STA REPORT

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP (LRiG)

Dabrafenib for treating unresectable, advanced or metastatic BRAF^{v600} mutation-positive melanoma [ID605]

This report was commissioned by the NIHR HTA Programme as project number 12/54

Completed 03 July 2014
Factual Error check update: 21 July, 2014

DOES CONTAIN CIC/AIC

UNIVERSITY OF LIVERPOOL

LIVERPOOL REVIEWS AND IMPLEMENTATION GROUP

A MEMBER OF THE RUSSELL GROUP

Dabrafenib for treating unresectable, advanced or metastatic BRAF^{v600} mutation-positive melanoma [ID605]

Produced by: Liverpool Reviews and Implementation Group (LR*i*G)

Authors:

Nigel Fleeman, Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

Adrian Bagust, Professor (Health Economic Modelling), LR*i*G, University of Liverpool

Sophie Beale, Research Associate (Decision Analysis), LR*i*G, University of Liverpool

Angela Boland, Assistant Director, LR*i*G, University of Liverpool

Rumona Dickson, Director, LRiG, University of Liverpool

Kerry Dwan, Research Fellow (Medical Statistician), LR*i*G, University of Liverpool

Marty Richardson, Research Fellow (Medical Statistician), LR*i*G, University of Liverpool

Yenal Dundar, Research Fellow (Clinical Effectiveness), LR*i*G, University of Liverpool

Helen Davis, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

Lindsay Banks, North West Medicines Information Centre, Pharmacy Practice Unit, Liverpool

Correspondence to: Professor Rumona Dickson, Liverpool Reviews and Implementation Group, University of Liverpool, Room 2.12, Whelan Building, The Quadrangle, Brownlow Hill, Liverpool L69 3GB

Date completed: 03 July,2014

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 12/54

Declared competing interests of the authors:
None

Acknowledgements:
We would like to thank Ernie Marshall (MacMillan Consultant in Medical Oncology, The Clatterbridge Centre NHS Foundation Trust, Liverpool) and Ruth Plummer (Consultant in Medical Oncology, The Newcastle Hospitals NHS Foundation Trust) for critical comment on clinical data. We also thank Susanna Dodd (Lecturer in Biostatics, University of Liverpool) for critical comment on methods for adjusting for treatment crossover.

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:

Fleeman N, Bagust A, Beale S, Boland A, Dickson R, Dwan K, Richardson M, Dundar Y, Davis H, Banks L. Dabrafenib for the treatment of unresectable, advanced or metastatic BRAF^{v600} mutation-positive melanoma[ID605]: A Single Technology Appraisal. Liverpool Reviews and Implementation Group,2014.

Contributions of authors:

Fleeman N	Project lead, drafted clinical results section and supervised the final report
Bagust A	Checking and validation of the economic model and critique
Beale S	Checking and reporting on the quality of economic modelling and critique
Boland A	Critical comment on economic and clinical data
Dickson R	Summary and critical appraisal of the clinical evidence
Dwan K	Critical appraisal of the statistical evidence
Richardson M	Critical appraisal of the statistical evidence
Dundar Y	Cross checking of the submission search strategy
Davis H	Critical appraisal of the submission
Banks L	Critical appraisal of the submission

Table of contents

1 SUMMARY ... 7
1.1 Scope of the submission ... 7
1.2 Critique of the decision problem in the manufacturer's submission ... 7
1.3 Summary of clinical effectiveness evidence submitted by the manufacturer ... 7
1.4 Summary of the ERG's critique of submitted clinical effectiveness evidence ... 8
1.5 Summary of cost effectiveness evidence submitted by the manufacturer ... 9
1.6 Summary of the ERG's critique of cost effectiveness evidence submitted ... 10
1.7 Summary of manufacturer's case for end of life criteria being met ... 11
1.8 ERG commentary on end of life criteria ... 11
1.9 ERG commentary on the robustness of evidence submitted by the manufacturer ... 12
1.10 Summary of exploratory and sensitivity analyses undertaken by the ERG ... 13
1.11 ERG conclusions ... 13
2 BACKGROUND ... 15
2.1 Critique of manufacturer's description of underlying health problems ... 15
2.2 Critique of manufacturer's overview of current service provision ... 16
3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM ... 18
3.1 Population ... 19
3.2 Intervention ... 19
3.3 Comparators ... 19
3.4 Outcomes ... 20
3.5 Economic analysis ... 20
3.6 Subgroups ... 20
3.7 Other relevant factors ... 20
4 CLINICAL EFFECTIVENESS ... 22
4.1 Critique of the methods of review(s) ... 22
4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these ... 24
4.3 Critique of trials identified and included in the ITC ... 35
4.4 Critique of the statistical approach employed for the ITC ... 41
4.5 Critique of the adverse events data ... 43
4.6 Critique of the health related quality of life data ... 49
4.7 Additional work on clinical effectiveness undertaken by ERG ... 49
4.8 Conclusions of the clinical effectiveness section ... 50
5 COST EFFECTIVENESS ... 52
5.1 Introduction ... 52
5.2 Overview of manufacturer's cost effectiveness literature review ... 52
5.3 ERG critique of the review ... 54
5.4 Overview of manufacturer's economic modelling ... 54
5.5 Detailed critique of manufacturer's economic model ... 67
6 ADDITIONAL WORK UNDERTAKEN BY ERG ... 79
7 END OF LIFE ... 88
8 KEY ISSUES ... 89
9 Overall Conclusions ... 91
9.1 Implications for research ... 91
10 REFERENCES ... 92
11 APPENDIX ... 97

List of tables

Table 1 Decision problem specified by NICE and addressed in the MS 18
Table 2 Location of key clinical information in the MS 22
Table 3 Eligibility criteria used for manufacturer's updated systematic review (October 2013) 23
Table 4: Summary of study characteristics of BREAK-3 25
Table 5: Results of comprehensive crossover adjustments for BREAK-3 OS data from the MS 31
Table 6: Characteristics and results of all dabrafenib studies 34
Table 7: Comparison of data-cuts 36
Table 8: Results of individual trials and inputs for the manufacturer's indirect treatment comparisons 37
Table 9: Characteristics of BREAK-3 and of BRIM-3 participants 39
Table 10: Results of indirect treatment comparison for PFS and OS (RPFSTM adjusted) for dabrafenib vs vemurafenib 40
Table 11: Results of unadjusted indirect treatment comparison for OS for dabrafenib vs vemurafenib 42
Table 12 Summary of general AE data for dabrafenib, dacarbazine and vemurafenib from Phase II and Phase III trials* 45
Table 13 Proportion (%) of all common, non-serious AEs in BREAK-3 and all-Grade AEs in BRIM-3 experienced by at least 15% of patients in any treatment arm* 46
Table 14 Proportion (%) of most common Grade 3+ adverse events reported in BREAK-3 and BRIM-3 experienced by at least 2% of patients in any treatment arm* 47
Table 15 Location of key economic information in the MS 52
Table 16 Data sources used to retrieve economic and quality of life evidence 53
Table 17 Economic evaluation search inclusion and exclusion criteria 53
Table 18 Drug doses used in the economic model 55
Table 19 Summary of PFS and OS evidence used in modelling 57
Table 20 Summary of quality of life values for cost effectiveness analysis 58
Table 21 Drug costs (28 day supply) 59
Table 22 Other drug related costs 60
Table 23 One-off costs associated with treatment 61
Table 24 Monthly health care costs 61
Table 25 Incidence of treatment related AEs 62
Table 26 Summary of mean (standard error [SE]) cost per adverse event used in the model 62
Table 27 Expected post-study anti-cancer therapy (PSACT) costs used in the model 63
Table 28 Base case results 63
Table 29 Summary of predicted (discounted) costs (£) for the base case 64
Table 30 Deterministic univariate sensitivity analysis results for dabrafenib vs dacarbazine (PAS prices) 65
Table 31 Deterministic univariate sensitivity analysis results for dabrafenib vs vemurafenib (PAS prices) 65
Table 32 Summary of PSA for dabrafenib (PAS prices) 66
Table 33 Most influential components of incremental costs and outcomes in manufacturer's base case decision model results 68
Table 34: Assessment of risk of bias of BREAK-3 and BRIM-3 conducted by the manufacturer 97

List of figures

Figure 1: Evidence network used to inform the ITC 35
Figure 2 Schematic of manufacturer's model 55
Figure 3 Cost effectiveness plane for pair-wise comparisons of dabrafenib vs vemurafenib and dacarbazine 66
Figure 4 Cost effectiveness acceptability curves for dabrafenib, vemurafenib and dacarbazine 67
Figure 5 Mortality rates in the manufacturer's base case analysis 71
Figure 6 PPS survival comparing dacarbazine patients who did and did not crossover to dabrafenib on disease progression 73
Figure 7 Matched PFS cumulative hazards plot comparing dacarbazine and dabrafenib patients in the BREAK-3 trial, relative to an assumption of a constant HR (Final PFS data-cut June 2012) 75
[redacted]) 75
Figure 9 Matched PFS cumulative hazards comparing dacarbazine and vemurafenib patients in the BRIM-3 trial, relative to an assumption of a constant HR (December 2010 data-cut) 75
Figure 10 Matched OS cumulative hazards comparing dacarbazine and vemurafenib patients in the BRIM-3 trial, relative to an assumption of a constant hazard ratio. (February 2012 data-cut censored at crossover) 76
Figure 11 Matched PFS cumulative hazards comparing dacarbazine and vemurafenib patients in the BREAK-3 trial, relative to an assumption of a constant hazard ratio. (December 2010 data-cut) 76
Figure 12 Matched OS cumulative hazards comparing dacarbazine and vemurafenib patients in the BRIM-3 trial, relative to an assumption of a constant hazard ratio. (February 2012 data-cut censored at crossover) 77
Figure 13 Indirect treatment comparisons to compare dabrafenib with vemurafenib 78
Figure 14 PFS BREAK-3 cumulative hazard Kaplan-Meier data, ERG long-term linear trends and the manufacturer's log-normal models 80
Figure 15 BREAK-3 Kaplan-Meier PFS data, ERG long-term exponential models and the manufacturer's log-normal models 80
[redacted] 82

1 SUMMARY

1.1 Scope of the submission

The remit of the Evidence Review Group (ERG) is to comment on the clinical and cost effectiveness evidence submitted to the National Institute for Health and Care Excellence (NICE) as part of the single technology appraisal (STA) process. Clinical and economic evidence have been submitted to NICE from GlaxoSmithKline in support of the use of dabrafenib (Tafinlar) for the treatment of advanced or metastatic $BRAF^{V600}$ mutation-positive melanoma. In Europe, dabrafenib is licensed 'for the treatment of adult patients with unresectable or metastatic melanoma with a $BRAF^{V600}$ mutation.'

1.2 Critique of the decision problem in the manufacturer's submission

The patient population and intervention addressed in the manufacturer's submission (MS) broadly reflects that in the NICE scope and for whom dabrafenib is licenced. However, the ERG notes that relevant comparative evidence only exists for treatment naïve patients and hence the manufacturer presents clinical and cost effectiveness evidence to support the use of dabrafenib for treatment-naïve patients with $BRAF^{v600}$ mutation-positive melanoma. Vemurafenib and dacarbazine are appropriate comparators to dabrafenib, being the current standard of care and previous standard of care, respectively. Outcomes addressed in the decision problem include overall survival (OS), progression-free survival (PFS), overall response rates (ORRs), adverse events (AEs) and health-related quality of life (HRQoL). These outcomes are standard in this disease area and are appropriate.

1.3 Summary of clinical effectiveness evidence submitted by the manufacturer

The manufacturer presents direct and indirect clinical effectiveness evidence to support use of dabrafenib. Direct evidence comes from BREAK-3, a phase III randomised controlled trial (RCT) which compares dabrafenib to dacarbazine in treatment-naïve patients with $BRAF^{v600E}$ mutation-positive metastatic melanoma. Compared with dacarbazine, a statistically significantly improvement in PFS (6.9 vs 2.7 months) but not OS (18.2 vs 15.6 months) was reported; however, OS data are not yet fully mature and a large proportion of patients (57%) crossed over from the control arm to receive dabrafenib on disease progression. Hence OS data adjusted for the effects of crossover using the Rank Preserving Structural Failure Time method (RPSFTM) were also reported. It was not possible to report the median RPSFTM adjusted OS in months but the manufacturer noted that there was no statistically significant difference between treatment arms (Hazard ratio [HR] 0.55,95% Confidence Interval [CI]: 0.21 to 1.43). This is to be expected as the RPSFTM maintains the p-value of the ITT analysis.

Supporting evidence for the outcomes of interest is presented from four additional studies, two RCTs (BRF113220 and Combi-d) and two non-RCTs (BREAK-2 and BREAK-MB). These include patients who were previously pre-treated with chemotherapy (BRF113220: 11%; BREAK-2: 80%) and cohorts of patients with brain metastases (BREAK-MB,48% of patients previously pre-treated for brain metastases). Estimates of PFS and OS varied across the studies, ranging from 3.8 months to 8.8 months (BREAK-MB and Combi-d respectively) for PFS and 7.2 months to 20.2 months (BREAK-MB and BRF113220 respectively) for OS. ORRs were similar across trials (51% to 59%) with the exception of patients with $BRAF^{v600K}$ mutation-positive melanoma in BREAK-2 (13%).

There is currently no direct evidence comparing dabrafenib with vemurafenib. Hence, the manufacturer undertook an indirect treatment comparison (ITC) to compare BREAK-3 with BRIM-3, a phase III RCT which compared vemurafenib to dacarbazine. No statistically significant differences in terms of PFS (HR 0.97,95% CI 0.59 to 1.60) or OS, using either RPSFTM adjusted OS (HR 0.86,95% CI: 0.32 to 2.29) or unadjusted OS data (HR 1.00,95% CI 0.62 to 1.62) were identified by the ITC.

Adverse events (AE) leading to treatment discontinuation were relatively rare for patients treated with dabrafenib across BREAK-3 and supportive studies (≤5%). Approximately 20% and 40% of patients in BREAK-3 who were treated with dabrafenib reported serious adverse events or Grade 3+ AEs respectively. In terms of specific all-Grade AEs, hyperproliferative skin toxicities (cutaneous squamous cell carcinoma and keratoacanthomas) and arthralgia appear to be more common in patients treated with vemurafenib (in BRIM-3) whereas pyrexia, hyperkeratosis and palmar-plantar erythrodysesthesia syndrome appear to be more common in patients treated with dabrafenib (BREAK-3 and supporting trials). No significant differences between dabrafenib and dacarbazine were reported in BREAK-3 in relation to HRQoL; it was not possible to compare HRQoL between dabrafenib and vemurafenib.

1.4 Summary of the ERG's critique of submitted clinical effectiveness evidence

The ERG is satisfied with the search strategy employed by the manufacturer to identify clinical effectiveness studies. It does not believe there are additional studies relevant to the decision problem which have not been identified by the manufacturer.

The ERG agrees that BREAK-3 presents supportive clinical evidence for the effectiveness of dabrafenib vs dacarbazine for treatment-naïve patients with $BRAF^{v600}$ mutation-positive melanoma in terms of PFS but not OS. The ERG is not convinced by the argument that treatment crossover explains a lack of OS benefit and does not believe the RPSFTM is an appropriate method of adjustment for crossover in BREAK-3 for three key reasons. Namely

(1) survival data from BREAK-3 are immature (the results are based on an interim analysis with few deaths) (2) it is questionable whether dabrafenib is the same when received on diagnosis as it is on disease progression (the assumption of a 'common treatment effect' is required for the appropriate use of the RPSFTM) and (3) subsequent treatments received on disease progression in the dabrafenib and dacarbazine arms differ. [redacted]

Both trials included in the ITC (BREAK-3 and BRIM-3) appear to be broadly similar in terms of patient eligibility criteria. Patients with V600K were inadvertently enrolled into BRIM-3. However, a potentially clinically meaningful baseline difference between the trials is that at baseline a greater proportion of patients had lactate dehydrogenase (LDH) levels above the upper limit of the normal range in both arms of BRIM-3 than either arm of BREAK-3. This may have had a negative impact on the prognosis of patients in BRIM-3 compared to patients in BREAK-3. Indeed, the ERG notes that the unadjusted median OS for the dacarbazine arm of BRIM-3 is lower than that of BREAK-3.

The ERG questions the validity of the approach used to conduct the ITC as key underlying assumptions are not met. Specifically, constant HRs for both PFS and OS data within both the BREAK-3 and BRIM-3 trials are not evident while comparison of dacarbazine OS data between the two trials indicates that constant proportion hazards are not supported ($p<0.001$). Hence the ERG does not believe there is sufficient evidence to support the results of the comparison between dabrafenib and vemurafenib.

1.5 Summary of cost effectiveness evidence submitted by the manufacturer

The manufacturer developed a de novo partitioned survival model. It was constructed in Microsoft Excel and structured using three patient health states (Progression-free, Post progression and Death). Variants of this model structure have been used in the modelling of metastatic oncology for a number of previous NICE Single Technology Appraisals. Three treatment options were considered - dabrafenib, vemurafenib and dacarbazine.

Progression-free survival and OS for dabrafenib were modelled using parametric survival functions fitted to BREAK-3 trial data. Hazard ratios, based on BREAK-3 trial data for dacarbazine and generated by the ITC for vemurafenib, were used to project survival up to the end of the 10-year model period. In the case of dacarbazine, the RPFSTM adjusted HR from BREAK-3 was used to project OS. For the remaining 20 years of the model timeframe, UK general population mortality data were used. Trial data have not been used directly in the

manufacturer's model. The utility values incorporated into the model were those derived from the BREAK-3 trial. Monthly health care costs were estimated by inflating the costs used in the NICE appraisal of ipilimumab for previously treated advanced (unresectable or metastatic) melanoma (TA268) to current costs.

A Patient Access Scheme (PAS) has been accepted. For the comparison of dabrafenib vs dacarbazine the manufacturer's incremental cost effectiveness ratio (ICER) per quality adjusted life year (QALY) gained is £49,019 and the ICER per life-year gained is £34,940. For the comparison of dabrafenib vs vemurafenib the ICER per QALY gained is £11,028 and the ICER per life year gained is £7,523.

The manufacturer carried out a wide range of deterministic sensitivity analyses for these two comparisons. The ICERs from the ten parameters that had the most influence on the ICER per QALY gained ranged from £8,286 to 'dominated' for dabrafenib vs dacarbazine and from £26,470 to 'dominated' for dabrafenib vs vemurafenib. The results of the manufacturer's probabilistic sensitivity analyses suggest that there are 6% and 56% chances respectively that the ICERs for dabrafenib vs dacarbazine and dabrafenib vs vemurafenib are less than £30,000 per QALY gained. The manufacturer also reports that there is a 56% chance that the ICER per QALY gained for dabrafenib vs dacarbazine is cost effective at a threshold of £50,000.

1.6 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG is satisfied with the search strategy employed by the manufacturer to identify cost effectiveness studies, and is reasonably confident that no other relevant published articles exist. In addition, the ERG agrees with the manufacturer's view that there is insufficient evidence to facilitate the development of an economic case for previously treated (second-line) patients.

The spreadsheet model submitted by the manufacturer is an adaptation of a generic cancer treatment model. The result is a model that is far more complex than would normally be considered necessary, or appropriate, to model the cost effectiveness of metastatic cancer. The complexity and over specification of the model mean that at times the ERG has struggled to trace the logic trails.

The main aspect of the manufacturer's model that gave cause for concern was the modelling of survival. The manufacturer has adopted a complex model approach involving three phases:

- Phase 1 (0 to 1.8 years) – a parametric survival curve is fitted to BREAK-3 dabrafenib data and HRs are used to adjust this curve to represent survival for

patients receiving dacarbazine and dabrafenib. For such an approach to be robust, a number of assumptions need to be valid:

  - The parametric curve gives a good fit to the dacarbazine data – the ERG has shown that using trial data followed by an exponential extrapolation gives a better fit;
  - Constant HRs are valid for both PFS and OS data within BREAK-3 (dabrafenib vs dacarbazine) – holds for OS but not PFS;
  - Constant HRs are valid for both PFS and OS data within BRIM-3 (vemurafenib vs dacarbazine) – holds for neither PFS nor OS;
  - PFS and OS data for dacarbazine within the two trials (BREAK-3 and BRIM-3) are broadly comparable and conform to a constant hazard relationship – holds for PFS but not OS.

- Phase 2 (1.8 to 10 years) – a log-logistic curve is fitted to 10-year American Joint Committee on Cancer data. This approach assumes that there is clinical or biological evidence to support use of the log-logistic curve; however, no such data are presented.
- Phase 3(10 to 30 years) - use of background UK mortality data assumes that long-term survivors of multiple myeloma are effectively cured; however, no evidence is presented to support this assumption.

The ERG is concerned about the discontinuities in mortality rates that occur in the model at 1.8 and 10 years. These abrupt changes have no clinical explanation and only relate to convenient choices made for modelling purposes. In addition, the manufacturer has used RPSFTM adjusted data as the basis for modelling OS but there is no evidence that post-progression survival in dacarbazine patients crossing over to receive dabrafenib is any different than that of dacarbazine patients who did not cross over.

1.7 Summary of manufacturer's case for end of life criteria being met

The manufacturer argues that similar to vemurafenib, dabrafenib meets the NICE end of life criteria. This is for three reasons: (1) based on BRIM-3 data, the patient population has a short life expectancy (i.e. under 24 months), with the median OS being 13.6 and 9.7 months for patients treated with vemurafenib and dacarbazine respectively (2) the patient population is small (maximum of 990 patients per annum) (3) data in the economic evaluation demonstrate an incremental OS for dabrafenib versus dacarbazine of 7.5 months (i.e. greater than the required 3 months).

1.8 ERG commentary on end of life criteria

The ERG notes that the figure used in the manufacturer's base case economic model employs a mean survival in the dacarbazine arm of [redacted] (i.e. over 24 months). Data reported from the BREAK-3 trial indicate the median OS to be 18.2 and 15.6 months a

difference of only 2.6 months. The ERG therefore points out that using the figures from the clinical data in the MS, the case for consideration as end of life treatment has not been met.

1.9 ERG commentary on the robustness of evidence submitted by the manufacturer

1.9.1 Strengths

- The majority of clinical effectiveness evidence is derived from two relatively large phase III trials (BREAK-3 for dabrafenib vs dacarbazine, BRIM-3 for vemurafenib vs dacarbazine), both of which are considered to be of good quality based on an assessment for risk of bias.
- Variants of the model structure used to derive the cost effectiveness evidence have been used previously in the modelling of metastatic oncology for a number of previous NICE STAs.

1.9.2 Weaknesses

- It has only been possible to present comparable RCT evidence and cost effectiveness evidence for treatment-naïve patients, a subgroup of the population for whom dabrafenib is licensed and which is also specified in the NICE scope.
- Both BREAK-3 and BRIM-3 allowed patients to cross over treatment from the control arm to the treatment arm, potentially confounding results.
- There is no direct evidence comparing dabrafenib to vemurafenib and so it has been necessary to compare these two treatments via an ITC. Unfortunately, the ERG believes the assumptions required in order for the ITC to present credible results are not met and so the findings from the ITC cannot be relied upon to inform either the clinical or cost effectiveness of dabrafenib vs vemurafenib.
- Aside from the use of RPSFTM adjusted data from the ITC to inform the decision model, the modelling of OS is further flawed due to discontinuity in mortality rates (as a result of using data from three different sources) and which also results in forecasting, in some patients, that metastatic cancer is cured after 10 years. The ERG is not aware of any published evidence to support this assumption.

1.9.3 Areas of uncertainty

- In order to adjust for treatment crossover, both BREAK-3 and BRIM-3 have employed the RPSFTM. It is a matter of some academic debate about which method is optimal for adjusting for crossover. Even assuming the RPSFTM is the optimal method in this instance, the ERG does not believe that the results will be reliable for three main reasons: (1) survival data from BREAK-3 are immature (the results are based on an interim analysis with few deaths) (2) it is questionable whether dabrafenib is the same when received on diagnosis as it is on disease progression (a key assumption required for the RPSFTM) and (3) treatments received on disease progression in BREAK-3 in the dabrafenib and dacarbazine arms differ. [redacted]
- It is also unclear if the BREAK-3 and BRIM-3 trials are sufficiently similar to be included in an ITC. Reasons for this uncertainty to some extent arise from the eligibility criteria which enables patients with $BRAF^{v600K}$ mutations to be included in BRIM-3 but not BREAK-3, although it is recognised the proportion of patients is small

(8%). Perhaps of greater clinical significance is the greater proportion of patients with LDH levels above the upper limit of the normal range in BRIM-3 which may also partly account for the difference in OS observed in the dacarbazine arms of the trials (15.6 months in BREAK-3 and 9.7 months in BRIM-3); the difference may also have been partly caused by differences in second-line treatment (3% of the dacarbazine arm received ipilimumab in BREAK-3 compared to 13% in BRIM-3; treatment crossover was 57% and 34% respectively)

1.10 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG employed the use of alternative interpretations of the BREAK-3 PFS and OS data. The ERG's re-analysis of PFS data resulted in an additional 7 to 8 days of PFS gain from use of dabrafenib. Using the ERG PFS estimates in the manufacturer's model results in a small increase in the incremental QALYs per patient (+0.16%). There is, however, a much larger increase in the incremental cost per patient (+6.32%). This is mainly due to additional days of treatment with dabrafenib. The ERG's alternative projection of OS reduced the estimated OS gain to [redacted] (from [redacted] in the manufacturer's base case).

The ERG also made two relatively minor alterations/corrections to the manufacturer's model, namely: re-calculating dacarbazine drug acquisition and administration costs and re-estimating the cost of additional lines of chemotherapy. These changes raise the manufacturer's ICER from £49,019 to £52,153 per QALY gained for dabrafenib vs dacarbazine.

Combining all of the ERG's changes results in an ICER of £112,727 per QALY gained for dabrafenib vs dacarbazine.

The ERG did not conduct exploratory analyses to examine the ICER for dabrafenib vs vemurafenib. This was because the ERG did not believe the results from the ITC produced credible results for the reasons outlined above in section 1.4 and 1.9.2.

