National Institute for Health and Care Excellence

Centre for Health Technology Evaluation

Pro-forma Response

ERG report

Trametinib in combination with dabrafenib for treating advanced (unresectable or metastatic) BRAF V600 mutation-positive melanoma [ID661]

You are asked to check the ERG report from Liverpool Reviews & Implementation Group (LRiG) to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies you must inform NICE by **5pm, 25th February 2016** using the below proforma comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the Evaluation report.

The proforma document should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Summary of factual inaccuracies in the ERG's critique

- The ERG have criticised the validity of the Combi-V study on the basis of the un-blinded, open label design, and have suggested that overall survival data derived from this study is subject to bias. As such, the ERG has rejected the pooling of both Comb-V and Combi-D studies. Overall survival is an objective measure which is not subject to investigator bias, therefore the ERG's criticism is unfounded. (see issue 2)

- The ERG have criticised the relevance of our submission to the NHS on the basis that immunotherapies were not included as comparators. This opinion from the ERG is contrary to the NICE scoping document, in which dabrafenib and vemurafenib were selected as comparators. To suggest that additional comparators should have been included within our submission undermines the entire NICE scoping process. (see issue 4)

- The ERG are proposing to use mean OS data to assess EOL criteria, which if accepted would invalidate prior determinations of EoL medicines (vemurafenib, dabrafenib); we ask the Committee to apply fair and consistent criteria across appraisals (see issue 7)

- The ERG propose a shortcoming in our model as we have not modelled second-line pembrolizumab treatment; we ask the Committee to consider whether this is a reasonable criticism given the timing of pembrolizumab's NICE recommendation and the absence of clinical evidence to inform this sequence (see issue 8)

- The ERG refer to a 'flaw' in the model structure; there is no flaw and the features of the model which Liverpool focus on are those of any and all partitioned survival models; we ask the Committee to consider whether it is reasonable to conclude that partitioned survival models are by definition flawed (see issues 10, 11 and 12)

- The method of extrapolation that Liverpool suggest is inappropriate (i.e. using KM data to an established cut point where patterns in the hazards change) is in fact one that they have suggested as valid in previous appraisals and in publications; we ask the Committee to consider the perversity of this criticism (see issue 14)

- The ERG have criticised overall survival extrapolation in spite of several alternative sources (including projections made by Liverpool themselves in previous appraisals) which validate the overall survival projections in this model (see issue 18)

- The ERG have suggested that patterns of hazards in projections for the BRAF monotherapy arm are implausible; whilst in fact they are entirely consistent with a shifting of the AJCC survival distribution such as that suggested by the ERG in their assessment of vemurafenib (see issue 19)

- The ERG have compared post-treatment discontinuation survival across the arms from COMBI-D and concluded that their comparability is evidence that OS projections from the model are implausible; the limitations of such comparisons are well described in literature and we highlight the factual inaccuracy of drawing conclusions on their basis (see issue 20)

- The ERG have proposed that TTD would be a more accurate basis by which to estimate treatment costs than PFS; a comprehensive analysis of this by Novartis suggests it is factually inaccurate and any concern relating to this is unfounded (see issue 21 and 22)

- Liverpool highlighted errors where there were none:

  - Novartis have correctly applied discounting such that annual discounting is achieved in a model with a weekly cycle (see issue 25)

  - In spite of Liverpool's suggestion to the contrary, Novartis have adjusted AJCC registry data to the M stage distribution across all arms of COMBI-d and COMBI-v combined (see issue 26)

Overall, the quality and tone of the critique of our submission from Liverpool suggests that whilst Novartis has attempted to learn from the history of prior appraisals (including the history of critique / suggestions / alternative approaches proposed by the ERG themselves), it would appear the ERG has not. This impacts unfairly on the manufacturer's case and manifests in several factual inaccuracies as detailed here.

