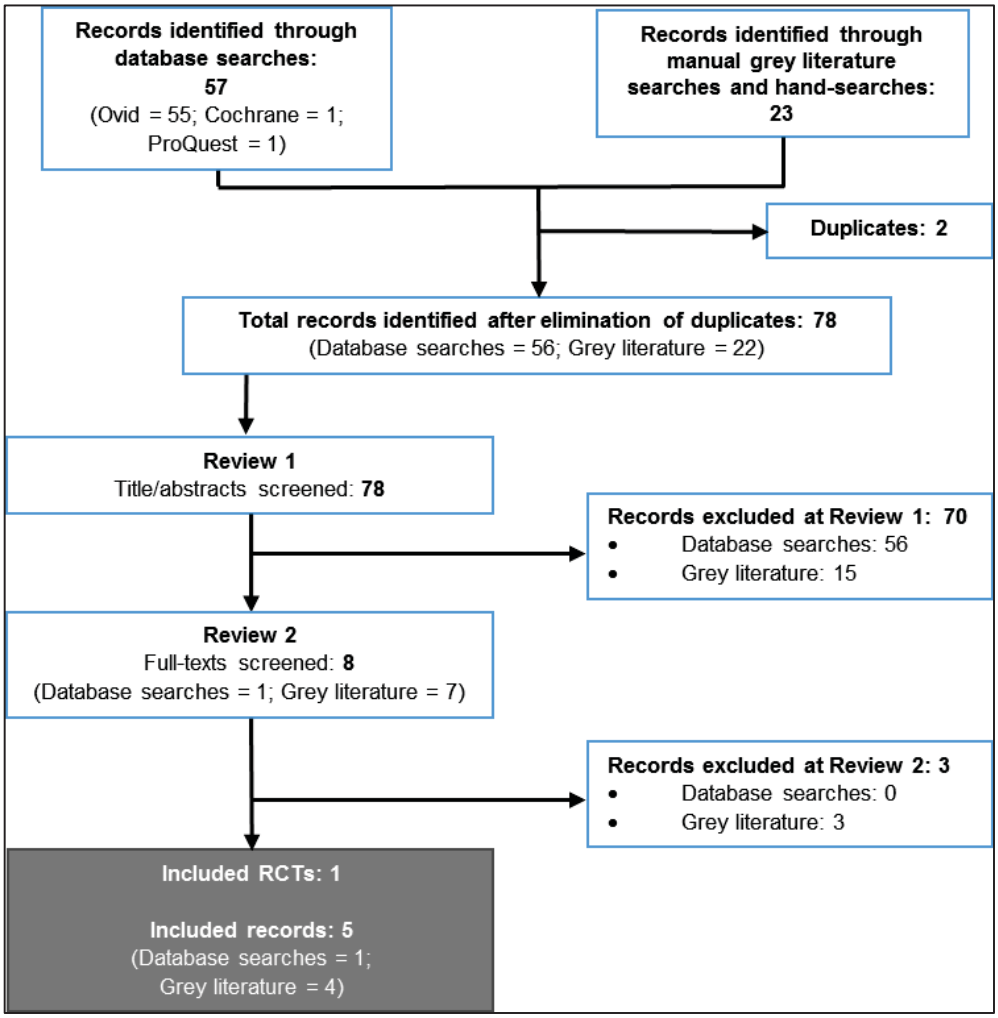


Figure 2 PRISMA diagram for systematic literature review of RCTs (search cut-off date: 8th or 9th September 2015; based on CS figure 2, page 40)

The trial journal publication[12] states that there were no significant differences in baseline characteristics between the study groups, but no p values were provided to support this. Both the trial publication[12] and the CS state that the characteristics of the patients at baseline were generally well balanced between the two study groups, with baseline characteristics provided in table format. Apart from ECOG performance-status score, the ERG would agree with this statement. While differences are said to not be statistically significant, 9% more participants in the intervention group were rated ECOG performance-status score 0 (defined as the patient is fully active and able to carry on all performance without restriction) and 9% fewer were rated ECOG performance-status score 1 (the patient is restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature) than in the control group. Further, 3% fewer participants in the intervention group were assessed at baseline as

metastatic status M1c compared to the intervention group and 4% more participants were assessed as unresectable stage IIIC (i.e. less advanced metastases) (Table 2) than in the control group. However, while these small imbalances might be expected to favour the placebo group, expert advice to the ERG did not consider that these differences would influence the improvement in clinical outcomes for those treated with cobimetinib.

3.1.3.1 Ongoing trials

In their response to a clarification question from the ERG, the company stated that a search for ongoing trials did not identify any studies which were expected to report results within the following 12 months. A search for ongoing trials conducted by the ERG identified one trial of potential interest (summarised in Table 3). The trial is comparing two different regimens of combination therapy: one with a vemurafenib monotherapy induction period, and one with a cobimetinib monotherapy induction period. However, as there is no comparison between combination therapy and monotherapy this trial is not fully relevant to the scope of the appraisal.

Table 2 Overview of baseline characteristics in the coBRIM trial

Parameter (data based on CS Table 14 , page 59 – 60)	**CoBRIM trial**	
	Vemurafenib + cobimetinib	**Vemurafenib + placebo**
Sample size, n	n=247	n=248
Age, mean years (range)	56 (23-88)	55 (25-85)
Male sex, n (%)	146 (59)	140 (56)
White, n (%)[a]	227 (92)	235 (95)
Geographic region n (%)		
Australia, New Zealand, or Israel	40 (16)	38 (15)
Europe (including Russia and Turkey)	182 (74)	184 (74)
North America	25 (10)	26 (10)
ECOG performance-status score n (%)[b]		
0	184/243 (76)	164/244 (67)
1	58/243 (24)	80/244 (33)
2	1/243 (<1)	0/244

Metastatic status n (%)		
Unresectable stage IIIC	21 (9)	13 (5)
M1a	40 (16)	40 (16)
M1b	40 (16)	42 (17)
M1c	146 (59)	152[c] (62)
Elevated lactate dehydrogenase n (%)	112/242 (46)	104/242 (43)
History of brain metastases n (%)	1 (<1)	2 (1)
BRAF-mutation genotype n (%)[d]		
V600E	170 (69)	174 (70)
V600K	24 (10)	32 (13)
Could not be evaluated	52[c] (21)	42 (17)

[a] Race was assessed by the investigator. [b] One patient randomly assigned to receive vemurafenib and cobimetinib had an ECOG performance-status score of 1 at randomisation but had an ECOG performance-status score of 2 after randomisation but before the first dose was received. [c] The CS reported fewer patients than the trial publication (n=153 and n=53 retrospectively). [d] After randomisation, tumour DNA was characterised to identify specific V600 mutations using next-generation sequencing. Cases that could not be evaluated were those in which either no tumour sample was provided or sequencing could not be performed on the tissue provided

Table 3 Ongoing trials

Trial identifier, sponsor	**Design, Country**	**Intervention, comparator, patient group**	**Expected end date**
NCT02427893; Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins/ Genentech	Phase III, open label RCT (parallel assignment) USA	Vemurafenib monotherapy for 10 days, followed by combination therapy by adding cobimetinib vs cobimetinib monotherapy for 10 days, followed by combination therapy by adding vemurafenib – dosages not specified (total n=200). Adult patients with advanced unresectable American Joint Committee on Cancer stage III or stage IV BRAF V600E/K mutant melanoma.	August 2015 - final data collection date for primary outcome measure April 2017

3.1.4 Description and critique of the approach to validity assessment

The CS includes a quality assessment of the CoBRIM trial in Table 15 (CS pages 60 – 61) and second quality assessment of the trial in the 'Quality Assessment of trials included in NMA' in Table 76 (CS pages 204 – 208).

The company's quality assessment of the trial is appropriate, using NICE recommended criteria (based on the Centre for Reviews and Dissemination (CRD) guidance).[16] However, the two quality assessments of the CoBRIM trial differed in their conclusions. The first quality assessment of the trial states that "The study was of high quality based on the respective responses for each category thus indicating low risk of bias in study conduct and design" (CS page 60), while the second assessment in the CS states that the overall risk of bias for the trial is unclear (CS Table 76, page 208).

While there were some differences between the ERG's and the company's assessments of the trial quality (Table 4), the ERG agrees with the CS second quality assessment, in that the CoBRIM trial appears of unclear quality due to some lack of clarity in reporting and an imbalance in drop-outs between the treatment groups.

Table 4: Company and ERG assessment of coBRIM trial quality

CRD quality assessment criteria for RCT[16]	**Judgements[1]**		
	CS judgement (Table 15): **CS judgement (Table 76):**		**ERG judgement**
1. Was the method used to generate random allocations adequate?	CS Table 15: CS Table 76:	Not clear Not clear	Unclear
Comment: The trial journal publication only states that patients were randomly assigned in a 1:1 ratio and stratified according to geographic region and metastasis classification. As stated in the CS (CS Table 15, page 60), additional information in the trial protocol states that a stratified, permuted-block randomisation scheme would be used for treatment allocation based on stratification factors (Geographic region: North America, Europe, Australia/New Zealand/others; Metastatic classification: unresectable Stage IIIc, M1a, and M1b; or M1c). While block randomisation could be considered a low risk of introducing bias if carried out			

adequately, the details of the random sequence generation method are not provided; hence the risk of selection bias is unclear.			
2. Was the allocation adequately concealed?	CS Table 15: CS Table 76:	Yes Not clear	Unclear
Comment: The company's two quality assessments for this question differ. The ERG concluded that, as no details of treatment allocation sequence generation were reported (as stated in the company's second quality assessment), there is an unclear risk of selection bias. While the trial protocol mentions that an interactive response system would be used for random assignment, which suggests possible separation of the study investigators from those doing the randomisation, it is not completely clear if this was used.			
3. Were the groups similar at the outset of the study in terms of prognostic factors, e.g. severity of disease?	CS Table 15: CS Table 76:	Yes Yes	Yes
Comment: The CS stated that the characteristics of the patients at baseline were generally well balanced between the two study groups. While there were differences in patient's baseline characteristics between the treatment groups as previously stated (e.g. ECOG performance status, and metastatic status),[12] it was stated that there were no statistically significant differences between treatment groups (no p value reported) and expert clinical advice to the ERG suggests that these differences would not influence the improvements in clinical outcomes for those treated with cobimetinib.			
4. Were the care providers, participants and outcome assessors blind to treatment allocation? If any of these people were not blinded, what might be the likely impact on the risk of bias (for each outcome)?	CS Table 15: CS Table 76:	Yes Not clear	Yes
Comment: The company's first quality assessment stated that there is a low risk of performance and detection bias in the trial as the investigator, patient, and sponsor were blinded to treatment assignment. However, this differs from their second assessment, which was that it is unclear. The trial was described as double-blinded, and the supporting text states that a blinded, independent central review of tumour assessments was performed. The ERG agrees with the company's first assessment, in that the trial appears to be at a low risk of performance and detection bias. The primary end-point was investigator-assessed PFS, but tumour assessments were reviewed by a blinded independent review committee (NB. there were some differences			

between the assessments, as reported in Section 3.3 of this report).			
5. Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	CS Table 15: CS Table 76:	No Not clear	Yes
Comment: The company's two quality assessments for this question differ. Firstly, the company suggests that the imbalances in drop-outs between the groups were not unexpected, given that cobimetinib was anticipated to extend survival. In the second assessment, the company suggests that there was an unclear risk of attrition bias, stating that this is not reported. The ERG disagrees with both assessments due to differences in patient withdrawal between the two treatment groups (higher withdrawals in the vemurafenib + placebo group) and judges that there is a potential risk of attrition bias in the trial (see Table 15).			
6. Is there any evidence to suggest that the authors measured more outcomes than they reported?	CS Table 15: CS Table 76:	Not clear No	Unclear
Comment: The company's two quality assessments for this question differ. The first assessment suggests that the risk of reporting bias is unclear and this seems to be because only the primary, secondary and safety outcomes were reported in the primary trial publication.[10] The second assessment by the company suggests that there is no reporting bias, with all outcomes stated in the methods also reported in the results section. The ERG agrees with the company's first assessment, that the risk of reporting bias is unclear, as not all of the outcomes in the protocol have been reported; primarily there is little data in the CS for the EQ-5D-5L. (NB. These data were later supplied to the ERG on request, see Section 3.3.4 of this report).			
7. Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	CS Table 15: CS Table 76:	Yes Yes	Yes
Comment: All the efficacy analyses were carried out in the ITT population, while the safety analyses were based on a modified ITT (mITT) population. A mITT population (all patients who underwent randomisation and received at least one dose of the study drug) omits those who may have had other relevant reasons for not receiving the drug. Modified ITT analyses for safety analysis are increasingly used in industry-sponsored trials[17].			

[1] The ERG used the CRD type responses (yes, no, unclear) to each question, although the CS approach differed slightly (yes, no, not clear).

3.1.5 Description and critique of the company's outcome selection

The outcomes specified in the CS are progression-free survival (PFS), overall survival (OS), response rates, adverse effects of treatment and HRQoL (CS Table 1, page 18). These are appropriate outcomes for evaluation of a cancer therapy, and are consistent with the final NICE scope.

The primary outcome is PFS, defined as the time from randomisation to either the first occurrence of disease progression, as assessed by the investigator according to Response evaluation criteria in solid tumours (RECIST) 1.1 criteria, or death from any cause (CS page 47).

Secondary outcomes are OS, response rates, PFS as assessed by independent review (two board certified radiologists), safety, and HRQoL (CS pages 47 - 50). Definitions provided in the CS are:

OS: The time from randomisation to death from any cause.

Overall response rate (ORR): For patients with measurable disease at baseline, investigator-assessed complete or partial response according to RECIST 1.1 criteria (CS page 48).

Best overall response rate (BORR): A complete response (CR) or partial response (PR) according to RECIST 1.1 criteria. Determined by two consecutive investigator assessments performed at least four weeks apart. For stable disease (SD), measurements had to meet SD criteria at least once post-randomisation at a minimum interval at least six weeks (CS page 48).

Duration of response (DOR): Evaluated in patients who satisfied BORR criteria. The time from a first occurrence of a documented CR until either disease progression as determined by investigator review (RECIST 1.1 criteria) or death from any cause (CS page 48).

Adverse events (AEs) were classified as: all AEs, drug related AEs, deaths, SAEs, drug-related SAEs, AESIs, and AEs leading to dose interruption/modification and to discontinuation of study treatment. The CS states that to classify AEs, the sponsor assigned preferred terms to the verbatim terms reported on the case report form, using the latest version of the Medical Dictionary for Regulatory Authorities (MedDRA 16.1) terminology. The CS does not state

whether AE classifications were checked or adjudicated independently. In addition to AEs, safety assessment included protocol-specified tests and vital signs. The AE classifications reported in the CS are generally consistent with those in the CSR, with some slight differences [redacted] (CS pages 48 – 49).

The instruments used to assess HRQoL were the European Organisation for Research and Cancer Quality of Life Questionnaire (EORTC QLQ-C30) and the EuroQol EQ-5D-5L questionnaires. These are both validated and widely-used HRQoL measures, and provide complementary information (QLQ-C30 is specific to cancer whilst EQ-5D is generic but provides utility estimates for use in the economic analysis) (CS page 49 - 50).

The ERG notes that although the QLQ-C30 is widely used in cancer research studies, it is not specific to skin cancer and might not necessarily be the most sensitive instrument for capturing effects of melanoma on patients' HRQoL. Several melanoma-specific instruments are available, including FACT-melanoma (Functional Assessment of Cancer Therapy-melanoma).[18]

In summary, the outcomes presented in the CS are appropriate for assessing effects of pharmacological therapy on melanoma and no important outcomes have been missed, although the cancer HRQoL instrument employed might not be the most sensitive of those available.

3.1.6 Description and critique of the company's approach to trial statistics

The rationale for the coBRIM trial sample size is provided on CS pages 50 - 51. The pre-specified number of progression events (206) required to give >95% power to detect a 5-month improvement in median PFS (from 6 months in the vemurafenib + placebo arm to 11 months in the vemurafenib + cobimetinib arm) was reached in May 2014 and the database was locked for PFS analysis on 10 July 2014. The final analysis of OS was to be performed after approximately 385 deaths had occurred, which would provide approximately 80% power to detect a 5-month improvement (from 15 months in the vemurafenib + placebo arm to 20 months in the vemurafenib + cobimetinib arm) in median OS (corresponding to a Hazard Ratio (HR) for death of 0.75).

The main analysis population for efficacy endpoints was the ITT population, defined as all randomised patients, regardless of whether or not study treatment was received (CS page 51).

Crossover between the treatment arms in the coBRIM trial was not permitted (CS Figure 3, page 43 and Table 12, page 57). The safety population was defined as all patients who received at least one dose of study treatment and analysed according to the treatment received (modified ITT, mITT). The population for analysis of HRQoL is referred to in the CS as the "PRO" population and was defined as all patients who had a baseline assessment and at least one post-baseline assessment. The CS (page 51) states that "the PRO population was analysed according to the treatment assigned at randomisation (i.e. ITT)". Dates for the first, second and final OS analysis are reported on page 52 and a summary of analysis dates for PFS, OS, response rates, and safety are reported in Table 8 (CS pages 42 - 43). There were no interim analyses for PFS.

The primary analysis compared PFS between the vemurafenib + cobimetinib and the vemurafenib + placebo arms of the coBRIM trial using a stratified log-rank test at overall significance 0.05 (2-sided). The HR for PFS was estimated using a Cox model stratified by geographic region (North America, Europe, Australia/New Zealand/others, where "others" appears to refer to Israel, whilst Russia and Turkey were included in the Europe subgroup according to the trial publication[12]) and metastatic classification (unresectable Stage IIIc, M1a, and M1b; or M1c). Median PFS for each treatment arm was estimated based on Kaplan-Meier (KM) methods. Data from patients who experienced disease progression or death were censored at the last tumour assessment date. Data from patients with no post-baseline tumour assessment were censored at the randomisation date (CS page 53). The CS (page 54) states that three sensitivity analyses on the PFS outcome were performed (non-stratified analysis, censoring for non-protocol anticancer therapy, and censoring accounting for missed visits). Results of these sensitivity analyses are not reported in the CS, but were provided in response to a clarification request by the ERG.

OS was compared between the two treatment arms (CS page 54) using a log-rank test stratified by geographic region and metastatic classification, with significance level 0.05 (2 sided). The HR for death was estimated using a stratified Cox model. A Lan-DeMets implementation of an O'Brien-Fleming boundary function was used (although not stated in the CS, this controls for the Type I error associated with interim analysis of accumulating data). Data for patients still alive at the time of analysis were censored at the date the patient was last known to be alive. Survival time for patients with no post-baseline survival information was censored on the date of

randomisation. The duration of OS was calculated as the date of death or censoring minus the date of randomisation plus one day.

BORR was compared between the two treatment arms (CS page. 55) using a Chi-square test with Schouten correction. 95% CI were calculated for the BORR using the Clopper-Pearson method (a common method for calculating binomial confidence intervals) and for the BORR difference between treatment arms using the Hauck-Anderson method (one of a number of methods used to compute two-sided confidence intervals for difference between independent binomial proportions).

DOR was estimated (CS page 55) only for patients who had a confirmed overall response or partial response (i.e. CR or PR) and, being based on a subgroup, was not used for formal hypothesis testing. Medians and interquartile ranges for DOR were estimated using KM methods, with 95% CI calculated using the Brookmeyer and Crowley method.

HRQoL was assessed at baseline, days 1 and 15 in cycles 1 and 2, (C1D1; C1D15; C2D1; C2D15 respectively) and every other cycle thereafter until patient withdrawal or end of study using the EORTC QLQ-C30 (CS page 67). Data were evaluable until Cycle 8 Day 1 (C8D1), after which too few patients remained enrolled in the vemurafenib + placebo treatment arm to allow for meaningful conclusions (<25% from baseline).[19] HRQoL analyses (CS page 55) were post-hoc and based on descriptive statistics. In addition, responder analysis summarised the frequency of patients experiencing “clinically meaningful” improvement in each scale of the EORTC QLQ-C30 (symptoms, functional impact, and health-related quality of life). Clinically meaningful improvement was defined as a ≥10-point change at ≥1 post-baseline assessment.[19]

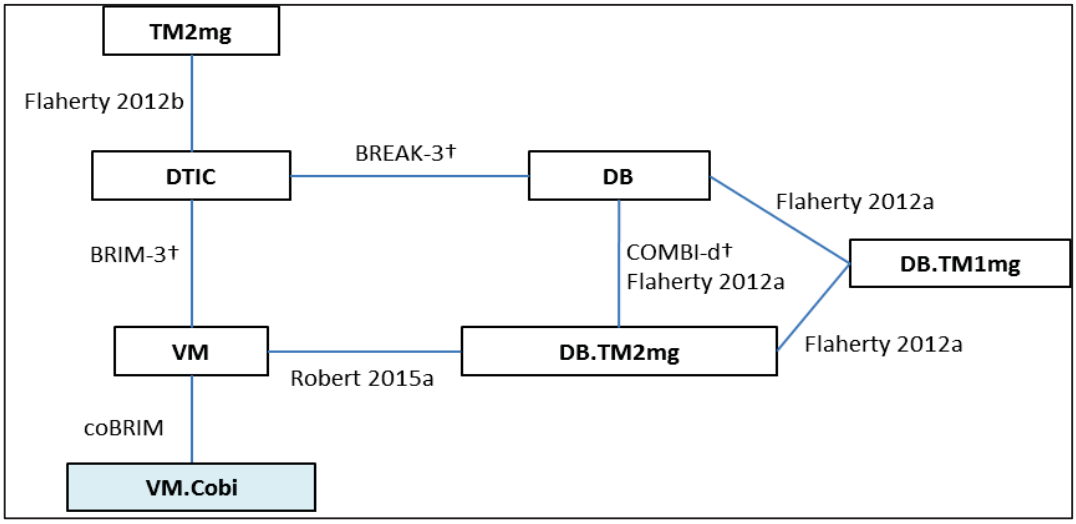
Exploratory pre-specified subgroup analyses for PFS and OS were conducted (see Section 2.3 of this report for details of the 12 subgroups).