1.11 ERG conclusions

The ERG's examination of the trial results and economic model provided by the manufacturer indicates a clinical benefit of dabrafenib over dacarbazine in a treatment-naïve population with statistically significant results related to PFS but not OS. The ERG identified a number of issues related to economic model assumptions and parameters. Additional analysis carried out by the ERG indicates that the incremental cost per QALY gained (dabrafenib vs dacarbazine) may in fact be as high as £112,727.

Comparison of clinical effectiveness data related to the use of dabrafenib and vemurafenib via the ITCs presented in the MS were not considered to be valid. The ERG is therefore unable to assess the clinical or cost effectiveness of dabrafenib compared to vemurafenib.

Implications for research

Ideally, a study directly comparing dabrafenib with vemurafenib is required. Alternatively, if and when findings from additional studies of dabrafenib and vemurafenib are available, an indirect or multiple treatment comparison may suffice. Given the problems identified by the ERG with the current ITC, it is important that any assumptions underlying such a comparison are tested and patients included in the analyses are sufficiently similar across studies.

2 BACKGROUND

2.1 Critique of manufacturer's description of underlying health problems

The manufacturer appropriately presents the key issues related to the underlying health problem, including epidemiology, diagnosis and prognosis, in section 2 of the MS.[3] A summary of these sections as stated in the MS[3] is presented in Box 1.

Box 1: Epidemiology and prognosis

Aetiology and epidemiology

- Although representing less than 5% of skin cancers overall, malignant melanoma is the most aggressive form of skin cancer and can be fatal, particularly if the disease is not detected and treated at an early stage.[4]
- Malignant melanoma is currently the fifth most common cancer in the UK, accounting for 4% of all new cancer cases.[5]
- Incidence rates [of malignant melanoma] in the UK have risen sharply from 11 cases per 100,000 population in 1999-2001 to 17 cases per 100,000 population in 2008-2010.[5]
- In 2010, there were 12,818 new cases of malignant melanoma diagnosed in the UK,[5] and in 2011 2,209 deaths from the disease, representing 1% of all cancer deaths (2% of all male deaths from cancer) and 80% of deaths from skin cancer.[5]
- Metastatic disease occurs in about 20% of patients with malignant melanoma.[6]
- $BRAF^{v600}$ mutations are found in approximately 50% of melanomas.[7]

Prognosis

- A melanoma often has a slow early growth phase, and if detected before it has spread, can be cured by surgical resection. Approximately 80-90% of melanomas are diagnosed as localised tumours (Stages I and II) without regional or metastatic spread and can be dealt with in this way.[6,8,9]
- An estimated 16.5% of those with primary melanoma relapse with regional or distant recurrence, and 2 to 8% of people already have metastatic disease when diagnosed,[10,11] most commonly the skin or subcutaneous soft tissues, lung, and brain.[8]
- Patients with metastases to the brain, lung and other visceral sites have a poorer survival than those with metastases in the skin or subcutaneous tissues.[8]
- For stage III melanoma (in which disease has spread to nearby lymph nodes),5-year survival rates can vary from 40 to around 80%, depending on the number of lymph nodes involved and the nodal metastatic burden.[8]
- Stage IV melanoma (in which disease has spread to other parts of the body) has an extremely poor prognosis and is largely refractory to conventional chemotherapy. Median survival times of 6 to 18 months after stage IV diagnosis are reported, depending on the organs involved.[8,9]
- Up to 50% of patients develop brain metastases during the course of their illness and 20% have them at first presentation of metastatic disease.[12,13] People with melanoma brain metastases have a life expectancy of only 3 to 5 months.[9] This therefore represents a unique subgroup of metastatic melanoma patients with a particularly poor prognosis.

2.2 Critique of manufacturer's overview of current service provision

The manufacturer appropriately outlines the current service provision for patients with malignant melanoma in section 2 of the MS.[3] This is summarised as stated in the MS[3] in Box 2.

Box 2: Current treatment and mutation screening

First-line treatment

- Systemic therapy is now considered the mainstay of treatment for patients with stage IIIc (inoperable regional metastases) or stage IV (distant metastases) disease.[9] The goals of such treatments are to delay disease progression and palliate symptoms whilst minimising toxicity and detriment to quality of life.[14,15]
- Since receiving a positive recommendation from NICE in December 2012, the BRAF inhibitor, vemurafenib, has become the standard of care for BRAF^{v600} mutation-positive unresectable or metastatic melanoma,[16] largely replacing dacarbazine. 92% of metastatic melanoma patients who tested BRAF positive receive first-line vemurafenib.[17]
- People with BRAF-wild type metastatic melanoma continue to receive first-line therapy with dacarbazine or may be entered into a clinical trial with an experimental therapeutic option.
- It is unclear at this stage how the first line indication for ipilimumab will affect clinical practice.

Second-line treatment

- Following the positive NICE recommendation for ipilimumab, all patients whose disease progresses on first-line treatment may be considered for ipilumumab[18] or other immunotherapies in the context of clinical trials.[9]

Dabrafenib is an alternative BRAF inhibitor to vemurafenib.[19-21] It has received marketing authorisation in Europe as monotherapy 'for the treatment of adult patients with unresectable or metastatic melanoma with a BRAF^{v600} mutation.'[22] The ERG notes that dabrafenib is currently available through the Cancer Drugs Fund. According to information provided by Melanoma Focus in its statement accompanying the MS,[3] the following criteria need to be met:

- The application should made by a consultant specialist specifically trained and accredited in the use of systemic anti-cancer therapy who should also be responsible for the first cycle of systemic anti-cancer therapy
- Patients must have advanced melanoma
- Patients must be BRAF^{v600E} mutation-positive
- Patients must have Eastern Cooperative Oncology Group (ECOG) performance status (PS)PS 0 or 1
- Patients must have severe intolerance of vemurafenib necessitating discontinuation within 2 months of initiating treatment with vemurafenib
- Patients must have an absence of disease progression whilst on full dose of vemurafenib
- Patients should not have received other previous systemic therapy for the treatment of melanoma other than vemurafenib

According to the manufacturer, if recommended,990 patients a year in England and Wales would be eligible for treatment (assuming all diagnosed patients are considered fit for treatment and all receive dabrafenib). The manufacturer states that this is comparable to the estimate of the number of patients considered to be eligible for vemurafenib (850 patients according to the manufacturer of vemurafenib in the Roche MS[23]).

The ERG also notes that in addition to second-line use, according to a press release from NICE on 13 June 2014,[24] NICE draft guidance now recommends ipilimumab for the first-line treatment for patients with unresectable stage III or IV malignant melanoma). It should be noted that patients treated with ipilimumab are not required to have BRAF^{v600} mutation-positive disease.

3 CRITIQUE OF MANUFACTURER'S DEFINITION OF DECISION PROBLEM

Table 1 displays the decision problem presented in the MS[3] and that addressed by the manufacturer (section 5, page 41).

Table 1 Decision problem specified by NICE and addressed in the MS

Parameter	Final scope issued by NICE	Decision problem addressed in the manufacturer's submission (and rationale for any differences)
Population	People with advanced or metastatic BRAF^{v600} mutation-positive melanoma	Adults with previously untreated advanced or metastatic BRAF^{v600} mutation-positive melanoma. There is insufficient evidence from the clinical trials with which to build an economic argument for the use of dabrafenib in people with BRAF^{v600} mutation-positive metastatic melanoma who have received prior therapy
Intervention	• Dabrafenib monotherapy	Dabrafenib monotherapy
Comparator(s)	For people with previously untreated melanoma: • Dacarbazine (or temozolomide for people whose melanoma has metastasised to the brain) • Vemurafenib For people with previously treated melanoma: • Dacarbazine (or temozolomide for people whose melanoma has metastasised to the brain) • Ipilimumab • Vemurafenib	In treatment-naïve patients with BRAF^{v600} mutation-positive metastatic melanoma: • Dacarbazine • Vemurafenib There is insufficient evidence from the clinical trials to build economic arguments for dacarbazine, ipilimumab or vemurafenib for people with previously treated melanoma There is insufficient evidence from the clinical trials to build economic arguments for temozolomide for people whose disease has metastasised to the brain
Outcomes	The outcome measures to be considered include: • Progression-free survival • Overall survival • Response rate • Adverse effects of treatment • Health-related quality of life	As per scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective. The availability of any patient access schemes for the intervention or comparator technologies should be taken into account.	As per scope

Parameter	Final scope issued by NICE	Decision problem addressed in the manufacturer's submission (and rationale for any differences)
Subgroups to be considered	None specified	There is insufficient evidence from the clinical trials with which to build an economic argument for the use of these treatments in people with BRAF^{v600} mutation-positive metastatic melanoma who have received prior therapy
Special considerations including equity or equality issues	Guidance will only be issued in accordance with the marketing authorisation. Cost of any additional mutational testing required for this treatment should be considered.	Evidence is derived from studies of treatment provided in accordance with the marketing authorisation. There are no equity or equality considerations

3.1 Population

The patient population specified in the scope is 'People with advanced or metastatic BRAF^{v600} mutation-positive melanoma'. A narrower patient population is addressed by the MS,[3] namely patients with stage III or IV melanoma with BRAF mutations who are treatment naïve (i.e. not previously treated with chemotherapy or BRAF inhibitors) for advanced or metastatic disease. This is because the majority of patients in the studies[1,2,25-28] from which evidence of effectiveness is derived (and all patients included in the two trials[1,2] within the manufacturer's indirect treatment comparison (ITC) and economic model) were similarly treatment naïve. It is further noted that the majority of patients in the studies[1,2,25-28] (and all patients in the pivotal BREAK-3[2] trial comparing dabrafenib to dacarbazine) had BRAF^{v600E} mutation-positive melanoma as opposed to BRAF^{v600K} mutation-positive melanoma. BRAF^{v600E} mutations are more prevalent (around 80%) than BRAF^{v600K} in clinical practice.[29] There is currently no evidence that the type of BRAF^{v600E} mutation effects prognosis or efficacy with any particular treatment (however, see also section 4.2.7 and 4.3.3).

3.2 Intervention

The intervention specified in the NICE scope is dabrafenib. This is the intervention addressed by the MS[3] where it is highlighted that dabrafenib inhibits the enzymatic activity of the mutated BRAF protein. The recommended dose of dabrafenib is 150 mg (two capsules of 75 mg) twice daily (corresponding to a total daily dose of 300 mg).

3.3 Comparators

Four comparators are specified in the NICE scope. Dacarbazine and vemurafenib are comparators for both those patients who are untreated and those who have had previous treatment for malignant melanoma. Ipilumumab is only specified as a comparator for patients previously treated for malignant melanoma. As noted in section 2.2, ipilimumab is recommended as a second-line treatment for patients with malignant melanoma; at the time of this appraisal, it was not recommended for first-line treatment (and was not specified in

the final NICE scope). Temozolomide is appropriately specified as a comparator for patients whose malignant melanoma has metastasised to the brain. Relevant evidence was only available to support the use of dacarbazine and vemurafenib.

3.4 Outcomes

Clinical evidence is reported in the MS[3] for all five outcomes specified in the scope: overall survival (OS),progression-free survival (PFS), response rate (reported as objective tumour response rate and disease control rate), adverse events (AEs) of treatment and healt-related quality of life (HRQoL).

3.5 Economic analysis

As specified in the final NICE scope, the cost effectiveness of treatments was expressed in terms of the incremental cost per QALY gained. Outcomes were assessed over a 30-year time horizon (equivalent to a lifetime horizon) and costs were considered from an NHS perspective.

3.6 Subgroups

No subgroups were specified by NICE. As noted in section 3.1, relevant evidence was only available for treatment naïve patients. Hence the MS[3] classifies adults with previously untreated advanced or metastatic mutation-positive melanoma to be a subgroup of the specified patient population. The manufacturer notes in section 2.3 of the MS that people with melanoma brain metastases represent a 'unique' subgroup of metastatic melanoma patients with a particularly poor prognosis. No RCT evidence was found for this subgroup.

3.7 Other relevant factors

Regarding the costs of additional testing, which is specified as a 'special consideration' in the NICE scope, the ERG notes that BRAF^{v600} testing is currently funded for patients with stage IIc disease and above by Roche, the manufacturer of vemurafenib. This funding scheme is due to end in December 2014. Testing for BRAF mutations is conducted using the Cobas v600 assay. This test was used in the BRIM-3[1] trial, a trial of vemurafenib vs dacarbazine. Although not explicitly specified in the MS,[3] the ERG notes from the clinical study report (CSR)[30] for BREAK-3[2] that an assessment of the analytical and clinical validity of a different companion diagnostic (the bioMerieux THxID BRAF assay) was planned. These results were not reported in the CSR[30] but the European Medicine's Agency Committee for Medicinal Products for Human Use European Assessment Report[31] notes that this assay has high specificity. It is therefore unclear if this test would supersede the Cobas v600 assay in clinical practice or, if not, if there would be any equity or equality issues (relating to cost and/or assay performance (e.g. with detecting BRAF^{v600E} and BRAF^{v600K}

mutation status) if different diagnostic tests were used for different drugs. However, the MS[3] does state that there are no known special considerations relating to equity or equality.

The ERG also notes that the manufacturer has submitted a patient access scheme (PAS) to the Patient Access Scheme Liaison Unit (PASLU) to support this submission. The impact of the proposed scheme has been incorporated into the manufacturer's base case assessments.

4 CLINICAL EFFECTIVENESS

This section provides a structured critique of the clinical evidence submitted by the manufacturer in support of the use of dabrafenib. Table 2 contains details of the location of key information within the MS.[3]

Table 2 Location of key clinical information in the MS

Key information	Page number	Key tables/figures
Decision problem	14 to 24	
Context	25 to 32	
Innovation	32 to 33	
Statement of the decision problem	34 to 36	
Identification of relevant to studies	37 to 38	
Eligibility criteria for inclusion of studies	38 to 40	Table 3
Information about included studies	41 to 57	Figure 3, Tables 5 to 8
Patient characteristics / disposition	58 to 61 / 74 to 79	Tables 9 and 10 / Figure 4
Trial outcomes and statistical analyses [adjustments for treatment crossover]	62 to 74 [71 to 73]	
Critical appraisal of studies	80 to 83	Table 22
Results of the relevant studies (pivotal trial : BREAK-3) [adjusted for treatment crossover]	84 to 94,104 to 106 (84 to 89) [93 to 94]	 (Tables 24 to 25, Figures 6 to 8) [Table 26]
Indirect treatment comparison (ITC): • information about pivotal trial : BRIM-3 • conduct of ITC • results from ITC	 96 to 98 99 to 102 102 to 103	 Tables 27 to 30 Figure 9, Tables 32 to 35 Tables 36 to 37, Figure 10
Adverse events	106 to 118	Tables 38 to 41, Table 44
Interpretation of clinical evidence	119 to 126	

4.1 Critique of the methods of review(s)

A systematic review was undertaken to identify all evidence relevant to the decision problem. This was conducted in October 2012 and updated in October 2013. A meta-analysis for any of the interventions/comparators was not possible as there was no more than one study for any given comparison. In order to compare dabrafenib with vemurafenib, an ITC was required.

4.1.1 Searches

Bibliographic databases including MEDLINE (1980 onwards and In process and other non-indexed citations), EMBASE (1980 onwards) and the Cochrane Library (1980 onwards) were searched on 22 October 2012. These searches were updated in October 2013 to identify additional published studies of relevance to the submission. No strategies for identifying ongoing studies are provided in the MS.[3] The specific search strategy (and number of hits) for each database is specified in Appendix 10.4.4 of the MS.[3] For all databases, a combination of free text and index terms were appropriately used.

4.1.2 Inclusion criteria

The inclusion and exclusion criteria used in the original systematic review (October 2012) are presented in detail in section 6.2.1 of the MS[3] (Table 3, page 38) and those for both the original and updated searches are included in Appendix 10.4 (Table 78, page 219). The criteria for the updated review are summarised in Table 3. In addition to the interventions and comparators specified by NICE, it is noted that trametinib, ipilimumab and fotemustine were also included. It is stated that these interventions were identified from clinical practice or from ongoing clinical trials for the treatment of metastatic melanoma. The ERG notes that as of October 2013, both dabrafenib monotherapy and dabrafenib + trametinib (dual therapy) were included in the NICE appraisal process.

Table 3 Eligibility criteria used for manufacturer's updated systematic review (October 2013)

Parameter	Included	Excluded
Population	Adults with advanced or metastatic melanoma	Studies of patients with other types of skin cancer were not included Studies with no subgroup data for the disease were not included, as these studies could introduce heterogeneity into the review Studies which enrolled a mixed population of stage I, II, III, and IV melanoma were only included if there was a subgroup analysis on the stage III and/or IV patient population
Interventions	Dabrafenib; trametinib; vemurafenib; dacarbazine	Other interventions
Comparators	Vemurafenib; dacarbazine	Other comparators
Outcomes	Efficacy: OS; PFS; ORR: CR + PR; proportion of patients with stable disease and progressive disease; TTR, DoR; HRQoL Safety: Incidence and severity of all AEs; incidence and severity of specific AEs; withdrawals due to AEs; withdrawals due to death; SAEs	Other outcomes
Study design	RCTs	All other types of studies, reviews, letters and commentaries
Language	English	Not English

AE=adverse event; CR=complete response; DoR=duration of response; HRQoL=health-related quality of life; ORR=overall response rate; OS-overall survival; PFS=progression-free survival; SAE=serious adverse event; TTR=time to treatment response

4.1.3 Data extraction

Details of the data extraction strategy are reported in Appendix 10.4.7 (page 222) in the MS.[3] Data from the included studies were appropriately extracted in parallel by two independent reviewers with any differences considered and reconciled by a third reviewer. Where more than one identified publication described a single trial, data were appropriately compiled into a single entry in the data extraction table to avoid the double counting of

patients. The data extraction table used is presented in the systematic review report.[32] The types of data extracted appear to be comprehensive.

4.1.4 Quality assessment

A descriptive critical appraisal of all the included studies was conducted using comprehensive assessment criteria based on the recommendations in the NICE STA guidelines.[33]

4.1.5 Evidence synthesis

All trials identified by the systematic review had different interventions and comparators and therefore it was not possible to conduct a meta-analysis. Findings were appropriately presented narratively. In addition, to compare interventions to comparators specified in the NICE scope, the manufacturer undertook an ITC. The ERGs critique of the manufacturer's ITC is presented in section 4.3.

4.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these

4.2.1 Identified studies

The manufacturer's systematic review identified two trials which provided evidence relevant to the comparisons of interest. These trials were BREAK-3,[2] which compared dabrafenib to dacarbazine, and BRIM-3,[1] which compared dacarbazine to vemurafenib. Therefore, it was possible to perform an ITC between dabrafenib and vemurafenib. As BREAK-3[2] was the only identified trial that examined the effects of dabrafenib, this is considered to be the key trial of interest. The characteristics of BRIM-3[1] and its comparability with BREAK-3[2] are explored in section 4.3.2. Supporting evidence was also provided from four other studies,[25-28] none of which were used in quantitative analyses. Further information on these trials is presented in section 4.2.7. The ERG is not aware of any additional studies of dabrafenib which are relevant to this appraisal. As no trials directly compared dabrafenib with vemurafenib, the ERG agrees with the manufacturer that it was not appropriate to conduct a meta-analysis.

4.2.2 Trial characteristics

BREAK-3,[2] is a phase III study conducted globally (across 12 countries, including: the U.S., European Union, Canada, Russian Federation and Australia); patients were randomised using a ratio of 3:1 to dabrafenib and dacarbazine respectively. The GlaxoSmithKline RAMOS (Registration and Medication Ordering System) interactive voice recognition system (IVRS) was used to assign patients to treatment groups, ensuring that allocation

concealment was achieved. Randomisation was stratified according to disease staging at study entry (unresectable III+IVM1a+IVb vs IVM1c). The ERG is satisfied that randomisation was carried out appropriately and allocations were adequately concealed in BREAK-3.[2]

The key trial characteristics are described in Table 7 of the MS[3] (page 49) and are summarised in Table 4.

Table 4: Summary of study characteristics of BREAK-3

Study Characteristic	BREAK-3[2]
Location	Multi-centre study involving 70 centres across 12 countries
Design	Randomised, open-label, multi-centre phase III pivotal study
Population	Subjects with treatment-naïve, stage IV or unresectable stage III $BRAF^{v600E}$ mutation-positive melanoma, with the exception of prior treatment with IL-2, surgery or radiotherapy
Duration of study	Ongoing. Enrolment between 23 December 2010 and 01 September 2011. Final data collection date for primary outcome measure: December 2011. Latest PFS data-cut June 2012. Follow-up for OS continues (latest data-cut; 18 December 2012)
Intervention(s) and comparator(s)	Oral dabrafenib 150mg twice daily, n=187 IV dacarbazine 1000mg/m2 every 3 weeks, n=63 Patients who had disease progression on dacarbazine (as confirmed by blinded, independent central reviewer) could crossover to receive dabrafenib 150mg twice daily. Patients were then followed for response, progression, survival, and further anti-cancer therapy
Primary outcomes (including scoring methods and timings of assessments)	Efficacy Investigator-assessed PFS
Secondary outcomes (including scoring methods and timings of assessments)	PFS by independent review OS ORR: CR + PR DoR HRQoL Safety Support of a BRAF mutation assay validation Pharmacokinetic analyses
Duration of follow-up	PFS, OS and response were analysed when a pre-specified number of PFS events (102) had accrued and the study was fully enrolled. A total of 118 PFS events had accrued at the time of December 2011 data-cut. The final analysis of OS will be conducted when 70% of subjects have died or been lost to follow-up As of the December 2011 data-cut: median follow-up time was 5.1 months (dabrafenib) and 4.8 months (dacarbazine); median OS was not reached As of the Jun 2012 data-cut: median follow-up time was 10.5 months (dabrafenib) and 9.9 months (dacarbazine) As of the December 2012 data-cut: median follow-up time was 15.2 months (dabrafenib) and 12.7 months (dacarbazine)

CR=complete response; DoR=duration of response; HRQoL=health-related quality of life; IL-2=Interlukin-2; ORR=objective response rate; OS=overall survival; PFS=Progression-free survival; PR=partial response

4.2.3 Participant characteristics

In total,250 patients were included in BREAK-3,[2] 187 in the dabrafenib arm and 63 in the dacarbazine arm. The participant characteristics of BREAK-3[2] are provided in Table 9 of the MS[3] (page 59). The MS[3] suggests that demographic and baseline disease characteristics were well balanced between treatment groups for the intention-to-treat (ITT) population in BREAK-3.[2] However, the ERG observed differences (≥5%) between treatment groups for the following characteristics: baseline lactate dehydrogenase (LDH), ECOG PS, visceral disease at baseline, and location of disease at baseline. The ERG does not believe that any of these differences are of particular concern. The ERG concludes that the patient population is similar to the population specified in the final scope issued by NICE.

4.2.4 Statistical approach

Sample size calculation

Details of the sample size calculation performed for BREAK-3[2] are reported in the study protocol. The ERG is satisfied with the pre-specified sample size calculation reported.

Protocol amendments

The ERG notes that overall response rate (ORR) was added as a secondary efficacy objective as part of the protocol amendments, which are listed in full in the CSR.[30] The ERG is satisfied that these amendments took place before recruitment into the trial began and so were not driven by early results. Therefore, the amendments are unlikely to be an important cause for concern.

Statistical analyses

In BREAK-3[2] the ITT population, which included all randomised patients, was used for efficacy analyses; in ITT analyses subjects are analysed according to the treatment group to which they were originally randomised. The crossover population consisted of all patients who were randomised to the dacarbazine arm but who switched to the dabrafenib arm on disease progression.

Analyses of PFS, OS, overall response rate (ORR), and duration of response (DoR) were performed for both the ITT and crossover populations separately. Kaplan-Meier curves were produced for both PFS and OS, and a log-rank test (stratified on disease staging) was conducted to compare treatment effects between the two treatment arms. Primary analyses of PFS were based on investigators' assessments, with additional analyses performed according to independent review. ORR (defined as subjects achieving a complete response [CR] or partial response [PR] per RECIST v1.1 criteria[34]) was calculated using both investigator assessment and independent review. Kaplan-Meier estimates were used to

calculate median DoR for patients in the ITT population who showed a CR or PR. The ERG is satisfied that the statistical analysis methods used were appropriate.

A number of pharmacokinetic analyses and BRAF assay validation analyses were also detailed in the statistical analysis plan (SAP) but not reported in the MS.[3] Safety and tolerability analyses were also specified, as measured by exposure, AEs, and changes from baseline in vital signs, laboratory parameters, and left ventricular ejection fraction; and the percentage of treatment-emergent non-melanoma skin lesions and second malignancies in each treatment group. Key AEs were reported in the MS (see section 4.5). HRQoL analyses were also specified and reported (see section 4.6).

Sensitivity analyses

Sensitivity analyses were pre-specified in the protocol for BREAK-3,[2] in order to assess robustness of treatment effect on PFS. These analyses are:

- Including symptomatic progression as PFS event
- Not censoring for new cancer therapy or extended loss to follow-up
- Cox regression stepwise selection of covariates: age, sex, presence of visceral disease, baseline LDH, baseline ECOG PS, number of disease sites.

There was also a post-hoc analysis to determine the potential effect of subjects coming in off-schedule for disease assessments.

The ERG observed that all sensitivity analyses results were reported sufficiently in the CSR,[30] although these were not presented in the MS.[3]

Subgroup analyses

The following subgroup analyses were pre-specified in the SAP for PFS:

- Number of metastatic disease sites [<3 or ≥3]
- ECOG PS [=0 or ≥1]
- Visceral disease [yes or no]
- Baseline LDH [≤ upper limit of normal (ULN) or >ULN]
- Age [<65 or ≥65]
- Sex
- Disease stage (IVM1c or III, IVM1a and IVM1b)
- Race: white, African-American/African heritage, Asian, other
- History of brain metastases (yes or no)

The manufacturer presented the results of subgroup analyses (Figure 22, page 242 of the MS[3]) and tests of interaction were provided as supplementary data. However, the results

from the subgroup analyses of race and history of brain metastases (as opposed to no metastases) were not reported. The ERG is unclear as to why these results were not reported in the MS[3] or CSR[30] particularly as patients with brain metastases have been identified as a 'unique' subgroup by the manufacturer (see section 3.6). Nevertheless, the ERG believes that these findings are unlikely to have a large impact on the conclusions of the study, unless the reason for non-reporting is based on the statistical significance or size or direction of the treatment effect estimate.