Issue 1 Misleading comparisons regarding BRAF inhibitors

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG has made misleading comparisons regarding BRAF inhibitors with respect to immunotherapies within the following sections of the report: **Page 21: Section 2.2.1** The ERG is aware that the majority of patients treated with a BRAF inhibitor achieve a rapid response; however, the response is often short-lived (approximately 6 months).[17] Response to treatment with ipilimumab may be more durable (survival rates plateau after 2 to 3 years although survival up to 5 years is reported);[17] however, it takes time for a response to be achieved and the proportion of patients who benefit is limited (10%).[18] Pembrolizumab has fewer side effects than ipilimumab and appears to benefit more patients (30 to 40%);[19] however, the data from the clinical trials in which pembrolizumab is compared with ipilimumab[19] and chemotherapy[20] are relatively immature and the long-term outcomes for patients treated with pembrolizumab are	Suggestion re-wording: **Page 21: Section 2.2.1** The ERG is aware that the majority of patients treated with a BRAF inhibitor monotherapy achieve a rapid response; in those who achieve a response, it typically lasts between 6-10 months. However, combination of a BRAF inhibitor with a MEK inhibitor has demonstrated high overall response rates in addition to increased duration of response in clinical trials. ~~Response to treatment with ipilimumab may be more durable (survival rates plateau after 2 to 3 years although survival up to 5 years is reported);17 however, it takes time for a response to be achieved and the proportion of patients who benefit is limited (10%).18 Pembrolizumab has fewer side effects than ipilimumab and appears to benefit more patients (30 to 40%);19 however, the data from the clinical trials in which pembrolizumab is compared with ipilimumab19 and chemotherapy20 are relatively immature and the long-term outcomes for patients treated with pembrolizumab are unclear at present.~~	The ERG are making invalid and unjustified cross trial comparisons using imprecise and inconsistent parameters, which do not take into account the different patient populations enrolled into the relevant studies. The statements made by the ERG are referring specifically to BRAF monotherapies. The “short lived” response referred to by the ERG is based upon a review article which quotes 6-7 months duration of response (not approximately 6 months). The review article references BREAK-3 and BRIM-3, both conducted at a time when treatment with BRAF inhibitors was in it’s infancy. The Overall Response Rates seen in Combi-V were 66% vs. 53%, (March 2015) and the median durations of response were 13.8 vs. 8.5 months for combination therapy vs vemurafenib respectively. The Overall Response Rates seen in Combi-D were 69% vs. 53% (January 2015), and the median durations of response were 12.9 vs. 10.6 months for combination therapy vs dabrafenib respectively. The Combi studies, which represent a	This is not a factual error. Section 2.2.1 is a broad summary of treatment options available to patients with advanced or unresectable BRAF V600 mutation-positive melanoma up to the time of the present appraisal.

unclear at present. **Page 63: Section 4.10** Second, whilst the majority of patients treated with a BRAF inhibitor achieve a rapid response, that response is often short-lived (approximately 6 months).17 In contrast, the proportion of patients who benefit from ipilimumab is limited (10%)18 but for those that do benefit, treatment can be durable (2 to 3 years). The ERG notes that treatment with pembrolizumab is associated with fewer side effects than ipilimumab and appears to benefit more patients (30% to 40%);57 however, longterm outcomes for patients treated with pembrolizumab are at present unclear.	**Page 63: Section 4.10** Second, the majority of patients treated with a BRAF inhibitor monotherapy achieve a rapid response, which typically lasts for 6-10 months (Ref to Combi-V, Combi-D, BREAK-3 & BRIM 3) However, combination of a BRAF inhibitor with a MEK inhibitor has demonstrated high overall response rates in addition to increased duration response in clinical trials. ~~In contrast, the proportion of patients who benefit from ipilimumab is limited (10%)18 but for those that do benefit, treatment can be durable (2 to 3 years). The ERG notes that treatment with pembrolizumab is associated with fewer side effects than ipilimumab and appears to benefit more patients (30% to 40%);57 however, longterm outcomes for patients treated with pembrolizumab are at present unclear~~	more recent and large body of data, show a duration of response for BRAF monotherapy of 8.5 – 10.6 months. The BRAF patient population of the Keynote 006 study, from which we assume the ERG are drawing their parallels, were a very different population from that enrolled into any of the BRAF inhibitor studies (Combi-D, Combi-V, BREAK-3 or BRIM-3). In addition, the statements about immunotherapy response appear to be based on the total population (whether BRAF wild-type or BRAF mutated) and not the specific population of patients with BRAF mutated tumours. The Keynote 006 (Pembrolizumab vs ipilimumab) protocol required the following for BRAF patients: -Patients must have testing for a BRAF mutation prior to study entry. Patients with BRAF V600E mutant melanoma may have received prior BRAF inhibitor therapy as first-line systemic therapy and be eligible for this study as second line treatment. -At the discretion of the investigator, patients with BRAF V600E mutant melanoma who have NOT received a BRAF inhibitor are also eligible for this study as first line treatment if they meet the following additional criteria: · LDH < local ULN · No clinically significant tumor related	

		symptoms in the judgment of the investigator · Absence of rapidly progressing metastatic melanoma in the judgment of the investigator (Robert C, et al. 2015) The Keynote 006 BRAF population had a substantially better prognosis than that enrolled into Combi-D, Combi-V, BREAK-3 or BRIM-3 studies. The comparison put forward by the ERG is therefore invalid and unjust.	