In summary, the ERG considers that the approaches to statistical analysis and data censoring are appropriate. Some of the information in the CS is referenced as from the CSR, but has not been marked AIC or CIC and is not available in the coBRIM trial publication.[12]

3.1.7 Description and critique of the company's approach to the evidence synthesis

Given that only one trial of vemurafenib + cobimetinib (the coBRIM trial) was included in the CS, a meta-analysis was not possible. The submission therefore provides a narrative summary of that trial. The coBRIM trial did not include a dabrafenib monotherapy arm, and therefore to permit comparisons between vemurafenib + cobimetinib and dabrafenib an indirect comparison was necessary. A network meta-analysis (NMA) was conducted to make the indirect comparison and is described in CS section 4.10. The NMA provides pairwise indirect comparison results for the outcomes of OS and PFS (CS Tables 21 and Table 22, respectively), and these are used to inform the economic model (see Section 4.3.5 of this report).

An accelerated failure time model (AFT) was used (see below in Section 3.1.7.4 for a discussion of this) with the results presented as the inverse of the acceleration factor (1/AFT). For both outcomes, vemurafenib + cobimetinib was more effective than dabrafenib (see Section 3.3 of this report for a summary of the results of the NMA, and below for a critique of the use of the AFT). Figure 3 provides an illustration of the network, Table 5 details the interventions and comparators of the included trials, and Table 6 provides the ERG's critical appraisal of the NMA.



TM2mg = trametinib 2mg; DTIC = dacarbazine, DB = dabrafenib, VM = vemurafenib, DB.TM2mg = dabrafenib + trametinib 2mg; DB.TM1mg = dabrafenib + trametinib 1mg; VM.Cobi = vemurafenib + cobimetinib † Trial has more than one publication contributing data

Figure 3 Illustration of the network meta-analysis presented in the company submission (reproduced from CS Figure 11).

Table 5 Summary of the trials included in the company NMA (adapted from CS Table 20)

Trial reference	Trial arm A	Trial arm B	Trial arm C
coBRIM[12]	Vemurafenib + cobimetinib	Vemurafenib + placebo	
BRIM-3[20]	Vemurafenib	Dacarbazine	
Flaherty 2012a[21]	Trametinib 1mg + dabrafenib	Trametinib 2mg + dabrafenib	Dabrafenib
Flaherty 2012b[22]	Trametinib 2mg	Chemotherapy (dacarbazine or paclitaxel)	
BREAK-3[23]	Dabrafenib	Dacarbazine	
COMBI-d[24]	Trametinib 2mg + dabrafenib	Dabrafenib	
Robert 2015a (COMBI-v)[25]	Trametinib 2mg + dabrafenib	Vemurafenib	

Table 6 ERG appraisal of NMA

APPRAISAL CRITERIA	
Rationale and searches	
Is the rationale for the NMA and the study objectives clearly stated?	Yes, in CS section 4.10.2.
Does the reported study follow conventional guidelines for systematic reviews, as well as use explicit search terms, time frames, and avoid ad hoc data?	Yes, details given in Appendix 4 and Appendix 5. Searches were conducted of key databases; inclusion criteria are stated; PRISMA flow chart is provided; tabulated details of the included studies are given.
Are inclusion/exclusion criteria adequately reported?	Yes, CS appendix 4 (Table 70).
Is quality of the included studies assessed?	Yes, CS appendix 5 (Table 76).
Methods – Model	
Is the statistical model described?	Yes, Bayesian framework, but with limited information given.
Has the choice of outcome measure used in the analysis been justified?	Yes. PFS and OS are key outcomes in the scope and were primary outcomes for all the included studies.
Has the choice of fixed or random-effects model been justified?	Yes (CS section 4.10.17). The Deviance information criterion (DIC) was used to determine choice of model. The DIC indicated that fixed-effects models provided improved fit. However, further discussion of the assumptions about the distribution of study effects would have also been informative.

Has a structure of the network been provided?	Yes (CS Figure 11)
Is any of the programming code used in the statistical programme provided (for potential verification)?	Yes – WinBUGS code is provided in CS Appendix 8.
Methods - Sensitivity analysis	
Does the analysis conduct sensitivity analyses?	No.
Results	
Are the results of the NMA presented?	Yes.
Does the study describe an assessment of the model fit?	Yes - The model fit of the fixed- and random-effects models conducted for each outcome was compared using the deviance information criterion (DIC).
If direct and indirect evidence is reported to be consistent, is the evidence combined and the results presented?	Yes, the results of the NMA reflect both direct and indirect evidence. Z values are presented to confirm no evidence of inconsistency.
Has there been any discussion around the model uncertainty?	No
Are the point estimates of the relative treatment effects accompanied by some measure of variance such as confidence intervals?	Yes – 95% credible intervals are given around the NMA point estimates.
Discussion	
Does the study discuss both conceptual and statistical heterogeneity and incoherence?	No.
Does the discussion flow from the results seen?	No, little discussion is given.
Have the authors commented on how their results compare with other published studies (e.g. NMAs)?	No.

3.1.7.1 Evidence included in the NMA

The company did a systematic literature review to inform their indirect comparison. As stated in Section 3.1.1, the ERG considers that the search strategies are reasonable, but notes that the search is now nearly a year out of date. The ERG therefore updated the search and no additional relevant references were identified meeting the scope of the appraisal. The NMA can therefore be considered to be up-to-date and comprehensive in terms of the available published evidence.

The inclusion criteria for the NMA (CS Table 70) specified the interventions and comparators listed in the scope, plus additional comparators (e.g. targeted therapies such as pembrolizumab, trametinib; immunotherapy such as ipilimumab; and chemotherapy). The submission explains that the NMA was conducted to support pricing and reimbursement submissions across all markets, hence inclusion of these additional comparators in the systematic review conducted to support the NMA. The NMA as presented in the submission is for a 'restricted' scenario based

only on studies of patients who are BRAF mutation positive (n=7 RCTs, Figure 3 and Table 5). It is not explicitly stated in the CS but it is the ERG's assumption that the results of the restricted NMA scenario are not influenced by the additional comparators in the 'broad NMA'.

The restricted network, despite its narrowed focus on BRAF mutation positive patients, remains broader than the scope of the current appraisal as it contains trials of trametinib and dabrafenib combination therapy and trametinib monotherapy[21,22,24,25] (NB. these are the subject of a separate NICE technology appraisal in progress, ID661). The ERG assumes that these trials have been included for completeness and with the intention of adding to the volume of evidence in the network (particularly for the dabrafenib monotherapy node, where there is no direct comparison – see Figure 3 and Table 5 thus increasing precision of the results.

As Figure 3 shows, the network contains only one trial for each pairwise comparison (with the exception of trametinib and dabrafenib combination therapy versus dabrafenib, which is not a comparison within the scope of this appraisal). The network can therefore be considered to be sparse and its results should be considered with caution given the limited number of trials available.

The results are also dependent on the methodological quality and risk of bias of the included trials. The CS provides a quality assessment of the included trials in Appendix 5 (Table 76) using the CRD criteria. Each trial is summarised in terms of its overall risk of bias (though the Cochrane Risk of Bias criteria are not actually used). The overall judgement of bias was high in five of the seven RCTs. In the remaining two trials, the judgement was low and unclear respectively. The submission does not comment on the overall assessment of methodological quality and risk of bias of the included studies in relation to the results of the NMA, other than to state that removal of the high risk studies would have left only two remaining studies and therefore the indirect comparison would not have been possible. The ERG has not conducted an independent assessment of the risk of bias of the trials included in the NMA (other than the coBRIM trial, see section 3.1.4). The ERG notes that the company's overall risk of bias judgements do not appear to be consistent with the individual domains of bias identified for each study and also seem inconsistent between studies. Furthermore, the quality assessment judgements for the coBRIM trial in Table 15 and Table 76 of the CS disagree on four of the seven domains assessed (as discussed earlier, see Table 4). The ERG considers the company's assessment of the quality of the studies in the NMA to be unclear. However, it

should be noted that most of the trials in the network appeared to be relatively large phase III licensing RCTs, and some of these have been used to inform previous NICE appraisals of treatments for melanoma (BREAK-3 informed the company submission for NICE TA321 (dabrafenib),[7] and BRIM-3 informed the company submission for NICE TA269 (vemurafenib).[26]

3.1.7.2 Assessment of heterogeneity

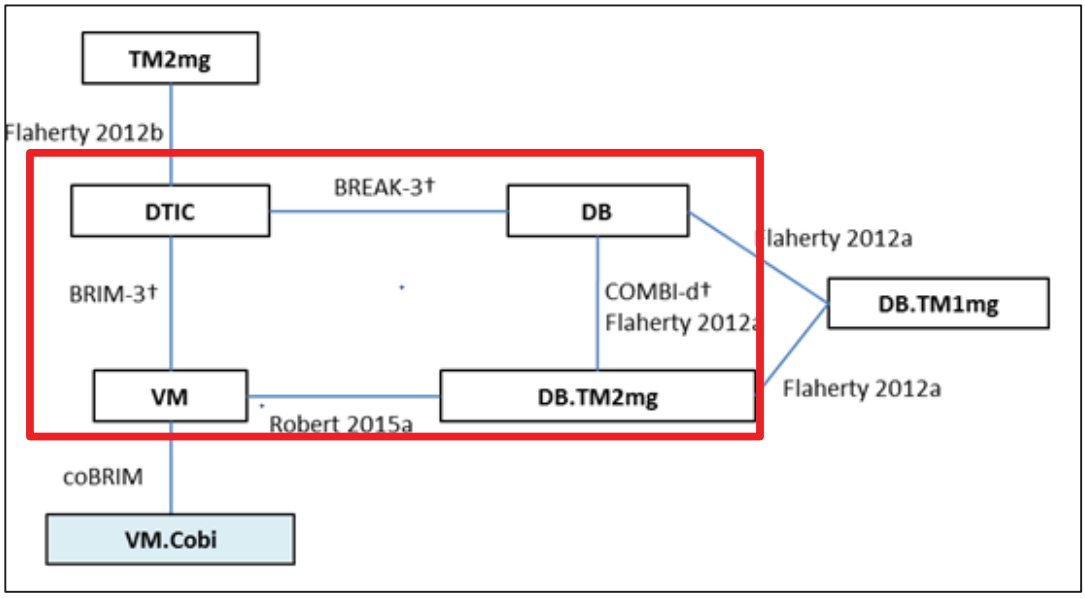
Heterogeneity is only briefly discussed in the CS. The submission states that each pairwise comparison was informed by a single trial, and therefore it was not possible to assess heterogeneity across studies per comparison (though as noted above, there was more than one study for the comparison with trametinib and dabrafenib combination therapy versus dabrafenib). The submission does not comment on conceptual heterogeneity amongst the trial network as a whole. In the ERG's assessment, the clinical trials in the network are broadly similar, as judged on the characteristics of the trials provided in CS Appendix 5 (Table 74). All trials included patients with unresectable stage III or IV melanoma with BRAF mutation, previously untreated with a BRAF or MEK (MAP extracellular signal-regulated kinases) inhibitor; generally had an ECOG status of 0 or 1 (approximately 60-70% of patients were classed as ECOG 0 across the trials); had adequate organ function (liver, kidney, cardiac); and the median age ranged from 49-58 years. These are selected trial characteristics from CS Table 74, however, there is no discussion in the CS about which of these factors (or others) may be regarded as treatment effect modifiers in patients with BRAF mutation positive advanced melanoma. Expert clinical advice to the ERG notes that LDH is an important prognostic and predictive factor, and that there was a higher proportion of patients with raised LDH (and therefore with an adverse prognosis) in the coBRIM trial than in the other major comparator studies (NB. The ERG has not checked the LDH values for the trials in the NMA as these data have not been provided in the CS).

The NMA was conducted according to the principle of a fixed effect. This assumes that the true treatment effect is common in all studies comparing the same treatments. In contrast, a random-effects model assumes that each study has its own true treatment effect, because study characteristics and the distribution of patient-related effect modifiers differ across studies. The submission reports that a fixed-effect model was chosen based on an assessment of model fit using the Deviance Information Criterion (DIC). The DIC is a standard method for assessing model fit in Bayesian models. The submission also justifies the use of fixed-effect because the network is small with a limited number of studies. A random-effects model, it is suggested in the

submission, would provide a poor estimate of the distribution of intervention effects (CS section 4.10.17). Other than this brief comment there is little further discussion about the assumptions to support the use of a fixed rather than a random-effects model. The issue of heterogeneity of effects across the studies is given little attention. Whilst use of the DIC to assess model fit is appropriate, the choice of model should also be guided by the plausibility of model assumptions. Methodological guidelines state that the assumptions of random-effects models are much more plausible than of fixed-effect models[27]. Notwithstanding the ERG's assessment above, that the studies in the network appear generally similar in selected patient and methodological characteristics, the submission provides little discussion of factors that might influence the distribution of effects between studies to support the justification of a fixed-effect model. Upon request from the ERG the company provided the random-effects NMA results for PFS and OS but only pairwise comparisons between dacarbazine and comparator drugs (clarification question A7). Whilst these results are not for all comparators in the NMA they illustrate that the point estimates for random-effects and fixed-effects were similar, though credible intervals (CrIs) for the former were wider (as would be expected), and in some cases (e.g. dacarbazine versus vemurafenib + cobimetinib) included 1, indicating greater uncertainty. Taking this into account, along with the recommendations of methodological guidelines,[27] the ERG's view is that the random effects model should be the primary analysis. However, we have not been able to incorporate random effect estimates into our exploratory economic analyses (Section 4.4) as the company has only provided data for comparisons with dacarbazine.

3.1.7.3 Assessment of consistency

The submission states that there is one closed independent loop in the network (where comparisons are informed by both direct and indirect evidence) incorporating four treatments (dacarbazine; vemurafenib monotherapy; dabrafenib monotherapy; dabrafenib and trametinib 2mg). This is illustrated by the box on the left hand side in CS Figure 14, reproduced here in Figure 4).



TM2mg = trametinib 2mg; DTIC = dacarbazine, DB = dabrafenib, VM = vemurafenib, DB.TM2mg = dabrafenib + trametinib 2mg; DB.TM1mg = dabrafenib + trametinib 1mg; VM.Cobi = vemurafenib + cobimetinib

Figure 4 Illustration of the network meta-analysis presented in the company submission showing closed loop (CS Figure 14).

The Bucher method was used to assess consistency between direct and indirect evidence in the network.[28] This method estimates inconsistency as the difference between direct and indirect evidence in each closed loop of the network. The z-test[29] was used to assess the assumption of consistency. The submission states that there was no evidence of inconsistency based on z scores. The ERG considers that this method for assessing inconsistency is appropriate for a network of this type. Upon request from the ERG the company provided PFS and OS 1/AFT estimates for the indirect and the direct evidence comparisons (in the closed loop – the box in Figure 4) as derived from the Bucher method, to allow comparison of consistency (clarification response A8, Tables 16 and 17). The ERG notes that there is slight variation in the point estimates between the direct and the indirect evidence comparisons. Furthermore in some cases, notably for OS, the CrIs around the indirect point estimates included 1, in contrast to the direct evidence comparison where the corresponding intervals did not include 1. There were also occurrences of the opposite, thereby indicating no consistent pattern.

The clarification response reported absolute error values (differences between direct and indirect evidence on a log scale), and associated variance and p-values, which did not demonstrate evidence of inconsistency.

A key assumption in the NMA is that the dacarbazine or paclitaxel treatment node is considered to interact in the same way as dacarbazine in other trials (CS Page 79). (NB. refer to the top left hand corner of
Figure 3 for a graphic illustration of this node, marked 'DTIC', in relation to the rest of the network). This assumption is necessary to allow the trametinib monotherapy arm of the trial by Flaherty and colleagues 2012b[22] to connect to the network, via its chemotherapy comparator arm (patients in this arm could receive either dacarbazine or paclitaxel). The node combines the dacarbazine or paclitaxel arm from this trial with the dacarbazine arms of the BREAK-3 and BRIM-3 trials. The CS does not cite any evidence in support of this assumption.

In response to a clarification question from the ERG the company cited phase II trial results showing comparable efficacy of paclitaxel to dacarbazine in the treatment of melanoma (clarification question A6). At the request of the ERG the company provided a sensitivity analysis in which the trial by Flaherty and colleagues 2012b[22] was removed from the NMA. These results were very similar to the base case results of the NMA. The ERG is therefore satisfied that the assumption that the dacarbazine or paclitaxel treatment node is considered to interact in the same way as dacarbazine in other trials, and does not introduce bias into the network. Furthermore, expert advice to the ERG suggested that dacarbazine and paclitaxel can be regarded as clinically similar.

3.1.7.4 Analysis methods in the NMA

A Bayesian framework was used for the NMA, though there is little elaboration on how the model was constructed and analysed. For example, there is no mention of the choice of the prior probability distribution.

Though proportional hazards (PH) models are common in cancer, the CS reports results from an accelerated failure time model (AFT) for the NMA. The ERG therefore investigated which approach (AFT or PH) appears to be better justified given the properties of the underlying data

and the assumptions of each model, and whether the use of the AFT was likely to introduce any bias in favour or against any of the treatment alternatives under consideration.

Whilst PH models do not require specification of any form of the baseline hazard function, they do rely on the assumption of proportionality in hazards.[30] This assumption may not hold in some survival studies, and using PH models in situations where the PH assumption may not hold could result in serious bias.[30] Accordingly, the CS cites a recent NMA (sponsored by Roche and presented as a conference poster)[31], which found that the AFT model, which allows for nonproportional hazards over time, provides a better fit for metastatic melanoma drug trials as compared to the PH model (NB. vemurafenib + cobimetinib combination therapy was not included in this NMA as the coBRIM trial appears to pre-date the analysis). However, it needs to be noted that AFT models are predicated on the choice of a statistical distribution (here lognormal) and that covariates impact on survival by a constant factor. The literature states that parametric methods, such as AFT, may work better than PH models if the functional form of the distribution has been determined appropriately.[32]

In order to test the assumptions of the PH and AFT models, the CS reports log-cumulative hazards plots for PH and Q-Q plots for AFT (CS Appendix 6). The ERG agrees that cumulative hazards plots indicate a violation of the PH assumption, though it is not entirely clear whether Q-Q plots indicate a much better fit of the AFT model. Nevertheless, the ERG concluded that there is no reason to assume that the use of the AFT model is inappropriate in this particular context. Though PH models are more widely reported in the medical literature, AFT models appear to be adequate for the purpose of analysing survival data, and they may even perform better than PH models if the assumption of proportional hazards is violated.

The CS does not state whether the data included in the NMA were adjusted to take into account patient crossover between trial arms following disease progression. If crossover occurred this would lead to underestimation of any differences in effects between treatments, and thereby could disadvantage the experimental intervention. Methods have been developed which adjust for the effect of crossover, subject to plausible assumptions and the availability of mature outcome data.[33] These include the Rank Preserving Structural Failure Time method, the Inverse Probability of Censoring Weighting method, and the Iterative Parameter Estimation method. There is academic debate about which method is optimal for adjusting for crossover.[34] Crossover was not permitted in the coBRIM trial, but it was allowed in the BREAK-3[21] and

BRIM-3[18] trials which compared dabrafenib and vemurafenib (respectively) to dacarbazine. The proportions of patients who crossed over from dacarbazine to targeted BRAF inhibitor therapy were 57% and 34% in the BREAK-3 and BRIM-3 trials, respectively.[34] From the outcome data provided in CS Table 75 (Appendix 5) it appears that unadjusted estimates for PFS and OS were used in the NMA for these trials. Therefore, the respective effects of dabrafenib and vemurafenib compared against dacarbazine in the NMA are likely to have been underestimated. In the current appraisal this is of most significance for the comparison of dabrafenib against vemurafenib + cobimetinib where only indirect evidence is available. The ERG is unable to adjust the analyses to take into account patient crossover, but considers that unadjusted estimates which fail to account for crossover would disadvantage dabrafenib compared to vemurafenib + cobimetinib. The ERG explores this by providing a scenario analysis in which the clinical effectiveness of dabrafenib and vemurafenib is equal (Section 4.4.2).

3.1.7.5 Summary of the ERG appraisal of the NMA

The strengths of the NMA include:

- A comprehensive literature search to identify relevant evidence. The ERG is not aware of any additional studies that could be included.

The potential limitations of the NMA include:

- The network is sparse with each comparison of relevance to this appraisal informed by only one trial.
- The risk of bias was judged by the company to be high in several of the included trials, however, there is uncertainty about the reliability of the company's quality assessment judgements.
- There is little discussion of conceptual heterogeneity between the trials in the network, and whether there are differences between the trials in potential effect modifiers. The ERG's assessment of the data provided is that the trials are broadly comparable, based on selected characteristics presented.
- It appears that estimates for dabrafenib and vemurafenib from the BREAK-3 and BRIM-3 trials have not been adjusted to take into account the effect of patient crossover in those trials, likely underestimating the effect of these treatments compared to dacarbazine, with a consequent influence on their comparisons with vemurafenib + cobimetinib.