4.2.5 Results

A summary of the primary and secondary outcome data from BREAK-3[2] are presented in the MS[3] (Table 23, page 85). The population utilised in the efficacy analyses is the ITT population.

Progression-free survival and overall survival

BREAK-3[2] results show a trend in favour of dabrafenib over dacarbazine in terms of both PFS (hazard ratio [HR] 0.37,95% confidence interval [CI]: 0.24 to 0.58; June 2012 data-cut), and OS (HR 0.76,95% CI: 0.48 to 1.21; December 2012 data-cut). However, this benefit was found to be significant for PFS only ($p<0.0001$).

PFS was also assessed by independent review. The results from this assessment are provided in the CSR,[30] although they are provided for the December 2011 data-cut and not for the June 2012 data-cut. The independent review results (HR 0.35,95% CI: 0.20 to 0.61) support the findings of the investigator assessment from the December 2011 data-cut (HR 0.30,95% CI: 0.18 to 0.51).

Results from the subgroup analyses are presented in Figure 7 in the MS[3] (page 88). Results from the tests for interaction for each of these subgroup analyses were provided as supplementary data. None of these tests generated a significant p-value.

Overall response rate

The ORR results for BREAK-3[2] suggest that dabrafenib improves ORR in comparison to dacarbazine, due to the substantial difference in response rates between treatment groups (59% vs 24%, respectively; June 2012 data-cut). Duration of response data are presented in the MS[3] in Table 24 (page 87). However, it is difficult to draw any conclusions regarding duration of response due to the small difference in median duration of response between dabrafenib (8.0 months,95% CI: 6.6 to 11.5) and dacarbazine (7.6 months,95% CI: 5.0 to 9.7) and the uncertainty around these estimates due to small sample size (dabrafenib n=110, dacarbazine n=15).

ORR was also assessed by independent review, and results from this assessment were provided in the CSR.[30] Again, these results were only available for the December 2011 data-cut. Confirmed ORRs were lower for independent review (50% vs 6%, dabrafenib vs dacarbazine) than those reported by investigator assessment (53% vs 19%, dabrafenib vs dacarbazine).

4.2.6 Results when adjusting for confounding due to treatment switching

Participants of the BREAK-3[2] study were allowed to crossover from the dacarbazine arm to the dabrafenib arm when their disease had progressed; patients in both arms also received other second-line treatment at disease progression. At the time of the analysis, it is reported in the MS[3] that 36 (57%) crossed over from dacarbazine to dabrafenib. In a clinical trial, when patients are allowed to crossover to a different treatment, a simple ITT analysis of OS may underestimate the incremental difference in effect between the intervention and the comparator arms if any effect of this additional treatment is not taken into account. In the MS[3] the manufacturer has highlighted that the optimal method to adjust for the effect of crossover in survival analyses in RCTs remains an area of academic debate, with each approach having different advantages and limitations.

The manufacturer employed a team from the University of Sheffield to consider the potential effect of crossover in the BREAK-3[2] study (results from a range of possible approaches are shown in Table 26 of the MS[3] and reproduced here in Table 5). The team concluded that the most appropriate crossover adjustment technique to use on the dacarbazine OS data from the BREAK-3[2] trial was the treatment group Rank Preserving Structural Failure Time method (RPSFTM). However, in a report prepared for the manufacturer by external consults,[35] it is concluded that the results of the RPFSTM analyses:

[redacted]

The ERG agrees with the main conclusions of the authors of this report.[35] The ERG highlights the three key points regarding the use of the RPSFTM model for consideration of the effect of crossover in the BREAK-3[2] study.

First, survival data from the BREAK-3[2] study are immature; the results are based on an interim analysis with few deaths.

Second, the manufacturer's RPSFTM analysis assumes that the effect of receiving dabrafenib is the same when received on diagnosis (i.e. in the dabrafenib group) as it is on disease progression (i.e. in dacarbazine crossover patients). The ERG questions the validity of this assumption as there appears to be no clinical effectiveness evidence to confirm or refute this claim.

Third, treatments received on disease progression in the dacarbazine arm and in the dabrafenib arm should be similar and typical of clinical practice. From trial data made available by the manufacturer at clarification, the ERG notes that 43% of patients in the dabrafenib arm received subsequent treatment whilst 21% of patients in the dacarbazine arm received subsequent treatment. A range of 21 different treatments were administered to patients in the dabrafenib patients including ipilimumab (20%), dacarbazine (11%) and vemurafenib (5%) and seven different treatments were administered to dacarbazine patients including ipilimumab (3%) and vemurafenib (15%). The ERG does not consider that all patients in the BREAK-3[2] study received similar treatments at the time of disease progression or that patients in both groups received treatments that are routinely available in UK clinical practice.

The ERG speculates that the Inverse Probability of Censoring Weighting (IPCW) method may be more appropriate than the RPSFTM, as this method can appropriately adjust for all treatment deviations while still estimating the causal estimate of interest. The IPCW method considers follow-up only up until the time of the first deviation from randomised treatment; as such all treatment changes are handled identically, as patients' follow-up is censored at the time of the first treatment change of any form. The potential selection bias introduced by such censoring is addressed by weighting the remaining uncensored patients' follow-up according to the assumption of no unmeasured confounders (NUC). Use of this method requires pre-planning, as it is necessary to decide during the design stage which factors are likely to be related both to treatment change and outcome (e.g. measure of disease severity or adverse reactions) in order to ensure appropriate data collection on all such confounders. However, the validity of the IPCW is dependent on sufficient numbers of events and thus the immature OS data may still invalidate the use of the IPCW at present. The ERG notes that the manufacturer intends to run IPCW adjustments on the mature OS data in future if this is advised by the external consultants who have advised them on adjusting for crossover.

However, the ERG also compared survival data from patients in the dacarbazine arm who did and did not crossover to receive dabrafenib. The results of this analysis appear to demonstrate that the patients who crossed over did not gain any significant benefit over the patients who did not crossover. The ERG's analysis is fully explained in section 0.

In summary, the ERG does not consider that the RPSFTM adjusted OS results for patients in the dabrafenib arm of the BREAK-3[2] study are sufficiently robust to support the manufacturer's main argument in support of dabrafenib. The ERG is of the opinion that the RPSFTM adjusted OS results should only be considered as a sensitivity analysis.

Table 5: Results of comprehensive crossover adjustments for BREAK-3 OS data from the MS

Method type	Description	BREAK-3[2] trial (dabrafenib monotherapy)	
		June 2012 (HR,95% CI)	December 2012 (HR,95% CI)
1. ITT	Intention-to-treat analysis (to provide a 'baseline' analysis)	0.75 (0.45 to 1.24)	0.76 (0.49 to 1.18)
2. Naive	Censor crossover patients at time of crossover	1.20 (0.54 to 2.65)	1.37 (0.66 to 2.85)
3. Naive	Exclude crossover patients	0.93 (0.42 to 2.05)	1.07 (0.52 to 2.22)
4. Naive	Include experimental treatment received as a time-dependent covariate	1.26 (0.58 to 2.76)	1.49 (0.72 to 3.08)
5. Naive	Include treatment crossover indicator as a time-dependent covariate	1.15 (0.52 to 2.54)	1.35 (0.65 to 2.80)
6. Complex	RPSFTM 'treatment group' analysis. This estimates an average treatment effect of being allocated to the experimental group, adjusted for crossover	0.52 (0.17 to 1.61)	0.55 (0.21 to 1.43)
7. Complex	RPSFTM 'on treatment' analysis. This estimates the average treatment effect of full treatment (no discontinuation) compared to no treatment	0.34 (0.05 to 2.19)	0.27 (0.03 to 2.19)
8. Complex	RPSFTM 'on treatment - observed' analysis. This estimates the average treatment effect of observed experimental treatment compared to no treatment, assuming an effect is only received while on treatment	0.57 (0.22 to 1.50)	0.55 (0.21 to 1.43)
9. Complex	IPE Algorithm 'treatment group' analysis. This estimates an average treatment effect of being allocated to the experimental group, adjusted for crossover	0.52 (0.18 to 1.55)	0.50 (0.15 to 1.64)
10. Complex	IPE Algorithm 'on treatment' analysis. This estimates the average treatment effect of full treatment (no discontinuation) compared to no treatment	0.36 (0.06 to 1.99)	0.22 (0.02 to 2.90)
11. Complex	IPE Algorithm 'on treatment - observed' analysis. This estimates the average treatment effect of observed experimental treatment compared to no treatment, assuming an effect is only received while on treatment	0.60 (0.26 to 1.41)	0.50 (0.15 to 1.63)

IPE=Iterative Parameter Estimation; RPSFTM=Rank Preserving Structural Failure Time method

4.2.7 Supporting evidence from other studies of dabrafenib

Alongside BREAK-3,[2] the manufacturer presents supporting evidence from four other studies of dabrafenib[25-28] as summarised in Table 6. Two RCTs (BRF113220[26] and Combi-d[27]) and two non-RCTs (BREAK-2[25] and BREAK-MB[28]) were identified by the manufacturer. The RCTs could not be included in a quantitative analysis because they compared dabrafenib to dabrafenib + trametinib (dual therapy) (see also section 4.2.1). The number of patients receiving dabrafenib monotherapy (or dabrafenib monotherapy+placebo) in the supportive studies ranged from 54[26] to 212[27] (530 in total).

The ERG notes that, of the supportive studies, only Combi-d[27] was a Phase III trial, with the other three studies[25,26,28] all being Phase II. The ERG further notes based on eligibility criteria and baseline characteristics, that there were some important differences in the patient populations across the studies. In particular, BRF113220[26] and BREAK-2[25] both included a proportion of patients previously treated with chemotherapy (11% and 80% respectively) whilst BREAK-MB[28] only included patients with brain metastases (all patients in Cohort B had been previously treated for brain metastases, i.e. 48% of study population). Therefore the patient population in Combi-d[27] was perhaps the most similar to the population in BREAK-3[2] but, as outlined in section 3.1, unlike BREAK-3,[2] Combi-d[27] also included patients with $BRAF^{v600K}$ mutations as did BRF113220[26] (17%), BREAK-2[25] (17%) and BREAK-MB[28] (9%, although these are excluded from the analyses of PFS and OS); the proportion of patients by mutation status for Combi-d[27] is not reported in the MS.[3] The ERG further notes from Table 10 of the MS[3] that, compared to the dabrafenib arm of BREAK-3,[2] in the dabrafenib plus placebo arm of Combi-d[27] there were a greater proportion of patients with ECOG PS ≥1 (73% vs 33%) and females (47% vs 40%). However, for other baseline characteristics (age, disease stage, tumour, node, metastases staging, LDH levels and number of organs with metastases involved) there were many similarities between these two studies.

Median PFS demonstrated in the BRF113220[26] study (approximately 6 months) was reported to be similar to the median PFS reported in BREAK-3[2] (approximately 7 months).[2] The manufacturer states that these results show the consistency of dabrafenib treatment effects across studies, and attributes any differences in median PFS to differences in the median length of follow-up between the two studies. However as noted above, the ERG believes that the patient population in Combi-d[27] is more like that of BREAK-3[2] than BRF113220.[26] In Combi-d,[27] PFS was reported to be higher (approximately 9 months) than for patients in BREAK-3[2] or BRF113220.[26] The ERG also notes that the median PFS was higher for patients with $BRAF^{v600E}$ than for patients with $BRAF^{v600K}$ mutations in BREAK-2.[25]

As expected in patients with brain metastases, the PFS was reported to be lower for this group of patients in BREAK-MB[28] than in any of the other studies but nevertheless, the results do suggest that dabrafenib has activity in this group of patients.

Median OS values for patients in the dabrafenib arm were slightly lower in the dabrafenib arm of BREAK-3[2] (approximately 18 months) than the dabrafenib arm of BRF113220 (approximately 20 months).[26] The manufacturer suggests that the median OS value from BRF113220[26] is likely to be confounded as high levels of patients crossed over from the dabrafenib monotherapy arm to dual therapy (83%). However, the ERG notes that 43% patients in the dabrafenib arm of BREAK-3[2] also received subsequent treatments on disease progression. The ERG further notes that the OS reported in these two RCTs is higher than the OS reported in BREAK-2[25] (approximately 13 months regardless of $BRAF^{v600}$ mutation status) and BREAK-MB[28] (approximately 7 months).[28] Given the high proportion of patients already pre-treated with chemotherapy in BREAK-2[25] and patients with brain metastases in BREAK-MB[28] (80% and 100% respectively), the lower OS is not unexpected in these studies.

The MS[3] notes similarities between ORR rates from the dabrafenib arms of the three RCTs (51% to 59%). The ERG notes that the ORR for the $BRAF^{v600E}$ population was the same as that reported in BREAK-3.[2] However, for the $BRAF^{v600K}$ population, the ORR was much lower (13%).

Overall, the ERG believes it is difficult to draw any firm conclusions from the supportive studies because of differences in the patient populations. All patient populations are nevertheless relevant to the NICE scope. However, only BREAK-3[2] provides comparative evidence to a relative comparator (dacarbazine).

Table 6: Characteristics and results of all dabrafenib studies

	Pivotal trial	Other trials		Non-RCTs	
	BREAK-3[2]	**BRF113220[26]**	**Combi-d[27]**	**BREAK-2[25]**	**BREAK-MB[28]**
Study design	Phase III double blind randomised	Phase II Open-label 4-part study: Part C (randomised)	Phase III double blind randomised	Phase II Open-label Single arm	Phase II Open-label Single arm
Patient population	Patients with treatment-naïve, $BRAF^{v600E}$ mutation-positive metastatic melanoma	Patients with $BRAF^{v600E/K}$ mutation-positive metastatic melanoma, with up to one prior chemotherapy for advanced/ metastatic disease	Patients with $BRAF^{v600E/K}$ mutation-positive metastatic melanoma	Patients with treatment-naïve or previously treated $BRAF^{v600}$ mutation-positive metastatic melanoma; patients were allowed to have received prior systemic therapy but not treatment with a BRAF inhibitor	Patients with $BRAF^{v600}$ mutation-positive melanoma metastatic to brain; Cohort A (patients with no prior local therapy for brain metastases), and Cohort B (patients who received prior local therapy for brain metastases)
Number of patients in dabrafenib arm	187	54	212*	92 $BRAF^{v600E}$ 76 $BRAF^{v600K}$ 16	172 Cohort A 89 Cohort B 83
Median follow-up in dabrafenib arm PFS assessment OS assessment	 10.5 15.2	 14.1† 24.0†	 9.0† Not applicable	 Not reported† 11.9	 Not reported† Not reported†
Median PFS (months)‡ 95% CI	6.9 5.2 to 9.0	5.8 4.6 to 7.4	8.8 5.9 to 10.9	$BRAF^{v600E}$ 6.3 4.6 to 7.7§ $BRAF^{v600K}$ 4.5 2.6 to 6.2§	Cohort A 3.7 3.6 to 5.1§ Cohort B 3.8 3.7 to 3.8§
Median OS (months) 95% CI	18.2 16.6 to NR	20.2 14.5 to 25.9	Data reported to be too immature	$BRAF^{v600E}$ 13.1¥ $BRAF^{v600K}$ 12.9¥	Cohort A: 7.6 5.9 to NR§ Cohort B:7.2 5.9 to NR§
ORR (%) 95% CI	59 51.4 to 66.0	54 39.6 to 67.4	51 44.5 to 58.4	$BRAF^{v600E}$ 59 48.2 to 70.3 $BRAF^{v600K}$ 13 0.0 to 28.7	Not applicable

CI=confidence interval; ORR=overall response rate; OS=overall survival; PFS=progression-free survival

* Monotherapy includes trametinib placebo for Combi-d[27]

† Median follow-up reported for trial as a whole for BRF113220[26] and Combi-d;[27] the median follow-up is not reported for PFS in BREAK-2;[25] updated OS (but not PFS) data are available from Ascierto et al 2013[25] where the median follow-up was reported to be 11.9 months; median follow-up in BREAK-MB[28] not reported, it is however stated in the published paper that all patients had at least 4 months of follow-up

‡ Investigator-assessed PFS

§ Median PFS for BREAK-2[25] and PFS and OS for BREAK-MB[28] are expressed in the MS[3] in weeks and have been converted to months by the ERG by diving by 52 and multiplying by 12; patients with $BRAF^{v600K}$ mutations are excluded from the PFS and OS analyses for BREAK-MB[28]

¥ 95% confidence intervals not presented

4.3 Critique of trials identified and included in the ITC

4.3.1 Identified studies

The BREAK-3[2] trial does not include vemurafenib. However, one further trial (BRIM-3[1]) was identified by the manufacturer's systematic review. This trial compared vemurafenib with dacarbazine and thus provided a possible link to allow an ITC to be carried out between dabrafenib and vemurafenib. The manufacturer's evidence network diagram for the ITC (MS Figure 9) is reproduced in Figure 1.

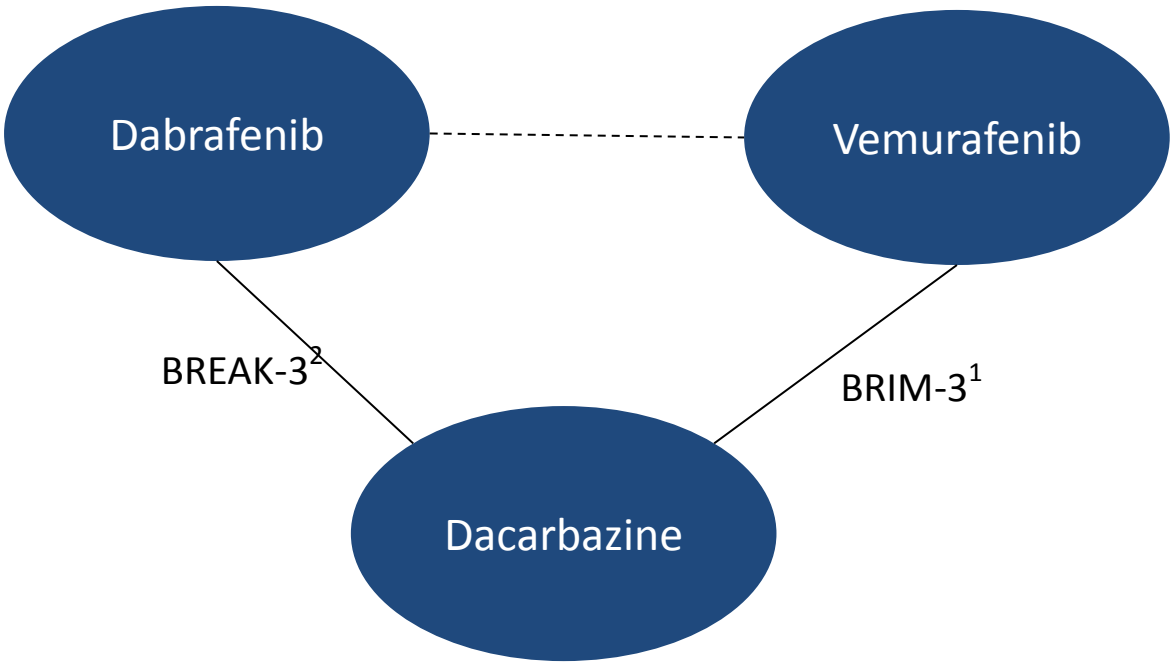


Figure 1: Evidence network used to inform the ITC

4.3.2 Methods and data sources used in the ITC

Indirect comparisons for PFS and OS were undertaken using the adjusted indirect comparison methodology described by Bucher.[36]

The manufacturer carried out ITCs for PFS and OS. The ITC for PFS included the most up-to-date data on this outcome from BRIM-3[1] (February 2012 data-cut) and the BREAK-3[2] study (June 2012 data-cut). Inputs for the ITC for OS were crossover adjusted results from BRIM-3[1] (February 2012 data-cut) and the BREAK-3[2] trial (December 2012 data-cut).

The ERG notes that the most up-to-date BREAK-3[2] trial data (January 2014 data-cut), despite being available, were not used in the ITCs. The manufacturer has provided two reasons for this:

1. The latest OS dabrafenib data from the BREAK-3[2] trial have not yet been adjusted for crossover

2. Using the earlier data-cuts ensures comparability of follow-up times across the two trials (see Table 7 for details).

Table 7: Comparison of data-cuts

	BREAK-3[2] **December 2012**	[redacted]	**BRIM-3[1]** **February 2012**
Proportion of patients who crossed over	57%	[redacted]	34%
Median follow-up (control arm) in months	12.7	[redacted]	9.5
Median follow-up (treatment arm) in months	15.2	[redacted]	12.5

Median OS values and the HRs used in the manufacturer's ITCs are presented in Table 8. The BRIM-3 OS HR used in the manufacturer's ITC is the uncensored result taken directly from the vemurafenib NICE Guidance (TA 269). The ERG also notes that the unadjusted median OS for the dacarbazine arms of BREAK-3[2] and BRIM-3[1] appear to differ. This may be indicative of differences in the patient population (see section 4.3.3) and/or differences in subsequent treatment received following disease progression in the two trials. Post-progression treatment is not reported for BRIM-3[1] but in a response to the first ACD for vemurafenib, the manufacturer of vemurafenib noted second-line use of ipilimumab in the dacarbazine arm was 19%.[37] In BREAK-3,[2] the ERG note that 3% of dacarbazine patients received subsequent ipilimumab.

Table 8: Results of individual trials and inputs for the manufacturer's indirect treatment comparisons

Outcome	BREAK-3[2] PFS data-cut: June 2012 OS data-cut: December 2012		BRIM-3[1] PFS and OS data-cut: February 2012	
	Dabrafenib (N=187)	**Dacarbazine (N =63)**	**Vemurafenib (n=337)**	**Dacarbazine (N=338)**
Investigator assessed median PFS months (95% CI)	6.9 (5.2 to 9.0)	2.7 (1.5 to 3.2)	6.9 (6.1 to 7.0)	1.6 (1.6 to 2.1)
ITT PFS HR (95% CI); p-value	0.37 (0.24 to 0.58); p<0.0001		0.38 (0.32 to 0.46); p<0·0001	
Median OS months (95% CI)	18.2 (16.6 to NR)	15.6 (12.7 to NR)	13.6 (12.0 to 15.2)	9.7 (7.9 to 12.8)
ITT OS HR (95% CI); p-value	0.76 (0.48 to 1.21); p=not significant		0.70 (0.57 to 0.87); p=0.0008	
OS HR (95% CI) input into the ITC	0.76 (0.49 to 1.18)		0.76 (0.63 to 0.93)	
RPSFTM adjusted HR (95% CI)	0.55 (0.21 to 1.43)		0.64 (0.53 to 0.78)	

CI=confidence interval; HR=hazard ratio; OS=overall survival; PFS=progression-free survival; RPSFTM=Rank Preserving Structural Failure Time Model

4.3.3 Characteristics of studies included in the ITC

The trial characteristics of the BREAK-3[2] trial have been described in section 4.2.2. Like BREAK-3,[2] BRIM-3[1] is a multi-centre, open label, phase III RCT. Full BRIM-3[1] trial details are provided in Tables 27-30 of the MS.[3] In this trial crossover was permitted prior to progressive disease and patients crossed from the dacarbazine arm to the vemurafenib arm. Crossover rates at each data-cut were as follows: December 2010,0%; Mar 2011,15%; Oct 2011,24%; and February 2012,34%.

To ensure the credibility of ITC results, the characteristics of the patients in the included trials must be comparable. The study population and eligibility criteria for BRIM-3[1] were similar to those for BREAK-3.[2] The characteristics of patients in BRIM-3[1] and BREAK-3[2] are presented in Table 29 in the MS[3] and are reproduced in Table 9. The ERG is of the opinion that most of the differences in the baseline characteristics of the patients in the two trials are unlikely to be of clinical importance. The possible exception, however, is level of LDH; a greater proportion of patients in the BRIM-3[1] trial had LDH levels above the upper limit of the normal range than patients in the BREAK-3[2] trial. As the level of LDH has been shown to result in worse PFS and OS,[8,38] patients in both arms of the BRIM-3[1] trial might be expected to have worse outcomes than those in the BREAK-3[2] trial. Indeed, as noted in section 4.3.2, median OS did differ in the dacarbazine arms of both trials, being lower in BRIM-3[1] (9.7 months) than BREAK-3[2] (15.6 months). The ERG further notes differences, albeit slight, in the proportions of patients with $BRAF^{v600E}$ and $BRAF^{v600K}$ mutations. As noted in section 4.2.7, there is some evidence from one phase II study (BREAK-2[25]) to suggest that patients with $BRAF^{v600E}$ mutations have poorer outcomes than patients with $BRAF^{v600K}$ mutations. However, contrary to this, BRIM-3[1] PFS and OS subgroup analyses, based on $BRAF^{v600}$ mutation status, show no significant differences between those with $BRAF^{v600E}$ and those with $BRAF^{v600K}$ mutations.

The manufacturer identified several differences in the characteristics of the BREAK-3[2] and BRIM-3[1] trials. These included differences in sample sizes, ratios of randomisation and different time points (data-cuts) for assessment. The manufacturer stated that the implications of these factors on the results of the ITCs were unclear.

Table 9: Characteristics of BREAK-3 and of BRIM-3 participants

	BREAK-3[2]		**BRIM-3[1]**	
	Dabrafenib (n=187)	**Dacarbazine (n=63)**	**Vemurafenib (n=337)**	**Dacarbazine (n=338)**
Median age (yr)	53.0 (22 to 93)	50.0 (21 to 82)	56 (21 to 86)	52 (17 to 86)
Male sex n (%)	112 (60)	37 (59)	200 (59)	181 (54)
White race n %	186 (100)	63 (100)	333 (99)	338 (100)
Geographic region n (%)				
Australia or New Zealand	NR	NR	39 (12)	38 (11)
North America	NR	NR	86 (26)	86 (25)
Western Europe	NR	NR	205 (61)	203 (60)
Other	NR	NR	7 (2)	11 (3)
$BRAF^{v600}$ mutation status n (%)*				
$BRAF^{v600E}$	187 (100)	62 (98)	304 (90)	312 (92)
$BRAF^{v600K}$	0	1 (2) †	33 (10)	24 (7)
$BRAF^{v600D}$	0	0	0	2 (1)
ECOG PS, n (%)				
0	124 (66)	44 (70)	229 (68)	230 (68)
1	62 (33)	16 (25)	108 (32)	108 (32)
Unknown	1 (<1)	3 (5)		
TNM staging at screening, n (%) §				
M0	6 (3)	1 (2)	20 (6)	13 (4)
M1a	23 (12)	10 (16)	34 (10)	40 (12)
M1b	34 (18)	12 (19)	62 (18)	65 (19)
M1c	124 (66)	40 (63)	221 (66)	220 (65)
Lactate dehydrogenase — no. (%)				
≤Upper limit of the normal range	119 (64)	43 (68)	142 (42)	142 (42)
>Upper limit of the normal range	67 (36)	19 (30)	195 (58)	196 (58)
Unknown	1 (<1)	1 (2)		

ECOG=Eastern Cooperative Oncology Group; NR=not reported; PS=performance status; TNM=tumour, node, metastases

* Data on $BRAF^{v600}$ mutation status for BREAK-3[2] has been taken from the CSR[30] and data for BRIM-3[1] has been taken from the published paper by McArthur et al.[1] Note: the numbers of patients reported with a $BRAF^{v600E}$ mutation in the McArthur et al paper appears to be incorrect; however, the number of patients reported for the other mutations does appear to be correct and hence the numbers for $BRAF^{v600E}$ mutations has been calculated by the ERG by subtracting the numbers with other mutations from the total number of patients in each arm

† One patient with $BRAF^{vv600K}$ mutation was reported to have been randomised in error but withdrawn prior to starting treatment (CSR,[30] Table 6, page 47)

§ Reported as 'Extent of metastatic melanoma' in MS[3]

4.3.4 Risk of bias of included studies

The manufacturer conducted assessments of the risk of bias for the BREAK-3[2] and BRIM-3[1] trials as recommended by the Cochrane Collaboration.[39] These assessments are presented in Table 22 of the MS[3] (pages 81 to 82 for BREAK-3[2]), in the Appendices (page 224 for BRIM-3[1]) and summarised here in Appendix 1. The manufacturer states that treatment allocation was not concealed in either study; however, as the interactive voice recognition system (IVRS) used for randomisation ensures concealment of treatment allocation the ERG believes that this is not a concern. The ERG concurs with the remainder of the manufacturer's risk of bias conclusions and agrees that both the BREAK-3[2] and BRIM-3[1] trials have a low risk of bias.

4.3.5 Results

Dabrafenib vs vemurafenib

The results of the PFS and OS ITCs for dabrafenib vs vemurafenib can be found in the MS[3] (Table 36 and Table 37) and are reproduced in Table 10. A HR of less than 1 indicates that the intervention (dabrafenib) is more effective than the comparator (vemurafenib). The results from the manufacturers ITCs show a trend in favour of dabrafenib over vemurafenib for both PFS and OS but the treatment difference was not found to be statistically significant for either outcome. The confidence intervals for both PFS and OS are noticeably wide, particularly for OS.

Table 10: Results of indirect treatment comparison for PFS and OS (RPFSTM adjusted) for dabrafenib vs vemurafenib

Outcome	Treatment	Control	HR	95% CI	
				Lower	Upper
PFS	Dabrafenib	Vemurafenib	0.97	0.59	1.60
OS	Dabrafenib	Vemurafenib	0.86	0.32	2.29

CI=confidence interval; HR=hazard ratio; OS=overall survival; PFS=progression-free survival

4.4 Critique of the statistical approach employed for the ITC

4.4.1 Main criticism of the ITC

The manufacturer claims that the method used to conduct the ITCs was appropriate. However, for the results of the ITCs to be valid, three assumptions need to hold:

1. Constant HRs for both PFS and OS data within the BREAK-3[2] trial (dabrafenib vs dacarbazine)
2. Constant HRs for both PFS and OS data within the BRIM-3[1] trial (vemurafenib vs dacarbazine)
3. PFS and OS data for dacarbazine within the two trials (BREAK-3[2] and BRIM-3[1]) should be broadly comparable and conform to a constant hazard relationship.

The analyses presented in more detail in section 5.3.1 show that:

- Analysis of PFS and OS data from BREAK-3[2] indicate a significant deviation from constant HR between the trial arms for PFS ($p<0.001$, chi-squared test) but not for OS ($p=0.91$).
- Analysis of PFS and OS data from BRIM-3[1] indicate significant deviations from the assumption of a constant HR between the trial arms ($p<0.001$ for PFS, $p<0.001$ for OS, chi-squared test).
- Comparison of dacarbazine PFS and OS data between the BREAK-3[2] and BRIM-3[1] trials indicates that PFS hazard profiles are not significantly different ($p=0.46$), but constant proportion hazards are not supported for OS ($p<0.001$).

Therefore, the results of the ITCs for dabrafenib vs vemurafenib presented by the manufacturer are unreliable and cannot be used to compare the clinical or cost effectiveness of the two treatments.

4.4.2 Additional criticisms

In the ITC for OS the manufacturer assumes that use of the Bucher ITC method[36] is appropriate and employs HRs for dacarbazine that have been adjusted using the RPSFTM method. Leaving aside the ERG's belief that assumptions regarding proportional hazards have not been met, the ERG is of the opinion that the unadjusted estimates should have been used in the primary analyses and the RPSFTM adjusted estimates used in a sensitivity analysis. The ERG requested the unadjusted analyses from the manufacturer during the clarification process and these were provided. The inputs into the unadjusted analyses are shown in Table 8 and the results from these analyses are shown in Table 11. With a resultant HR of 1, the unadjusted OS results suggest that dabrafenib and vemurafenib may

be of equal clinical efficacy. In addition, although the confidence interval using the unadjusted data is narrower than that resulting from using the RPSFTM adjusted data, the results still suggest that either treatment may be approximately 50% better than the other.

Table 11: Results of unadjusted indirect treatment comparison for OS for dabrafenib vs vemurafenib

	Treatment	Control	HR	95% CI	
				Lower	Upper
Unadjusted OS	Dabrafenib	Vemurafenib	1.00	0.62	1.62

CI=confidence interval; HR=hazard ratio; OS=overall survival

On page 100 of the MS[3] the manufacturer notes that the estimate of the crossover-adjusted 95% CI for BRIM-3[1] may be incorrect. This conclusion was reached because the RPSFTM method should increase the width of the 95% CI but the reported 95% CIs for the adjusted and unadjusted HRs are similar. The narrower 95% CI serves to underestimate the uncertainty surrounding the HR. To correct this, additional data would need to be obtained from the manufacturers of the drugs used in BRIM-3.[1] The ERG agrees that this error cannot be rectified during this appraisal.

4.4.3 ERG conclusions on the credibility of results of the manufacturer's ITC

The ERG is of the opinion that there is no robust evidence available to ensure a credible comparison of the clinical and therefore cost effectiveness of dabrafenib vs vemurafenib.

4.5 Critique of the adverse events data

A summary of general AE data from all relevant dabrafenib[2,25-28] and vemurafenib[1] studies (RCTs and non-RCTs) is presented in Table 12. Most patients receiving dabrafenib, dacarbazine or vemurafenib experienced at least one AE. The majority of AEs were deemed to be treatment-related, particularly those experienced by patients receiving dabrafenib or vemurafenib. Adverse events leading to treatment discontinuation were relatively rare in all studies, except in the vemurafenib arm of BRIM-3[1] where these were reported to occur in around a quarter of all patients at the original data-cut (December 2010) presented in the MS for vemurafenib.[23] However, the ERG notes that a conference presentation[40] from an later data-cut (March 2011), cited by in the GlaxoSmithKline MS,[3] reports the treatment discontinuation rate for vemurafenib to be much lower (7%). Dose modification was not required for the majority of AE cases for any drug and permanent discontinuations due to AEs were infrequent. The highest incidence of Grade 3+ AEs was reported for patients in the BRIM-3[1] trial who received vemurafenib (50%). The ERG notes that, compared with BREAK-3[2] a lower proportion of Grade 3+ AEs were reported for patients in the BRIM-3[1] trial who were treated with dacarbazine (30% vs 42%, respectively). The incidence of Serious adverse evebts (SAEs) were reported to be highest for vemurafenib (33%) and lowest for dacarbazine (16%), both in BRIM-3.[1] Serious adverse events related to treatment were only reported in the dabrafenib studies,[2,25-28] being higher in the dabrafenib arms (15% to 23%) than the dacarbazine arm (3%). Fatal AEs were rare and only one fatal AE possibly related to study treatment was reported. Within the MS,[3] the manufacturer also cites its systematic review report[32] which includes findings from an ITC of AE data in which dabrafenib was compared with other treatments, including vemurafenib. No significant differences between dabrafenib and vemurafenib were reported for any of the general AEs summarised in Table 12.

The most common (≥5%) non-serious AEs from BREAK-3[2] reported in Table 39 of the MS[3] are summarised in Table 13 alongside those from BRIM-3[1] extracted by the ERG. The ERG cautions that the different types of AEs reported (all common, non-serious vs all-Grade) mean that the incidences of AEs are not directly comparable. Nevertheless, based on these data, the ERG notes that the most common (≥25%) AEs associated with both dabrafenib and vemurafenib were hyperkeratosis, arthralgia, headache, alopecia, nausea and skin papilloma. Pyrexia was experienced by ≥25% of patients with dabrafenib but not ≥25% of patients with vemurafenib whereas the opposite was true for fatigue, photosensitivity, rash, diarrhoea and pruritis. The most common AEs (≥25%) associated with dacarbazine were nausea, vomiting, fatigue, asthenia and constipation. Results from the ITC showed that dabrafenib was associated with significantly fewer incidences of diarrhoea, nausea, and photosensitivity/phototoxicity in comparison to vemurafenib.[32]

Specific types of SAEs are not reported for BREAK-3[2] in the MS[3] but are reported in the CSR.[30] The most common SAEs were pyrexia (4%) and squamous cell carcinoma (SCC) (4%) in patients treated with dabrafenib.

Grade 3+ AEs from BREAK-3[2] are reported in Table 40 of the MS[3] and are summarised in Table 14 alongside those from BRIM-3[1]. Squamous cell carcinoma and back pain were the most common (≥2%) Grade 3+ AEs associated with dabrafenib, whilst cutaneous SCC of the skin, keratoacanthoma, rash and arthralgia were the most common Grade 3+ AEs associated with vemurafenib. Neutropenia and thrombocytopenia were the most common (≥2%) Grade 3+ AEs associated with dacarbazine. No significant differences in Grade 3+ AEs between dabrafenib and vemurafenib were identified from the results of the ITC.[32]

The manufacturer noted that 'Treatment-related pyrexia was reported with dabrafenib (16% Grade; n=30,3% Grade 3; n=5), while for vemurafenib there was 18% Grade pyrexia, but no cases of Grade 2 or above were reported' (MS,[3] page 121). The ERG concurs that both non-serious pyrexia and Grade 3+ pyrexia appear to be more common in patients treated with dabrafenib compared to those treated with vemurafenib and notes, from the SAE data across published dabrafenib trials,[2,25,26,28] that this AE is commonly reported to be treatment-related.

Table 12 Summary of general AE data for dabrafenib, dacarbazine and vemurafenib from Phase II and Phase III trials*

Adverse event (AE)	Incidence AEs (%)								
	Dabrafenib + placebo	Dabrafenib				Crossover from dacarbazine to dabrafenib	Dacarbazine		Vemurafinib
	Combi-d[27] (n=211)	BREAK-2[25] dabrafenib (n=92)	BREAK-MB[28] dabrafenib (n=83)†	BRF113220[26] (n=53)	BREAK-3[2] (n=187)	BREAK-3[2] (n=28)	BREAK-3[2] (n=59)	BRIM-3[1] (n=282)	BRIM-3[1] (n=336)
Any AE	98	93	92	100	99	93	92	90	97
AEs related to study treatment	88	86	82	NR	88	82	73	69	94
AEs leading to treatment discontinuation	5	1	2	2	3	0	3	5‡	26‡
AEs leading to dose reduction	13	14	14	NR	18	7	17	NR	NR
AEs leading to dose interruption	30	29	32	28	27	21	27	16¥	38¥
Any Grade 3+ AE	NR	36	40	43	42	30	42	30	50
Any SAE	30	27	30	25	23	21	22	16	33
SAEs related to study treatment	23	18	17	19	15	14	3	NR	NR
Fatal AEs	0	0	2	0	1	4	0	1§	2§
Fatal SAEs related to study treatment	0	0	0	0	0	0	0	NR	NR

AE=adverse event; NR=not reported; SAE=serious adverse event

* All AE data are reported in the MS[3] except for Grade 3+ AEs for BREAK-2[25] extracted from the published paper[25] and all AE data for BRIM-3[1] which are taken from a previous MS submitted by Roche for vemurafenib (interim data-cut, December 2010**);**[23] data for BRF113220[26] are taken from the most recent data-cut (May 2012); data for BREAK-3[2] are reported in the MS[3] for the initial data-cut (December 2011) with the exception of any Grade 3+ AEs which are reported for the most recent data-cut (December 2012)

† Data are provided in the MS[3] for both cohort A (not previously treated for brain metastasis) and cohort B (previously treated for brain metastasis); the data are presented here for all patients in both cohorts

‡The ERG notes that the proportions of patients ceasing treatment due to AEs differs in a more recent data-cut reported in McArthur et al 2014[1] where the proportion is reported to be lower in both arms: 7% and 2% respectively

¥ Described as 'AEs that led to dose modification/interruption'

§ Deaths were based on the all-treated population (n=336 for vemurafenib; n=289 for dacarbazine) where it is stated that the vast majority of deaths were due to progression (35/42 and 63/66 respectively); it is assumed by the ERG that all the remaining deaths are from AEs

Table 13 Proportion (%) of all common, non-serious AEs in BREAK-3 and all-Grade AEs in BRIM-3 experienced by at least 15% of patients in any treatment arm*

Adverse event (AE)	Incidence of AEs			
	BREAK-3[2] non-serious AEs		BRIM-3[1] all-Grade AEs	
	Dabrafenib (N=187)	Dacarbazine (N=59)	Dacarbazine (N=282)	Vemurafenib (N=336)
Hyperkeratosis	39	2	<1	29
Arthralgia	35	2	4	56
Headache	35	9	10	33
Pyrexia	31	14	10	21
Alopecia	29	5	2	48
Nausea	28	51	45	38
Skin papilloma	25	2	<1	28
Fatigue	23	24	35	46
Asthenia	20	15	10	14
PPE syndrome	20	2	1	9
Rash	19	0	2	41
Vomiting	19	25	28	21
Nasopharyngitis	16	7	3	13
Back pain	16	7	7	13
Cough	16	7	7	13
Myalgia	15	0	2	15
Diarrhoea	14	12	13	36
Constipation	14	14	25	14
Pain in extremity	13	10	7	21
Decreased appetite	12	10	8	22
Dry skin	11	0	<1	23
Erythema	10	2	2	17
Odema, peripheral	9	10	5	20
SCC/Keratoacanthoma	6*	0*	NR	NR
- Cutaneous SCC of skin	NR	NR	<1	19
- Keratoacanthoma	NR	NR	<1	11
Pruritis	5	2	2	25
Photosensitivity reaction	3	7	5	41
Neutropenia	3	17	12	<1

AE=adverse event; NR=not reported; PPE=palmar-plantar erythrodysesthesia; SCC=squamous cell carcinoma

* All data taken from MS[3] (Table 39) for BREAK-3[2] (Most recent data-cut: December 2012) and appendix to McArthur 2014[1] for BRIM-3[1] (data-cut: February 2012)

Table 14 Proportion (%) of most common Grade 3+ adverse events reported in BREAK-3 and BRIM-3 experienced by at least 2% of patients in any treatment arm*

Adverse event (AE)	Incidence AEs (%)			
	BREAK-3[2]		BRIM-3[1]	
	Dabrafenib (n=187)	Dacarbazine (n=59)	Dacarbazine (n=282)	Vemurafenib (n=336)
SCC	6	0	NR†	NR†
Back pain	4	0	<1	<1
Gamma-glutamyltransferase increased	3	2	0	4
Hyperglycaemia	3	0	<1	0
Pyrexia	3	0	<1	<1
Alanine aminotransferase increased	2	0	<1	2
Hypertension	2	2	<1	1
PPE syndrome	2	0	0	<1
Abdominal pain	1	0	<1	2
Pulmonary embolism	<1	2	<1	1
Anaemia	<1	3	2	1
Constipation	1	0	0	<1
Thrombocytopaenia	<1	5	2	<1
Neutrophil count decreased	0	2	2	0
Neutropenia	0	14	6	<1
SCC/Keratoacanthoma				
- Cutaneous SCC of skin	2	0	<1	19
- Keratoacanthoma	0*	0*	<1	10
Rash	0*	0*	0	9
Arthralgia	1*	0*	1	6
Photosensitivity reaction	0*	0*	0	4
Fatigue	1*	0*	2	3
Rash maculo-papular	0*	0*	0	3
Basal cell carcinoma	<1*	0*	<1	3
Malignant melanoma	1*	0*	0	2
Nausea	<1*	0*	2	2
Diarrhoea	<1*	0*	<1	2
Blood alkaline phosphate increased	0*	0*	0	2
Dyspnoea	1*	0*	2	1
Pain in extremity	<1*	0*	2	<1

NR=not reported; SCC=squamous cell carcinoma

All data for BREAK-3[2] taken from MS[3] (Table 40) (Most recent data-cut: December 2012) except * taken from CSR[30] (Most recent data-cut: December 2011); all data for BRIM-3[1] taken from appendix to McArthur 2014[1] (Data-cut: February 2012)

† AEs for SCC of skin are however reported

Regarding the other dabrafenib studies, the manufacturer argues that: 'The safety findings from BREAK-2[25] and BREAK-MB[28] largely reflect the safety data demonstrated by the BREAK-3[2] study.' (MS,[3] page 107) Similarly: 'the number and severity of events [in Combi-d[27] and BRF113220[26]] did not appear to show any inconsistency with the findings from the BREAK-3[2] study, aside from hypertension and anaemia which have arisen, although in small numbers, within subjects participating in these trials' (MS,[3] page 113). The ERG notes that fatigue appeared to be more commonly reported in BRF113220[26] (40%), Combi-d[27] (35%) and Cohort B of BREAK-MB[28] (30%) than BREAK-3.[2]. It was not possible to compare Grade 3+ AEs across trials from the data presented in the MS[3] as SAEs are presented for BRF113220,[26] BREAK-2[25] and BREAK-MB[28] and only Grade AEs for Combi-d[27]). However, compared with the published papers, it is noted that Grade 3+ fatigue (6%) and cutaneous SCC (including keratoacanthoma) (17%) were more common in BRF113220[26] than BREAK-3[2] and Grade 3+ cutaneous SCC (8%), anaemia (4%) and hypophosphatemia (4%) were more common in BREAK-2[25] than BREAK-3.[2] The most common SAEs (≥4%) reported in the published studies of dabrafenib[25,26,28] appear to be similar to the most common SAEs reported in the CSR[30] for BREAK-3,[2] i.e. pyrexia and SCC, with the addition of SCC of the skin in BRF113220[26] and basal cell carcinoma and anaemia in BREAK-2.[25]

The manufacturer concludes that whilst treatment with dabrafenib and vemurafenib share some AEs in common, the safety profile for dabrafenib can be differentiated from vemurafenib in a number of ways; in particular, the incidence of photosensitivity which it is argued is 'a notable side effect with vemurafenib' and which 'can have a significant effect on quality of life' (MS,[3] page 117) is much lower for dabrafenib patients. Additionally, the manufacturer argues that certain types of hyperproliferative skin toxicities (cutaneous SCC and keratoacanthomas) and arthralgia appear to be more common, and more severe, for patients treated with vemurafenib than for patients treated with dabrafenib. The ERG notes that in BRIM-3,[1] keratoacanthomas (but not cutaneous SCC) were most common in patients with BRAF^{v600K} mutations, a group of patients not included (and therefore not treated with dabrafenib) in BREAK-3.[2] Hyperkeratosis, palmar-plantar erythrodysesthesia (PPE) syndrome and treatment-related pyrexia appear to be more common with dabrafenib treatment than with vemuarfenib treatment. The ERG agrees that the data appear to support these conclusions, despite the limitations recognised by the manufacturer from the lack of head-to-head data for dabrafenib vs vemurafenib.

4.6 Critique of the health related quality of life data

4.6.1 Dabrafenib vs dacarbazine

The MS[3] states that, in BREAK-3, the HRQoL endpoint was change from baseline in HRQoL, based on patient self-reported scores collected using the EORTC QLQ-C30 (which measures functionality) and the EQ-5D (which measures subjects' perceived levels of problem) questionnaires. However, the ERG notes that only descriptive statistics for EQ-5D Visual Analogue Scale (VAS) data are presented in the MS[3] (Table 25). A poster presentation by Grob et al 2012[14] reports no statistically significant differences between patients treated with dabrafenib and dacarbazine for HRQoL. It is however stated that differences in functionality and symptoms at disease progression between the arms were clinically meaningful in favour of dabrafenib.

4.6.2 Dabrafenib vs vemurafenib

In the MS,[3] it is argued that photosensitivity can have a 'significant effect' (Page 11) on HRQoL. The MS[3] notes that this is lower and typically less severe in patients treated with dabrafenib (all being Grade 1 in BREAK-3[2]) compared with those treated with vemurafenib, implying a better HRQoL for dabrafenib patients. No direct HRQoL data are reported for BRIM-3[1] in the MS.[3] Furthermore, HRQoL were not presented by the manufacturer of vemurafenib in a previous submission.[23] However, the ERG notes that in BRIM-3,[1] HRQoL data were collected via the FACT-M questionnaire, making it impossible to compare across trials even if the data were available.

4.7 Additional work on clinical effectiveness undertaken by ERG

None undertaken.

4.8 Conclusions of the clinical effectiveness section

The manufacturer has presented direct clinical effectiveness evidence treatment naïve patients with BRAF^{v600} mutation-positive metastatic melanoma of dabrafenib vs dacarbazine. Treatment naïve patients are a subgroup of the population specified in the NICE scope while dacarbazine is one of the relevant comparators, being the standard of care until the recommendation of vemurafenib for patients with BRAF^{v600} mutation-positive metastatic melanoma. This evidence is from a good quality trial, BREAK-3,[2] which shows a statistically significant benefit in for dabrafenib over dacarbazine for PFS (of 4.2 months) but not for OS. It is speculated by the manufacturer that the lack of benefit in OS may be due to treatment crossover from dacarbazine to dabrafenib. For this reason, the manufacturer has adjusted for crossover using the RPSFTM. However the ERG does not consider that the RPSFTM adjusted OS results to be sufficiently robust because the results are based on an interim analysis with few deaths. Furthermore, the use of the RPSFTM requires the assumption that the effect of receiving dabrafenib is the same when received on diagnosis as it is on disease progression to be true. The ERG questions the validity of this assumption. [redacted]

Based on current clinical practice, vemurafenib is a more appropriate comparator for dabrafenib. Comparisons have been made to this BRAF inhibitor by means of an ITC enabling the results from BREAK-3[2] to be compared to the results of another good quality trial, BRIM-3.[1] Both trials appear to be broadly similar in terms of the patients included although it is noted that a minority (8%) of patients in BRIM-3[1] have BRAF^{v600K} mutations whilst all patients in BREAK-3[2] have the more common BRAF^{v600E} mutation. Furthermore, it is noted a greater proportion of patients had LDH levels above the upper limit of the normal range in both arms of the BRIM-3[1] trial than in BREAK-3[2] which is speculated by the ERG may result in a worse prognosis. Indeed, it is noted that the unadjusted median OS in the dacarbazine arm of BRIM-3[1] is lower than that in BREAK-3.[2] No statistically significant differences in terms of OS or PFS were identified by the ITC presented by the manufacturer, either using adjusted OS (based on estimates from RPSFTM in both trials) or unadjusted OS. However, the ERG questions the validity of the approach used to conduct the ITC as key underlying assumptions are not met. Specifically, constant HRs for both PFS and OS data within both the BREAK-3[2] and BRIM-3[1] are not evident while comparison of dacarbazine OS data between the two trials indicates that constant proportion hazards are not supported ($p<0.001$).

Evidence is also presented from four supportive studies reporting the findings for the dabrafenib arms only of two RCTs (BRF113220[26] and Combi-d[27]) and two non-RCTs (BREAK-2[25] and BREAK-MB[28]). The latter two studies include a large proportion of patients pre-treated with chemotherapy (80%) and brain metastases (48% were pre-treated for brain metastases, the remaining 52% of patients also had brain metastases) which constitute additional subgroups of the population specified in the NICE scope. However, the ERG believes it is difficult to draw any firm overall conclusions from these studies because of differences in the patient populations and because none of the four studies[25-28] provide comparative evidence to a relevant comparator.

While dabrafenib and vemurafenib have some AEs in common, there are notable differences between the two drugs for a number of AEs. Specifically, hyperproliferative skin toxicities (cutaneous SCC and keratoacanthomas) and arthralgia appear to be more common and more severe with vemurafenib whereas pyrexia, hyperkeratosis and PPE syndrome appear to be more common with dabrafenib. A separate analysis[32] not formally included within the MS[3] has also reported dabrafenib to be associated with significantly fewer incidences of diarrhoea, nausea, and photosensitivity/phototoxicity than vemurafenib. There are no comparative quality of life data for dabrafenib and vemurafenib.