3.2 Summary statement of company's approach to evidence synthesis

Overall, the CS is reasonable in its quality of methodology and reporting, although the company does not appear to have applied quality assessment criteria consistently (Table 7).

Table 7 Quality assessment (CRD criteria) of CS review

CRD Quality Item: score Yes/ No/ Uncertain with comments	
1. Are any inclusion/exclusion criteria reported relating to the primary studies which address the review question?	Yes, consistent with the decision problem. For the NMA the population was restricted to patients treatment-naïve for metastatic disease. In response to a clarification request by the ERG (clarification question A9) the company stated that this restriction was to ensure study populations in the NMA were consistent with the coBRIM trial and the comparator studies.
2. Is there evidence of a substantial effort to search for all relevant research? i.e. all studies identified	Yes. The searches for studies in the NMA were approximately 1 year out of date. The ERG updated this search but did not identify any additional relevant evidence.
3. Is the validity of included studies adequately assessed?	Uncertain. Whilst the methods used to assess validity were appropriate, the conclusions about validity were different for the coBRIM trial depending on whether its validity was assessed in the CS as a single RCT or whether its validity was assessed as part of the NMA process. Most of the trials in the NMA were classed as being at high risk of bias, but the justification for this conclusion is unclear.
4. Is sufficient detail of the individual studies presented?	Yes for the pivotal coBRIM trial, but limited characteristics of patients in the trials included in the NMA were reported (only sex, age and performance status).
5. Are the primary studies summarised appropriately?	Yes, since only one RCT met the inclusion criteria a summary across primary studies is not applicable.

Eligibility criteria for the company's systematic review of clinical effectiveness are reported in CS Table 6 (page 38) and are appropriate for the decision problem. However, in their systematic review of studies for the network meta-analysis (NMA) the company limited inclusion criteria to patients who were treatment naïve when presenting with metastatic disease (CS Table 19, pages 74-75), without providing a justification.

The CS presents an extensive search for all relevant evidence using three major bibliographic databases (MEDLINE, Embase, and the Cochrane Library), and scrutiny of the proceedings of three international oncology conferences (CS page 37) (for complete details see Section 3.1.1 of this report). The systematic review of studies for the NMA was conducted in April 2015 and so nearly a year out of date. However, as stated earlier (Section 3.1.7) an update of this search by the ERG did not identify any additional relevant evidence. The clinical effectiveness review methods are clearly reported and appear adequate for minimising the risks of errors and bias in the evidence synthesis process (although the rationale for conclusions regarding risk of bias is unclear – see below). Study selection was conducted by two independent reviewers, whilst the data extraction and quality assessment steps were each conducted by one reviewer with checks made by a second reviewer.

The validity of the single included clinical effectiveness RCT (i.e. the coBRIM trial) was assessed using the CRD criteria (CS Table 15, pages 60 - 61). The validity of the eight trials included in the NMA was also assessed by the company using CRD criteria (CS Appendix 5, pages 204 - 208). Although the CRD questions are appropriate, the CS gives inconsistent answers for the coBRIM trial when it was assessed as the primary clinical effectiveness trial (Table 15 – overall low risk of bias) and when it was assessed as a contributor to the NMA (Appendix 5 – overall unclear risk of bias).

In Appendix 5 the company concludes that six of the eight trials included in the NMA were at high risk of bias but the CS does not provide a clear justification for this conclusion, and the CS does not mention study quality when considering the results of the NMA.

3.3 Summary of submitted evidence

3.3.1 Summary of results for progression-free survival

The CS (CS Table 16, page 63) presents results for three PFS analyses. These are the planned primary analysis of investigator-assessed progression with a data cutoff of 9 May 2014; a planned secondary analysis with the same cutoff date but with progression assessed by an independent review facility; and an unplanned post-hoc analysis of investigator-assessed progression based on a data cutoff of 16 January 2015. The planned analyses are specified in the published study protocol,[12] whilst the post-hoc analysis is reported in a conference presentation[15] (Table 8). According to the CS (page 52), no interim analyses of PFS were planned or performed. Median follow-up in the vemurafenib + cobimetinib group was 7.3 months in the pre-planned analysis and 14.9 months in the post-hoc analysis.

The planned and post-hoc analyses of PFS all favour vemurafenib + cobimetinib over vemurafenib + placebo, with HRs that are statistically significant (p<0.001 and/or the HR 95% confidence interval excludes 1). The median survival difference between the cobimetinib and placebo arms ranged from 3.7 months (HR=0.51; primary analysis, investigator-assessed progression) to 5.3 months (HR=0.60; secondary analysis, independent review board assessed progression). The primary analysis gives a more conservative estimate of PFS than the independent review board assessment, though both were reported to have been blinded to treatment allocation. The former analysis is preferred by the ERG as it was designated as the primary analysis. Results of the post-hoc update analysis are similar to those of the planned secondary analysis (Table 8), although did not include a p value.

Table 8 PFS: primary, secondary and updated analyses

Analysis	Vemurafenib + cobimetinib (n=247)	Vemurafenib + placebo (n=248)
Primary: PFS assessment by investigator - Data cutoff 9 May 2014[12]		
PFS events, n (%)	Not reported	Not reported
Median follow-up, months	7.3	7.3
Median duration, months (95% CI)	9.9 (9.0 – NR)	6.2 (5.6 – 7.4)
HR for death or disease progression	0.51 (0.39 – 0.68);[a]	Reference

(95% CI)	p<0.001	
Secondary: PFS assessment by independent review facility - Data cutoff 9 May 2014[12]		
PFS events, n (%)	Not reported	Not reported
Median follow-up, months	Not reported	Not reported
Median duration, months (95% CI)	11.3 (8.5 to NR)	6.0 (5.6 – 7.5)
HR for death or disease progression (95% CI)	0.60 (0.45 – 0.79);[a] p<0.001	Reference
Post-hoc update: PFS assessment by investigator - Data cutoff 16 January 2015[15]		
PFS events, n (%)	143 (57.9)	180 (72.6)
Median follow-up, months	14.9	13.6
Median duration, months (95% CI)	12.3 (9.5 – 13.4)	7.2 (5.6 – 7.5)
HR for death or disease progression (95% CI)	0.58 (0.46 – 0.72)[a]	Reference

NR, Not reached
[a] Analysis was stratified according to geographic region and metastasis classification.

3.3.1.1 Sensitivity analyses on PFS

The following sensitivity analysis results (Table 9) are not reported in the CS; they were provided by the company in response to a clarification request from the ERG. These analyses were conducted based on the primary PFS analysis of investigator-assessed progression with a cutoff date of 9 May 2014. The CS states that the company planned three sensitivity analyses: a non-stratified analysis; an analysis in which patients who received non-protocol therapy were censored; and an analysis in which patients with missed visits were censored. The company's response to the clarification request shows that the second analysis (censoring for non-protocol anti-cancer therapy) was in fact split into stratified and non-stratified analyses (Table 9). The company's response also states that the proposed third sensitivity analysis (censoring accounting for missed visits) was not conducted due to low patient numbers.

Table 9 Results of sensitivity analyses on PFS

Median PFS (95% CI)	Cobimetinib + vemurafenib (n=247)	Vemurafenib + placebo (n=248)	HR for progression (95% CI)
Reference analysis (from	9.9 (9.0 – NE)	6.2 (5.6 – 7.4)	0.51 (0.39 – 0.68);

Table 8 above)			p<0.001
Non-stratified analysis	9.9 (9.0 – NE)	6.2 (5.6 – 7.4)	0.51 (0.39 – 0.68[b]); p<0.001
PFS censored for non-protocol anti-cancer therapy: stratified analysis[a]	11.1 (9.0 – NE)	6.2 (5.6 – 7.4)	(0.51 (0.38 – 0.67); p<0.0001
PFS censored for non-protocol anti-cancer therapy: non-stratified analysis	11.1 (9.0 – NE)	6.2 (5.6 – 7.5)	(0.51 (0.38 – 0.68); p<0.0001

NE, Not evaluable

[a] stratification by geographical region and metastasis classification

[b] Company clarification text states upper limit of 95% CI is 0.89 but company clarification table states upper limit is 0.68

As shown in Table 9, the non-stratified analysis had no impact on the results. The analyses censoring for non-protocol anti-cancer therapy increased median PFS in the vemurafenib + cobimetinib group only, but this has no discernible impact on the hazard ratio. The ERG would expect that varying these analyses would have influenced the hazard ratios and their 95% CIs, but the company provided no explanation for the results being nearly identical.

3.3.1.2 NMA results for PFS

As discussed earlier (Section 3.1.7.4) the NMA estimated the effect of the comparisons using an AFT model. Results are expressed as the inverse of the acceleration factor (1/AFT). Accordingly:

- 1/AFT of 1 indicates there is no difference between the treatment and control
- 1/AFT < 1 favours treatment
- 1/AFT >1 favours control

CS Table 22 provides results for PFS for all comparators in the NMA. The fixed-effect 1/AFT for vemurafenib + cobimetinib compared to dabrafenib was 0.60 (95% CrI 0.46 to 0.77) and compared to vemurafenib it was 0.63 (95% CrI 0.53 to 0.76), indicating that combination therapy is more beneficial in terms of PFS than both comparators.

3.3.2 Summary of results for overall survival

The CS (pages 64 - 65) presents results for three OS analyses. These are the first interim analysis with a data cutoff of 9 May 2014; the second interim analysis with a data cutoff of 16 January 2015; and the planned final analysis which had a data cutoff of 28 August 2015 (Table 10). Although the CS specifies OS as being a secondary endpoint (CS page 64), criteria are given for ensuring an adequate sample size for 80% statistical power: the final analysis of OS was to be performed after approximately 385 deaths had occurred (see ERG section 3.1.6). The ERG notes that the total number of events which had occurred in both combined study groups at the final analysis data cutoff was 255 (Table 10), meaning that this outcome analysis would likely have been underpowered.

Table 10 OS: interim and final analyses

Analysis	**Vemurafenib + cobimetinib (n = 247)**	**Vemurafenib + placebo (n = 247)**
Interim analysis: Data cutoff 9 May 2014[12]		
OS at 9 months, % (95% CI)	81 (75 - 87)	73 (65 - 80)
Median OS duration, months (95% CI)	NR	NR
HR for death (95% CI)	0.65 (0.42 – 1.00)[a]	Reference
p-value	0.046	Reference
Interim analysis: Data cutoff 16 January 2015 (CS Table 17 and draft SmPC)		
OS at 12 months, % (95% CI)	74.9 (69.3 – 80.5)	63.0 (56.8 – 69.3)
Median OS duration, months (95% CI)	NE (20.7 – NE)	17.0 (15.0 – NE)
HR for death (95% CI)	0.65 (0.49 – 0.87)[a]	Reference
p-value	Not reported	Not reported
Final analysis: Data cutoff 28 August 2015[13]		
OS events, n (%)	114 (46.2)	141 (56.9)
Median OS duration, months (95% CI)	22.3 (20.3 - NE)	17.4 (15.0 - 19.8)
Hazard ratio for death (95% CI)	0.70 (0.55 - 0.90)[a]	Reference
p-value	0.005	Reference

NE = not evaluable; NR = not reached.
[a] Patients were stratified according to geographic region and metastasis classification
NB HR for updated analysis (16 January 2015) 0.65 (95% CI 0.49 – 0.87); reported in CS but source unclear

All three OS analyses appear to favour vemurafenib + cobimetinib over vemurafenib + placebo (Table 10). In the published interim analysis,[12] the proportion surviving to 9 months was marginally higher (8%) in the vemurafenib + cobimetinib group, and although the HR appears favourable (0.65) it has high uncertainty, with an upper CI of 1. In the final analysis, despite the planned number of events not being reached, meaning that the analysis was likely underpowered, patients in the vemurafenib + cobimetinib group had a median survival 4.9 months longer than in the vemurafenib + placebo group with the HR (0.70) being statistically significant (p = 0.005).

3.3.2.1 Sensitivity analyses on OS

The CS does not mention sensitivity analyses for OS. However, a sensitivity analysis on OS was provided in the company's response to the ERG's clarification request about PFS analyses (clarification request question A2). This compared stratified and non-stratified analyses for investigator-assessed OS with a data cutoff of 28th August 2015 (i.e. the planned final analysis data set). This analysis had no impact on the results other than to slightly alter the 95% CI of the HR (stratified: 0.55 – 0.90; unstratified: 0.54 – 0.89). The company's clarification also stated that a planned sensitivity analysis of OS censored for subsequent anti-cancer therapy was not conducted because there were no patients in the vemurafenib + placebo group who had crossed over to cobimetinib at the time of the final OS analysis. The reason for this statement seems unclear, given that a similar analysis was reported by the company for PFS and also that crossovers were not permitted, and therefore would not have been expected, in the coBRIM trial.

3.3.2.2 NMA results for OS

CS Table 21 provides results for OS for all comparators in the NMA. The fixed-effect 1/AFT for vemurafenib + cobimetinib compared to dabrafenib was 0.63 (95% CrI 0.47 to 0.86), and compared to vemurafenib it was 0.73 (95% CrI 0.59 to 0.90), indicating that combination therapy is more beneficial in terms of OS than both comparators.

3.3.3 Summary of results for response rates

The CS (pages 65 - 67) presents results for two analyses of response rates. These are for the published planned analysis which had a data cutoff of 9 May 2014 and a post-hoc updated analysis with a data cutoff of 16 January 2015 (Table 11). The proportion of patients who experienced an objective (i.e. complete or partial) response was significantly higher in the vemurafenib + cobimetinib group (68%) compared to the vemurafenib + placebo group (45%), although the ERG notes that the majority of these patients had a partial rather than complete response.

Table 11 Response rates outcomes: primary and updated analyses

Analysis	**Vemurafenib + cobimetinib (n = 247)**	**Vemurafenib + placebo (n = 248)**
Planned analysis: Data cutoff 9 May 2014[12]		
Complete response, n (%) Partial response, n (%) Stable disease, n (%) Progressive disease, n (%) No complete response or progressive disease, n (%) Could not be evaluated, n (%)	25 (10) 142 (57) 49 (20) 19 (8) 0 12 (5)	11 (4) 100 (40) 105 (42) 25 (10) 1 (<1) 6 (2)
Complete or partial response Percent of patients (95% CI) p-value	(n = 167) 68 (61 – 73) <0.001	(n = 111) 45 (38 – 51) Reference
Median duration of response, months (95% CI)	NR (9.3 – NR)	7.3 (5.8 – NR)
Post-hoc updated analysis: Data cutoff 16 Jan 2015[15]		
Complete response, n (%) Partial response, n (%)	39 (16) 133 (54)	26 (11) 98 (40)
ORR (complete or partial response) Percent of patients (95% CI) Difference in ORR, % (95% CI)	(n = 172) 69.6 (63.49 - 75.31) 19.64 (10.95-28.32)	(n = 124) 50.0 (43.61 - 56.39) Reference

Duration of response (DOR)		
Patients with event, n (%)	84 (48.8)	73 (58.9)
Median DOR, months (95% CI)	12.98 (11.10 - 16.62)	9.23 (7.52 - 12.78)
Range	2.86 - 20.11	1.77 - 17.68

NR, Not reached.
[a] Response could not be evaluated for patients who withdrew consent, were withdrawn by site investigator, died or started new anticancer therapy before the first tumour assessment.

3.3.4 Summary of results for health related quality of life

The CS presents a brief, narrative description of patients' HRQoL (CS pages 67 - 69), based on the EORTC QLQ-C30 as applied within the coBRIM trial, and reported in a poster by Dréno and colleagues.[19] The only quantitative HRQoL results reported in the poster are line charts showing percentage changes from baseline in different subscales of the QLQ-C30, and a histogram showing the percentage of patients with clinically meaningful improvement in scores for each QLQ-C30 subscale.[19] Only the histogram is included in the CS and this is reproduced in Figure 5 below. The CS points out that the analysis was considered exploratory, with patients deemed to have experienced a clinically meaningful improvement in HRQoL if the QLQ-C30 score improved by at least 10 points (a ≥10-point increase in scores for global health status and functioning scales, or a ≥10-point decrease in scores for symptom scales) at one or more post-baseline assessments. The CS does not clarify whether this 10-point threshold is arbitrary or based on validation studies.

The CS points out that the QLQ-C30 scores do not capture patients' HRQoL post disease progression since few patients completed the QLQ-C30 questionnaire upon discontinuation of study treatment (CS page 69). However, it is not reported how many patients did complete the questionnaire upon discontinuation and whether their responses were analysed.

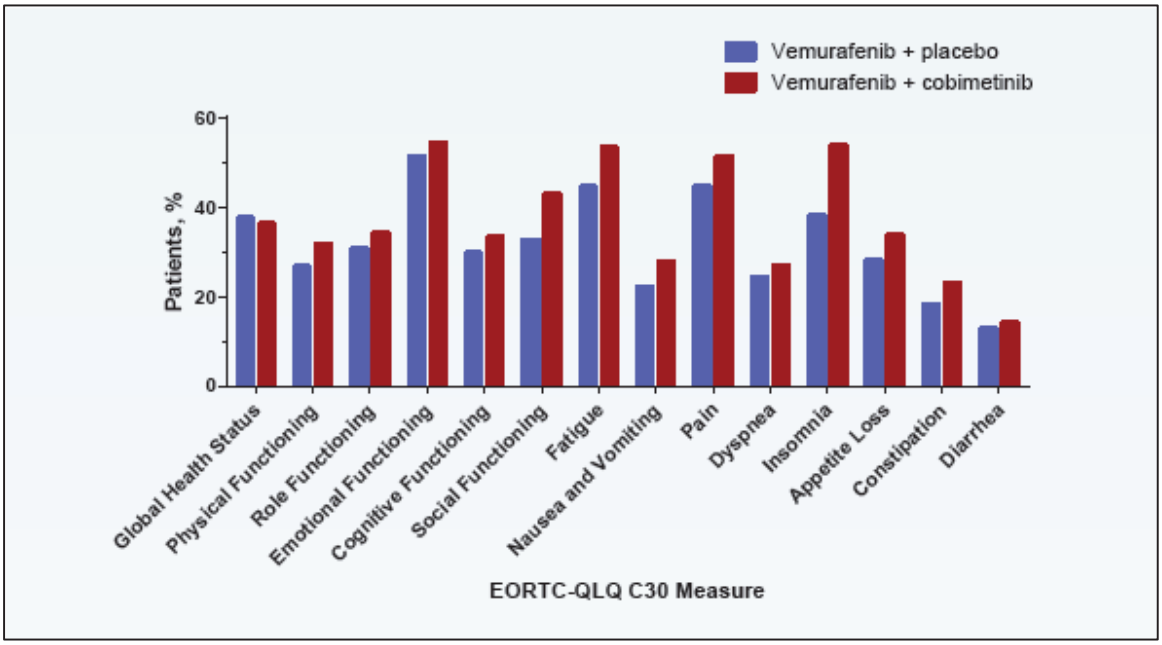


Figure 5 Clinically meaningful improvement in HRQoL (QLQ-C30 score change ≥10 points relative to baseline)

As shown in Figure 5, for many scales of the QLQ-C30 the frequency of patients with clinically meaningful improvement in the QLQ-C30 subscales was higher in the vemurafenib + cobimetinib group, with the largest difference being for the insomnia subscale. However, it is unclear whether these differences between study groups would be statistically significant and the analysis provides no indication of the durability of these improvements across the course of patients' therapy.

The CS concludes that patients experienced clinically meaningful improvement in insomnia relative to baseline only in the vemurafenib + cobimetinib group. The CS also concludes that patients in the vemurafenib + cobimetinib group experienced clinically meaningful worsening of diarrhoea from baseline to C2D15 but not in subsequent cycles, without any concomitant change in diarrhoea scores in the vemurafenib + placebo group. These conclusions are consistent with line charts presented in the poster by Dréno and colleagues[19] (not reproduced here). However, the ERG cautions that these results are uncertain since all comparisons reported in the line charts[19] have very wide (unspecified) error bars. Also, the number of patients providing QLQ-C30 scores at each sampling point is not reported.

In response to a request by the ERG (clarification question B2), the company provided details of the EQ-5D values obtained from the coBRIM trial. Utilities were collected in coBRIM using the EQ-5D-5L and were mapped to their equivalent EQ-5D-3L values using a crosswalk method. The ERG considers this approach to be appropriate, as the EQ-5D-3L is widely used and is the standard utility estimation method in other related NICE technology appraisals. The EQ-5D was completed mostly by patients who had not progressed (see Section 4.3.6.1 of this report).

In their clarification response, the company provided the estimates of mean and median EQ-5D-3L scores and the 95% CI of the mean for day 1 of the first and second cycles and then for day 1 of every other cycle up to cycle 22, as well as for the end of study treatment (clarification response Table 19). The mean utility estimates were summarised graphically (Figure 6 in the clarification response) and are reproduced below in [redacted].

[redacted]

The ERG notes that the HRQoL analysis in coBRIM would be statistically under-powered for detecting differences between the treatment arms and according to the company's response (clarification Table 19), the 95% CI of the mean utility estimates for both arms (not shown in [redacted]) would overlap for all sampling times. However, there appears to be a trend for improved HRQoL in the vemurafenib + cobimetinib arm from cycle 2 to cycle 18. These results

are increasingly uncertain after cycle 18 due to the low number of utility estimates available (for vemurafenib + cobimetinib these are n=27 at C18D1; n=9 at C20D1, and n=2 at C22D1).

Overall, taking together the QLQ-C30 and EQ-5D results there appears to be some pre-progression HRQoL benefit associated with vemurafenib + cobimetinib compared to vemurafenib + placebo, but this is based on an analysis that was unable to demonstrate statistical significance. It is unclear whether the slightly higher utility estimates in the vemurafenib + cobimetinib arm reflect HRQoL improvements resulting from less insomnia and/or other factors (e.g. the incidence of non-melanoma skin cancers which was lower in the vemurafenib + cobimetinib arm). The trend for slightly higher pre-progression utilities in the vemurafenib + cobimetinib arm, and the uncertainty associated with these, is captured in the economic model through these utility estimates and their 95% CIs (see Section 4 of this report).

3.3.5 Summary of results for sub-group analyses

The CS reports results of 12 subgroup analyses (as listed above, section 2.3). These are presented as forest plots showing nine subgroup analyses for PFS (CS Figure 9, page 71) and three subgroup analyses for OS (CS Figure 10, page 71) (the CS does not justify why the same subgroups were not analysed for each outcome). Hazard ratios are consistently below 1.0 for all the subgroup analyses reported, although for some subgroups with small sample sizes the 95% CI includes 1.

Overall, all the subgroups analysed gave broadly similar responses to the vemurafenib + cobimetinib therapy when compared against vemurafenib + placebo. However, given that the 12 subgroup analyses were split (9 analysed for PFS, 3 analysed for OS) rather than all subgroups being analysed for both outcomes, there is some uncertainty as to whether these results are fully representative and generalisable across both outcomes.

3.3.6 Summary of adverse events

The safety population was analysed according to the study treatment received; however, eight patients assigned to the vemurafenib + placebo group received investigational cobimetinib instead of placebo as a result of dispensing errors. Two patients (one in each study group) were excluded from the safety analysis as they did not receive the assigned study drug.

The CS reports AEs according to their frequency (CS Table 23, page 86), consistent with the trial journal publication,[12] and also lists those AEs that led to discontinuation or dose modification (CS Table 24, pages 87 - 88), partly consistent with a poster by Dréno and colleagues.[14] In response to a clarification request by the ERG the company provided: a table showing the most frequent serious AEs; a table showing grade ≥3 AEs that had the largest differences (≥2%) between the study arms; and a table showing the primary reasons why patients discontinued from the coBRIM trial. Due to the different ways these AEs are grouped in each table it is not always easy to see clear patterns. We have reproduced these tables below, considering first the frequency of occurrence of AEs and then the factors that led to patient discontinuation.

3.3.6.1 Frequency of adverse events and differences between trial arms

The most frequent AEs are shown separately by grade (1 - 4) in Table 12 (taken from CS Table 23, page 86 and consistent with the trial journal publication[12]). No individual grade 4 AEs exceeded 1% frequency, except for grade 4 elevation of creatine kinase which occurred in 4% of patients in the vemurafenib + cobimetinib arm. The most frequent grade 3 AEs were the elevated enzymes: alanine aminotransferase (11% and 6% of patients in the cobimetinib and placebo arms respectively); aspartate aminotransferase (8% and 2% respectively); and creatine kinase (7% and 0% respectively). The most frequent AEs occurring at grades 1 - 2 were diarrhoea, nausea and rash.

In Table 12 the row for patients experiencing “Any adverse event” has a specific interpretation in the trial publication and CS and should be not interpreted to literally mean “any” AE (see Table 12 footnote). If a patient had multiple AEs these are counted in the “Any adverse events” row only once, in the column for which the highest grade of event occurred. For example, in the vemurafenib + cobimetinib arm, the “Any adverse event” row shows that 19 patients (7%) had AEs no greater than grade 1; 66 patients (26%) had AEs up to and including grade 2; 125

patients (49%) had AEs up to and including grade 3; whilst 34 patients (13%) had AEs up to and including grade 4 (Table 12). Only grade 2 AEs were higher in the vemurafenib + placebo group (29%) than those in the vemurafenib + cobimetinib group.

Table 12 Most frequent adverse events (occurring in ≥20% of patients in either group)

Adverse event, n (%)	Vemurafenib + cobimetinib (n = 254)				Vemurafenib + placebo (n = 239)			
	Grade 1	Grade 2	Grade 3	Grade 4	Grade 1	Grade 2	Grade 3	Grade 4
Any adverse event	19 (7)	66 (26)	125 (49)	34 (13)	21 (9)	70 (29)	117 (49)	22 (9)
Diarrhoea	99 (39)	29 (11)	16 (6)	0	51 (21)	16 (7)	0	0
Nausea	75 (30)	22 (9)	2 (1)	0	43 (18)	12 (5)	2 (1)	0
Vomiting	41 (16)	10 (4)	3 (1)	0	21 (9)	6 (3)	2 (1)	0
Rash	55 (22)	29 (11)	13 (5)	2 (1)	46 (19)	27 (11)	12 (5)	0
Photosensitivity reactions	48 (19)	18 (7)	6 (2)	0	25 (10)	12 (5)	0	0
Hyperkeratosis	23 (9)	3 (1)	0	0	49 (21)	14 (6)	5 (2)	0
Fatigue	48 (19)	24 (9)	9 (4)	0	42 (18)	24 (10)	7 (3)	0
Pyrexia	49 (19)	13 (5)	4 (2)	0	43 (18)	10 (4)	0	0
Arthralgia	54 (21)	23 (9)	6 (2)	0	53 (22)	31 (13)	12 (5)	0
Alopecia	33 (13)	1 (<1)	1 (<1)	0	55 (23)	14 (6)	1 (<1)	0
Increased alanine amino-transferase	16 (6)	15 (6)	28 (11)	1 (<1)	17 (7)	11 (5)	14 (6)	1 (<1)
Increased aspartate amino-transferase	17 (7)	18 (7)	21 (8)	0	15 (6)	10 (4)	4 (2)	1 (<1)
Increased creatine kinase	23 (9)	27 (11)	17 (7)	9 (4)	6 (3)	1 (<1)	0	0

Multiple occurrences of a specific adverse event for a patient were counted once at the highest grade of the occurrence (based on the National Cancer Institute Common Terminology Criteria for Adverse Events, version 4.0), e.g. if a patient had two episodes of a specific toxic event, one grade 3 and one grade 4, the patient was counted only once in the grade 4 column. This applied also to "Any adverse events", e.g if a patient had three separate events of grade 1, 3, and 4, the patient was counted only once in the grade 4 column.

The CS states (page 87) that the majority of first grade ≥3 AEs occurred early in treatment, with a median time to onset among patients with grade ≥3 AEs of 0.53 months for vemurafenib + cobimetinib and 0.79 months for vemurafenib + placebo. After the first cycle (28 days), the incidence of grade ≥3 AEs decreased over time. The median time to resolution for grade ≥3 AEs that occurred in the first 28 days was 0.5 months for both arms.

Overall, six deaths were attributed to AEs in the vemurafenib + cobimetinib arm and three deaths in the vemurafenib + placebo arm (CS page 89). In the vemurafenib + cobimetinib arm,

the AE was recorded as the primary cause of death for two patients (cardiac arrest and pneumonia). An additional two patients had primary cause of death recorded as “other” (unexplained; asthenia and fatigue). The remaining two patients had disease progression recorded as the primary cause of death.

Of the three deaths in the vemurafenib + placebo arm, cardiac failure was the reported cause for one patient and disease progression was documented as the cause of death for the other two patients.

Two patients, one in each treatment arm, died as a result of AEs that were considered by the investigator to be related to study treatment: One patient in the vemurafenib + cobimetinib arm (fatigue and asthenia) and one patient in the vemurafenib + placebo arm (cardiac failure).

The most frequent serious adverse events (SAEs) are shown in Table 13 (taken from the company’s clarification response).

[redacted]*

Table 13 [redacted]

[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

A clearer indication of how AEs differed between the cobimetinib and placebo arms of the coBRIM trial is shown in Table 14 (taken from the company's clarification response).

Table 14

Cob: cobimetinib; NR, not reported; Vem: vemurafenib; ↑ elevated
NB. This table has been redrawn from Table 2 and Table 3 in the company's clarification response to show AEs ordered according to their difference between the trial arms.

Although retinal detachment was relatively uncommon, it appears to have only affected patients in the cobimetinib arm (Table 13 and Table 14). However, the company in their clarification response did not provide the frequency of retinal detachment for the latest data cutoff (Table 14).

3.3.6.2 Reasons for treatment discontinuation

The primary reasons for treatment discontinuation, i.e. factors including but not limited to AEs, are not presented in the CS but were provided in the company's clarification response (Question A1) and are reproduced below in Table 15.

Table 15

[redacted]

[redacted]

The AEs which led to patient discontinuation are reported in the CS (CS Table 24, pages 87 - 88), and in a poster reported by Dréno and colleagues,[14] and are reproduced below in Table 16.

Table 16 Adverse events leading to discontinuation or dose modification

AE, n (%) (n=493)	Incidence (all grades)		Discontinuation		Reduction[a]		Interruption	
	Vem + Cob	Vem + placebo	Vem + Cob	Vem + placebo	Vem + Cob	Vem + placebo	Vem + Cob	Vem + placebo
Diarrhoea	144 (57)	67 (28)	0	1 (<1)	6 (2)	0	14 (6)	6 (3)
Photosensitivity	105 (41)	75 (31)	0	0	0	0	3 (1)	1 (<1)
Rash	182 (72)	157 (66)	8 (3)	2 (1)	10 (4)	9 (4)	13 (5)	17 (7)
Elevated CPK	76 (30)	8 (3)	1 (<1)	0	0	2 (1)	4 (2)	0
Liver laboratory value abnormalities	101 (40)	76 (32)	8 (3)	3 (1)	4 (2)	3 (1)	9 (4)	5 (2)
Serous retinopathy	61 (24)	5 (2)	3 (1)	0	2 (1)	0	9 (4)	0
Cutaneous	10 (4)	42 (18)	0	0	0	0	0	3 (1)

neoplasm[b]								

Cob: cobimetinib; Vem: vemurafenib
[a] Interruption or reduction of both drugs for AEs occurred in 42% of patients in the vemurafenib + cobimetinib arm and in 33% of patients in the vemurafenib + placebo arm
[b] Implied in the CS (page 87) that this refers to secondary cutaneous squamous cell carcinoma and keratoacanthoma

The CS lists 7 AEs as leading to discontinuation: diarrhoea, photosensitivity, rash, elevated creatine phosphokinase (CPK), liver laboratory abnormalities, serous retinopathy, and cutaneous neoplasm, as shown in Table 15. These categories are slightly different to those reported elsewhere in the CS and the company's clarification response. It is unclear whether the category "serous retinopathy" includes both retinal detachment and chorioretinopathy.

There appear to be discrepancies between the discontinuations due to AEs reported in the company's clarification response (ERG Table 15), those reported in the CS (ERG Table 16) and a poster presentation reported by Dréno and colleagues.[14]

[redacted] However, in the CS (as reproduced in ERG Table 16) the total number of discontinuations in the vemurafenib + cobimetinib arm is only 20. The poster presentation[14] states that in the vemurafenib + cobimetinib arm 45 patients had an AE leading to drug withdrawal, of which 32 (71%) withdrew both drugs, whereas 10 (22%) and 3 (7%) withdrew cobimetinib or vemurafenib alone, respectively. Similarly, the numbers of patients withdrawing vemurafenib and/or placebo in the comparator arm differs between Table 15 and Table 16 and the poster, which reports 35 patients who had an AE leading to drug withdrawal, of whom 30 (86%) withdrew both drugs, whereas 3 (9%) and 2 (6%) withdrew placebo or vemurafenib alone, respectively.[14] It is unclear whether these differences are related to different analysis times.

According to the poster presentation by Dréno and colleagues,[14] in both arms, most patients who discontinued study treatment for AEs withdrew from both study treatments at the same time. The poster also states that of 10 patients in the vemurafenib + cobimetinib arm who withdrew cobimetinib alone, 6 cases of withdrawal were at least partly a result of serous retinopathy (SR). The poster also states that no AE of any single preferred term led to discontinuation of either drug in >4.0% of patients.

3.4 Overall summary of clinical effectiveness

The ERG considers that the CS presents a generally unbiased estimate of the treatment effect of vemurafenib + cobimetinib combination therapy for patients with advanced melanoma within the stated scope of the decision problem. The company's systematic review of clinical effectiveness followed standard procedures and is of good quality. The ERG is not aware of any additional relevant published trials that could be included. The key RCT, coBRIM, is generally well-designed and provides an appropriate evidence base to inform the assessment of clinical and cost-effectiveness in this appraisal. The trial showed statistically significant differences in favour of vemurafenib + cobimetinib compared to vemurafenib monotherapy in terms of measures of survival and treatment response, with a generally favourable safety profile. The NMA enabled indirect comparisons to be made between vemurafenib + cobimetinib and dabrafenib. This showed that vemurafenib + cobimetinib was more favourable on measures of survival. However, the evidence network was sparse, with only one trial informing each comparison, and there was no discussion of clinical heterogeneity between the trials in the network.

4 COST EFFECTIVENESS

4.1 Overview of company's economic evaluation

The company's submission to NICE includes:

i) a review of published economic evaluations of vemurafenib + cobimetinib compared with vemurafenib and with dabrafenib for patients with advanced melanoma.

ii) a report of an economic evaluation undertaken for the NICE STA process. The cost effectiveness of vemurafenib + cobimetinib is compared with vemurafenib and with dabrafenib in patients with advanced melanoma and the BRAF-mutation.

4.2 Company's review of published economic evaluations

A systematic search of the literature was conducted by the company to identify economic evaluations of cobimetinib for patients with advanced (unresectable or metastatic) BRAF V600 mutation positive melanoma. See section 3.1 of this report for the ERG critique of the search strategy.

The inclusion and exclusion criteria for the systematic review are listed in Appendix 9 of the CS, page 223. The inclusion criteria state that full economic evaluations of BRAF inhibitors (cobimetinib in combination with vemurafenib compared to any other BRAF inhibitor) would be included. The exclusion criteria state that studies that only considered V600 mutation-negative melanoma would be excluded.

Eight studies were identified from screening 114 titles and abstracts. All eight studies were excluded, as none included cobimetinib. No studies therefore were included for full review.

The ERG found one study that compared the clinical effectiveness and cost effectiveness of seven drugs used for the treatment of advanced melanoma in the Norwegian setting[35]. The drugs included cobimetinib, dabrafenib, vemurafenib, ipilimumab, nivolumab, pembrolizumab and trametinib. The study conducted a systematic literature review and compared treatments with NMA using direct and indirect evidence with dacarbazine as a common comparator. The study conducted a probabilistic discrete-time Markov cohort model to compare the cost effectiveness of the treatments over a 10 year time horizon with monthly cycles. Excluding those treatments not included within the scope of this appraisal, vemurafenib was dominated and the ICER of vemurafenib + cobimetinib compared to dabrafenib was 2,837,207 NOK (Norwegian Krone) (£234,000 per QALY).

4.3 Critical appraisal of the company's submitted economic evaluation

4.3.1 NICE reference case

The ERG considered the NICE reference case requirements during the critical appraisal of the submitted economic evaluation (Table 17).

Table 17 NICE reference case requirements

NICE reference case requirements:	Included in submission	Comment
Decision problem: As per the scope developed by NICE	Yes	CS Table 1, page 17
Comparator: As listed in the scope developed by NICE	Yes	Consistent with NICE scope but does not include immunotherapies which have been recommended for this patient group.

Perspective on costs: NHS and PSS (personal social services)	Yes	
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	
Perspective on outcomes: All direct health effects, whether for patients or, when relevant, carers	Yes	
Type of economic evaluation: Cost utility analysis with fully incremental analysis	?	Cost utility analysis. Results not presented incrementally for all comparators, rather separate analyses are presented for the intervention compared to vemurafenib and dabrafenib.
Synthesis of evidence on outcomes: Based on a systematic review	Yes	Inclusion criteria reported in CS Table 6, page 38.
Time horizon: Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes	30 year time horizon
Measuring and valuing health effects: Health effect should be expressed in QALYs. The EQ-5D is the preferred measure of health related quality of life.	Yes	QALYs are used in the analysis. EQ-5D-5L used from the coBRIM trial for PFS and standard gamble used for progressed disease from Beusterien et al.[36]
Source of data for measurement of health related quality of life: Reported directly by patients and/or carers.	?	Patients from the coBRIM trial completed EQ-5D for PFS, but general public sample estimated utility directly for progressed disease.
Source of preference data: Representative sample of the UK population	?	The EQ-5D values are derived using a Crosswalk study of which a small sample are from the UK.
Equity considerations: An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit.	Yes	
Discount rate: 3.5% pa for costs and health effects	Yes	

Notes: ? = uncertain; N/A=not applicable

In general, the company model is in line with the NICE reference case. However there are two aspects to note: firstly, the company has not presented a fully incremental analysis, rather, pairwise analyses have been presented for vemurafenib + cobimetinib compared to

vemurafenib and to dabrafenib; secondly, the source of quality of life used for patients with progressed disease is not in line with the NICE reference case.

4.3.2 Model Structure

The company developed a partitioned survival model with three health states: progression-free survival (PFS), progressed disease (PD) and death. A schematic of the company's model is presented in Figure 6. The model was developed in Microsoft Excel with a 1-week cycle length. Costs, QALYs and life years were presented as outputs of the model. The CS states that the model takes the perspective of NHS England and PSS. The time horizon is set at 30 years and the CS states that this is consistent with prior melanoma appraisals and appropriate for the average age of patients in the model (CS Table 45). Both costs and outcomes are discounted at 3.5% per annum, as recommended by the NICE Methods Guide for Technology Appraisals.[37]

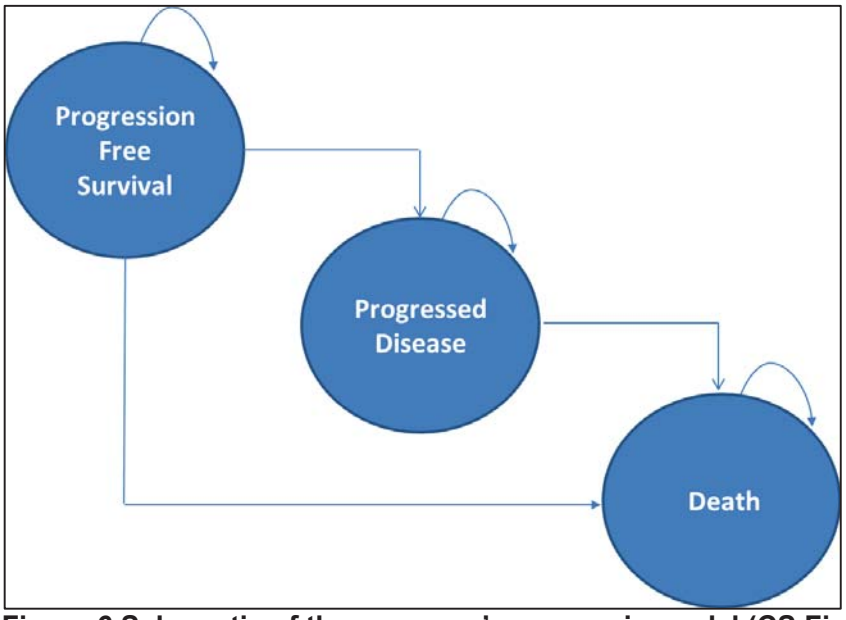


Figure 6 Schematic of the company's economic model (CS Figure 15)

All patients enter the model in the PFS state. Patients progress to the death state according to an estimate of OS from the coBRIM trial and the number of patients remaining in the PFS state is determined by the PFS estimated from the coBRIM clinical trial. The number of patients in the progressed states is then calculated as the difference between the PFS and OS curves. Patients remain on treatment whilst they are in the PFS health state, as per the marketing authorisation for vemurafenib and cobimetinib. Treatment duration is taken from the time on treatment data from the coBRIM study. The model assumes no subsequent lines of anti-cancer

therapy, following progression on the intervention or a comparator. The CS states that there is a lack of data to allow modelling of subsequent treatments. Furthermore, the CS states that the frequency and type of subsequent treatments did not differ by trial arm in coBRIM and that this approach is consistent with prior melanoma technology appraisals. The ERG considered that this was a reasonable approach and that the exclusion of subsequent treatments is unlikely to have an impact on the model results as the time spent in the progressed disease state is similar in all treatment arms. We conduct an illustrative scenario that shows the impact of including subsequent treatment in the economic model in section 4.4.

The CS justifies the model structure and health states by stating that they are closely aligned with the clinical pathway (NICE Melanoma Pathway) and are consistent with approaches taken in earlier technology appraisals for metastatic melanoma (TA269,[26] TA319, [6] TA321,[7] TA366[9]) The ERG agrees that the approach taken and structure adopted is appropriate for the disease pathway. Furthermore the ERG suggests that a survival model is an intuitive structure as it directly utilises the clinical trial survival data. The ERG also considers that the time horizon chosen is appropriate as only a small proportion of patients would still be alive after 30 years (~5% of patients).

4.3.3 Population

The economic model includes treatment-naïve adult patients with unresectable or metastatic BRAF600 mutation positive melanoma. The characteristics of the patients in the model are based upon the patients in the coBRIM trial and are described in CS Table 14. The CS states that the coBRIM study included a small number of UK patients (n=29 from 11 centres) and that the outcomes seen from the study are expected in UK patients (CS Table 45). Expert clinical advice to the ERG was that the patients in the coBRIM trial are similar to those seen in the UK. The patient population is consistent with the marketing authorisation and the population specified in the NICE scope.