5 COST EFFECTIVENESS

5.1 Introduction

This section provides a structured critique of the economic evidence submitted by the manufacturer in support of dabrafenib for the initial treatment of patients with unresectable or metastatic BRAF^{v600} mutation-positive malignant melanoma. The manufacturer states that there are insufficient clinical data with which to construct an economic case for previously treated (second-line) patients.

The two key components of the economic evidence presented in the MS[3] are (i) a systematic review of the relevant literature and (ii) a report of the manufacturer's de novo economic evaluation. Table 15 contains details of the location of key information within the MS.[3] The manufacturer has also provided an electronic version of their economic model which was developed in Microsoft Excel.

Table 15 Location of key economic information in the MS

Key information	Page numbers	Tables/figures
Details of the systematic review of the economic literature	127 to 130	Table 47
De novo analysis	131 to 134	Table 48*
Clinical parameters and variables	135 to 146	Tables 49 to 50, Figures 12 to 15
Measurement and valuation of health effects	147 to 155	Tables 51 to 53
Resource identification, measurement and valuation	156 to 166	Tables 54 to 62
Methods of sensitivity analysis	167 to 174	Tables 63 to 64
Results-clinical outcomes from the model	175 to 180	Tables 65 to 69
Results-base-case analysis	181 to 184	Tables 70 to 71
Results-sensitivity analysis	184 to 192	Tables 72 to 73, Figures 16 to 19
Validation	193	
Subgroup analysis	194 to 195	
Interpretation of economic evidence	196 to 198	
Assessment of factors relevant to the NHS and other parties	199 to 201	Tables 74 to 77

* Also diagrammatical representation of model structure

5.2 Overview of manufacturer’s cost effectiveness literature review

5.2.1 Objective of the manufacturer's cost effectiveness literature review

The manufacturer's search was designed to capture economic literature which focussed on treatments for metastatic melanoma. The initial searches were conducted on 10 April 2012 and were updated on the 24 and 25 October 2013. The databases were searched from start of database (1960) to 24/25 October 2013.

Details of the search strategies employed by the manufacturer are provided in Appendix 10.15 of the MS (p295-302).[3] Table 16, extracted from the MS (p295),[3] displays the data sources searched by the manufacturer.

Table 16 Data sources used to retrieve economic and quality of life evidence

Data source	Service Provider
MEDLINE®	Embase.com; http://www.embase.com/
Embase®	
MEDLINE® In-process	PubMed®(in-process citations); http://www.ncbi.nlm.nih.gov/sites/entrez
Cochrane (NHS EED)	Cochrane library; http://mrw.interscience.wiley.com/cochrane/cochrane_search_fs.html
Cochrane (CENTRAL and Method Studies)	
EconLIT®	AEAweb.org interface

The manufacturer states that a search of the CRD database was also undertaken to identify relevant technology appraisals.

5.2.2 Inclusion and exclusion criteria used in study selection

The inclusion/exclusion criteria used in study selection are presented in Table 17.

Table 17 Economic evaluation search inclusion and exclusion criteria

Parameter	Selection criteria
Inclusion criteria	
Population	Adults (males and females) of any age or race with unresectable advanced or MMM (Stage IIIc or IV)
Study type and design	Cost studies/surveys/analyses; database studies collecting cost data; resource surveys; cost effectiveness analyses; cost utility analyses; cost benefit analyses; cost minimisation analyses; budget impact models; cost consequences studies
Intervention	All pharmacological treatments; all treatments including adjuvant therapy; surgery, radiotherapy or isolated limb perfusion; any other treatment
Countries	All
Language restriction	English only
Publication timeframe	No date restriction for database searches; Year 2009 to 2011 for conference proceedings
Exclusion criteria	
Subgroup analyses	Subgroup data for the disease of interest, adult population, disease stage
Language	Languages other than English

MMM=metastatic or malignant melanoma

Nine cost effectiveness studies met the inclusion criteria for the systematic review but only one was deemed relevant by the manufacturer for discussion. In addition, the manufacturer

identified one relevant NICE single technology appraisal, Vemurafenib for the treatment of locally advanced or metastatic BRAF^{v600} mutation-positive malignant melanoma (NICE technology appraisal 269). Details relating to this appraisal are available from the NICE website.[16,23]

5.3 ERG critique of the review

The manufacturer's searches were run in April 2012 and October 2013. No systematic searching has been undertaken since October 2013, meaning that some more recent studies may have been missed. The search strategies for economics studies included details of the specific databases searched, the dates when the searches were conducted and the spans of the searches. The literature search appears to be appropriate although lacking in specificity as intervention-specific terms were not included. This means that the searches were likely to pick up many references that were irrelevant. The ERG does not believe that the lack of specificity will have led to any relevant studies being overlooked.

The ERG agrees with the manufacturer that the nine cost effectiveness studies that met the inclusion criteria for the systematic review were not directly relevant to this appraisal. The manufacturer's only relevant study is the NICE appraisal of vemurafenib treatment for locally advanced or metastatic BRAF^{v600} mutation-positive malignant melanoma.[16] The manufacturer has included a brief summary of this appraisal in the main body of the MS[3] and a more detailed summary in the appendix (MS,[3] Appendix 10.1, pages 211 to 213).

The ERG notes that the NICE appraisal of ipilimumab for previously treated advanced (unresectable or metastatic) melanoma (TA268)[41] meets the manufacturer's inclusion criteria but has not been considered in the manufacturer's review of the literature. Ipilimumab is a relevant second-line comparator specified in the NICE scope. However the population under consideration in TA268[41] was adults with advanced (unresectable or metastatic) melanoma regardless of BRAF^{v600} mutation status and hence is a different patient population.

5.4 Overview of manufacturer's economic modelling

5.4.1 Description of manufacturer's economic model

A schematic of the manufacturer's model is provided in the MS[3] and reproduced in Figure 2. It comprises three health states described as 'Progression-free', 'Post progression' and death. All patients enter the model in the 'Progression-free' health state and are at risk of disease progression or death over time. Patients who progress are assumed to discontinue therapy, move to the 'Post progression' health state and stay in that health state until death. The model has been developed in Microsoft Excel and employs a 1-week cycle length.

Variants of this model structure have been used in the modelling of metastatic oncology for numerous NICE STAs including NICE TA269,[16] NICE TA295,[42] and NICE TA310.[43]

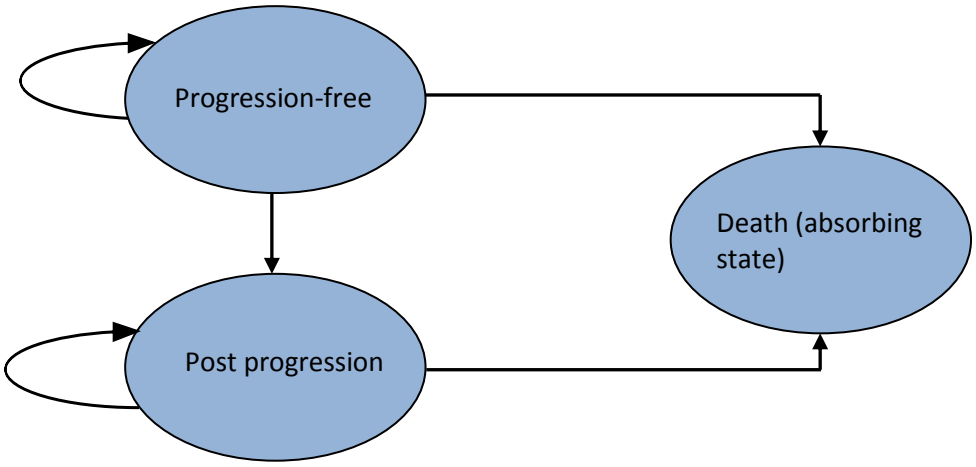


Figure 2 Schematic of manufacturer's model

5.4.2 Population

The economic model includes previously untreated adult patients with BRAF^{v600} mutation-positive metastatic melanoma. The value for body surface area (BSA) used in the model was based on the population recruited to the BREAK-3[2] trial, i.e. 1.93m^2. Age is not a variable in the model. The manufacturer considered that there was insufficient evidence to facilitate the development of an economic case for previously treated (second-line) patients.

5.4.3 Intervention and comparator technology

The interventions are implemented in the model in accordance with their current, or anticipated, marketing authorisations and doses (See Table 18). Patients are treated until progression and the model includes a relative dose intensity (RDI) multiplier which is designed to align average on-treatment dosing with trial experience (BREAK-3[2]).

Table 18 Drug doses used in the economic model

Drug	Dose	Source
Dabrafenib	150mg (2 x 75mg capsules) twice daily	BREAK-3 CSR[30]
Vemurafenib	960mg (4 x 240mg tablets) twice daily	Chapman 2011[44]
Dacarbazine	1000mg/m^2 every 3 weeks	BREAK-3 CSR[30]

CSR=clinical study report

5.4.4 Perspective, time horizon and discounting

The manufacturer states that the economic appraisal is undertaken from the perspective of the NHS. Outcomes are expressed in terms of quality adjusted life years (QALYs) and life-years gained. The time horizon is set at 30 years and, in line with the NICE Methods Guide to Technology Appraisal,[45] both costs and outcomes are discounted at 3.5%.

5.4.5 Treatment effectiveness

Modelled PFS and OS survival distributions have been segmented into two periods, the trial period and the projection period.

Progression-free survival

Progression free survival for dabrafenib and dacarbazine for the trial period are based on independent parametric survival distributions fitted to individual patient data (IPD). The IPD is investigator-assessed PFS from the BREAK-3[2] trial.(June 2012, the final data-cut for this endpoint). The trial periods for dabrafenib and dacarbazine were defined as 71.1 weeks and 53.1 weeks respectively based on the last censored or observed failure time for investigator-assessed PFS in BREAK-3.[2]

The manufacturer reports that curve-fitting was carried out using standard methodology and that exponential, Weibull, log-logistic, log-normal and Gamma distributions were considered. The manufacturer found that the curves and fit statistics for all of the two-parameter models were similar; however, the log-normal was marginally the best. The log-normal distribution was therefore used to model PFS during the trial and projection periods for both dabrafenib and dacarbazine.

In the base case, proportional hazards were used to model PFS for vemurafenib. The HR from the indirect treatment comparison of dabrafenib vs vemurafenib was used and this was applied throughout the entire model timeframe.

Overall survival

Overall survival for dabrafenib during the trial period was based on an independent parametric survival distribution fit to IPD from the BREAK-3[2] trial (December 2012 data-cut) using AFT regression. The duration of the trial period was defined as 96 weeks based on the maximum censor or failure time for dabrafenib OS in BREAK-3.[2] Standard approaches to curve-fitting were used, with exponential, Weibull, log-logistic, log-normal and Gamma distributions being considered. The log-normal distribution seemed to provide the best fit to the data and was used to model OS during the trial period.

The manufacturer considered that, for dacarbazine the RPSFTM analysis (rather than the ITT analysis) should be used to account for the confounding effect of crossover on OS. However, as such data were sparse, the manufacturer used the dabrafenib vs dacarbazine RPSFTM adjusted HR from BREAK-3[2] to project dacarbazine OS survival during the trial period (96 weeks), after which it was assumed that there was no further effect.

In the base case, proportional hazards were used to model OS for vemurafenib. The HR from the indirect treatment comparison of dabrafenib vs vemurafenib was applied during the trial period (96 weeks), after which it was assumed there was no further effect of treatment.

Data in Table 19 have been extracted directly from the MS[3] (p 193) and summarises the PFS and OS evidence used in the manufacturer's model.

Table 19 Summary of PFS and OS evidence used in modelling

	Dabrafenib (BREAK-3) vs			**Dacarbazine vs**			**Vemurafenib (BRIM-3) vs**		
	HR	**95% CI**		**HR**	**95% CI**		**HR**	**95% CI**	
PFS									
Dabrafenib	1.00	---	---	2.70	1.72	4.35	1.03	0.62	1.69
Dacarbazine	0.37	0.23	0.58	1.00	---	---	0.38	0.32	0.46
Vemurafenib	0.97	0.59	1.60	2.63	2.17	3.13	1.00	---	---
OS									
Dabrafenib	1.00	---	---	1.82	0.70	4.76	1.16	0.44	3.10
Dacarbazine	0.55	0.21	1.43	1.00	---	---	0.64	0.53	0.78
Vemurafenib	0.86	0.32	2.29	1.56	1.28	1.89	1.00	---	---

Shaded cells represent direct evidence, unshaded cells relate to indirect treatment comparisons; OS was adjusted for confounding due to treatment switching

CI=confidence interval; HR=Hazard ratio; OS=overall survival; PFS=progression free survival

Projection post trial -period

Overall survival for dabrafenib from the end of the trial follow-up out to 10 years was based on OS data from the American Joint Committee on Cancer (AJCC) registry reported by Balch et al[8] in the form of survival data by stage and site of metastases. Survival for dabrafenib was estimated based on the relative proportions of patients in the BREAK-3[2] trial in the corresponding stages. A log-logistic distribution was found to provide a good fit. Proportional hazards were used to model survival for dacarbazine and vemurafenib; the HRs used previously to model survival for the trial period were also applied for this period.

Beyond 10 years, survival was based on UK general population mortality data.

5.4.6 Health related quality of life

The utility values incorporated into the model are those derived from the BREAK-3[2] trial.The manufacturer reports that these are comparable to those reported in Beusterien et al,[46] the results from which have been used in models informing previous NICE technology appraisals.[23,47] In the absence of equivalent data for patients receiving vemurafenib it was assumed that the quality of life of patients receiving vemurafenib was the same as that for patients receiving dabrafenib.

Table 20 Summary of quality of life values for cost effectiveness analysis

	Dabrafenib and vemurafenib	Dacarbazine
Progression-free survival	0.767	0.750
Post-progression survival	0.677	0.677

5.4.7 Resources and costs

Drug costs

The model assumes that a patient will be dispensed a 28 day supply of dabrafenib (2 x 75mg tablets twice daily). The number of packets administered depends on the number of patients in PFS at the beginning of each 28 day cycle. Relative dose intensity (RDI) multipliers were used to adjust medication and administration costs to take into account dose reductions, treatment interruptions and differences between the distribution of time to discontinuation and the distribution of PFS. These were calculated based on IPD from the BREAK-3[2] trial and are the ratio between the estimated actual dose (or administration) to the estimated planned dose (or administration). Further details are provided in the MS (page 160)[3] and are discussed in the ERG's critique of the manufacturer's model.

The model also assumes that vemurafenib (4 x 240mg tablets twice daily) is administered every 28 days. The manufacturer states that due to an absence of data, the RDIs for vemurafenib are the same as those for dabrafenib.

Dacarbazine is administered at a dose of 1000mg/m^2 and the mean BSA was assumed to be 1.93m^2. Dacarbazine can be purchased in 1000mg vials and the cost used in the model, extracted from BNF62,[48] is £31.80 per vial. The cost used in the model is £61 every 21 days, which equates to £82 every 28 days. The RDI multipliers for dacarbazine were estimated using the same approach used to generate RDI multipliers for dabrafenib (see above).

BRAF test costs

As dabrafenib and vemurafenib are indicated solely for patients who are BRAF^{v600} mutation-positive, mutation testing is required. A cost of £206.34 per treated patient is included in the

model. The source for this value is not provided. The model assumes that patients receiving dacarbazine have not had their BRAF status tested.

Dispensing costs

The manufacturer has assumed that the time taken to dispense dabrafenib or vemurafenib would be equal to that taken to dispense capecitabine (another oral chemotherapy agent), i.e. 12 minutes (Millar 2008).[49] The cost associated with the dispensing time has been extracted from 'Unit costs of health and social care'[50] (£67 per hour for a hospital pharmacist to perform patient-related activities, including overheads, qualifications and salary on-costs). The manufacturer's base case estimates that the cost of dispensing dabrafenib or vemurafenib is £13.40 (every 28 days).

The manufacturer's model does not include a cost for dispensing dacarbazine.

Administration costs

Dabrafenib and vemurafenib are both oral therapies and, therefore, do not have an associated administration cost.

Administration of dacarbazine requires a chemotherapy suite chair time. This has been included in the model using the NHS Reference Cost SB12Z – Deliver simple parenteral chemotherapy at first attendance (outpatient) applied every 3 weeks that patients spend in the PFS state.[51]

Costs are summarised in Table 21 and Table 22.

Table 21 Drug costs (28 day supply)

	List price		PAS price	
	Price	**Detail**	**Price**	**Details**
Dabrafenib	£5600	28 x 75mg: £1400; 28 x 50mg: £933.33	[redacted]	[redacted]
Vemurafenib	£7000	56 x 240mg: £1750	[redacted]	[redacted]
Dacarbazine	£82	1000mg vial @£31.80 per vial Dacarbazine is administered once every 21 days. Cost per 28 days calculated by multiplying 21-day cost (£61) by 28/21	N/A	N/A

Table 22 Other drug related costs

Costs	Value		
	Dabrafenib Vemurafenib	**Dacarbazine**	**Source**
BRAF testing (one off cost per patient)	£206.34	N/A	MS[3]
Dispensing (every 28 days)	£13.40 every 28 days	N/A	Millar,2008[49] Curtis 2012[50]
Administration (every 28 days)		£206.13	NHS Reference Costs[52] (SB12Z Deliver parenteral chemotherapy at first attendance)

Health care costs

One-off costs for initiation of therapy and death were included in the model (see Table 23). The ERG notes that the treatment initiation figure actually reported in the manufacturer's technology appraisal submission for ipilimumab was £365 (95% CI: £191 to £538)[47] and that the terminal care cost used in this appraisal has been updated using the PSSRU Pay & Prices index 2011/12.[50]

Table 23 One-off costs associated with treatment

Health state	Cost	Source
Treatment initiation	£337	BMS (Ipilimumab NICE submission),2011[47]
Terminal care cost	£6,177	Addicott,2008[53]

Monthly costs by state were estimated by combining monthly resource use estimates from the NICE appraisal of ipilimumab with unit cost estimates from NHS Reference Costs 2011-12.[51] A summary of values used in the manufacturer's model is presented in Table 24.

Table 24 Monthly health care costs

Health State	Service	Cost	Reference
PFS	Outpatient visit (GP and plastic surgeon)	£5	NHS Reference Costs 2011-2012[51] (resource use from NICE ipilimumab submission)[47]
	Inpatient stay	£15	
	Laboratory tests	£7	
	Radiological exams	£100	
	Total	£127	
PPS	Outpatient visit (largely medical oncologist)	£211	NHS Reference Costs 2011-2012[51] (resource use from NICE ipilimumab submission)[47]
	Inpatient stay	£15	
	Laboratory tests	£7	
	Radiological exams	£100	
	Total	£333	

PFS=progression-free survival; PPS=post-progression survival

Adverse event costs

The model includes the cost of treating Grade 3+AEs with a difference in incidence of ≥5% for treatment vs control in BREAK-3[2] or BRIM-3[1] and/or those that were considered important from a clinical or economic perspective based on the opinion of physicians participating in an advisory board meeting. Estimates of the costs of treating neutropenia, rash and squamous cell carcinoma were based on results of a cost-of-illness study commissioned by the manufacturer.[54] In addition, the following assumptions were made:

- Grade 3+ keretocanthoma was assumed to have the same cost as Grade 3+ squamous cell carcinoma
- Grade 3+ PPE was assumed to cost the same as Grade 3+ rash

The cost of treating each AE is calculated by multiplying the incidence of the AE by the expected cost of the event. Estimates of the incidence of AEs are displayed in Table 25 with each cost per event in Table 26.

All AEs are assumed to occur during the first year of treatment and are applied as a one-off, undiscounted, cost.

Table 25 Incidence of treatment related AEs

Adverse Event	Number of patients experienced event		
	Dabrafenib (N=187)	**Dacarbazine (N=59)**	**Vemurafenib (Grade 3+ incidence) (N=336)[1]**
Source	BREAK-3[2]	BREAK-3[2]	BRIM-3[1]
PPE	4 (2.1%)	0	40 (11.9%)*†
Photosensitivity	0	0	Incidence not reported, assumed=0
Pyrexia	6 (3.2%)	0	40 (11.9%)†
SCC‡	11 (5.8%)	Incidence not reported, assumed=0	40 (11.9%)
Neutropenia	0	8 (13.6%)	1 (0.3%)
Rash	0	Incidence not reported, assumed=0	28 (8.3%)
Keratoacanthoma	0	Incidence not reported, assumed=0	20 (5.9%)

AE=adverse event; PPE=Palmar-plantar erythrodysesthesia; SCC=Squamous cell carcinoma

* Grade 2+

† The ERG notes that, compared with figures reported in McArthur et al (appendix) the manufacturer appears to have substantially overestimated the incidences of PPE and pyrexia for vemurafenib

‡ The ERG notes that only SCC of skin are reported for both BRIM-3[1] and BREAK-3[2] (BRIM-3:[1] vemurafenib (19%) and dacarbazine (<1%); BREAK- 3: dabrafenib (2%) and dacarbazine (2%)

Table 26 Summary of mean (standard error [SE]) cost per adverse event used in the model

Adverse event	Cost per event	
	Mean	**SE**
Grade 3+		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
Grade 1 or 2		
[redacted]	[redacted]	[redacted]

PPE=palmar-plantar erythrodysesthesia

* Note that the figure used in the model is £13,380.80

Other costs

Costs of post-study anti-cancer therapy (PSACT) were also included in the model. For each treatment group in BREAK-3[2] the expected cost of PSACT per patient was calculated by summing the product of the proportion of patients receiving various PSACTs and the corresponding estimates of the cost of a single course of each therapy. Summary expected costs are presented inTable 27 with further details provided in the MS[3] (Appendix 10.22, p323-324).[3] In the absence of vemurafenib specific data, the cost of PSACTs for patients receiving vemurafenib was assumed to be the same as that for those receiving dabrafenib.

Table 27 Expected post-study anti-cancer therapy (PSACT) costs used in the model

Comparators	Expected cost	Standard error
Dabrafenib	£3,013	£2,461
Vemurafenib	£3,013	£2,461
Dacarbazine	£6,044	£3,283

The manufacturer explained that they did not include personal social services costs in their model as they anticipated that, not only would they not be significant, but also that they would be similar for all comparators.

5.4.8 Model validation

The external workbook used for the evaluation was validated by the developers by entering model inputs into an older version of the model which had been used in prior economic evaluations and which had been validated by Programs for Assessment of Technology in Health (PATH) and Research Institute and EcoStat Consulting UK Ltd. This older model was also validated by York Health Economics Consortium.

5.4.9 Results included in manufacturer's submission

The results included in the MS[3] have been generated using PAS costs for dabrafenib and vemurafenib. The base case incremental results generated by the manufacturer's model are presented in Table 28. The ICERs for dabrafenib vs dacarbazine are £49,019 per QALY gained and £34,940 per life-year gained, whilst those for dabrafenib vs vemurafenib are £11,028 per QALY gained and £7,523 per life-year gained.

Table 28 Base case results

Tech	Total costs (£)	Total LYG	Total QALYs	Inc. costs (£)	Inc. LYG	Inc. QALYs	ICER (£) cost/ LY	ICER (£) cost/ QALY	ICER (£) Inc. QALYs
Dac	[redacted]	[redacted]	[redacted]	-	-	-	-	-	
Vem	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	44,088	60,980	Extended dominance
Dab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	34,940	49,043	49,019

Dac=dacarbazine; Dab=Dabrafenib; ICER=incremental cost effectiveness ratio; INC=incremental; LYG=life years gained; QALYs=quality-adjusted life years; Tech=Technology; Vem=vemurafenib

Table 29 Summary of predicted (discounted) costs (£) for the base case

	Dabrafenib	**Dacarbazine**	**Vemurafenib**	**Dabrafenib vs. Dacarbazine**	**Dabrafenib vs. Vemurafenib**
Medication	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Administration	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Adverse Events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Diagnostic Testing	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Progression-Free	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Post-Progression	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

5.4.10 Sensitivity analyses

Deterministic sensitivity analyses

The manufacturer carried out a wide range of deterministic sensitivity analyses. Results for the ten parameters showing the greatest variability are presented for the comparisons of dabrafenib with dacarbazine and dabrafenib with vemurafenib in Table 30 and Table 31 respectively.