4.3.4 Interventions and comparators

The economic model provides two separate analyses: i) vemurafenib + cobimetinib compared to vemurafenib and ii) vemurafenib + cobimetinib compared to dabrafenib. These comparators correspond to NICE's scope. The interventions are implemented in the model in accordance

with their current marketing authorisation and doses (see section 4.3.7). The ERG notes that other potential comparator melanoma treatments have not been included within the NICE scope, for example ipilimumab and pembrolizumab. These are now recommended by NICE to treat advanced melanoma, including BRAF mutation positive melanoma and we have been advised that these treatments are often prescribed for BRAF mutation positive melanoma patients. The ERG also notes that an alternative BRAF inhibitor and MEK inhibitor combination therapy, dabrafenib + trametinib, is currently being appraised by NICE in a separate technology appraisal (ID661).

4.3.5 Treatment effectiveness and extrapolation

For the comparison between vemurafenib + cobimetinib and vemurafenib, the clinical effectiveness estimate is based upon the survival curves from the coBRIM trial. For the comparison between vemurafenib + cobimetinib and dabrafenib, an indirect treatment comparison was conducted via an NMA and the inverse AFT values were applied to the vemurafenib + cobimetinib survival curves to obtain the survival curve for dabrafenib. As discussed earlier in this report (Section 3.1.7), the ERG considered the NMA to be comprehensive in terms of its search for relevant studies, but was limited by a small number of included trials, and uncertainties about heterogeneity between the trials in potential effect modifiers.

4.3.5.1 Progression free survival (PFS)

The company fitted parametric distributions to the observed PFS from the coBRIM trial. The log-logistic, Weibull, lognormal, gamma, Gompertz and exponential distribution were assessed using the Akaike Information Criterion (AkIC), Bayesian Information Criterion (BIC) and visual assessment of each fitted curve against the observed data. The company considered the log-logistic distribution to be the most appropriate based upon AkIC, BIC and visual inspection. The fitted PFS data for both treatment arms compared to the observed data is shown in Figure 7. Alternative distributions are shown in CS Appendix 10 and model results using these distributions shown in Table 27 of the ERG report. The company further justifies the use of the log-logistic by demonstrating that the log-logistic assumption is satisfied (CS Figure 17). It is also possible to use the KM data with a fitted parametric curve. The CS does not report the results from these analyses but the ERG observes that for each parametric distribution, there

was little difference in the results from using a parametric distribution for the whole survival curve or using the KM data with the parametric distribution for the tail.

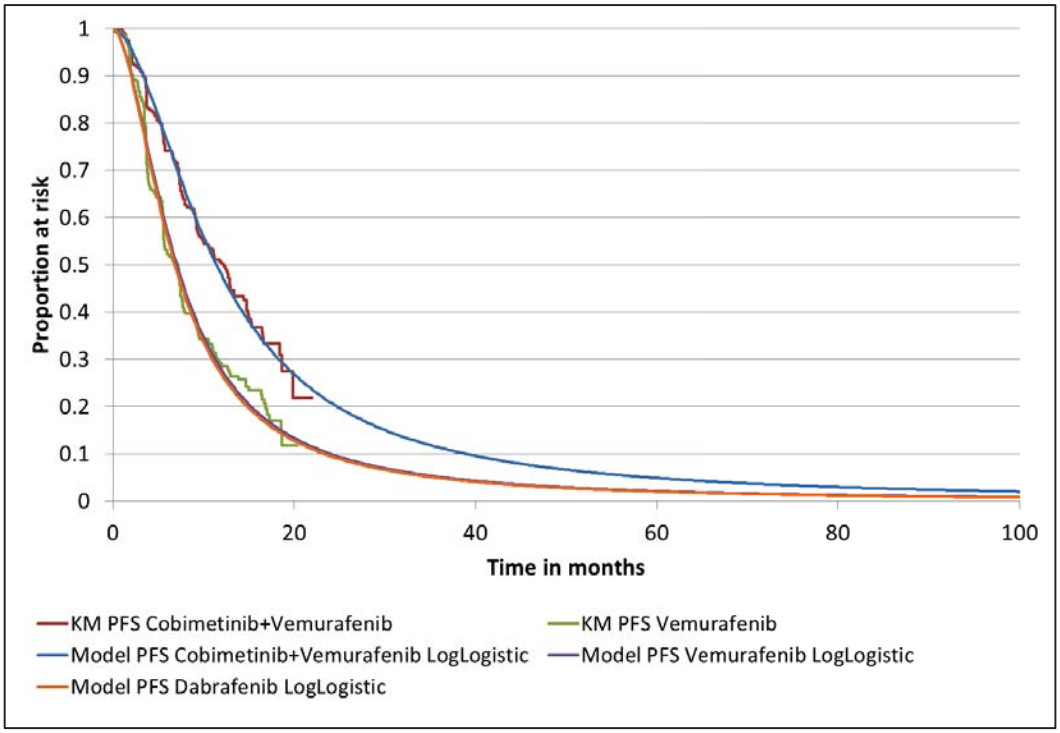


Figure 7 PFS observed data from coBRIM trial and the Log-logistic fitted curves used in the economic model (CS Figure 16)

The ERG considers that the company's choice of parametric distribution is reasonable and provides a good fit to the trial survival data based upon AkIC/BIC and visual interpretation (Figure 7). In addition, the long-term extrapolation appears to be similar to that seen in a phase I trial of vemurafenib where 3 out of 48 patients (6%) were treated for more than four years.[38] We agree that the conditions required for the use of the log-logistic distribution have been satisfied. Furthermore we consider that the lognormal and gamma distributions would produce a similar fit to the log-logistic and result in higher ICERs for vemurafenib + cobimetinib vs vemurafenib of between £157,377 and £164,485 per QALY respectively. The differences between these PFS curves is small and therefore, given the uncertainty in extrapolating beyond the trial data, the ERG considers both these curves would also be a reasonable choice for PFS.

The NMA used the AFT model in preference to a PH model. The ERG's critique of this approach is in section 3.1.7. The PFS for vemurafenib + cobimetinib was adjusted according to

the inverse AFT for the PFS curve for dabrafenib. The inverse AFT results for vemurafenib + cobimetinib vs. dabrafenib were 0.599 (95% CrI 0.47, 0.86) for PFS.

4.3.5.2 Overall survival (OS)

Extrapolation of the OS curve beyond the observed coBRIM trial data was a two stage process. Firstly, the survival data were adjusted, using a 'mixture cure rate methodology', assuming that a small proportion of patients had a low risk of cancer-related death. Secondly, parametric models were fitted to the adjusted trial data.

The CS assumes that the risk of death for patients with metastatic melanoma declines with time, with the risk inversely proportional to the time since diagnosis. The CS cites evidence from the US Surveillance, Epidemiology, and End Results (SEER) cancer registry[39] that shows this decline in risk over time (CS Figure 19). Furthermore, the CS notes that prior NICE melanoma appraisals (TA357[10] and TA366[9]) have also included this assumption.

Using a mixture cure-rate methodology, the overall risk is estimated by combining (via a weighted average) the risk for patients with low risk of cancer related death and those with high risk of cancer related death.

The trial population survival is expressed as S(t),

$$S(t) = S_b(t)\pi + (1-\pi)S_b(t)S_c(t)$$

where the 'cure fraction' is expressed as π, the patients at high risk of cancer-related death [$S_c(t)$], and the patients at low risk [$S_b(t)$].

The 'cure fraction' represents the proportion of patients for whom their disease is stable and the risk of death attributable to cancer is equal to their risk of death from other causes. Therefore the risk of death for these patients is assumed to be the same as for the general population.

The cure fraction is estimated using data from the SEER registry. The resulting cure fraction is [redacted]. The OS curves fitted to coBRIM trial data with and without the cure rate for patients treated with vemurafenib is shown in Figure 8.

The median OS follow-up for the coBRIM trial was 18 months. The ERG was unable to find any longer term studies for patients treated with vemurafenib + cobimetinib. The ERG identified a

study by Puzanov and colleagues[38] that reported 4 year melanoma specific survival rate in patients treated with vemurafenib. The ERG has compared data from this study with the base case OS from the company's model for vemurafenib (which includes the cure rate), as seen in Figure 8. From a visual inspection of this figure, the ERG suggests that this study supports the use of a cure rate and the OS curve from Puzanov and colleagues[38] is seen to be similar to that used in the company's model.

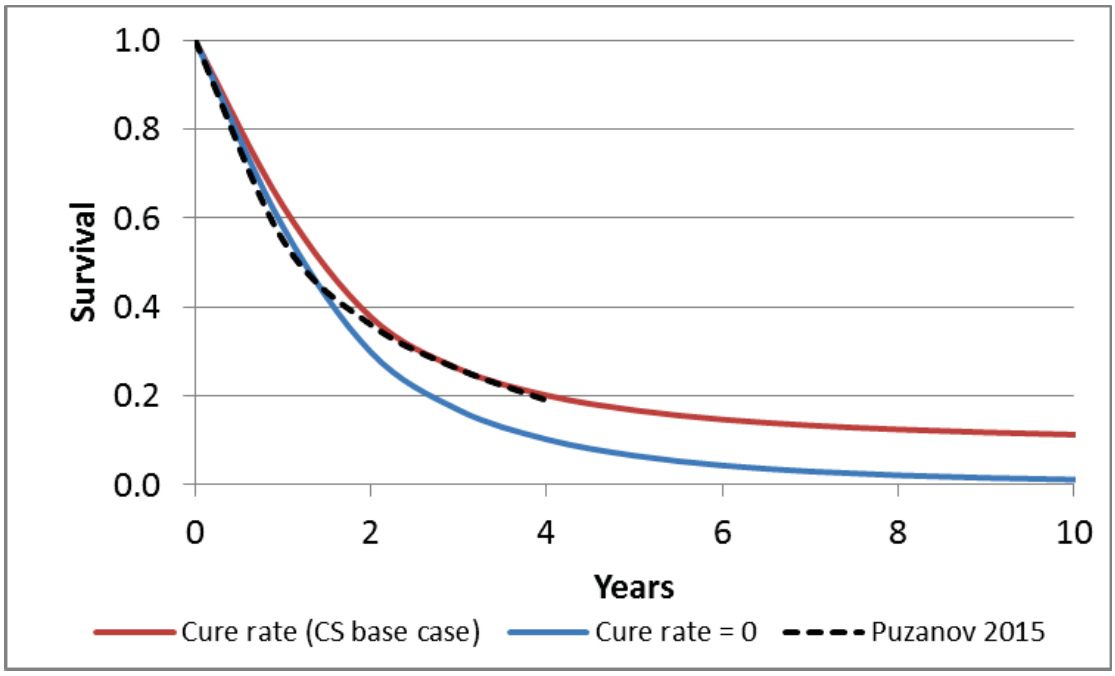


Figure 8 OS for the economic model and Puzanov et al study for patients treated with vemurafenib

The company fitted exponential, Weibull, log-logistic, lognormal, Gompertz, gamma and generalized gamma distributions to the trial arms of the coBRIM trial. The parametric models incorporated the cure function. According to the AkIC / BIC, the [redacted] distribution was selected as the most appropriate fit. The CS presents the fit for the lognormal distribution in CS Figure 23 (see **Error! Reference source not found.** below) and the fit for the other distributions in Appendix 10. The company conducted sensitivity analyses using the alternative distributions (CS Table 60). The model results are sensitive to the parametric distribution chosen and the ICER varies up to £253,766 per QALY for the Gompertz distribution for vemurafenib + cobimetinib compared to vemurafenib (Table 26).

The OS for vemurafenib + cobimetinib was adjusted according to the Inverse AFT for the OS curve for dabrafenib. The inverse AFT results for vemurafenib + cobimetinib vs. dabrafenib were 0.635 (95%CrI 0.46, 0.77) for OS.

The ERG considers that the company's choice of parametric distribution is reasonable and provides a good fit to the trial OS data based upon AkIC/BIC and visual interpretation (CS Figure 23). The ERG considers that the log-logistic distributions would also produce a similar fit to the [redacted]. In addition, the company provides an option for analyses using the KM trial data with a fitted tail. For this option, the KM data are used for the period until the time at which 20% of patients were still at risk, and then the fitted parametric curve is used thereafter We consider that using the KM data with a fitted [redacted] tail would also be a reasonable approach and results in a higher ICER of £176,358 per QALY.

4.3.5.3 Time on treatment

Patients are treated until disease progression or unacceptable toxicity for both cobimetinib and vemurafenib, as per their marketing authorisations. The CS stated that as a small proportion of patients discontinued treatment due to AEs or toxicity, the actual treatment duration in the trial is

shorter than the time until disease progression (CS Table 45). The model analysis for vemurafenib + cobimetinib compared to vemurafenib uses the treatment duration observed in the coBRIM trial. The company fitted parametric curves to the time on treatment distributions observed in the coBRIM trial for vemurafenib + cobimetinib and vemurafenib treatment arms. Visual inspection of the fitted curves indicated that the parametric distributions were a poor fit for the beginning of the KM data and the company therefore used the KM data in the model with a parametric tail. According to the AkIC and BIC treatment values, the Weibull provided the best fit for both cobimetinib and vemurafenib in the treatment arm and the log-logistic was the best fit to the vemurafenib arm. Parametric curves were fitted from month 15 in both the intervention and comparator arms, based on the time point at which 20% of patients were still at risk.

The ERG notes that the company used different parametric curves for the intervention and comparator arms and this choice causes a convergence in the TOT curves for the intervention and comparator arms which favours the vemurafenib + cobimetinib (Figure 9). The ERG considers it is more reasonable to use the log-logistic parametric curve for both the intervention and comparator (Figure 10). Furthermore, this would also be consistent with the choice of the log-logistic distribution for the PFS curve. The ERG investigates the use of using the KM data with a log-logistic tail for both the intervention and the comparator arm results in Section 4.4.

For the comparison with dabrafenib, PFS was used for a proxy for time on treatment for the vemurafenib + cobimetinib and dabrafenib treatment arms. The company justifies the use of PFS as a proxy for TOT for dabrafenib by stating that there was an absence of available data for this comparator. Using this approach gives higher treatment costs for this comparison compared to the comparison between vemurafenib + cobimetinib and vemurafenib. For this reason the company presents two pairwise comparisons between vemurafenib + cobimetinib compared to vemurafenib and dabrafenib, rather than a fully incremental analysis.

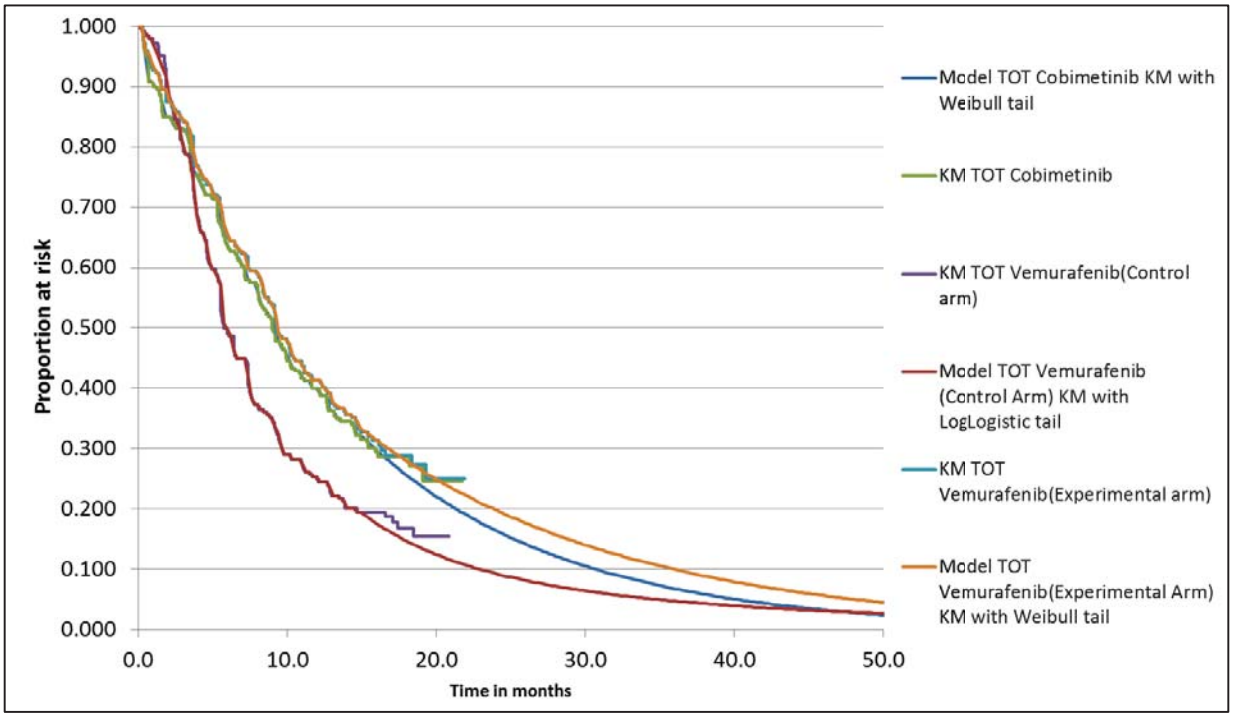


Figure 9 TOT for the economic model using the company base case parametric curves (CS Figure 26)

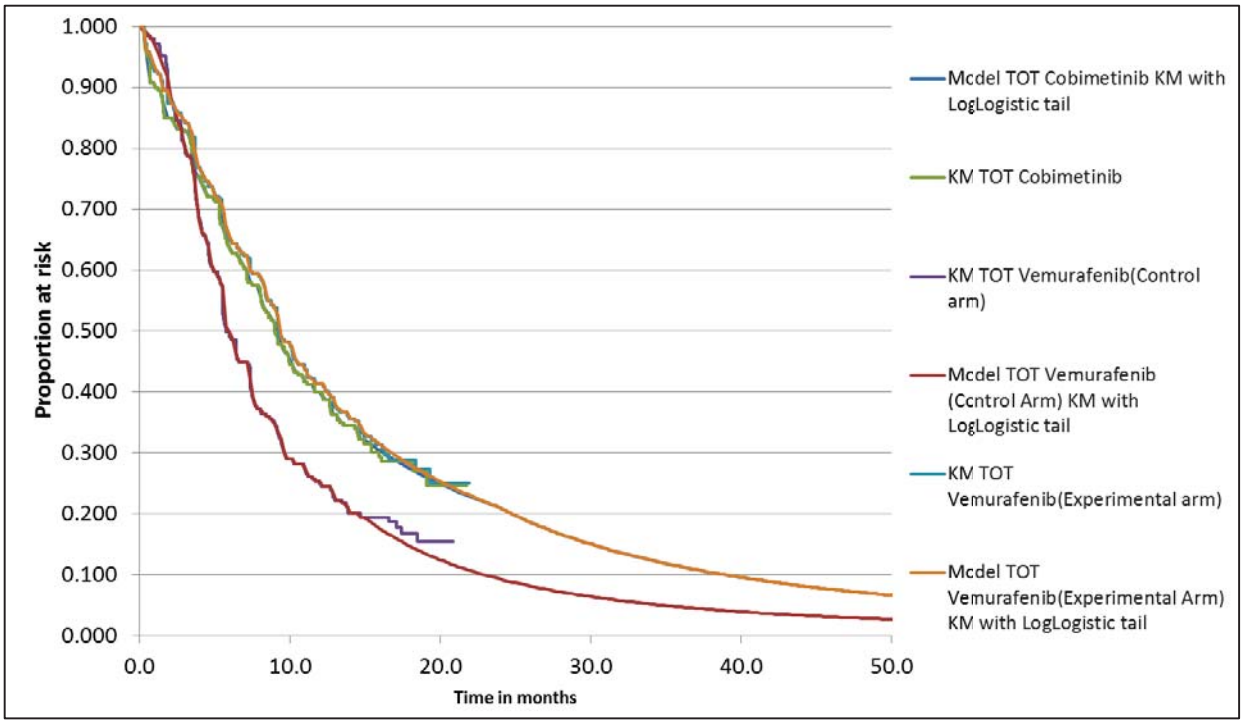


Figure 10 TOT for the economic model using the KM log-logistic parametric curve for the vemurafenib + cobimetinib treatment arm

4.3.5.4 Adverse events

The CS model includes AEs for grade 3 or 4 at an incidence of 3% or more from the coBRIM trial for the vemurafenib + cobimetinib and vemurafenib arms. The incidence of the AEs is shown in Table 14 (CS Table 23). The costs of treating AEs have been included in the economic model. More details on the AEs in the coBRIM trial are given in Section 3.3.6.

4.3.6 Health related quality of life

HRQoL outcomes were quantified in the cost-effectiveness model by assigning utility weights to the two alive-states of the decision model (PFS and PD), a method that is commonly used in cancer models. Whilst for PFS, HRQoL weights were estimated directly from patient data collected within the coBRIM trial, the HRQoL weights for the PD state were from the literature. The HRQoL values used in the model are shown in Table 18.

Table 18 Summary of utility values for the company's economic model (adapted from CS Table 36)

State	Utility value: mean (standard error)	95% confidence interval	Source of utility data
PFS (vemurafenib + cobimetinib)	0.837	0.830, 0.844	coBRIM trial
PFS (vemurafenib)	0.819	0.812, 0.827	
PFS (dabrafenib)	0.819	0.812, 0.827	Assumed to be equivalent to vemurafenib arm of coBRIM trial
PD <5 years	0.590	0.578, 0.602	Literature review (Beusterien et al., 2009)[36]
PD ≥5 years	0.770	0.755, 0.785	

4.3.6.1 Individual patient data collection for HRQoL within the coBRIM trial

PFS was the primary endpoint of the coBRIM study, but HRQoL-outcomes were collected as secondary end points using both the EORTC QLQ-C30 and EQ-5D-5L instruments. The EORTC QLQ-C30 is a cancer specific HRQoL questionnaire which consists of 30 questions with a 4-point response format, moving from "not at all" to "very much".[40] The generic EQ-5D instrument was used in its 5-level format, and both questionnaires were administered in local language on the first day of cycle 1 before initiating treatment and at each study visit before undertaking any further study related assessments[12]. Whilst for PFS, total observations reached 1323 in the vemurafenib + cobimetinib and 1103 in the vemurafenib + placebo group,

observations in the progressed disease state are scarce, with only 57 in total. It is not clear how observations from individual patient data were pooled to estimate state weights for HRQoL, whether results were properly adjusted for baseline utility in both arms of the trial, nor whether observations for some patients were missing, and if so, how missing data were handled.

4.3.6.2 Systematic review of relevant HRQoL literature

As stated in Section 3.1.1 of this report, a systematic literature review was conducted by the company in March 2015 and updated in December 2015 in order to identify health state utility weights for patients with advanced (unresectable or metastatic) melanoma. The initial search was not restricted by intervention, study design, country, or date and aimed at identifying studies published in English language which reported utility weights elicited through time trade off (TTO), standard gamble (SG), generic preference-based HRQoL instruments or mapping studies which allow cross-walking towards such generic preference-based measures for patients with advanced melanoma. It is not clear whether any measures to assess study quality were applied, but accordance with the NICE reference case was explicitly assessed. The literature search identified 17 papers, and the CS states that six of these 17 studies reported health state utility weights using the EQ-5D. However, the CS states that none of these 6 studies were deemed appropriate to estimate health state utility weights for the economic model as the reported EQ-5D data were related to different pharmacological treatments for melanoma (rather than the treatments in this appraisal) or the treatment was not reported. The CS further states that data from five of the six publications reporting EQ-5D utilities were not consistent with the NICE reference case (or information was limited so that it was not possible to assess this). Nevertheless, the SLR also identified a study by Beusterien and colleagues [36] which has been used previously to estimate utility weights in the technology appraisal on vemurafenib for treating locally advanced or metastatic BRAF V600 mutation-positive malignant melanoma (NICE TA269),[26] This study has subsequently been used to estimate utility weights for the PD state of the model.

4.3.6.3 Utility weights for progression free survival (PFS)

For the PFS state of the model, patient-level EQ-5D-5L data from the coBRIM study were used to estimate utility weights. The recently developed EQ-5D-5L has advantages over its 3L predecessor as it is deemed more responsive to changes in health status and avoids ceiling or floor effects which have been observed in the 3L version.[41] However, in order to use the EQ-5D

5L data from the coBRIM trial in the economic model, it was necessary to obtain respective health state utility weights. Two previously validated methods were used for obtaining such utility weights: first, a crosswalk method[41] to estimate utility values for the EQ-5D-5L from available 3-level general population value sets; and second, an algorithm developed by the Office of Health Economics (OHE)[42] which utilised both TTO and discrete choice experiments (DCE) in order to estimate EQ-5D-5L utility weights for the general population in England.[42] Resulting utility weights from both methods were then discussed with expert advisors, and subsequently those relating to the OHE algorithm were dropped for the base case but tested in scenario analysis as they were deemed unrealistically high by experts and exceeded average population norms for the equivalent age group of the coBRIM trial. The ERG believes that the approach taken by the sponsor is reasonable and well justified. Though the OHE dataset is now validated and endorsed by the EuroQoL Group to obtain EQ-5D 5L weights for England, these values are not consistent with EQ-5D 3L weights used in previous technology appraisals. Further, though not clearly stated in the CS, it is reasonable to assume that the crosswalk values are UK specific as the converter used by the sponsor allows estimation of UK specific EQ-5D 3L utility weights from available EQ-5D 5L data. Therefore, using the crosswalk method to obtain equivalent 3L values so to ensure comparability with other technology appraisals in the base case and testing OHE values as a form of scenario analysis is a reasonable approach taken in the CS.

Adverse events were not explicitly accounted for when estimating utility weights for the PFS state of the model. Rather, as all utility values were estimated directly from individual patient data obtained within the coBRIM trial, any AEs are assumed to be reflected in the observed utility values. The ERG believes that this assumption is reasonable given the nature of the data available, but it does not allow explicit assessment of disutility from AEs within the economic model. This is of particular importance as the utility weight for the vemurafenib + cobimetinib group (0.837) was higher compared to the vemurafenib or dabrafenib group respectively (0.819).

Though the ERG agrees with the CS that utility weights for PFS (Table 18 above, CS Table 36) are broadly similar to previous NICE technology appraisals for melanoma (Table 19, CS Table 35), particular concerns relate to the difference in utility weights between intervention and comparator arms.

Table 19 Comparison of utility values from prior NICE technology appraisals relevant to scope population (CS Table 35)

	Drug					
Health state	**Vemurafenib**	**Dabrafenib**	**Ipilimumab**	**Pembrolizumab**	**Nivolumab**	**Cobimetinib +vemurafenib**
PFS	0.85	0.77	-	-	-	0.832
PD	0.59	0.68	-	-	-	0.798
PFS≥ 30 days					0.8018	
PFS< 30 days					0.7795	
PD≥ 30 days					0.7277	
PD< 30 days					0.7054	
12+ months until death			0.885	0.82		
9-12 months until death			0.880	0.71		
6-9 months until death			0.854	0.66		
3-4 months until death			0.810	0.66		
1-3 months until death			0.739	0.57		
<1month until death			0.631	0.33		

According to the CS, the HRQoL for patients in the PFS state for vemurafenib + cobimetinib is higher than for those patients receiving vemurafenib or dabrafenib (Table 18). A potential justification could be, for instance, that AEs are generally lower in the combined treatment group, which would then lead to higher HRQoL for these patients in PFS. However, this is neither confirmed through the data reported on AEs in the coBRIM trial nor was the disutility from adverse events explicitly assessed in the CS which makes any assessment of potential reasons for the difference in utility weights by the ERG uncertain. In addition, the CS does not provide details on the statistical methods used to calculate HRQoL weights from patient data or the frequency or handling of missing values. For instance, if values are not missing at random, this could already explain why the HRQoL weight in the combined treatment group was found to be higher compared to the vemurafenib arm of the coBRIM trial. The ERG therefore suggests re-running the model with identical HRQoL values for PFS across the three therapies under assessment, for example by using the HRQoL weight estimated for vemurafenib from patient

data in the coBRIM trial (0.819). The ERG has conducted this as a scenario analysis in Section 4.4.

4.3.6.4 Utility weights for the progressed disease (PD) state

Patient data from the coBRIM trial was scarce for PD, due to the fact that observations are only available until 12 weeks after the end of treatment. Consequently, utility weights for the PD state in the model are based on a published study by Beusterien and colleagues.[36] This study was identified through the systematic literature review on HRQoL as summarised above, and it has been used to estimate state weights in a previous technology appraisal for vemurafenib by the same company (NICE TA269).[26]

The study by Beusterien and colleagues[36] elicited utilities from 140 respondents in the UK (n = 63) and Australia (n = 77) related to advanced melanoma whilst explicitly capturing both intended clinical response and unintended toxicities related to treatment. Standard gamble was used to elicit utilities and the study sample was recruited from the general population. The authors estimated utilities for four different clinical response states (partial response (0.85); stable disease (0.77); progressive disease (0.59); and best supportive care (0.59)).

The utility weight for progressive disease (0.59) was then used for the PD state in the model without taking into account utility decrements from toxicities. This assumption is reasonable as treatment should have discontinued when patients enter the PD state. Further, the utility weight for stable disease (0.77) was applied to patients in the PD state for more than 5 years, which follows the assumption that patients with long term survival have a higher HRQoL as their disease state has stabilised. The same assumption has been made in a previous technology appraisal on vemurafenib for treating locally advanced or metastatic BRAF V600 mutation-positive malignant melanoma (NICE TA269).[26] In addition, different AEs from toxicities were assessed and estimated as utility decrements.

The ERG has some concerns with respect to the choice of the study by Beusterien and colleagues[36] to estimate utility weights for the PD state of the model. Both the method used in this study (standard gamble), which is choice based, and a sample for valuing health states, which was taken from the general public, accord with the NICE reference case. However, according to the reference case, the actual assessment of health states should be based on patients, and this was not the case in the study by Beusterien and colleagues. Further, a later

appraisal of nivolumab for patients with advanced (unresectable or metastatic) melanoma (NICE TA384)[4] used higher utility values for progressed disease (CS Table 35). In this appraisal, utility weights were estimated from patient-level observations on EQ-5D from the CheckMate 066 trial.[43] EQ-5D data from 1540 visits involving 362 study patients was used to estimate health state utility weights for both pre-progression and post-progression, and in each state for patients being either more or less than 30 days from death. Utility weights for progressed disease were estimated as 0.7277 for patients ≥30 days from death and 0.7054 for patients <30 days from death respectively. The ERG believes that this study may be more in accord with the NICE reference case as the EQ-5D was used in a patient population with advanced (unresectable or metastatic) melanoma. Accordingly, the ERG suggests using alternative utility weights for PD in a scenario analysis using the weights from the NICE TA384 on nivolumab. The ERG has done this scenario analysis in Section4.4.

4.3.6.5 Bottom-line summary of ERG view on utilities

The ERG has several concerns related to the utility weights used in the model by the CS. For PFS, it is not clear why utility weights for combined therapy are higher compared to vemurafenib or dabrafenib only. Fewer AEs from toxicities in the combined treatment arm may be a potential reason, but this cannot be confirmed by the data on AEs reported from the coBRIM trial. The lack of reporting of statistical methods to estimate utility from the data and the existence and handling of missing values make it impossible to assess potential reasons for higher utility weights in the combined treatment arm of the coBRIM trial. The ERG therefore suggests a scenario analysis using identical utility weights in PFS for combined therapy and monotherapy with vemurafenib or dabrafenib respectively.

The ERG has further concerns regarding the utility weight for PD as taken from Beusterien and colleagues.[36] This study does not accord the NICE reference case as health states were assessed from a general public sample, and TA 384 estimated higher utility weights in PD using EQ-5D for patients with advanced (unresectable or metastatic) melanoma. The ERG therefore suggests scenario analysis using a utility weight of 0.77 for PD as taken from NICE TA384.

4.3.7 Resource use and costs

Costs considered in the model include drug cost and administration costs, treatment of AEs, weekly supportive care (health state cost), and diagnostic costs. Drug costs are based on

published list prices, assumptions about existing patient access schemes and average doses taken by patients observed in the coBRIM trial. Administration costs were estimated in accordance with previous technology appraisals for vemurafenib and dabrafenib (NICE TA269;[26] TA321[7]). The cost of AEs was estimated from different sources, including previous technology appraisals and Healthcare Resource Groups (HRG) reference cost (unit cost estimates). Weekly supportive care (health state cost) were taken from a published study[44] which was identified through a systematic literature review on cost and resource use for advanced (unresectable or metastatic) melanoma, and this estimate was also consistent with previous technology appraisals (NICE TA269;[26] NICE TA319;[6] NICE TA357 [10]). Finally, diagnostic costs were estimated in accordance with the prior vemurafenib technology appraisal (NICE TA269[26]), and these costs were applied both to the intervention and both comparators.

4.3.7.1 Systematic literature review to identify resource use and unit cost estimates

The company conducted a systematic literature search to identify resource use and unit cost estimates for the model (see Section 3.1.1). Studies were included if they were published in English language between 1st January 2000 and 9th December 2015, and only if they reported on resource use or cost drivers related to advanced (unresectable or metastatic) melanoma. No restriction was placed on country or study design, though papers in which the economic perspective was unclear were excluded. Studies were also excluded if they related to early stage melanoma or reported on indirect costs (such as productivity costs) only.

The systematic review identified only two relevant papers for estimating resource use and unit costs for the model. One was a published study[44] and one a conference poster.[45]

Johnston and colleagues[44] performed a retrospective analysis of hospitalisations, hospice care, and outpatient visits recorded during a multinational observational study (the MELODY study). The aim was to assess resource utilisation and costs in patients with advanced melanoma (stage III or IV) who either received active treatment or best supportive care. Patients from the UK, Italy, and France were recruited between July 2005 and June 2006 and medical records were used to assess patient and disease characteristics, treatment patterns, respective outcomes and resource use from the time of diagnosis until May 2008 or until the patient had died. Resource use data were then multiplied with national unit cost estimates to calculate per patient cost from a UK NHS perspective. Unit cost data referred to the year 2009.

The study by Vouk and colleagues[45] also took a UK NHS perspective, and reported on the economic burden of adverse effects associated with metastatic melanoma treatment. Based on a literature review, 29 AEs with all severity grades were selected. Associated resource use was then estimated through a Delphi panel consisting of 4 clinicians, and unit costs (2011-2012) were obtained from the literature to estimate the cost per event per patient.

4.3.7.2 Drug acquisition costs

Drug acquisition costs in the model consist of drug cost and administration cost. For vemurafenib + cobimetinib, weekly drug costs were estimated from published list prices and the actual dose taken by patients within the coBRIM trial. Due to dose modification, this dosage was generally lower than the label dose, and clinical advisors were consulted to judge whether observed doses are likely to match those used in clinical practice. The costs in the model are based on the list price in the base case with further analyses including the confidential PAS which exists for dabrafenib and vemurafenib. The drug cost assumed in the model are summarised in Table 20 below.

In contrast to the above, the cost for dabrafenib was estimated from the label doses. The company justified this assumption by stating that the coBRIM trial did not include dabrafenib so that actual doses could not be estimated. However, the ERG believes that this is an inconsistency which may bias results of the economic model. To be consistent, one could either use label doses to estimate weekly drug cost for all three treatments, or estimate drug cost for dabrafenib with modified doses. The ERG has tested these scenarios in Section 4.4.

Table 20 Weekly vemurafenib + cobimetinib drug costs at according to average dose taken in coBRIM study & weekly dabrafenib cost based on label doses.

	Daily dose according to label dose	Actual dose taken in coBRIM trial	Per cycle cost (week)
Cobimetinib	3 tablets of 20 mg days 1 - 21 of a 28 day treatment cycle	2.602 tablets per day	£1236*
Vemurafenib + cobimetinib	8 tablets of 240mg	7.062 tablets per day	£1545
Vemurafenib	8 tablets of 240mg	6.99 tablets per day	£1529
Dabrafenib	150mg twice daily (equivalent to a total daily dose of 300 mg)	75mg x 28 = £1400	£1400

*during treatment periods only

Besides dosage and prices, another important factor to estimate drug cost in the model is the time on treatment. In general, time on treatment was assumed to last until disease progression or until toxicity reaches an unacceptable level. The company's approach to estimating time on treatment is discussed earlier in Section 4.3.5.3.

Finally, a pharmacy charge of £13 was applied to both the intervention and each comparator respectively to reflect administration cost for each 28 day cycle of therapy. The ERG agrees that this approach is consistent with previous STAs and reasonable given that all drugs under assessment are oral products which do not require access to a chemotherapy suite and only have to be dispensed by a pharmacist once per cycle.

4.3.7.3 Health state cost (weekly supportive care)

The company used data from the study by Johnston and colleagues,[44] which was identified in the systematic review described above. This study has also been used as a basis for estimating health state cost in a number of previous NICE STAs (TA269;[26] TA319;[6] TA357;[10] and TA366 [9]). In line with these technology appraisals, the company therefore set monthly cost of best supportive care for both PFS and PD at £378 (£87 per week). The ERG believes that this approach is generally reasonable and justified for the base case analysis, especially as the same health state cost was assumed for the intervention and comparator technologies. We consider there is also reason to believe that health state costs reduce in patients with long term stable disease. The ERG has investigated the use of lower health state costs for progressed disease by using the values shown in CS Table 41 from discussion with clinical experts (PD year 1 £87.23/week; PD year 2-3 £20.25 / week; PD year 4-6 £12.17/week; and PD year 6+ £6.51/week) and found that using these values has a minimal impact on the model ICERs.

4.3.7.4 Adverse events costs

The coBRIM trial was used for combination therapy and vemurafenib monotherapy for estimating cost related to the treatment of AE, incidence data from. Unit costs were used from prior STAs or from HRG unit cost (2014-15)[46] for valuing the resource use associated with treating AEs. Only AE with an incidence of at least 3% and grades 3 or 4 were considered for the economic model. The study by Vouk and colleagues,[45] which was included in the systematic review, was not used for the model as the sponsor deemed treatments considered there not specific enough to match either the intervention or comparator.

Table 21 (CS Table 43) summarises AEs with their respective cost assumed in the economic model.

Table 21 List of adverse events and summary of costs in the economic model

Adverse reactions	Items	Value	Reference in submission
Liver function test abnormality*	Outpatient cost, % No treatment, % Mean cost per patient	£158.54, 20% £0, 90% £31.71	HRG service code 370, medical oncology
Arthralgia	Outpatient cost	£139.52	HRG service code 191, pain management
Basel cell carcinoma		£198.66	JC41Z: outpatient major skin procedure
Diarrhoea	Inpatient cost, %	£838.46, 50%	TA366 2015
	Outpatient cost, %	£144.05, 50%	TA366 2015
	Mean cost per patient	£491.26	TA366 2015
Fatigue	Inpatient cost, %	£596.38, 10%	TA366 2015
	Outpatient cost, %	£156.84, 90%	TA366 2015
	Mean cost per patient	£200.79	TA366 2015
Hypertension		£287.04	EB04Z: outpatient procedure
Hyponatraemia	Cost assumed £0	Cost assumed £0	TA366 2015
Keratoacanthoma		£198.66	JC41Z: outpatient major skin procedure
Pain in extremity	Outpatient cost	£139.52	HRG service code 191, pain management
Rash		£137.31	TA269 2012
Rash-maculo popular		£137.31	TA269 2012
Squamous cell carcinoma of skin		£198.66	JC41Z: outpatient major skin procedure

* Alanine aminotransferase increase, aspartate aminotransferase increase, gamma-glutamyltransferase increase, blood alkaline phosphatase increase, blood creatine phosphokinase increase

AE costs were then estimated as the total weekly AE cost per patient using incidence data from the coBRIM trial and unit cost estimates from previous STAs and HRG unit costs. These costs were estimated to be £3.20 for vemurafenib + cobimetinib and £3.90 for vemurafenib. For dabrafenib, safety data were not available at the same level of detail so that AE cost were assumed to be the same as the lower vemurafenib + cobimetinib cost. However, the CS states that the difference in AE costs between intervention and vemurafenib was mainly due to the

higher incidence of squamous cell carcinoma in the vemurafenib group. This was also one of the very few AEs for which the difference in incidence between vemurafenib + cobimetinib and vemurafenib was statistically significant. The ERG therefore concludes that the use of differential AE cost between intervention and vemurafenib appears to be justified for the base case analysis. However, the same AE cost for all therapies should be tested within scenario analysis. The ERG tested the effect of using the same AE for all therapies and found that the model ICERs were not significantly affected by this change.

4.3.7.5 Other costs assumed in the model

Apart from the costs summarised above, the only additional cost considered in the economic model was for BRAF mutation testing. However, as this is part of routine management for patients with advanced melanoma in the UK, the diagnostic cost at a rate of £95 was applied to both the intervention and comparators and does not impact on the incremental cost in the model.

Further, the company assumed that post-progression would be the same for the intervention and comparators so no costs were included for active treatment after progression.

4.3.7.6 Bottom-line summary of ERG view on resource use

Resource use considered in the economic model consists of drug cost, dispensing cost, treatment of AEs, weekly supportive care (health state cost), and diagnostic cost. Resource use has been estimated either directly from the coBRIM trial, the published literature or previous STAs on advanced melanoma in the UK. Unit costs for resources were obtained from previous STAs and HRG reference costs. The cost year varies by resource item between 2009 and 2014/15. In general, the approach to estimate resource use and cost appears to be justified and reasonable.

However, the ERG has a few concerns over the costing methodology to estimate drug cost, time on treatment, health state cost, and the cost of AEs. First, drug cost should be estimated using either label doses throughout all treatment options or by reflecting dose modification for all drugs. Further to that, the ERG suggests scenario analysis with decreasing health state cost for long term stable PD as there is evidence suggesting that healthcare cost decrease in these patients. Finally, a scenario should be tested which assumes identical AE costs between

intervention and both comparators, as only very few differences in the incidence of AEs between treatment arms were observed in the coBRIM trial.

4.3.8 Cost effectiveness results

Deterministic results from the economic model are presented (CS Section 5.7, page 138) as incremental cost per QALY gained for vemurafenib + cobimetinib compared with vemurafenib and with dabrafenib. Results are also reported for total life years. Base case results are shown for the drug list price of the treatment with further results presented in CS Appendix 14 with PAS discounts included for vemurafenib and dabrafenib.

For the base case, an incremental cost per QALY gained of £150,514 is reported (see Table 22) for vemurafenib + cobimetinib compared to vemurafenib and £209,942 for vemurafenib + cobimetinib compared to dabrafenib. As these analyses have been conducted using different assumptions, the company has not provided an analysis that compares all three treatments with each other incrementally. It is not possible to combine the two analyses in one incremental analysis because the analyses have used different assumptions.