Table 30 Deterministic univariate sensitivity analysis results for dabrafenib vs dacarbazine (PAS prices)

Parameter	Low value ICER/QALY	High value ICER/QALY
Base case value: £49,019		
HR for dabrafanib vs dacarbazine (95% CI)	£26,470	Dominated
Disutility of PFS vs perfect health (95% CI)	£42,637	£54,877
Disutility of PPS vs PFS (95% CI)	£46,274	£56,495
PPS cost per month (±50%)	£47,197	£50,841
PSACT costs (±50%)	£47,497	£50,560
Administration costs (±50%)	£48,399	£49,640
PF cost per month (±50%)	£48,646	£49,392
AE costs (±50%)	£48,648	£49,392
Terminal care cost (±50%)	£48,841	£49,198
Testing costs (±50%)	£48,915	£49,123
HR for PFS for dabrafenib vs dacarbazine (95% CI)	£48,551	£48,738

AE=adverse events; PF=progression free PFS=progression free survival; PPS=post progression survival; PSACT=post-study anti-cancer therapy; HR=hazard ratio; ICER=incremental cost effectiveness ratio; QALY=quality adjusted life years

Table 31 Deterministic univariate sensitivity analysis results for dabrafenib vs vemurafenib (PAS prices)

Parameter	Low value ICER/QALY	High value ICER/QALY
Base case value: £11,028		
HR for dabrafenib vs vemurafenib (95% CI)*	Dominant	£67,220
PPS cost per month (±50%)	£8,286	£13,771
Disutility of PPS vs PFS (95% CI)	£10,124	£13,771
Disutility of PFS vs perfect health (95% CI)	£10,371	£11,902
Terminal care cost (±50%)	£10,841	£11,215
PF cost per month (±50%)	£10,959	£11,907
AE costs (±50%)	£10,959	£11,097
Administration costs (±50%)	£11,020	£11,037
PSACT costs (±50%)	£11,024	£11,033
HR for OS for dabrafenib vs vemurafenib (95% CI)	£6,422	£4,276

AE=adverse event; PFS=progression free survival; PPS=post progression survival; PSACT=post-study anti-cancer therapy; HR=hazard ratio; ICER=incremental cost effectiveness ratio; QALY=quality-adjusted life year

* Note that the ERG is unclear whether this relates to the HR for PFS or to that for OS

Probabilistic sensitivity analyses

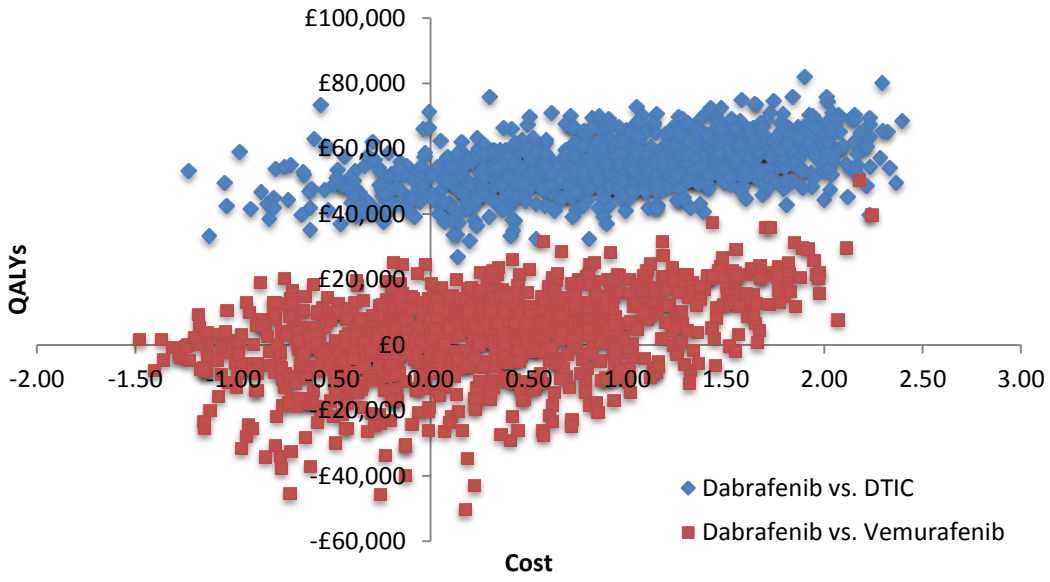
The manufacturer undertook probabilistic sensitivity analysis (PSA) to derive the mean ICERs per QALY gained for dabrafenib vs vemurafenib and dabrafenib vs dacarbazine. The distributions used in the PSA have not been provided in the MS.[3] The PSA results (see Table 32) indicate that, at a threshold of £30,000 per QALY gained, when dabrafenib was compared with dacarbazine and vemurafenib, it was found to be cost-effective in 6% and 56% of cases respectively. At a threshold of £50,000 per QALY gained, when dabrafenib was compared with dacarbazine, it was found to be cost-effective in 43.5% of cases.

Table 32 Summary of PSA for dabrafenib (PAS prices)

QALY threshold (£/QALY)	Probability that dabrafenib monotherapy is cost effective vs comparator	
	Dacarbazine	Vemurafenib
£30,000	6%	56%
£50,000	43.5%	*NR*

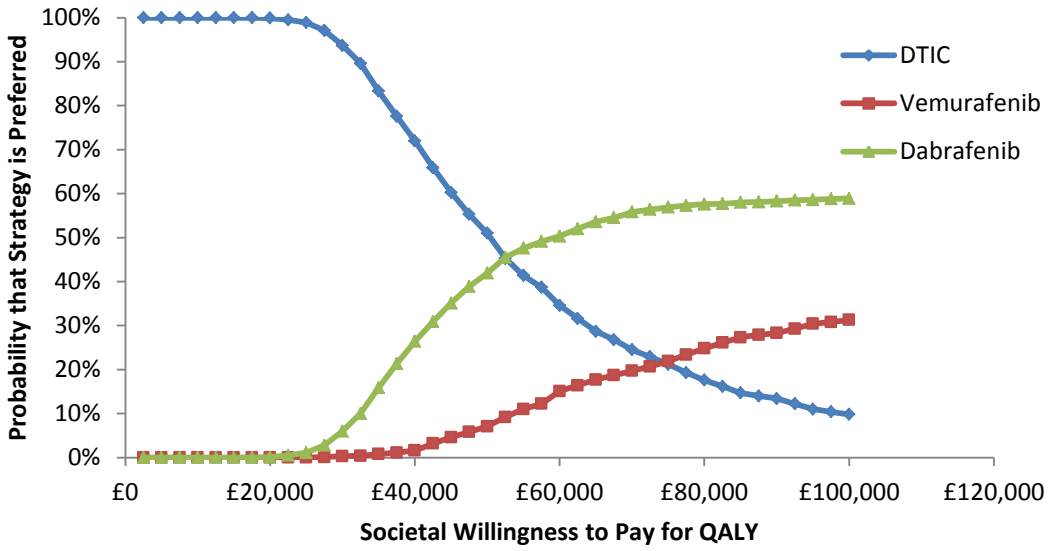
PAS=Patient access scheme; PSA=probabilistic sensitivity analysis; QALY=quality adjusted life year

The cost effectiveness plane and cost effectiveness acceptability curves included in the MS[3] are reproduced in Figure 3 and Figure 4 respectively.



DTIC=dacarbazine

Figure 3 Cost effectiveness plane for pair-wise comparisons of dabrafenib vs vemurafenib and dacarbazine



DTIC=dacarbazine

Figure 4 Cost effectiveness acceptability curves for dabrafenib, vemurafenib and dacarbazine

5.5 Detailed critique of manufacturer's economic model

The spreadsheet model submitted by the manufacturer for this appraisal is structured as a generic cancer treatment model to be able to accommodate multiple regimens of treatment simultaneous, each of which may require several component drugs to be administered. For each regimen, and each component drug, separate model variables with unique parameters are required. The result of this flexibility is that the model construction and logic is more complex than would normally be considered necessary, or appropriate, for the relatively simple logical structure of a three-state transition model. In particular, the model includes a variable table allowing nearly 10,000 separate variable parameter, together with their individual uncertainty characteristics, to be specified.

As a consequence of this over-specification of the model the ERG has at times struggled to trace the logic trails for normally straightforward aspects of such a model. Another difficulty with this structure is that if only allows results for a single regimen to be generated at a time. Incremental comparisons and incremental cost-effectiveness ratios (ICERs) can only be obtained by generating and saving costs and outcomes separately for each regimen, before they can be combined to give the final result.

5.5.1 Time to event outcomes

Importance of post-progression survival

In the manufacturer's model the majority of the estimated incremental QALYs arise from estimated life-years *after* disease progression (see Table 33). For dabrafenib vs dacarbazine 62% of the QALYs are gained in post-progression survival (PPS) vs 38% in PFS, whilst for dabrafenib vs vemurafenib 93% of QALYs are gained in PPS vs 7% in PFS. For dabrafenib vs dacarbazine nearly all of the incremental costs (99.3%) are drug related costs that occur during PFS. For dabrafenib vs vemurafenib incremental costs are divided almost equally between PFS and PPS with 44.1% of additional costs coming from drug costs during PFS and 46.3% coming from additional health care costs during PPS. Thus the methods used to estimate PFS, OS and PPS, based on data from the BREAK-3[2] trial, are central to determining the most reliable assessment of cost effectiveness. All other components of incremental costs and outcomes are of relatively minor importance.

Table 33 Most influential components of incremental costs and outcomes in manufacturer's base case decision model results

Model result	Phase	Dabrafenib vs dacarbazine	Dabrafenib vs vemurafenib
Incremental life-years	PFS	**34.9%**	6.2%
	PPS	**65.1%**	**93.8%**
	OS	*100%*	*100%*
Incremental QALYs	PFS	**38.2%**	6.9%
	PPS	**61.9%**	**93.1%**
	OS	*100%*	*100%*
Incremental costs	Drugs (PFS)	**99.3%**	**44.1%**
	PPS care	0.4%	**46.3%**
	Other costs	0.3%	9.6%
	Total	*100%*	*100%*

OS=overall survival; PFS=progression free survival; PPS=post progression survival; QALY=quality-adjusted life year

The importance of the large estimated values for PPS gain ascribed to dabrafenib becomes clear from the sensitivity of the manufacturer's ICER to alternative assumptions: if only 50% of the PPS gain in the manufacturer's base case is considered realistic, the ICER increases by about 45% to over £70,000 per QALY gained; if there is no additional survival advantage from dabrafenib beyond disease progression, the ICER more than doubles to over £120,000 per QALY gained.

Complex model structure for survival

To provide lifetime estimates of cost effectiveness, the manufacturer of dabrafenib has adopted a complex model structure for representing patient survival involving three phases:

- Phase 1 - from randomisation until 1.8 years (96 weeks) a parametric model of BREAK-3[2] trial dabrafenib survival data based on a log-normal distribution is applied, modified by a HR (1.16 for vemurafenib and 1.82 for dacarbazine) to represent the survival associated with the comparator treatments
- Phase 2 - from 1.8 years to 10 years a parametric model, based on a log-logistic distribution calibrated from 10 year AJCC registry data,[8] is applied to survivors at time 1.8 years and is applied equally to dabrafenib, vemurafenib and dacarbazine treated patients
- Phase 3 - from 10 to 30 years, survival is estimated separately for each treatment arm by applying UK life table mortality rates to those surviving at 10 years.

This complex approach to modelling long-term survival involves a number of assumptions.

Critique of Phase 1 assumptions:

- The use of a log-normal distribution to represent the trial survival data is based on comparison of Akaike and Bayesian information criteria scores (AIC/BIC) and correspondence of AUC estimates (Tables 130 & 131 of MS). However, such a comparison is not definitive as the basis for choosing a parametric function. The significance of small AIC/BIC score differences is difficult to interpret, and also assumes that the 'true' distribution is included in the set of five candidate distributions being compared (exponential, Weibull, log-normal, log-logistic and gamma). For neither AUC nor AIC/BIC is a single functional form clearly superior across all comparisons in the manufacturer's analyses, and the log-normal distribution is known to exhibit a 'long tail' which tends to exaggerate long-term survival. A more fundamental problem is that using the full data set as the basis for projecting the experience of the minority of patients still alive (at risk) at the end of trial fails to recognise the strong influence of survivor bias altering the case-mix and risk profile of long-term survivors compared to the original randomized population. This can only be overcome by concentrating curve-fitting on the data (events) towards the end of the trial.
- The use of the dabrafenib parametric model as the basis for modelling survival in the comparator arms is also vulnerable to challenge as this approach relies heavily on belief in a) the suitability of the parametric function chosen to model dabrafenib and b) the ability of this function to represent survival for the comparator treatments. The approach also assumes that hazards are proportional across all three treatments.

- In addition, the reliability of the representation of dacarbazine is strongly affected by the efficacy of the assumptions made to support the use of the RPSFTM adjustment, a technique to remove the supposed influence of crossover treatment in the BREAK-3[2] control arm by altering the trial HR.

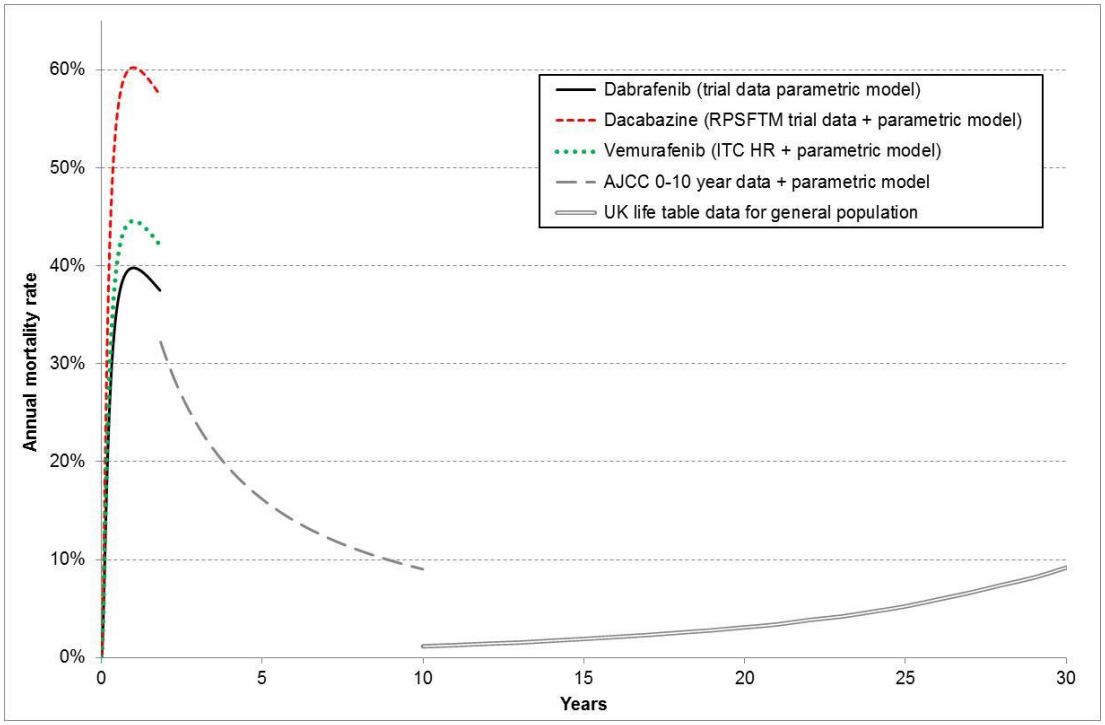
Critique of Phase 2 assumptions:

- The manufacturer's model assumes that a log-logistic parameterisation of AJCC registry data[8] is appropriate but offers no evidence to support its clinical or biological plausibility. In addition, the AJCC published data[8] relate specifically to a North American population which may not be representative of a UK population.

Critique of Phase 3 assumptions:

- The submitted model defaults to the use of background mortality rates based on UK life table statistics to represent the experience of all survivors from 10 to 30 years. Implicitly, this assumes that long-term survivors are effectively cured of metastatic cancer. This is a very strong assumption for which no supporting objective evidence has been offered (e.g. from comparing melanoma recurrence rates among long-term survivors of chemotherapy for malignant melanoma with rates in the general population).

Figure 5 displays the changes in annual mortality rates over the duration of the manufacturer's model in the base case scenario. Of particular concern are the two time points (1.8 and 10 years) at which distinct discontinuities occur. Such effects can occur in clinical trials but are always the consequence of a significant disease or treatment event, such as the start or termination of treatment, or a serious alteration in a patient's condition (such as disease progression in cancers, or a major cardiac event in heart disease). Here the abrupt changes in mortality rates have no clinical explanation and only relate to convenient choices made for modelling purposes. However, the consequence of these choices is to bias the analysis in favour of dabrafenib and against the comparators, especially dacarbazine. The use of the RPSFTM adjusted survival data from the BREAK-3[2] trial establishes a large survival difference at the end of the trial data (after 1.8 years). The subsequent use of reduced mortality rates in Phases 2 and 3 has the effect of extending the duration of this benefit. The result is that PPS gains are elongated and the ICER per QALY gained for dabrafenib compared with dacarbazine is substantially reduced.



AJCC=American Joint Committee on Cancer; HR=hazard ratio; ITC=indirect treatment comparison; RPSFTM=Rank preserving structural time failure model

Figure 5 Mortality rates in the manufacturer's base case analysis

Updated data and interpretation

To minimise uncertainty it is important that the model results available to the NICE Apparaisal Committee (AC) should be based on the most complete data available. The manufacturer's PFS analyses have not been altered as the existing data were defined in the trial protocol as being based on the final PFS analysis. The OS data from the BREAK-3 clinical trial[2] used in the submitted model were obtained from the December 2012 data-cut. In response to the ERG clarification letter, the manufacturer has provided evidence relating to a more recent analysis (January 2014 data-cut) which extends the maximum follow-up time [redacted]. However, the RPSFTM calculations have not yet been updated for these new data.

A related issue concerns the appropriate censoring rule to be applied to incomplete survival data. The ERG requested survival analyses be rerun using a revised censoring rule with the date of data-cut rather than the date of last observation used for censoring, the date of last observation being that conventionally used in clinical trial reports. The requested approach was designed to identify any bias affecting Kaplan-Meier analyses due to informative censoring as this effect has been observed in various previous appraisals (notably that of vemurafenib for the treatment of locally advanced or metastatic BRAF^{v600} mutation-positive malignant melanoma)[29] and can misrepresent survival projections when used to calibrate

parametric survival functions. The manufacturer refused this request, considering that '*the approach suggested may introduce bias in some instances*' preferring to state that '*the conventional approach to censoring is to censor at last observation or event.*' The manufacturer went on to state that '*We are unaware of any examples where this alternative approach (censor at data-cut off) has been used for the purpose of informing economic evaluations conducted as part of a NICE appraisal.*' The ERG is of the view that the suggested endpoint bias (if real) is outweighed by the bias observed in multiple cancer trials, and demonstrated by simulation study.[55] Furthermore, similar re-analysis using date of data-cut has been previously requested in NICE appraisals and was used by the ERG in preparing their report for the appraisal of afatinib for treating EGFR mutation-positive locally advanced or metastatic non-small cell lung cancer [TA310].[56]

Effects of crossover within the clinical trial

The manufacturer's base case analysis employs an RPSFTM adjustment to the dacarbazine OS results prior to projecting survival. This is based on an assumption that the effectiveness of dabrafenib exhibited in the intervention arm of the BREAK-3[2] trial will also apply to the dacarbazine arm for those patients who crossover at the time of disease progression to receive dabrafenib as a second-line treatment. Although this argument at first sight appears plausible and attractive, it requires strong justification since it can have the effect of radically changing the model estimates of cost effectiveness. In particular, patients receiving dacarbazine whose disease has progressed are unlikely to be in a similar physical condition (and therefore have similar survival prognosis) as those starting dabrafenib as first-line chemotherapy. So, even if crossover from dacarbazine to dabrafenib can be confirmed to have some efficacy, it is not certain that the recipients will receive 100% of the efficacy experienced in first-line therapy since those patients suffering early progression are likely to be those with the least propensity to benefit from further treatment.

To consider the extent of supporting evidence within the key BREAK-3[2] trial, the ERG requested an analysis of PPS comparing those 37 dacarbazine treated patients (69%) who elected to crossover to dabrafenib with the 17 patients (31%) who did not. Despite this being a non-randomised comparison, similar proportions died (70%,71%) and remained in ongoing follow-up at data-cut (24%,24%). The time from disease progression in terms of median (50% survival), first quartile (75% survival) and third quartile (25% survival) are shown in Figure 6 together with confidence intervals. It is apparent that these data offer no convincing evidence of any significant survival difference, and therefore the null hypothesis (that there is no crossover effect favouring dabrafenib) cannot be rejected.

[redacted]

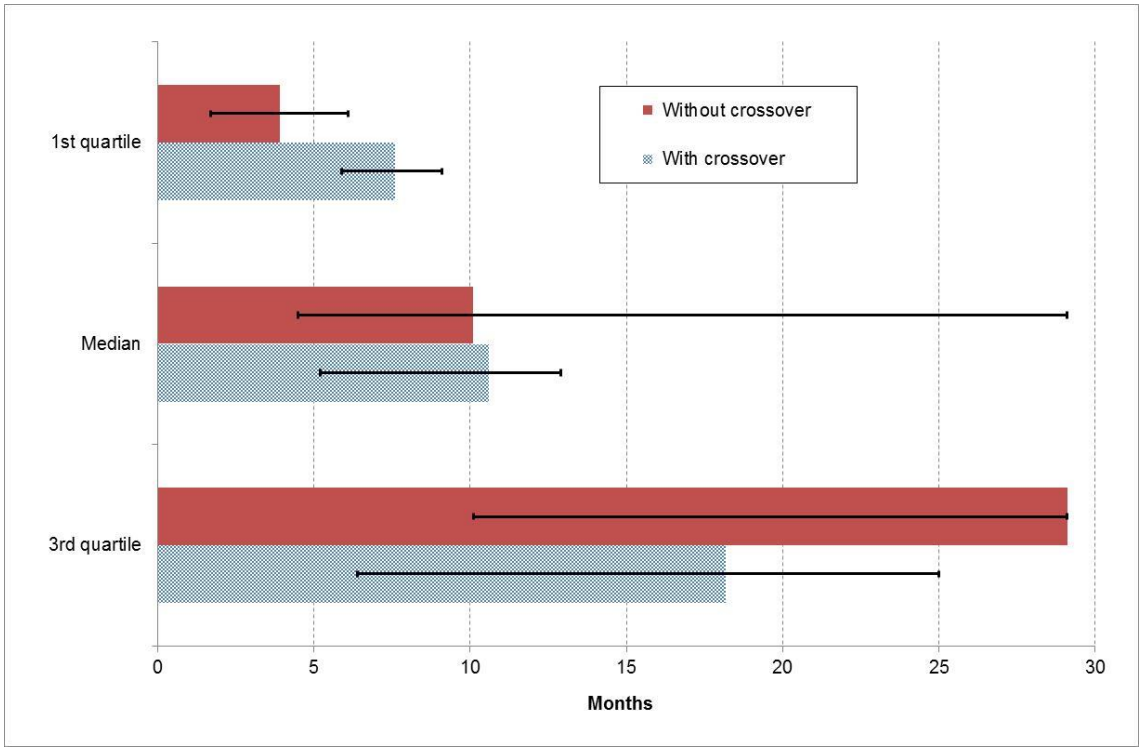


Figure 6 PPS survival comparing dacarbazine patients who did and did not crossover to dabrafenib on disease progression

Reliability of hazard ratios

The manufacturer's model generates survival curves for PFS and OS by fitting log-normal parametric curves to the dabrafenib data from the BREAK-3[2] trial, and then applying a constant HR to the dabrafenib survival curve to estimate the survival of patients receiving dacarbazine. The model extends this approach to vemurafenib by applying a second constant HR derived from an indirect treatment comparison between the BREAK-3[2] and BRIM-3[1] trial results to the BREAK-3[2] trial dabrafenib estimates. Thus there are three distinct assumptions underlying this approach to modelling and projecting survival data:

1) Constant HRs are valid for both PFS and OS data within the BREAK-3[2] trial (dabrafenib vs dacarbazine)
2) Constant HRs are valid for both PFS and OS data within the BRIM-3[1] trial (vemurafenib vs dacarbazine)
3) PFS and OS data for dacarbazine within the two trials (BREAK-3[2] and BRIM-3[1]) are broadly comparable and conform to a constant hazard relationship.

Assumption 1: Analysis of PFS data (Figure 7) and OS data ([redacted]8) from BREAK-3[2] indicate a significant deviation from constant HR between the trial arms for PFS (p<0.001, chi-squared test) but not for OS (p=0.91).

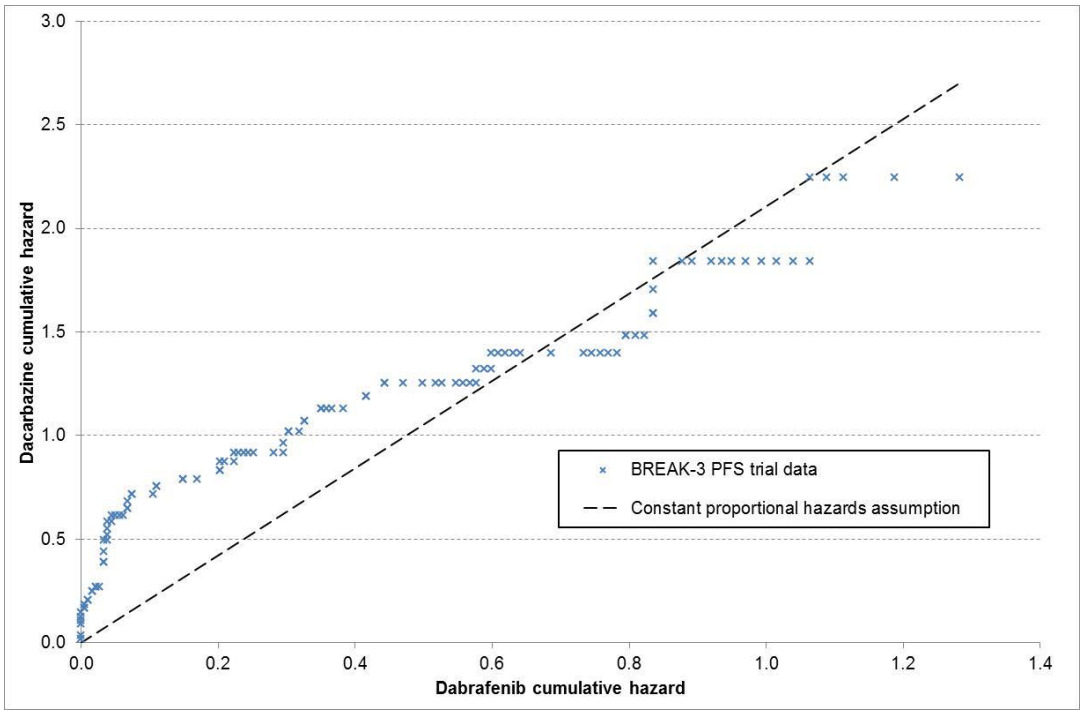


Figure 7 Matched PFS cumulative hazards plot comparing dacarbazine and dabrafenib patients in the BREAK-3 trial, relative to an assumption of a constant HR (Final PFS data-cut June 2012)

[redacted]8 [redacted] (confidential therefore removed)

Figure 9 Matched PFS cumulative hazards comparing dacarbazine and vemurafenib patients in the BRIM-3 trial, relative to an assumption of a constant HR (December 2010 data-cut)

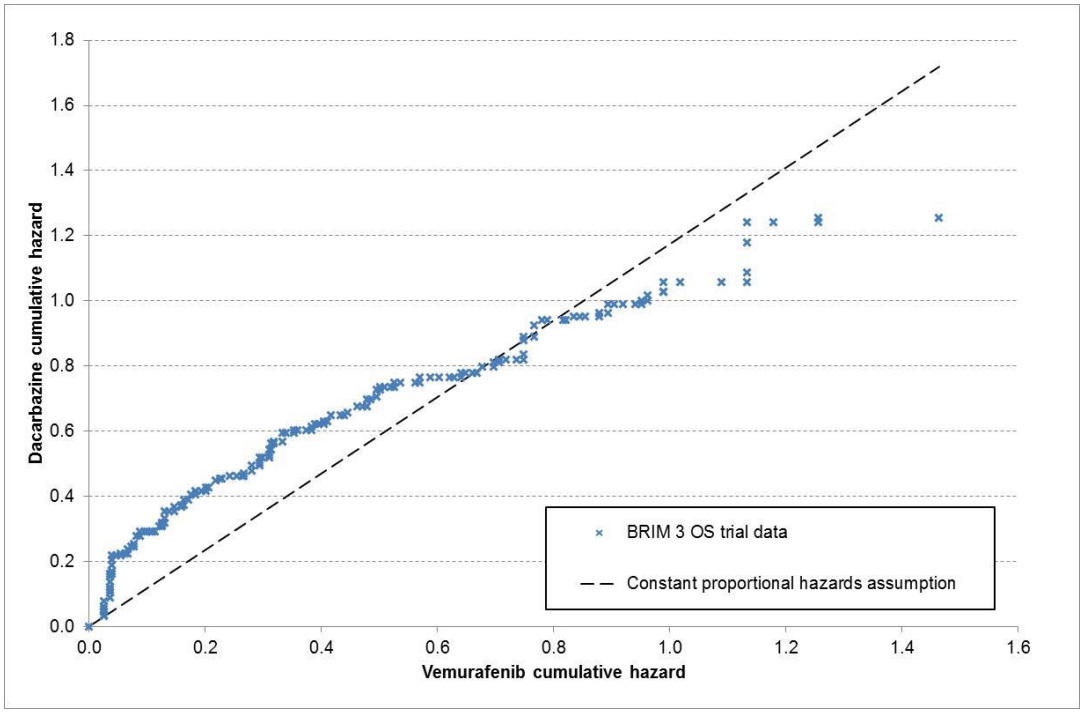


Figure 10 Matched OS cumulative hazards comparing dacarbazine and vemurafenib patients in the BRIM-3 trial, relative to an assumption of a constant hazard ratio. (February 2012 data-cut censored at crossover)

Assumption 3: Comparison of dacarbazine PFS and OS data between the BREAK-3[2] and BRIM-3 trials[1] (Figure 11 and Figure 12) indicates that PFS hazard profiles are not significantly different (p=0.46), but constant proportion hazards are not supported for OS (p<0.001).

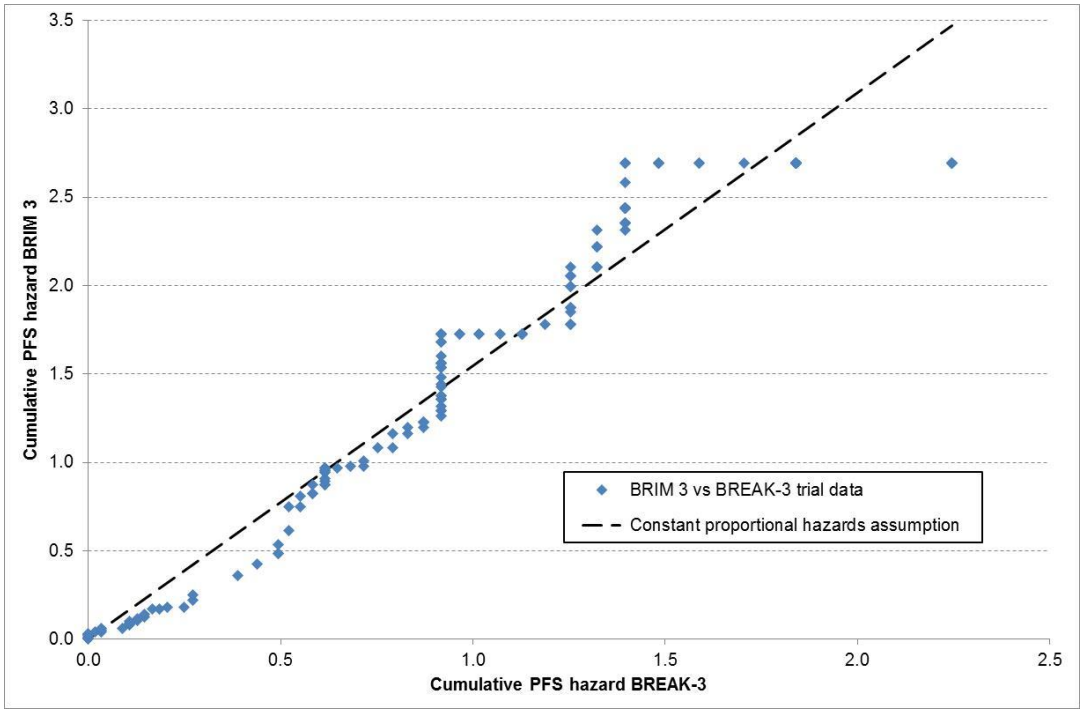


Figure 11 Matched PFS cumulative hazards comparing dacarbazine and vemurafenib patients in the BREAK-3 trial, relative to an assumption of a constant hazard ratio. (December 2010 data-cut)

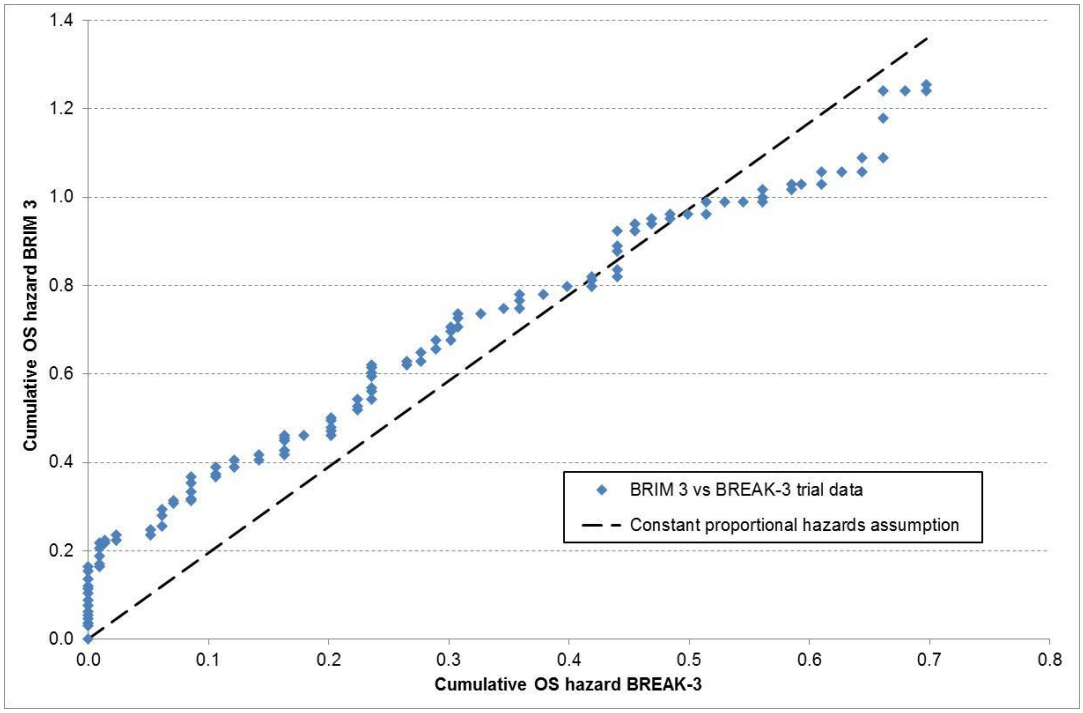
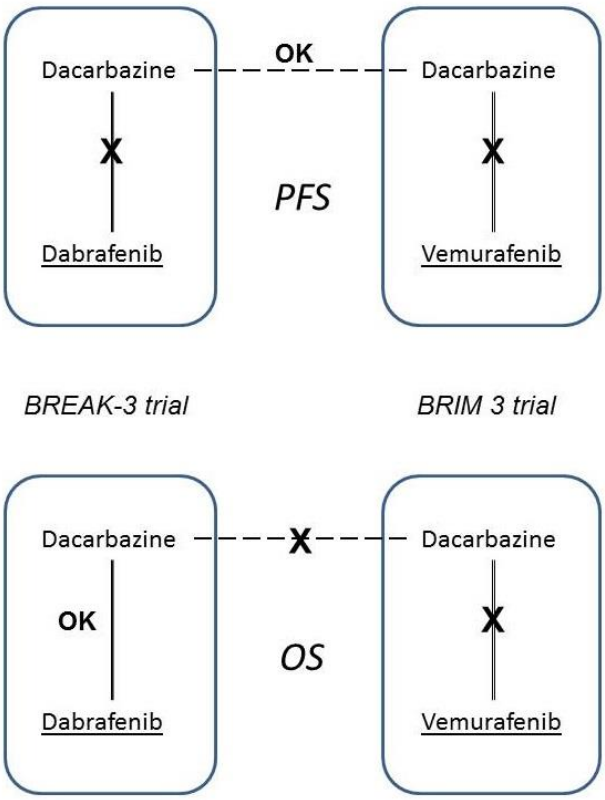


Figure 12 Matched OS cumulative hazards comparing dacarbazine and vemurafenib patients in the BRIM-3 trial, relative to an assumption of a constant hazard ratio. (February 2012 data-cut censored at crossover)

The ERG therefore concludes that all three necessary assumptions are not validated from the two key clinical trials[1,2] for either PFS or OS (Figure 13), and therefore the results of the ITCs of dabrafenib vs vemurafenib presented by the manufacturer are inappropriate and unreliable. Any assessment of the relative cost effectiveness of dabrafenib compared with vemurafenib based on derived HRs cannot be considered meaningful.



Statistically significant violations of the constant proportional hazard assumptions are indicated by **X**, otherwise relationships are marked **OK**

Figure 13 Indirect treatment comparisons to compare dabrafenib with vemurafenib

6 ADDITIONAL WORK UNDERTAKEN BY ERG

6.1.1 Revising PFS projection

In the absence of receiving the formal re-analysis of the BREAK-3[2] PFS data requested by the ERG, the ERG carried out a careful examination of the timing of censoring compared to the timing of PFS events to determine the extent of possible informative censoring bias. Patients with the longest time in study who were censored at protocol-prescribed assessment times were considered for adjustment by adding back the estimated missing exposure time between date of last assessment and date of data-cut. In total, censoring times for six dacarbazine patients and 49 dabrafenib patients were modified. Re-analysis of the modified data indicated that adjustments resulted in relatively minor alterations to the PFS estimates which were limited to the final few events in each trial arm.

The PFS data were then converted into cumulative hazards (Figure 14) and examined for long-term trends. This analysis was undertaken to inform the choice of the most suitable parametric function with which to model BREAK-3[2] PFS data in order to allow reliable long-term projection of PFS trends. The ERG noted that, for both dabrafenib and dacarbazine, it appears that there is a short period (between 8 and 12 weeks) free of events, followed by a period of high event risk modulated by a regular pattern related to the regular protocol-determined patient assessments. Finally, in both trial arms the cumulative hazard trend settles to a steady near-linear pattern. The ERG therefore concluded that the most appropriate method of representing lifetime PFS trends was to use the 'area under the curve' (AUC) approach to capture PFS in the early variable segment, and then fit a simple linear cumulative hazard trend, equivalent to an exponential projective function for the latter period.

The results are shown in Figure 14 and Figure 15, and compared with both the trial Kaplan-Meier data and the manufacturer's log-normal models. It is evident that the manufacturer's approach tends to over-estimate the long-term hazard rate (i.e. under-estimate PFS) in both the dabrafenib arm and the dacarbazine arm. This is a result of using the whole data set (including the early variable phase) for model calibration and the adoption of a common proportional hazards assumption.

The resulting lifetime estimates of expected PFS are compared with the manufacturer's base case scenario in Table 24 and show that when the ERG's approach to projective modelling is applied there is an additional 7 to 8 days of PFS gain from use of dabrafenib.

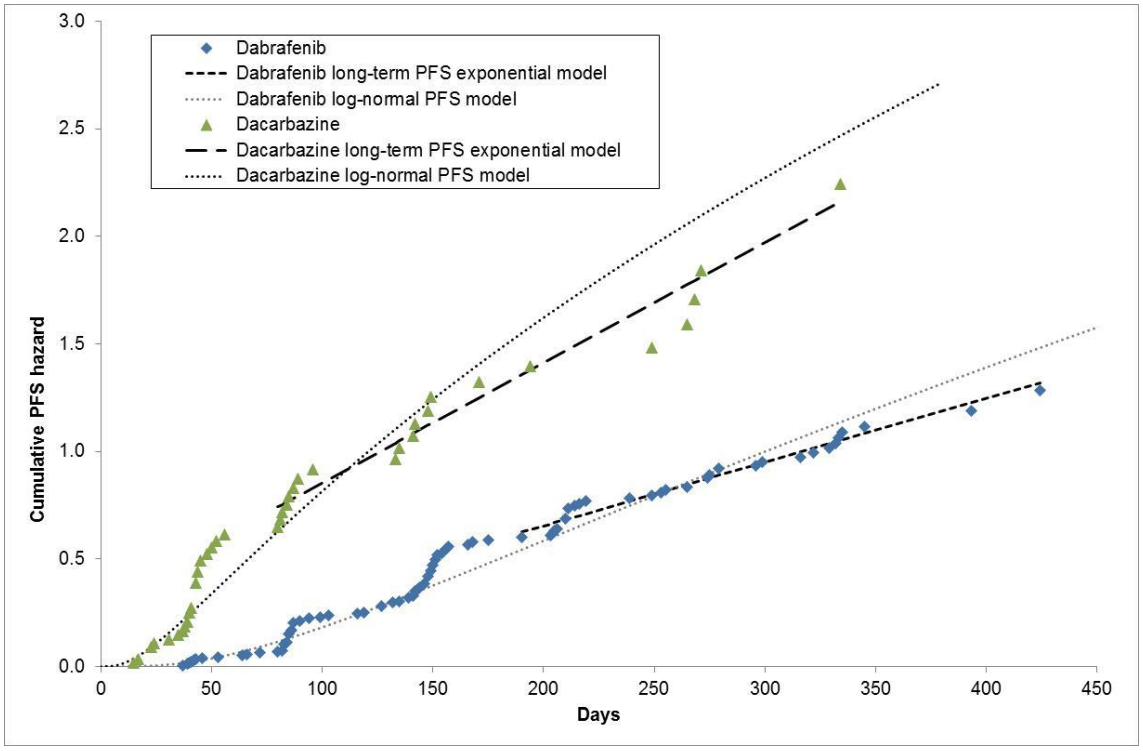


Figure 14 PFS BREAK-3 cumulative hazard Kaplan-Meier data, ERG long-term linear trends and the manufacturer's log-normal models

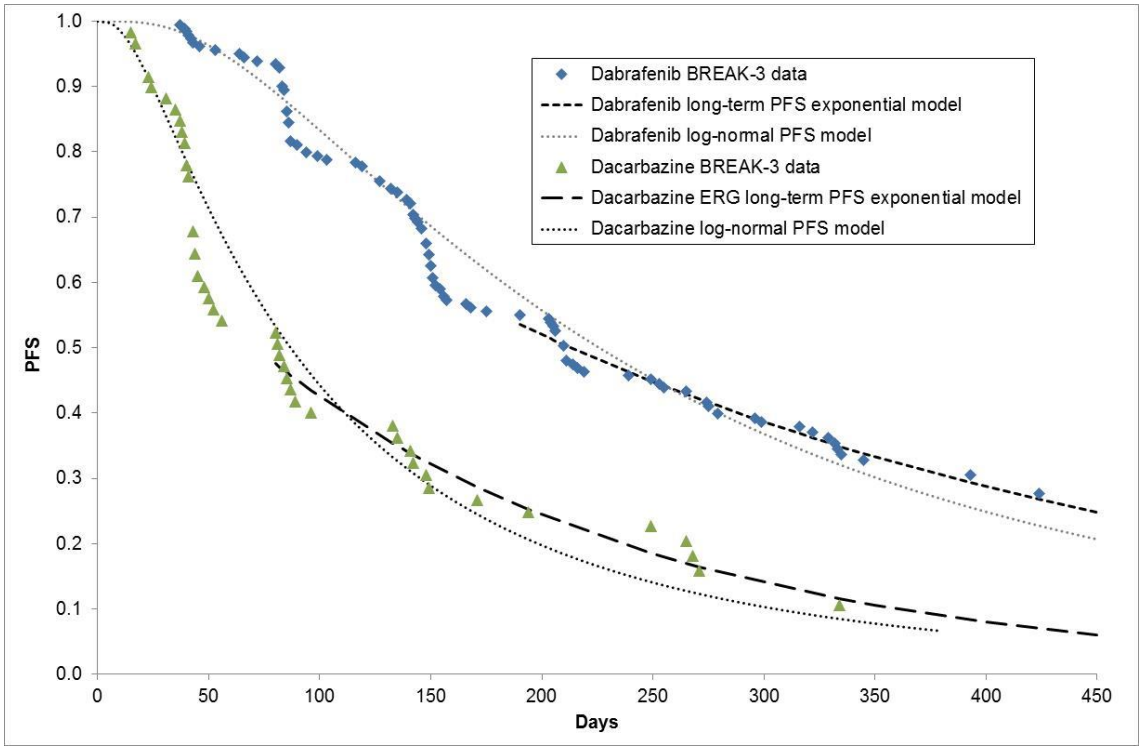


Figure 15 BREAK-3 Kaplan-Meier PFS data, ERG long-term exponential models and the manufacturer's log-normal models

Table 24 Comparison of estimated lifetime PFS between the manufacturer's base case and the ERG long-term projection method

	Manufacturer's base case PFS estimates (months)		**ERG long-term projection PFS Estimates (months)**	
Treatment	Dabrafenib	Dacarbazine	Dabrafenib	Dacarbazine
AUC PFS	-	-	6.69	3.31
Projected PFS	10.69	4.73	4.28	1.46
Total PFS	10.69	4.73	10.97	4.76
PFS gain	**+5.96**	**-**	**+6.21**	**-**

PFS=progression free survival

Using these ERG PFS estimates in the manufacturer's model results in a small increase in the incremental QALYs per patient (+0.16%). There is, however, a much larger increase in the incremental cost per patient (+6.32%). This is mainly due to additional days of treatment with dabrafenib. The consequence is that the ICER for dabrafenib vs dacarbazine increases from £49,019 to £52,035 per QALY gained.

6.1.2 Updating OS trial data

The manufacturer's decision model has been calibrated using OS data from the December 2012 data-cut, with maximum follow-up of 21 to 22 months. In response to the ERG clarification requests, the manufacturer has supplied updated OS results (data-cut January 2014) with maximum follow-up [redacted]. Unfortunately, the submitted economic model has not yet been revised to include the new data, and the calculations of crossover RPFSTM adjustments for these data are not yet available. The ERG has explored some of the implications of incorporating the new evidence in the model, and has also considered alternative interpretations of the data.

On the basis of the considerations outlined in section 5.5.1, the ERG does not consider that the BREAK-3[2] data can be combined with BRIM-3[1] trial results to assess the cost effectiveness of dabrafenib in relation to vemurafenib. The ERG also considers that no convincing evidence has been advanced to support the suggestion that patients who crossed over from dacarbazine to dabrafenib received any additional health benefit compared with second-line treatments available to patients who did not crossover. Therefore, the analyses presented in this section are restricted to the comparison of survival in patients treated with dabrafenib or dacarbazine, excluding any hypothesised RPFSTM adjustment for a crossover effect.

ERG OS projection model

The ERG has digitised the OS Kaplan-Meier results from the latest data-cut included in the manufacturer's response to the clarification questions. On examination of the marked censoring times and the number of patients remaining in follow-up in each arm the ERG has concluded that the risk of bias introduced through the choice of censoring method is small, and could only possibly affect the last few recorded events in each trial arm. The analysis, therefore, was carried out without any post-hoc adjustments.

Figure 16 displays the cumulative hazard trends; after an initial period of 2 to 3 months in which mortality was very rare (probably due to protocol exclusion criteria), both trial arms present near-linear trends, corresponding to exponential survival curves.

[redacted] 16 [redacted] confidential therefore removed

In Table 25 the lifetime estimate of mean OS from the manufacturer's base case model scenario and the ERG's simple exponential model are compared. If the ERG projective model is used the estimated OS gain is reduced to [redacted] (from [redacted] in the manufacturer's base case) and this corresponds to an approximate three-fold increase in the estimated ICER per QALY gained.

Table 25 Comparison of estimated mean lifetime OS between the manufacturer's base case (including RPFSTM crossover adjustment using December 2012 data-cut) and the ERG long-term projection method (unadjusted using January 2014 data-cut)

	Manufacturer's base case OS estimates (months)		[redacted]	
Treatment	Dabrafenib	Dacarbazine	[redacted]	[redacted]
AUC OS	-	-	[redacted]	[redacted]
Projected OS	[redacted]	[redacted]	[redacted]	[redacted]
Total OS	[redacted]	[redacted]	[redacted]	[redacted]
OS gain	[redacted]	[redacted]	[redacted]	[redacted]

Compound OS projection model

The ERG has also calibrated an alternative parametric model formulation, which was developed previously for the vemurafenib STA appraisal.[29] This was based on clinician observations that patients with metastatic melanoma appear to be divided between a large

group who have a very poor prognosis (usually only a few months) and a small subgroup who may survive as long as 8 to 10 years. Therefore a 2-group model was proposed in which each group was subject to a constant mortality risk, but risks differ markedly between the groups. This formulation was tested on Surveillance, Epidemiology, and End Results SEER program[57] and AJCC registry data[8] and found to be very accurate in representing the long-term prospects in malignant melanoma patients.

When this model was fitted to the updated BREAK-3 trial data,[2] it estimated that about 10% of trial patients (9.8% in the dacarbazine arm and 10.2% in the dabrafenib arm) fell into the good prognosis subgroup for whom the estimated mean survival was 100 months (8.3 years) in both trial arms. By contrast, for the remaining 90% of patients the prognosis is much poorer and differed markedly by treatment: 17.2 months mean OS for dacarbazine, and 24.4 months for dabrafenib. Overall, the net gain in mean survival attributable to dabrafenib is estimated by this method to be **6.82 months**, slightly less than with the basic ERG model.

The meaningful differences between these formulations only occur within the first year, suggesting that the beneficial effect of dabrafenib may only apply to the majority of patients with very short life expectancy. The ERG has therefore not used this alternative more complex model in re-estimating cost-effectiveness, which nonetheless serves to illustrate how the ERG model has clinical plausibility despite its apparent simplicity.

Substituting these ERG estimates for OS in the manufacturer's model results in a large reduction in the incremental life years (-68.3%) and incremental QALYs gained per patient (-35.7%%), but a smaller decrease in incremental cost per patient (-11.4%). The net result is that the ICER for dabrafenib vs dacarbazine increases from £49,019 to £99,560 per QALY gained.

6.1.3 Minor issues

Dacarbazine drug acquisition and administration costs

In estimating the acquisition cost of dacarbazine the manufacturer has assumed that BSA conforms to a log-normal distribution with a mean of 1.926 and standard deviation of 0.224.

This is incorrect as BSA within the general population and in cancer chemotherapy populations (Sacco[58] paper) conforms to a normal distribution, unlike body weight which is best represented by the log-normal function. In addition, the manufacturer has failed to recognise that separate analysis by gender is an important factor when calculating the average dose and related cost of medication The ERG has re-estimated the cost of dacarbazine assuming the same gender balance as in the BREAK-3[2] trial (101 females: 149

males) and average UK patients' chemotherapy BSA values for a similar age range (females: mean BSA 1.726, standard deviation 0.182; males: mean BSA 1.929, standard deviation 0.197[58]). In addition, the most recently reported eMIT[59] prices for dacarbazine have been applied in the most efficient manner across the population BSA range to achieve a more accurate overall mean cost per dose of £35.79 in place of the manufacturer's estimate of £58.36 per dose.

The cost of dacarbazine administration has been updated using the most recent NHS Reference Costs 2012/13,[31] resulting in an increase from £203.16 to £214.16 per dose.

The net effect of these parameter value changes is to increase the incremental cost by £86 per patient, and the ICER for dabrafenib vs dacarbazine by £87 per QALY gained.

Cost of additional lines of chemotherapy

The manufacturer's base case analysis features different mean costs of subsequent lines of chemotherapy (PSACT) for patients treated with dabrafenib and dacarbazine, based on the number of such treatments recorded in the BREAK-3[2] trial.The totals in the two trial arms are small (38 among 248 patients), and do not suggest a statistically significant difference. Appendix 10.2 of the MS[3] records:

> *'The validity of the approach outlined above to incorporating PSACT costs into the model was explored with experts during an advisory board meeting. They commented that it would be appropriate not to include any PSACT costs given that in UK practice there is unlikely to be any relationship between the type of PSACT (and therefore the cost) and the first line treatment that was received.'*

The ERG concurs with this view and therefore considers that equal costs for PCAST should be used for dacarbazine and dabrafenib. Amending the model leads to the incremental cost of dabrafenib vs dacarbazine increasing by £3,031 per patient, and the ICER increasing by £3,047 per QALY gained.

6.1.4 Summary of ERG model amendments

The impact that each of the ERG model amendments has on the modelled cost effectiveness of dabrafenib compared with dacarbazine is shown in Table 26. Three of these changes result in relatively minor increases in the ICER. However, the alternative modelling of OS has a major effect leading to a doubling of the ICER to nearly £100,000 per QALY gained. When taken together with the other three ERG amendments the ICER becomes £112,727 per QALY gained.