Table 22 Deterministic base-case results, (list prices), for direct treatment comparison (actual TOT used) (CS Table 46)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/ (QALY gained)
Vemurafenib + cobimetinib	£163,974	3.034			
Vemurafenib	£81,984	2.489	£81,990	0.545	£150,514

Table 23 Deterministic base-case results, (list prices) for indirect treatment comparison (PFS as surrogate for TOT) (CS Table 47)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/ (QALY)
Vemurafenib + cobimetinib	£208,047	3.034			
Dabrafenib	£78,392	2.417	£129,655	0.618	£209,942

4.3.9 Assessment of uncertainty

The company conducted a range of sensitivity analyses including deterministic sensitivity analyses and PSA.

4.3.9.1 Probabilistic sensitivity analysis

The company performed PSA with the distributions used for the input parameters shown in CS Table 44. The company used the gamma distribution for the utilities, the multivariate normal distribution for the parametric survival curves and log normal distribution for the costs. The ERG considers a more standard approach is to use the beta distribution for the utilities. The ERG re-ran the PSA using 1000 simulations which took approximately 2.5 minutes. We note that the utilities for PFS and progressed disease have been varied independently. However, it would be more correct if these utilities were correlated for each health state. The PSA results are relatively consistent with 1000 iterations and we consider this number of iterations is sufficient.

The PSA results are shown in CS Tables 57 and 58 (Table 24 and Table 25 of this report) and give similar results to the deterministic base case results. The CS summarises the results of the PSA stating that there is a 0% probability of vemurafenib + cobimetinib being cost-effective, relative to vemurafenib at a threshold willingness to pay of £50,000 per QALY gained. Furthermore, the CS stated that for all simulations the cost of vemurafenib + cobimetinib treatment was higher than vemurafenib treatment and the total QALYs was higher for patients treated with vemurafenib + cobimetinib than those treated with vemurafenib. A scatterplot of the PSA results is given in CS Figure 32 (Figure 11 of this report).

Table 24 Mean results of PSA compared to base case (list prices) for vemurafenib + cobimetinib vs. vemurafenib (CS Table 57)

	Costs	QALYs	ICER
Vemurafenib + cobimetinib	£164,636	3.028	£151,668
Vemurafenib	£81,615	2.480	

Table 25 Mean results of PSA compared to base case (list prices) for vemurafenib + cobimetinib vs. dabrafenib (CS Table 58)

	Costs	QALYs	ICER
Vemurafenib + cobimetinib	£210,076	3.028	£215,264
Dabrafenib	£79,472	2.421	

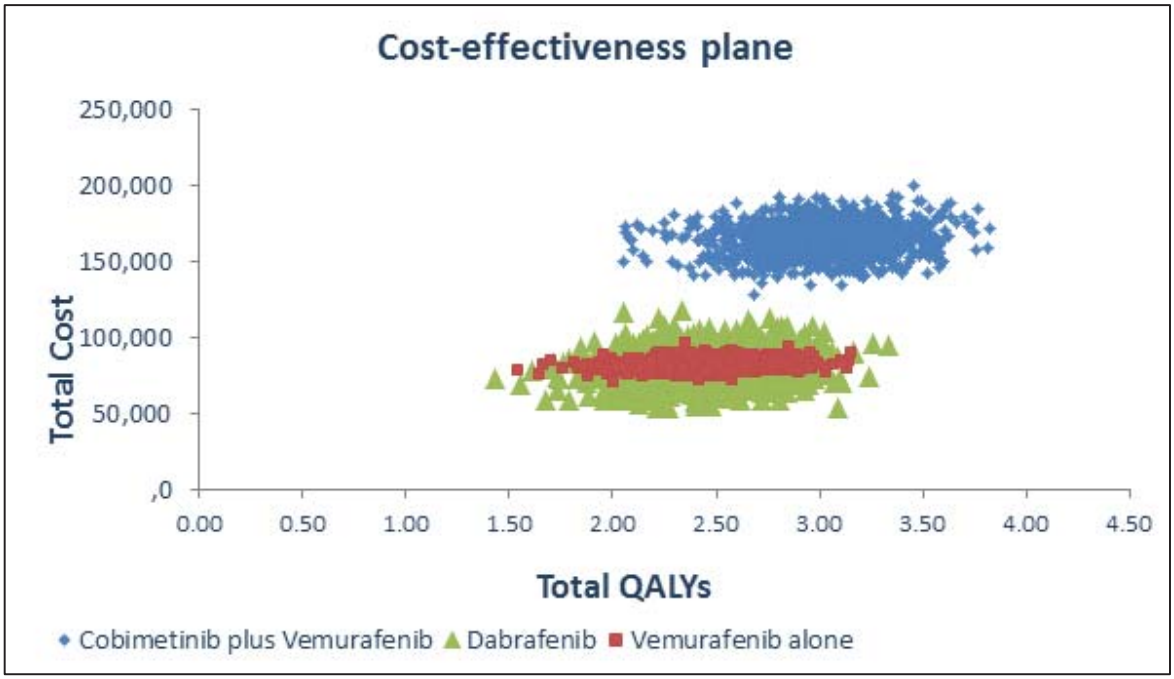


Figure 11 Scatterplot of PSA results for cost effectiveness plane (CS Figure 32)

4.3.9.2 Deterministic sensitivity analyses

The CS provides deterministic sensitivity analyses for six input parameters that are stated to have the greatest impact on percentage increment in costs or QALYs. These are shown in CS Table 59 and shown in this report in Table 26. The CS does not comment on the ranges used for the sensitivity analyses. The ranges were based upon the 95% CIs for the PFS utility state, but arbitrary ranges appear to be chosen for the cost input parameters and utility for PD. The ERG found that there was an error in the calculation of the sensitivity analysis for the weekly cost of vemurafenib, as the cost for vemurafenib was only altered in the vemurafenib + cobimetinib arm and not altered in the vemurafenib only arm. The ERG's corrected values for this parameter varied between £126,111 and £174,916. The parameters that have the largest impact on the model results are for the utility for PFS and the cost of cobimetinib. However, the sensitivity analyses for PFS utility should be treated with caution, as the PFS utilities of the two treatment arms have been varied independently, whereas the ERG considers they would be highly correlated. The company has not included sensitivity analyses for the effectiveness of the treatment or the cure rate. The ERG conduct these sensitivity analyses in Section 4.4

Table 26 Deterministic sensitivity analyses (CS Table 59) (List prices)

Variable	Range (% of base case)	Resulting ICER using lower value	Resulting ICER using higher value
Weekly cost cobimetinib	±50%	£101,933	£199,082
Weekly cost vemurafenib	±50%	£60,248[a]	£240,568[a]
Supportive costs (PFS)	±50%	£144,209	£156,820
Utility PFS vemurafenib + cobimetinib	25th percentile: 0.736; 1	£209,002	£103,686
Utility PFS vemurafenib	25th percentile 0.735;1	£131,025	£221,505
Utility PD	0.5; 1	£176,090	£90,579

[a] The ERG's calculated ICERs for weekly cost of vemurafenib is £126,111; £174,916

4.3.9.3 Scenario analysis

The company conducted scenario analyses to assess the uncertainty around structural assumptions of the model. Results are shown in CS Table 60 (and Table 27 of this report) for the following scenarios: alternative OS and PFS parametric distributions; utility values; dose / treatment duration assumptions; discount rate; time horizon; zero drug cost for cobimetinib.

The CS states that without PAS discounts there are no conditions at which the ICER is below the acceptable threshold. The results show that at zero cost for cobimetinib, the ICER for vemurafenib + cobimetinib compared to vemurafenib is £53,358 per QALY. Vemurafenib + cobimetinib is not cost effective at zero price because patients are treated until disease progression and so the treatment costs are increased because of longer PFS in the vemurafenib + cobimetinib arm compared to the vemurafenib + placebo arm.

The model results were most sensitive to changes in the parametric distributions used for OS and the assumption used for TOT. The ERG has conducted additional scenario analyses for the case where there is no cure rate fraction incorporated and changing the TOT extrapolation curve for vemurafenib + cobimetinib to KM with a log-logistic tail in Section 4.4.

Table 27 Scenario analysis results for vemurafenib + cobimetinib vs vemurafenib or dabrafenib (List prices)

Scenario	Base case	Analyses	ICER intervention vs. vemurafenib	ICER intervention vs. dabrafenib
Base case	n/a	n/a	£150,514	£209,942
	OS parametric distribution			
1	Lognormal	Exponential	£161,902	£219,912
2		Weibull	£229,890	£269,146
3		Log-logistic	£175,592	£212,255
4		Gompertz	£253,766	£269,898
5		Gamma	£217,135	£262,084
	PFS parametric distribution			
6	Log-logistic	Exponential	£166,292	£193,165
7		Weibull	£157,072	£169,530
8		Gamma	£164,485	£191,655
9		Lognormal	£157,377	£203,455
10		Gompertz	£152,215	£164,942
	Utilities			
11		Alternative health state utilities using the OHE value set for EQ-5D-5L valuation	£143,536	£200,778
12		Alternative health states utilities using one value (0.59) for all PD	£157,952	£219,640
	Dose / treatment duration			
13	Weibull	KM with Exponential tail for vemurafenib + cobimetinib (when in combination), TOT	£137,839	£209,942
14	Log-logistic	KM with Lognormal tail for vemurafenib (when monotherapy), TOT	£159,817	£209,942
15		PFS as a proxy for TOT for vemurafenib + cobimetinib vs. vemurafenib	£221,732	£209,942
16		Dosing as per label for vemurafenib + cobimetinib vs. vemurafenib	£170,305	£254,301
	Discount rate: effects and costs			
17	3.5%	1.5%	£140,198	£203,763
	Time horizon			
18	30	20	£152,911	£209,811
19		10	£169,632	£217,655
	Drug costs			

20		£0 cobimetinib	£53,358	£90,977

4.3.10 Model validation

Internal consistency

The company reported (CS page 155) that clinical and health economic experts were consulted in the construction of the economic model to validate the methodological and clinical assumptions and to verify that the model outputs were clinically plausible. Aspects discussed with experts included: the model structure and health state; the mix model for OS extrapolation; derivation of the utilities from the coBRIM trial; resource use within the model.

The company reported that internal consistency of the model had been conducted by York Health Economics Consortium (YHEC) and consisted of a number of tests, including using extreme values.

The economic model is coded in Microsoft Excel and is fully executable. The model is well presented and intuitive to use. The ERG has not undertaken a comprehensive check of all cells in the model, rather, internal consistency checks have been performed and random checking of the model has been done for some of the key equations in the model. The ERG have performed a detailed checking of all model inputs reported in the CS (white box testing); changing the parameter values produced intuitive results (black box testing) and from random checking the 'wiring' of the model appears to be accurate. The ERG was able to replicate the results presented in the CS and the deterministic sensitivity analyses, as reported in CS Tables 59 and Tables 60. The ERG views the model as a reasonable approach to modelling the cost effectiveness of advanced melanoma.

External consistency

The CS does not report a comparison of the results from the economic model with coBRIM trial. The ERG has compared the results from the economic model for PFS and OS and considers that the model provides comparable outcomes with those from the coBRIM trial.

The CS does not compare results from its analysis with analyses from other NICE technology appraisals for other BRAF inhibitors. The ERG compared the results for previous NICE

technology appraisals that reported the results for BRAF mutation positive patients treated with vemurafenib (Table 28). The results for the current appraisal are compared to the previous NICE appraisals TA319 of ipilimumab,[6] TA366 for pembrolizumab,[9] and TA384 for nivolumab.[4] The costs for patients treated with vemurafenib are similar between appraisals. The QALYs and life years vary between appraisals, with higher benefits in the CS compared with other appraisals. These variations are likely to be due to differences in assumptions used in the extrapolation of OS beyond the clinical trial data.

Table 28 Comparison of results patients treated with vemurafenib in the CS with previous NICE technology appraisals (list price)

Results for vemurafenib	Costs	QALYs	Life years
Current company submission	£81,984	2.49	3.39
TA319 ipilimumab	£80,658	2.13	2.98
TA384 nivolumab	Not available	1.69	2.37
TA366 pembrolizumab	£83,384	1.73	2.74

Results are not available for TA269 vemurafenib; TA321 dabrafenib

4.4 Additional work undertaken by the ERG

This section details the ERG's further exploration of the issues and uncertainties raised in the review and critique of the CS cost effectiveness analyses. This consists of additional sensitivity analyses for parameters that have not been adequately explored by the company and scenario analyses with alternative assumptions.

4.4.1 Sensitivity analyses

The ERG has completed the following sensitivity analyses:

i) Cure rate fraction

ii) Treatment effectiveness vemurafenib + cobimetinib vs vemurafenib (PFS, OS)

iii) Treatment effectiveness vemurafenib + cobimetinib vs dabrafenib (PFS, OS)

For each analysis, the ERG varied the parameter between its upper and lower 95% CIs/CrIs and the results are shown in Table 29.

Table 29 ERG deterministic analyses using 95% CIs/CrIs (using list prices)

Comparator / Scenario	Parameter value			ICER (£/QALY)	
	Base case	Low 95% CI/CrI	High 95% CI/CrI	Low 95% CI/CrI	High 95% CI/CrI
Vemurafenib					
Cure rate fraction	[redacted]	[redacted]	[redacted]	£149,166	£151,770
Vemurafenib PFS[a]	0.58	0.46	0.72	£161,467	£158,733
Vemurafenib OS[a]	0.70	0.55	0.90	£114,084	£310,905
Dabrafenib					
Dabrafenib PFS[b]	0.599	0.46	0.77	£219,956	£195,405
Dabrafenib OS[b]	0.635	0.47	0.86	£155,149	£404,450

[a] HR from coBRIM trial; [b] 1/AFT from NMA

Of these analyses, the sensitivity analyses for OS have most effect on the model results: changing the effectiveness of vemurafenib led to an ICER of between £114,084 and £310,905 per QALY gained.

4.4.2 Scenario analyses

The ERG conducted the following scenario analyses:

i) Cure rate fraction removed
ii) TOT extrapolation curve changed to KM with log-logistic tail
iii) Changes to utility values
iv) Consistency in dosing between vemurafenib + cobimetinib and dabrafenib
v) Shorter treatment duration
vi) Inclusion of subsequent treatment costs
vii) Assuming equal efficacy between vemurafenib and dabrafenib for OS
viii) Combination analysis of scenarios ii), iii) and iv)

The results of the scenario analyses are shown in Table 30 and discussed below.

Table 30 ERG scenario analyses (using list prices)

Comparator / Scenario	Base case	Value used in analysis	vs. vemurafenib ICER (£/QALY)	vs. dabrafenib ICER (£/QALY)
Base case results:	-	-	£150,514	£209,942
i) Cure rate fraction	[redacted]	0	£137,928	£190,964

ii) TOT for vemurafenib + cobimetinib	KM with Weibull tail	KM with log-logistic tail	£204,340	£209,942
iii) Utility values PFS vemurafenib + cobimetinib	0.837	0.819	£158,414	£219,603
iii) Utility values PD	0.59; 0.77	0.73	£154,717	£211,447
iv) Dabrafenib dose	£1400 per week	£1232 per week	£150,514	£223,277
v) Shorter treatment duration	Treat until disease progression	Treat for a maximum 2 years	£123,478	£139,532
vi) Subsequent treatment	No subsequent treatment	Subsequent treatment £1400 / week	£149,669	£219,201
vii) Dabrafenib OS	0.635	0.7	£150,514	£243,836
viii) Combination analysis (scenarios ii,iii & iv)[a]	See scenarios above	See scenarios above	£210,046	£224,877

[a] The combination analysis includes changes for scenario iii only for PD utilities and not for PFS utilities.

i) Cure rate fraction removed

The company modelled OS by assuming a proportion of patients will have a low risk of cancer mortality represented by a cure rate fraction. For illustration this scenario explores the effect of removing the cure rate fraction, so that all patients are at a the same high risk of cancer mortality. This results in a reduction in the ICER to £137,928 per QALY gained for vemurafenib + cobimetinib compared to vemurafenib and a reduction in the ICER to £190,964 compared to dabrafenib.

ii) TOT extrapolation curve changed to KM with log-logistic tail

The ERG considers that the TOT should be modelled for both treatment arms with KM data with a log-logistic tail to allow consistency between treatment arms, rather than the company's approach which uses KM data with a Weibull tail for the intervention arm (as discussed in Section 4.3.5.3). Changing the TOT extrapolation results in an increased ICER of £204,340 per QALY for vemurafenib + cobimetinib compared to vemurafenib. The ICER compared with dabrafenib is unchanged because this analysis using a different method to estimate TOT.

iii) Changes to utility values

The ERG considered alternative utility values for the PFS and PD state. We changed the utility value for the PFS health state to be the same in both arms (0.819) and changed the utility value for the PD health state to 0.73 (as used in the NICE appraisal of nivolumab, TA384[4]), (as discussed in more detail in Section 4.3.6). The model results were not sensitive to changes in these utility values, with the ICER increasing to £158,414 per QALY gained compared to vemurafenib when changing the PFS utility value and £154,717 per QALY compared to vemurafenib when changing the PD utility value. Similar reductions in the ICER were seen in the comparison with dabrafenib.

iv) Consistency in dosing between vemurafenib + cobimetinib and dabrafenib

The ERG noted that in the comparison between vemurafenib + cobimetinib and dabrafenib, the company used the mean dose for vemurafenib + cobimetinib given in the coBRIM trial and the planned (label) dose for dabrafenib (which was not included in the trial). However, the mean dose of vemurafenib + cobimetinib was lower than the planned dose in the trial. For consistency, we used the same reduction in dosage for dabrafenib as seen in the trial for vemurafenib + cobimetinib (12%). This resulted in an increased ICER of £223,277 per QALY gained.

v) Shorter treatment duration

The ERG conducted an illustrative scenario whereby patients could be treated for a shorter duration, contrary to the current marketing authorisation of treating until disease progression. For simplicity, we conservatively assumed that there would be no loss in treatment effect due to the shorter treatment duration. Under the hypothetical assumption of maximum treatment duration of two years, the ICER reduced to £123,478 per QALY for vemurafenib + cobimetinib compared to vemurafenib and reduced to £139,532 compared to dabrafenib. However, we advise caution on the interpretation of this scenario as it is unclear what effect shorter treatment duration would have on health outcomes. Furthermore, expert clinical advice to the ERG was that in clinical practice responding patients would be unlikely to stop treatment before disease progression if they have acceptable toxicity.

vi) Inclusion of subsequent treatment costs

The CS did not include subsequent anti-cancer drug treatment in the economic model. The ERG considered an illustrative scenario whereby all patients received subsequent anti-cancer treatment when their disease progressed. This was assumed to be at the same price as dabrafenib, i.e. £1400 per week for as long as patients remained alive. This illustrative scenario showed that including subsequent anti-cancer drug treatment in the economic model would have had a minimal impact on model results.

vii) Assuming equal efficacy between vemurafenib and dabrafenib for OS

As discussed in Section 3.1.7.4 of this report, patient crossover from dacarbazine to dabrafenib occurred in the BREAK-3 trial and the effect of this does not appear to have been taken into account in the company's NMA. We consider that this would underestimate the clinical effectiveness of dabrafenib. We are not able to perform the necessary adjustment and therefore provide an exploratory analysis that assumes similar effectiveness for dabrafenib as for vemurafenib, as accepted in the previous NICE technology appraisal of dabrafenib.[7] Using the same OS estimates for dabrafenib and vemurafenib increases the ICER for vemurafenib + cobimetinib compared dabrafenib to £243,836 per QALY.

viii) Combination analysis of scenarios ii), iii) and iv)

The scenario combines ERG scenarios ii), iii) and iv), i.e. changes to the TOT for vemurafenib + cobimetinib; changes in utility values for the PD state (NB. there are no changes to the utility values for the PFS health state); and alterations to the dosing schedule for dabrafenib. The combination scenario provides the basis for the ERG's base case, discussed in the following section. The scenario increases the ICERs for vemurafenib + cobimetinib compared vemurafenib to £210,046 per QALY and increases the ICER compared to dabrafenib to £224,877 per QALY.

4.4.3 SHTAC base case

Based upon our critique of the company's economic model, the ERG suggests an alternative base case (Table 31). Our base case analysis was conducted using a replication of the

company model based on the NMA for the comparison with vemurafenib + cobimetinib and vemurafenib and with dabrafenib (the company's model is informed by the coBRIM trial for the comparison with vemurafenib, and the NMA for the comparison with dabrafenib).