The major difference between the ERG OS long-term projections and those generated by the manufacturer's submitted decision model lies in the estimation of long-term survival in the post-progression period. The ERG PFS estimates are very similar to those in the manufacturer's base case, albeit suggesting a slightly greater benefit for dabrafenib, but the manufacturer's base case analysis suggests three times as much advantage during PPS as that estimated by the ERG.

The contrasting elements of these two methods of analysis are as follows:

- The submitted model is based on fitting a log-normal parametric model to the whole BREAK-3[2] Kaplan-Meier data and then using this to represent the whole period of trial follow-up. The ERG used a more recent BREAK-3[2] data-cut and used the actual (AUC) data initially (i.e. until a long-term trend had been established) before applying the long-term exponential function thereafter.
- At the end of the available trial data (1.8 years) the submitted model applies hazard rates derived from a log-logistic parametric representation of AJCC registry data[8] to patients remaining alive in each arm of the model until 10 years have elapsed. The ERG continues to apply the long-term exponential projection model derived from the BREAK-3[2] trial throughout this period.
- After 10 years have elapsed, the submitted model applies hazard rates obtained from UK life tables for the general population without any adjustment until 30 years have elapsed. In the ERG formulation effectively all patients have died by 10 years, so no further modelling is required.
- The submitted model relies on the assumption of constant proportional hazards within both the BREAK-3[2] and BRIM-3[1] trials and between the control arms of the two trials. The ERG concluded that this assumption is rejected by the trial evidence and that it is therefore not meaningful to rely on an indirect treatment comparison of PFS and OS between dabrafenib and vemurafenib. Furthermore, the ERG has not used this assumption when projecting PFS and OS beyond the BREAK-3[2] trial data.
- The anomalous discontinuities in mortality rates in the submitted model (highlighted in Figure 6) cannot occur when employing the ERG method as a single parametric function is used throughout the lifetime of patients.
- The submitted model includes a large alteration to the control (dacarbazine) arm data of the BREAK-3[2] trial on the assumption of equivalent dabrafenib crossover effectiveness as that experienced by those receiving first-line dabrafenib treatment. The ERG did not find any supporting evidence for this hypothesis, and therefore employed the BREAK-3[2] outcome data without modification.

Table 26 Cost-effectiveness results with ERG modifications to manufacturer's base case

Model scenario & revisions	Dabrafenib			Dacarbazine			Incremental			ICER
	Cost	QALYs	Life years	Cost	QALYs	Life years	Cost	QALYs	Life years	£/QALY
Manufacturer's base case	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£49,019**
Dacarbazine treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£49,106
Common PSACT costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,066
ERG PFS model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£52,035
ERG OS model	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£99,560
Base case + cost revisions	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£52,153**
Base case + cost and survival revisions	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	**£112,727**

ICER=incremental cost effectiveness ratio; QALY=quality-adjusted life year

6.1.5 Conclusions of consideration of cost-effectiveness evidence

The manufacturer of dabrafenib has submitted a complex decision model involving several important assumptions, but the ERG considers these lack reliable evidential support.

Firstly, in order to meet the requirements of the NICE the manufacturer has employed a simple indirect treatment comparison based on two clinical trials (BREAK-3[2] and BRIM-3[1]) to facilitate a cost-effectiveness comparison between dabrafenib and vemurafenib. This relies on strong assumptions about constant proportional hazards both within and between the two trials, and on examination of the evidence the ERG concludes that such an indirect comparison is not valid, and the results should not be considered meaningful or reliable.

Secondly, the manufacturer has employed a complex three-phase survival model combining clinical trial data with registry data and UK general population mortality data in a fashion which generates unexpected and unjustifiable sudden changes in mortality rates which serve only to artificially extend expected patient survival times and thereby exaggerate the apparent post-treatment survival benefit attributed to dabrafenib.

Thirdly, the manufacturer has chosen to apply an exploratory analysis of a possible effect of patients crossing over from dacarbazine to dabrafenib in order to diminish the trialsurvival outcomes for dacarbazine patients, thus increasing the apparent incremental survival gain attributed for dabrafenib. The ERG has seen no objective evidence that such crossover has generated any measurable effect on survival outcomes, nor evidence that patients electing for crossover at disease progression to receive dabrafenib are in a similar condition, and there have an equivalent propensity to benefit, to those receiving dabrafenib at randomisation. In the absence of such evidence the ERG concludes that the unadjusted BREAK-3[2] trial results should be used, and any crossover analyses can only be considered speculative and optimistic.

Fourthly, the ERG has identified some errors in the costing of dacarbazine therapy, post-progression subsequent treatments and the administration cost of chemotherapy which have a minor effect on cost-effectiveness estimates.

The ERG has applied the latest OS data from the BREAK-3[2] trial to test the sensitivity of the model results to a less complex model structure, based solely on the BREAK-3[2] trial and concludes that the best estimated ICER for dabrafenib compared with dacarbazine is £112,727 per QALY gained. This amended model estimates a net survival advantage for dabrafenib of more than 7 months, compared with an expected survival of less than 22 months for patients treated with dacarbazine.

7 END OF LIFE

The manufacturer makes a case that the previous appraisal of vemurafenib met the criteria set by NICE for end of life treatment. Namely:

- The life expectancy of the patient population was short (< 24 months)
- The number of patients who would be eligible for the treatment is small
- The increase in OS is >3 months

The manufacturer argues that dabrafenib meets these criteria (page 183 of the MS[3]). They indicate that the patient population has a short life expectancy. Citing data from BRIM-3[1] that reports median OS of 13.6 and 9.7 months for patients treated with vemurafenib and dacarbazine respectively. The MS[3] highlights that there is a small patient population which would be a maximum of 990 patients per annum. The manufacturer also point to the fact that the data in the economic evaluation demonstrates an incremental OS for dabrafenib versus dacarbazine of 7.5 months.

The ERG notes however that the figure used in the MS economic model base case employs a mean survival in the dacarbazine arm of [redacted]. Also the OS data for the BREAK-3[2] trial indicates median OS of 18.2 and 15.6 months respectively for dabrafenib and dacarbazine, a difference of just 2.6 months. The ERG therefore points out that using the figures from the clinical data in MS,[3] the case for consideration as end of life treatment has not been met.

8 KEY ISSUES

Critical assessment of the MS[3] for this appraisal is challenging for a number of reasons. Although trial data are available from two key RCTs, BREAK-3[2] and BRIM-3,[1] there is a lack of appropriate data that allows the clinical and cost effectiveness of dabrafenib to be assessed for the entire patient population or against all of the comparators outlined in the final scope issued by NICE.

One limitation is that the manufacturer was only able to address directly the effectiveness of dabrafenib vs dacarbazine in treatment-naïve patients using data from a single RCT. However, even within this subgroup of patients, the data are limited in that BREAK-3[2] allowed patients to cross over from the comparator arm (dacarbazine) at the time of disease progression to receive dabrafenib. In order to adjust for crossover, the manufacturer has commissioned external consultants to explore the potential effects of using a number of methodological approaches and chosen to employ the treatment group RPFSTM to adjust OS data in the MS.[3] [redacted]. Importantly, however, this analysis assumes there will be a treatment effect from crossover and does not take into consideration the fact that there is no evidence of clinical effectiveness related to the use of dabrafenib in patients whose disease has already progressed. Analysis by the ERG, of data proved by the manufacturer, indicates that there is no convincing evidence of any significant survival difference between the patients who crossed over to receive dabrafenib after disease progression and those that did not. The MS[3] also failed to consider whether the subsequent treatments received by patients in both arms were sufficiently similar and/or reflected current practice, important requirements underlying the RPFSTM since subsequent treatments other than the treatment of interest can also influence OS.

A second limitation was that there are no RCT data available that allow the direct comparison of dabrafenib vs vemurafenib. As noted by the manufacturer, vemurafenib has replaced dacarbazine as the standard of care in clinical practice. Hence a comparison with vemurafenib is of greater clinical significance than a comparison with dacarbazine. The manufacturer did conduct an ITC in order to attempt to present evidence for the clinical effectiveness of dabrafenib vs vemurafenib. However, as shown from evidence presented by the ERG, the assumptions required for the ITC to yield credible results are not met. This means that the estimated HRs from the ITC analyses cannot be relied upon to inform either the clinical or cost effectiveness of dabrafenib vs vemurafenib.

The economic model provided by the manufacturer was structured as a generic cancer treatment model and therefore was complex and built in three phases. The use of these three phases causes a discontinuity in the mortality rates employed in the model. The ERG has highlighted a number of assumptions inherent in these three phases and explains why they do not hold. In addition, the manufacturer used the RPSFTM analysis in the economic model and, as noted above, these estimates can only be considered as exploratory. A key feature of the manufacturer's model is that the majority of the estimated incremental QALYs arise from estimated life years after disease progression. The ERG carried out additional analysis to explore PFS and PPS using BREAK-3[2] trial data. The PFS analysis indicates that there is an additional 7 to 8 days of PFS gain from the use of dabrafenib rather than dacarbazine. In terms of OS, the ERG's estimated OS gain is reduced to [redacted] rather than the [redacted] months used in the manufacturer's model.

As noted above the ERG demonstrated a lack of validity of the methods used in the ITC and therefore was not able to assess the cost effectiveness of dabrafenib vs vemurafenib.

9 OVERALL CONCLUSIONS

The ERG's examination of the trial results and economic model provided by the manufacturer indicates a clinical benefit of dabrafenib over dacarbazine in a treatment-naïve population with statistically significant results related to PFS but not OS. The ERG identified a number of issues related to economic model assumptions and parameters. Additional analysis carried out by the ERG indicates that the incremental cost per QALY gained (dabrafenib vs dacarbazine) may in fact be as high as £112,727.

Comparison of clinical effectiveness data related to the use of dabrafenib and vemurafenib via the ITCs presented in the MS[3] were not considered to be valid. The ERG is therefore unable to assess the clinical or cost effectiveness of dabrafenib compared to vemurafenib.

9.1 Implications for research

Ideally, a study directly comparing dabrafenib with vemurafenib is required. Alternatively, if and when findings from additional studies of dabrafenib and vemurafenib are available, an indirect or multiple treatment comparison may suffice. Given the problems identified by the ERG with the current ITC, it is important that any assumptions underlying such a comparison are tested and patients included in the analyses are sufficiently similar across studies.

.

10 REFERENCES

1. McArthur GA, Chapman PB, Robert C, Larkin J, Haanen JB, Dummer R, *et al.* Safety and efficacy of vemurafenib in BRAF^{V600E} and BRAFV^{600K} mutation-positive melanoma (BRIM-3): extended follow-up of a phase 3, randomised, open-label study. Lancet Oncol. 2014; 15:323-32.

2. Hauschild A, Grob JJ, Demidov LV, Jouary T, Gutzmer R, Millward M, *et al.* Dabrafenib in BRAF-mutated metastatic melanoma: A multicentre, open-label, phase 3 randomised controlled trial. Lancet. 2012; 380:358-65.

3. GlaxoSmithKline UK. Melanoma (unresectable/metastatic BRAF^{V600} mutation positive) - dabrafenib. Manufacturer's Submission to NICE Appraisal Process. 2014 April.

4. Algazi A, Soon C, Daud A. Treatment of cutaneous melanoma: current approaches and future prospects. Cancer Manag Res. 2010; 2:197-211.

5. Cancer Research UK. Skin Cancer statistics - UK. 2011; Available from: http://info.cancerresearchuk.org/cancerstats/types/skin/ (Accessed: 3 July 2014)

6. Lacy KE, Karagiannis SN, Nestle FO. Advances in the treatment of melanoma. Clin Med. 2012; 12:168-71.

7. Long G, Menzies A, Nagrial A, Haydu L, Hamilton A, Mann G, *et al.* Prognostic and clinicopathologic associations of oncogenic BRAF in metastatic melanoma. J Clin Oncol. 2011; 29:1239-46.

8. Balch C, Gershenwald J, Soong S, Thompson J, Atkins M, Byrd D, *et al.* Final version of 2009 AJCC melanoma staging and classification. J Clin Oncol. 2009; 27:6199-206.

9. Garbe C, Peris K, Hauschild A, Saiag P, Middleton M, Spatz A, *et al.* Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline - Update 2012. Eur J Cancer. 2012; 48:2375-90.

10. Francken AB, Accortt NA, Shaw HM, Wiener M, Soong SJ, Hoekstra HJ, *et al.* Prognosis and determinants of outcome following locoregional or distant recurrence in patients with cutaneous melanoma. Ann Surg Oncol. 2008; 15:1476-84.

11. Reed KB, Cook-Norris RH, Brewer JD. The cutaneous manifestations of metastatic malignant melanoma. Int J Dermatol. 2012; 51:243-9.

12. Carlino MS, Fogarty G, Long GV. Treatment of melanoma brain metastases: A new paradigm. The Cancer Journal. 2012; 18:208-12.

13. Falchook GS, Long GV, Kurzrock R, Kim KB, Arkenau TH, Brown MP, *et al.* Dabrafenib in patients with melanoma, untreated brain metastases, and other solid tumours: A phase 1 dose-escalation trial. Lancet. 2012; 379:1893-901.

14. Grob J, Algarra S, Amonkar M, Demidov LV, Goodman V, Grotzinger, *et al.* Dabrafenib versus dacarbazine (DTIC) in patients with BRAF^{V600+} advanced and metastatic melanoma in BREAK 3: Quality of life analysis. Society for Melanoma Research 2012 Congress; California, USA 2012.

15. Kiebert GM, Jonas DL, Middleton MR. Health-related quality of life in patients with advanced metastatic melanoma: Results of a randomized phase III study comparing temozolomide with dacarbazine. Cancer Invest. 2003; 21:821-9.

16. National Institute for Health and Clinical Excellence. Vemurafenib for the treatment of unresectable locally advanced or metastatic BRAF^{V600} mutation-positive malignant melanoma 2012; Available from: http://www.nice.org.uk/guidance/TA269 (Accessed: 3 July 2014).

17. IMS Oncology Analyzer. Vemurafenib monotherapy usage in 1L BRAF mutation positive stage IV metastatic melanoma patients. Q4 2012.

18. National Institute for Health and Clinical Excellence. Ipilimumab for previously treated advanced (unresectable or metastatic) melanoma. NICE technology appraisal guidance 268. 2012; Available from: www.nice.org.uk/nicemedia/live/14001/61847/61847.pdf (Accessed: 3 June 2014).

19. Menzies AM, Long GV, Murali R. Dabrafenib and its potential for the treatment of metastatic melanoma. Drug Des Devel Ther. 2012; 6:391-405.

20. Walkington LA, Lorigan P, Danson SJ. Advances in the treatment of late stage melanoma. BMJ. 2013; 346:f1265.

21. Srivastava K, Pathak P, Amonkar M, Stapelkamp C, Swann RS, Casey M, *et al.* Systematic review and indirect treatment comparison (ITC) of dabrafenib (Dab) and trametinib (Tram) versus first-line (1L) treatments for metastatic melanoma (MM). J Clin Oncol. 2013; 31:e20012.

22. European Medicines Agency. Tafinlar (dabrafenib) EPAR summary for the public 2013; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/002604/WC500149674.pdf (Accessed: 3 July 2014).

23. Roche. Vemurafenib for the treatment of locally advanced or metastatic BRAF V600 mutation positive malignant melanoma. London: National Institute for Health and Care Excellence; 2012; Available from: http://www.nice.org.uk/guidance/ta269/resources/melanoma-braf-v600-mutation-positive-unresectable-metastatic-vemurafenib-roche-products4 (Accessed: 3 July 2014).

24. National Institute for Health and Care Excellence. NICE proposes ipilimumab as a first treatment for advanced skin cancer (Press Release, 13 June 2014). Available from: http://www.nice.org.uk/News/Press-and-Media/nice-proposes-ipilimumab-as-a-first-treatment-for-advanced-skin-cancer (Accessed: 2 July 2014).

25. Ascierto PA, Minor D, Ribas A, Lebbe C, O'Hagan A, Arya N, *et al.* Phase II trial (BREAK-2) of the BRAF inhibitor dabrafenib (GSK2118436) in patients with metastatic melanoma. J Clin Oncol. 2013; 31:3205-11.

26. Flaherty KT, Infante JR, Daud A, Gonzalez R, Kefford RF, Sosman J, *et al.* Combined BRAF and MEK inhibition in melanoma with $BRAF^{V600}$ mutations. N Engl J Med. 2012; 367:1694-703.

27. GlaxoSmithKline. NCT01584648: A Study Comparing Trametinib and Dabrafenib Combination Therapy to Dabrafenib Monotherapy in Subjects With BRAF-mutant Melanoma. Available from: http://clinicaltrials.gov/show/NCT01584648 (Accessed: 3 June 2014).

28. Long GV, Trefzer U, Davies MA, Kefford RF, Ascierto PA, Chapman PB, *et al.* Dabrafenib in patients with Val600Glu or Val600Lys BRAF-mutant melanoma metastatic to the brain (BREAK-MB): A multicentre, open-label, phase 2 trial. Lancet Oncol. 2012; 13:1087-95.

29. Dickson R, Bagust A, Beale S, Blundell M, Proudlove C. Vemurafenib for the treatment of locally advanced or metastatic BRAF V600 mutation positive malignant melanoma. 2012; Available from: http://www.nice.org.uk/guidance/ta269/resources/melanoma-braf-v600-mutation-positive-unresectable-metastatic-vemurafenib-evidence-review-group-report2 (Accessed: 3 July 2014).

30. GlaxoSmithKline. A Phase III randomized, open-label study comparing GSK2118436 to DTIC in previously untreated subjects with BRAF mutation positive advanced (Stage III) or metastatic (Stage IV) melanoma (CSR). 2011. p. 1-1768.

31. Department of Health. NHS Reference costs guidance 2012-13. London 2013; Available from: https://www.gov.uk/government/publications/reference-costs-guidance-for-2012-13--2 (Accessed: 3 July 2014).

32. Heron Evidence Development. Systematic review in metastatic malignant melanoma, Version 2.0. 2014.

33. National Institute for Health and Clinical Excellence. Specification for manufacturer/sponsor submission of evidence, June 2012. 2012.

34. Eisenhauer E. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer. 2009; 45:228-47.

35. Latimer N, Abrams K. Adjusting for treatment crossover in the BREAK-3 and METRIC clinical trials - Stage 1 feasibility analysis results, February 2013 update: School of Health and Related Research (ScHARR), University of Sheffield 2013.

36. Bucher H, Guyatt G, Griffith L, Walter S. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1991; 50:683-91.

37. Roche. Roche response to first ACD (7 July 2012); Available from: http://www.nice.org.uk/guidance/ta269/resources/melanoma-braf-v600-mutation-positive-unresectable-metastatic-vemurafenib-roche-products2 (Accessed: 2 July 2014).

38. Gray MR, Martin del Campo S, Zhang X, Zhang H, Souza FF, Carson WE, 3rd, *et al.* Metastatic melanoma: lactate dehydrogenase levels and CT imaging findings of tumor devascularization allow accurate prediction of survival in patients treated with bevacizumab. Radiology. 2014; 270:425-34.

39. Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011] The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org (Accessed: 2 July 2014).

40. McArthur G, Hauschild A, Robert C, Haanen JB, Ascierto P, Lee RJ, *et al.* Vemurafenib improves overall survival compared to dacarbazine in advanced $BRAF^{V600E}$-mutated melanoma: Updated survival results from a phase III randomised, open-label, multicentre trial. Eur J Cancer. 2011; 47:14.

41. Dickson R, Boland A, Bagust A, Blundell M, Massey G, Dundar Y, *et al.* Ipilimumab for previously treated unresectable malignant melanoma: A Single Technology Appraisal. 2011; Available from: http://www.nice.org.uk/guidance/ta268/resources/melanoma-stage-iii-or-iv-ipilimumab-evidence-review-group-report3 (Accessed: 3 July 2014).

42. Novartis. Everolimus (Afinitor) in combination with exemestane for the treatment of advanced or metastatic HER2 negative, hormone receptor positive breast cancer after prior endocrine therapy. London: National Institute for Health and Care Excellence; 2013; Available from: http://www.nice.org.uk/guidance/TA295 (Accessed: 3 July 2014).

43. Boehringer Ingelheim. Submission of evidence by Boehringer Ingelheim on the use of afatinib in adult patients with locally advanced or metastatic non-small cell lung cancer (NSCLC) with activating Epidermal Growth Factor Receptor (EGFR) mutation(s). London: National Institute for Health and Care Excellence; 2013; Available from: http://www.nice.org.uk/guidance/ta310/resources/lung-cancer-non-small-cell-egfr-mutation-positive-afatinib-evaluation-report2 (Accessed: 3 July 2014).

44. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, *et al.* Improved survival with vemurafenib in melanoma with BRAF^{V600E} mutation. N Engl J Med. 2011; 364:2507-16.

45. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. 2008; Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf (Accessed: 30 March 2012).

46. Beusterien KM, Szabo SM, Kotapati S, Mukherjee J, Hoos A, Hersey P, *et al.* Societal preference values for advanced melanoma health states in the United Kingdom and Australia. Br J Cancer. 2009; 101:387-9.

47. Bristol-Myers Squibb. Ipilimumab for previously treated unresectable malignant melanoma: Single technology appraisal submission to NICE 2010; Available from: http://www.nice.org.uk/guidance/ta268/resources/melanoma-stage-iii-or-iv-ipilimumab-bristol-myerssquibb2 (Accessed: 3 July 2014).

48. Joint Formulary Committee. British National Formulary. London 2011; Available from: www.bnf.org (Accessed: 3 July 2014).

49. Millar C, P, Hill M, Pulfer A. A service evaluation to compare secondary care resource use between XELOX and FOLFOX-6 regimes in the treatment of metastic colorectal cancer from a UK National Health Service (NHS) perspective. ISPOR; Athens 2008.

50. Curtis L. Unit costs of health and social care 2012. 2012; Available from: http://www.pssru.ac.uk/project-pages/unit-costs/2012/ (Accessed: 7 May 2014).

51. Department of Health. NHS Reference costs guidance. London 2012; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012 (Accessed: 13 November 2013).

52. Department of Health. National schedule of reference costs 2011-12. Department of Health; 2012; Available from: https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012 (Accessed: 3 July 2014).

53. Addicott R, Dewar S. Improving choice at end of life; a descriptive analysis of the impact and costs of the Marie Curie Delivery Choice Programme in Lincolnshire. London: Kings Fund 2008.

54. GlaxoSmithKline. GSK data on file2013.

55. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: An alternative approach. Med Decis Making. 2014; 34.

56. Fleeman N, Bagust A, Beale S, Dwan K, Boland A, Greenhalgh J, *et al.* Afatinib for treating epidermal growth factor receptor mutation positive locally advanced or metastatic non-small cell lung cancer. 2014; Available from: http://www.nice.org.uk/guidance/ta310/resources/lung-cancer-non-small-cell-egfr-mutation-positive-afatinib-evaluation-report2 (Accessed: 3 July 2014).

57. Xing Y, Chang G, Hu C, Askew R, Ross M, Gershenwald J, *et al.* Conditional survival estimates improve over time for patients with advanced melanoma. Cancer. 2010; 116:2234-41.

58. Sacco J, MacBeth F, Bagust A, Clark P. The average body surface area of adult cancer patients in the UK: A multicentre retrospective study. PLoS ONE. 2010; 5:e8933.

59. Commercial Medicines Unit - Department of Health. Electronic Market Information Tool (eMit). 2014; Available from: http://cmu.dh.gov.uk/electronic-market-information-tool-emit/ (Accessed: 3 July 2014).

11 APPENDIX

Table 34: Assessment of risk of bias of BREAK-3 and BRIM-3 conducted by the manufacturer

Critical appraisal criterion	BREAK-3[2]	BRIM-3[1]
Was randomisation carried out appropriately?	Yes. All BRAF^{v600} mutation-positive patients meeting eligibility criteria were randomised 3:1 (open label) to either dabrafenib or dacarbazine. Patients were randomised centrally using the GlaxoSmithKline RAMOS IVRS. Method of randomisation and allocation concealment was adequate.	Randomisation in this study was appropriately carried out. All BRAF^{v600} mutation-positive patients meeting eligibility criteria were randomised 1:1 (open label) to either vemurafenib or dacarbazine.
Was the concealment of treatment allocation adequate?	Treatment allocation was not concealed as this was an open label study.	Treatment allocation was not concealed as this was an open label study
Were the groups similar at the outset of the study in terms of prognostic factors, for example, severity of disease?	Yes. Treatment groups were well balanced at baseline in terms of demographic and disease characteristics (e.g. age, gender, histology, organs involved, number of metastatic sites, ECOG PS, LDH levels).	The groups randomised to vemurafenib and dacarbazine were comparable
Were the care providers, participants and outcome assessors blind to treatment allocation?	BREAK-3[2] was an open label study, however, the study team was blinded from the efficacy data and any review of these data by the study team including statistics and programming was performed on data with scrambled subject numbers to prevent identification of study treatment. Additionally, the independent review of tumour response was blinded to treatment, investigator-assessed response and progression.	This was an open label study; not blinded
Were there any unexpected imbalances in drop-outs between groups? If so, were they explained or adjusted for?	No. Reasons for withdrawal of patients were reported adequately. Patients mainly withdrew for the following reasons: disease progression; death; adverse events; investigator's discretion; subject decision.	There were no unexpected imbalances between the groups
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No. The authors reported all the outcomes as specified in the protocol of the study.	There is no evidence to suggest that more outcomes were measured than reported
Did the analysis include an intention-to-treat analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Efficacy and safety analyses were performed using appropriate ITT and modified ITT populations, respectively. Since the duration of treatment for a given subject depends on efficacy and tolerability, the duration of follow-up will vary between subjects. All available time-to-event data was analysed using appropriate statistical methods; subjects with shorter treatment and follow-up due to the natural history of their disease or medical necessities of the treatment of their disease were not considered to have missing data. Consequently there was no imputation for missing time-to-event data.	The analyses available for BRIM-3[1] include both ITT and analyses which have been adjusted for the impact of crossover of control patients to the vemurafenib arm

ECOG=Eastern Cooperative Oncology Group; ITT=intention-to-treat; IVRS =interactive voice recognition system; LDH=lactate dehydrogenase; PS=performance status