Table 31 Base case specification

	Company base case	SHTAC base case
Analysis	Pairwise comparison for vemurafenib + cobimetinib vs vemurafenib; vemurafenib + cobimetinib vs dabrafenib.	Fully incremental analysis comparing vemurafenib + cobimetinib, vemurafenib and dabrafenib.
PFS	Loglogistic distribution fitted to coBRIM trial data for direct comparison. The indirect treatment comparison used the parametric distribution for vemurafenib + cobimetinib adjusted using the AFT.	The same approach was used as in the company base case. In the incremental analysis, PFS estimates for vemurafenib were derived from the NMA rather than the coBRIM trial. This was done for methodological consistency to enable a fully incremental analysis to be conducted.
OS	[redacted] distribution fitted to the coBRIM trial data for direct comparison. The indirect treatment comparison used the parametric distribution for vemurafenib + cobimetinib adjusted using the AFT.	The same approach was used as in the company base case. In the incremental analysis, OS estimates for vemurafenib were derived from the NMA rather than the coBRIM trial. This was done for methodological consistency to enable a fully incremental analysis to be conducted.
Time on treatment (TOT)	Time on treatment estimated with fitted parametric curves to KM data for time on treatment for direct comparison. KM data used with parametric extrapolation ('tail') with Weibull for vemurafenib + cobimetinib and loglogistic for vemurafenib. For indirect comparison, the PFS curve was used a proxy for time on treatment.	In the incremental analysis, TOT was estimated using the NMA. The PFS curve was used as a proxy for time on treatment. (ERG scenario ii) In the incremental analysis the drug costs were reduced by 7% to allow for the differences between actual TOT and time on treatment predicted using PFS.
Costs	Health state costs based upon MELODY study. Adverse event costs included based upon those in the coBRIM trial.	We used the same costs as the company except for the dabrafenib costs which were reduced by 12% for consistency with the drug dosages used in the coBRIM trial (ERG scenario iv).
Utilities	HRQoL taken from coBRIM trial	HRQoL taken from the coBRIM trial for

	for PFS and from Beusterien et al. for progressed disease.	PFS and from nivolumab NICE TA384 for progressed disease. (ERG scenario iii).
Time horizon	Lifetime	Lifetime
Discounting	3.5% per year for cost and effects	3.5% per year for cost and effects

The company did not provide a fully incremental analysis comparing vemurafenib + cobimetinib to both vemurafenib and dabrafenib monotherapies. Rather, two separate pairwise comparisons were provided. These analyses are not easily combined to form an incremental analysis as different assumptions have been used, i.e. differences in the methods used to extrapolate survival and to estimate time on treatment. The ERG conducted a fully incremental analysis by using comparative effect estimates for vemurafenib and for dabrafenib from the NMA (notwithstanding our caveats about the validity of the NMA, see Section 3.1.7), along with the assumptions from the SHTAC preferred base case. We used the same assumptions for extrapolation of survival and estimation of time on treatment for the comparisons with vemurafenib and dabrafenib.

Our preferred base case includes changes to the extrapolation of TOT (scenario ii), changes to the utility values for PD (scenario iii) and a consistent dose estimation for vemurafenib + cobimetinib and for dabrafenib (scenario iv). The extrapolation of TOT using the KM with log-logistic tail allows a consistent approach for extrapolation between study arms. Similarly there was an inconsistency in the approach in the company's model for the treatment dosage between vemurafenib + cobimetinib and dabrafenib and we have used consistent dose estimation for both treatment arms.

In addition to the changes for scenario ii, iii and iv (combination scenario viii), another adjustment was made for the vemurafenib and dabrafenib analyses. The company's base case indirect treatment comparison uses drug costs, drug administration and AE costs based on PFS as a proxy which overestimates these costs compared to those observed in the trial analysis by about 7%. These costs have been multiplied by 0.93 to give them a better approximation.

The analysis is shown in Table 32 using the drug list prices and in a separate confidential appendix for the NICE Appraisal Committee with the PAS discounted prices.

Table 32 ERG analysis: Fully incremental analysis (using list prices)

	QALY	Cost	Incremental QALY	Incremental Cost	ICER (£/QALY)
Dabrafenib	2.479	£65,908			
Vemurafenib	2.576	£77,846	0.10	£11,938	£123,072
Vemurafenib + cobimetinib	3.092	£193,295	0.52	£115,449	£223,738

4.5 Overall summary of cost effectiveness

The structure of the economic model was appropriate, comprehensive and reflected the clinical pathway for patients with advanced melanoma. The economic model, developed in Microsoft Excel, was well structured and intuitive. The ERG did not find any errors in the coding of the model structure.

The methods chosen for the analysis were generally appropriate and conformed to NICE methodological guidelines. However the treatments were not compared in a fully incremental analysis, due to differing methods used for the comparison of vemurafenib + cobimetinib with vemurafenib and dabrafenib.

The CS base case analyses for vemurafenib + cobimetinib compared to vemurafenib had an ICER of greater than £150,000 per QALY (using list prices) and the CS stated that vemurafenib + cobimetinib would not be cost effective at a threshold of £50,000 per QALY even if the price of cobimetinib was reduced to zero. Vemurafenib + cobimetinib is not cost effective at zero price because patients are treated until disease progression and so the treatment costs are increased because of longer PFS in the vemurafenib + cobimetinib arm compared to the vemurafenib + placebo arm. To illustrate, the vemurafenib treatment cost for patients in the vemurafenib + cobimetinib arm is £98,285 compared to £71,699 in the vemurafenib monotherapy arm.

In general, the ERG considers that the choice of parameters used in the model and choice of parametric curves used for extrapolating beyond trials data were reasonable. However we identified several areas where we suggest alternative data sources or parametric curves would be more appropriate. Given the uncertainty over extrapolating beyond clinical trial data, other parametric curves may also be plausible and in all cases these would result in a less favourable ICER. The ERG had reservations about the parametric curve used to fit time on treatment for

vemurafenib + cobimetinib. The ERG suggests using the KM data followed by a log-logistic tail would be consistent with the curve used for PFS and the vemurafenib + placebo arm.

The ERG noted that in the comparison between vemurafenib + cobimetinib and dabrafenib, the company used the actual dose for vemurafenib + cobimetinib and the planned dose for dabrafenib. However the actual dose was lower than the planned dose in the trial. For consistency, the ERG suggests using the same reduction in dosage for dabrafenib as seen for vemurafenib in the coBRIM trial.

Based upon the ERG's analyses the preferred base case ICER for vemurafenib + cobimetinib compared to vemurafenib is likely to be at least £223,738 per QALY.

5 END OF LIFE

CS Section 4.13.2 (CS Table 26) provides the company's justification for why vemurafenib + cobimetinib combination therapy meets all of NICE's end of life criteria. Two of the three criteria relate to the condition under appraisal (melanoma) and have been accepted as meeting the end of life criteria in previous NICE appraisals of treatments for advanced (unresectable or metastatic) melanoma (e.g. NICE TA269[26]; NICE TA384[4]). That is, the treatment is indicated for patients with a short life expectancy, normally less than 24 months, and is licensed for small patient populations. In terms of the latter, the CS estimates that there would be 576 patients eligible for vemurafenib + cobimetinib treatment in England and Wales. This is based on the assumption that 34% of patients with malignant melanoma would have the mutated BRAF V600 gene, though other sources suggest that around 50% of cutaneous melanomas harbour the gene.[12] As stated earlier in this report, expert clinical advice to the ERG is that the majority of BRAF mutation positive patients would be considered for immunotherapy as first line treatment. All of these factors suggest that the estimate of 576 patients is uncertain.

The third criterion, that the treatment offers an extension to life normally of at least an additional three months, is reported to have been met by median OS from the coBRIM trial showing a statistically significant increase of nearly five months. As Table 10 of this report shows, the difference in median OS between combination therapy and vemurafenib monotherapy was 4.9 months and therefore in excess of NICE's threshold (NB. an equivalent difference between combination therapy and dabrafenib monotherapy is not reported in the submission, though the NMA did report a survival benefit in favour of combination therapy in this comparison). In

summary, the ERG agrees with the company's assertion that combination therapy meets the end of life criteria.

6 INNOVATION

The CS provides a rationale for why vemurafenib + cobimetinib should be considered an innovative therapy (CS Section 2.5). It describes how the addition of cobimetinib to vemurafenib offers inhibition of MEK which reduces the possibility of drug resistance to monotherapy. The clinical efficacy and safety results of the coBRIM trial appear to support this mechanism, and hence the rationale for combination therapy. The ERG notes that the benefits of this innovation will be more apparent in patients for whom BRAF inhibitor therapy is considered to be the most appropriate first line treatment for advanced disease.

Vemurafenib + cobimetinib are taken orally, and are therefore more convenient than other forms of administration. It is noted in the CS that the advantages that this brings to patients social and working lives is not captured in the QALY calculations. The ERG notes that dabrafenib is also taken orally, but that the currently available immunotherapies (ipilimumab, pembrolizumab, and nivolumab) are administered intravenously.

7 DISCUSSION

7.1 Summary of clinical effectiveness issues

The rationale for combination therapy in BRAF mutation positive advanced melanoma patients, as stated in the CS, is to reduce the likelihood of resistance to BRAF inhibition monotherapy. The combination of vemurafenib and cobimetinib is reported to simultaneously target mutated BRAF V600 proteins and MEK proteins in melanoma cells, resulting in stronger inhibition of intracellular signaling, decreased tumour cell proliferation and thereby limit mechanisms of resistance to BRAF inhibition by vemurafenib monotherapy. The coBRIM trial found that median PFS was increased by between 3.7 and 5.3 months (depending on which analysis used) by vemurafenib + cobimetinib versus vemurafenib. Similarly, median OS was also increased by just under five months. Likewise, the company's NMA reported superior efficacy for vemurafenib + cobimetinib compared with dabrafenib in terms of PFS and OS, though the ERG has noted caveats about the validity of the indirect comparison. The results therefore provide some degree

of support the assertion that this combination therapy is more efficacious than BRAF inhibitor monotherapy, in terms of survival outcomes.

Combination therapy appears to be generally well tolerated, though in the coBRIM trial Grade 4 AEs appeared to be more common in the vemurafenib + cobimetinib group.

[redacted]There also appears to be a HRQoL advantage associated with combination therapy, with improvements resulting from less insomnia and/or other factors (e.g. the incidence of non-melanoma skin cancers). However, it is unclear whether these analyses would be statistically significant, and the durability of improvements over the treatment period and beyond is uncertain.

The coBRIM trial included centres in United States, Canada, Australia, New Zealand, Europe, Russia, Turkey, and Israel. There were a small number of patients from the UK and in terms of generalisability the CS suggests that the outcomes of the study were as would be expected in UK patients. Expert clinical advice to the ERG was that the patients in the coBRIM trial are similar to those seen in the UK. It is noted that the inclusion criteria excluded patients with poorer prognosis (e.g. patients with life expectancy <12 weeks were not eligible), but expert clinical advice to the ERG was that these patients would still be treated with targeted therapy. Further, as stated earlier in this report, the trial did not include patients who had previously been treated for advanced disease, even though the scope of this NICE appraisal permitted previously treated patients. As noted earlier, many BRAF mutation positive patients would be treated with immunotherapy first line before switching to BRAF inhibitor and MEK inhibitor treatment as necessary. The CS does not take this into account in its modelling of cost-effectiveness, though expert clinical advice to the ERG is that there are no data to suggest that outcomes would be worse for second line treatment.

7.2 Summary of cost effectiveness issues

The CS includes evidence on the cost effectiveness of cobimetinib in combination with vemurafenib compared to vemurafenib and dabrafenib for BRAF mutation positive patients with advanced melanoma. The economic evaluation generally conforms to the NICE reference case and the model structure and model parameter inputs are consistent with the clinical disease pathways and the available trial evidence. The CS used direct evidence from the coBRIM trial in

the comparison between vemurafenib + cobimetinib and vemurafenib and an indirect comparison using evidence from an NMA for the comparison between vemurafenib + cobimetinib and vemurafenib.

The company's base case analysis for vemurafenib + cobimetinib compared to vemurafenib was £150,514 per QALY gained using list prices. The company performed a wide range of sensitivity analyses, including deterministic, probabilistic and scenario analyses to assess model uncertainty. However in all analyses the cost effectiveness estimates were higher than £50,000 per QALY gained. Furthermore the company's PSA analyses showed a 0% probability that vemurafenib + cobimetinib would be cost effective at a willingness to pay of £50,000 per QALY. The ERG suggests an alternative approach for modelling time on treatment and this would further increase the ICER to £204,340 per QALY gained for vemurafenib + cobimetinib compared to vemurafenib. Additional sensitivity and scenario analyses conducted by the ERG resulted in ICERs above £100,000 per QALY gained for all analyses.

8 REFERENCES

1. National Institute for Health and Care Excellence. Melanoma: assessment and management (NG14). https://wwwniceorguk/guidance/ng14/chapter/introduction 2015.
2. Erdei E, Torres SM. A new understanding in the epidemiology of melanoma. Expert review of anticancer therapy 2010;**10**(11):1811-23.
3. National Institute for Health and Care Excellence. Treating stage IV melanoma. http://pathwaysniceorguk/pathways/melanoma Pathway last updated: 17 February 2016 2016.
4. National Institute for Health and Care Excellence. Nivolumab for treating advanced (unresectable or metastatic melanoma) TA384. Secondary Nivolumab for treating advanced (unresectable or metastatic melanoma) TA384 2016.
5. National Institute for Health and Care Excellence. Ipilimumab for previously treated advanced (unresectable or metastatic) melanoma. (TA268). Secondary Ipilimumab for previously treated advanced (unresectable or metastatic) melanoma. (TA268) 2012.
6. National Institute for Health and Care Excellence. Ipilimumab for previously untreated advanced (unresectable or metatstatic) melanoma. TA319. Secondary Ipilimumab for previously untreated advanced (unresectable or metatstatic) melanoma. TA319 2014.
7. National Institute for Health and Care Excellence. Dabrafenib for treating unresectable or metatstaic BRAF V600 mutation-positive melanoma. TA321. Secondary Dabrafenib for treating unresectable or metatstaic BRAF V600 mutation-positive melanoma. TA321 2014.
8. National Institute for Health and Care Excellence. Vemurafenib for treating locally advanced or metastatic BRAF V600 mutation-positive malignant melanoma. NICE technology appraisals (TA269). https://wwwniceorguk/guidance/ta269.

9. National Institute for Health and Care Excellence. Pembrolizumab for advanced melanoma not previously treated with ipilimumab. TA366. Secondary Pembrolizumab for advanced melanoma not previously treated with ipilimumab. TA366 2015.
10. National Institute for Health and Care Excellence. Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab. TA357. Secondary Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab. TA357 2015.
11. European Medicines Agency. Cotellic (cobimetinib): EPAR - Product Information. http://wwwemaeuropaeu/docs/en_GB/document_library/EPAR_-_Product_Information/human/003960/WC500198563pdf;**24/9/2015**. **24/2/2016**.
12. Larkin J, Ascierto PA, Dréno B, et al. Combined Vemurafenib and Cobimetinib in BRAF-Mutated Melanoma. New England Journal of Medicine 2014;**371**(20):1867-76.
13. Atkinson V, Larkin J, McArthur GA, et al. Improved Overall Survival With Cobimetinib + Vemurafenib in Advanced BRAFV600-mutated Melanoma and Biomarker Correlates of Efficacy. Oral presentation at SMR congress, November 2015 2015.
14. Dréno B, Ribas A, Larkin J, et al. Incidence, course and management of toxicities associated with vemurafenib and cobimetinib in patients with melanoma: Insights from the CoBRIM study. Society for Melanoma Research – 11th International Melanoma Congress, 2014 Poster presentation 2014.
15. Larkin J, Yan Y, McArthur GA, et al. Update of progression-free survival and correlative biomarker analysis from coBRIM: Phase 3 study of cobimetinib plus vemurafenib in advanced BRAF-mutated melanoma. Journal of Clinical Oncology 2015;**33**((suppl; abstr 9006)).
16. Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. York Publishing Services Ltd., 2009.
17. Montedori A, Bonacini MI, Casazza G, et al. Modified versus standard intention-to-treat reporting: Are there differences in methodological quality, sponsorship, and findings in randomized trials? A cross-sectional study. Trials 2011;**12**:58.
18. Hibler BS, Rossi AM. Health-related quality of life in skin cancer patients. Am J Orthop 2014;**94**(5):215-16.
19. Dréno B, Bartley K, Ascierto PA, et al. Quality-of-life assessment in patients with metastatic melanoma receiving vemurafenib and cobimetinib. Journal of Clinical Oncology 2015;**33**((suppl; abstr 9021)).
20. Chapman PB, Hauschild A, Robert C, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. New England Journal of Medicine 2011;**364**(26):2507-16.
21. Flaherty KT, Infante JR, Daud A, et al. Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. New England Journal of Medicine 2012;**367**(18):1694-703.
22. Flaherty KT, Robert C, Hersey P, et al. Improved survival with MEK inhibition in BRAF-mutated melanoma. New England Journal of Medicine 2012;**367**(2):107-14.
23. Hauschild A, Grob JJ, Demidov LV, et al. Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. Lancet (London, England) 2012;**380**(9839):358-65.
24. Long GV, Stroyakovskiy D, Gogas H, et al. Dabrafenib and trametinib versus dabrafenib and placebo for Val600 BRAF-mutant melanoma: a multicentre, double-blind, phase 3 randomised controlled trial. Lancet (London, England) 2015;**386**(9992):444-51.
25. Robert C, Karaszewska B, Schachter J, et al. Improved Overall Survival in Melanoma with Combined Dabrafenib and Trametinib. New England Journal of Medicine 2015;**372**(1):30-39.
26. National Institute for Health and Care Excellence. Vemurafenib for treating locally advanced or metastatic BRAF V600 mutation-positive malignant melanoma. TA269. Secondary

Vemurafenib for treating locally advanced or metastatic BRAF V600 mutation-positive malignant melanoma. TA269 2012.
27. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2014;**17**(2):157-73.
28. Bucher HC, Guyatt GH, Griffith LE, et al. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. Journal of Clinical Epidemiology 1997;**50**(6):683-91.
29. Nikolakopoulou A, Chaimani A, Veroniki AA, et al. Characteristics of networks of interventions: a description of a database of 186 published networks. PloS one 2014;**9**(1):e86754.
30. Orbe J, Ferreira E, Nunez-Anton V. Comparing proportional hazards and accelerated failure time models for survival analysis. Statistics in medicine 2002;**21**(22):3493-510.
31. The Comparative Efficacy of First-Line Treatments for Stage IIIC and Stage IV Melanoma: Results of a Systematic Review and Network Meta-Analysis. . European Society for Medical Oncology (ESMO) congress Poster presentation 1095P; 2014 2014.
32. Bradburn MJ, Clark TG, Love SB, et al. Survival Analysis Part II: Multivariate data analysis - an introduction to concepts and methods. British Journal of Cancer 2003;**89**(3):431-36.
33. Latimer NR, Abrams K, Lambert P, et al. Adjusting for treatment switching in randomised controlled trials - A simulation study and a simplified two-stage method. Statistical Methods in Medical Research 2014.
34. Fleeman N, Baghurst A, Beale S, et al. Dabrafenib for treating unresectable, advanced or metastatic BRAFv600 mutation-positive melanoma [ID605]: A Single Technology Appraisal. Liverpool: Liverpool Reviews and Implementation Group (LRiG), University of Liverpool, 2014.
35. Pike E, Torkilseng E, Saeterdal I, et al. A health technology assessment of the new drugs for inoperable or metastatic malignant melanoma patients. Health Technology Assessment. Oslo: Norwegian Knowledge Centre for the Health Sciences, 2015.
36. Beusterien KM, Szabo SM, Kotapati S, et al. Societal preference values for advanced melanoma health states in the United Kingdom and Australia. British Journal of Cancer 2009;**101**(3):387-9.
37. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. NICE, London 2013.
38. Puzanov I, Amaravadi RK, McArthur GA, et al. Long-term outcome in BRAF(V600E) melanoma patients treated with vemurafenib: Patterns of disease progression and clinical management of limited progression. European Journal of Cancer (Oxford, England : 1990) 2015;**51**(11):1435-43.
39. Xing Y, Chang GJ, Hu CY, et al. Conditional survival estimates improve over time for patients with advanced melanoma: results from a population-based analysis. Cancer 2010;**116**(9):2234-41.
40. Giesinger JM, Kieffer JM, Fayers PM, et al. Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. Journal of Clinical Epidemiology 2016;**69**:79-88.
41. van Hout B, Janssen MF, Feng YS, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 2012;**15**(5):708-15.
42. Devlin NvH, B.;. An EQ-5D-5L value set for England. Secondary An EQ-5D-5L value set for England 2014. https://www.ohe.org/events/eq-5d-5l-value-set-england.

43. Robert C, Long GV, Brady B, et al. Nivolumab in previously untreated melanoma without BRAF mutation. New England Journal of Medicine 2015;**372**(4):320-30.
44. Johnston K, Levy AR, Lorigan P, et al. Economic impact of healthcare resource utilisation patterns among patients diagnosed with advanced melanoma in the United Kingdom, Italy, and France: results from a retrospective, longitudinal survey (MELODY study). European Journal of Cancer (Oxford, England : 1990) 2012;**48**(14):2175-82.
45. Economic burden of adverse effects associated with metastatic melanoma (MM) treatments in the United Kingdom. ISPOR 19th Annual International Meeting; 2014; Atlanta, USA. Value in Health.
46. Department of Health. NHS reference costs 2014-2015. http://wwwdhgovuk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_122803 2015 12/16/15 AD. (accessed 1/3/2016).

Cobimetinib in combination with vemurafenib for treating advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma

Changes needed to the model to estimate the Southampton Health Technology Assessments Centre (SHTAC) base case incremental analysis

The SHTAC base case makes the following changes:

i) Changes to the TOT for vemurafenib + cobimetinib to KM + log-logistic tail;
ii) Changes to utility values for PD to 0.73;
iii) Changes to dabrafenib dose, reduce weekly cost by 12%;
iv) Adjustment to the cost of PFS drug costs, adverse event costs and drug administration costs to allow for difference between PFS proxy and TOT. Reduce by 7%. (see model inputs!g414)

In order to compare treatments in an incremental analysis, the NMA data estimates are used for vemurafenib and dabrafenib vs vemurafenib + cobimetinib.

For cobimetinib + vemurafenib the actual costs from the direct trial evidence are used (ie from Results Table).

For dabrafenib, this analysis is already completed in the indirect treatment comparison worksheet and so the steps i-iv) above are completed.

For vemurafenib, the indirect treatment comparison sheet is used for vemurafenib instead and dabrafenib, by changing the PFS/OS AFT factors in the 'ITC parameter sheet' and the drug cost to that for vemurafenib and then steps i-iv above are completed. Results are shown in Results Table (ITC) sheet.