Issue 2 Implied bias and uncertainty from open-label nature of combi-v

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
The ERG's incorrect assumptions have had a significant impact on many aspect s of the ERG report i.e. the pooling of Combi-D and Combi-V studies, cost effectiveness, and end of life criteria. The ERG ought to reconsider and re-submit an amended report. **Page 17: Section 1.8.2** The COMBI-v trial was an open-label, unblinded RCT that did not include a blinded independent assessment of radiological outcomes; the extent to which this may have caused trial results to be biased is unclear	Suggested rewording: **Page 17: Section 1.8.2** The COMBI-v trial was an open-label, unblinded RCT that did not include a blinded independent assessment of radiological outcomes; ~~the extent to which this may have caused trial results to be biased is unclear~~ **Page 67: Section 4.11**	Overall Survival (OS) is the gold standard measure of efficacy (*U.S Department of health and humans services. Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics 2017*) OS is an objective measure which is not affected by a study being placebo-controlled or open label (Gisele Chvetzof et.al *2003).* There will have been no bias in the Combi-V study with respect to the overall survival results, which are used for the pooling and economic models.	This is a matter of opinion, not a factual error. One of the ERG's roles is to highlight areas of uncertainty and the ERG considers that elements of the design of the COMBI-v trial (namely lack of blinding and the absence of a placebo) may have had an effect on patient and clinician behaviour. Furthermore, these limitations may also have had an effect on both objective measures and patient reported outcomes (quality of life). The reasons why there may be

Page 67: Section 4.11 However, the COMBI-d trial was blinded and placebo controlled, whilst the COMBI-v trial was neither blinded nor placebo controlled. The key issue is the extent to which both the lack of blinding and the lack of a placebo may have affected the results from the COMBI-v trial. **Page 68: Section 4.11** Second, the COMBI-d trial was a blinded and placebo controlled trial, whilst the COMBI-v trial was an open-label trial that did not involve a placebo. The methodology employed in the COMBI-d trial is, therefore, considered to be "gold standard", whilst the method used in the COMBI-v trial is recognised as having limitations and potential bias. The ERG considers that pooling data from a high quality blinded trial (COMBI-d) with data from a trial with outcomes that are subject to bias (COMBI-v) may actually reduce the reliability of the available trial results. **Page 69: Section 4.11** The ERG considers that these differences are likely, in part, to be driven by the lack of a blinding and the lack of a placebo in the COMBI-v trial and may have had an impact on the duration of patient OS	~~However, the COMBI-d trial was blinded and placebo controlled, whilst the COMBI-v trial was neither blinded nor placebo controlled. The key issue is the extent to which both the lack of blinding and the lack of a placebo may have affected the results from the COMBI-v trial.~~ **Page 68: Section 4.11** ~~Second, the COMBI-d trial was a blinded and placebo controlled trial, whilst the COMBI-v trial was an open-label trial that did not involve a placebo. The methodology employed in the COMBI-d trial is, therefore, considered to be "gold standard", whilst the method used in the COMBI-v trial is recognised as having limitations and potential bias. The ERG considers that pooling data from a high quality blinded trial (COMBI-d) with data from a trial with outcomes that are subject to bias (COMBI-v) may actually reduce the reliability of the available trial results.~~ **Page 69: Section 4.11**		concerns about the pooling of data from the COMBI-d and COMBI-v trials and the impact this could have on OS and utility values are clearly stated within the ERG report. The company have referred to the Chvetzof[1] study to support their view that OS is not affected by a study being placebo-controlled or open label. The ERG, however, notes that the conclusion of the Chvetzof[1] review conflicts with the company's assertion: *Although our review suggests limited improvement of patients with cancer who are receiving placebos in randomized trials, we strongly support the use of a double-blind, placebo- controlled design for randomized trials whenever this can be done with minimal impact on patients [or on recruitment to the trial]. In an open comparison of BSC plus active treatment with BSC alone, the effects in the active treatment group might be inflated, not only by a true placebo effect but also by biased expectations of the investigator, more encouragement to the patient to report benefit, patient's desire to please, and other influences. A placebo-controlled trial has the advantages of ensuring equal conditions in both arms and of minimizing various types of bias.* The ERG also notes that the

Page 69: Section 4.11 In conclusion, the ERG considers that results from a high quality blinded and placebo controlled trial (COMBI-d) are more likely to show the true effectiveness of a drug than results from an unblinded, non-placebo controlled trial (COMBI-v), or results generated from pooling data from a blinded and placebo controlled trial with those from an unblinded, non-placebo controlled trial. The rationale behind the company's use of pooled OS data in their model may be based on pragmatism. Analyses undertaken by the ERG show that use of pooled OS data in the company model generates cost effectiveness results that make trametinib+dabrafenib appear much more cost effective (compared to either dabrafenib or vemurafenib monotherapy) than those generated using OS data only from the COMBI-d trial. **Page 71: Section 4.14** The company used pooled data from the COMBI-d and the COMBI-v trials in the cost effectiveness analysis. The ERG considers that results from a high quality blinded and placebo controlled trial (COMBI-d) are more likely to show the true effectiveness of a drug than results from an unblinded, non-placebo controlled trial (COMBI-v), or results generated from pooling data from a blinded and placebo controlled trial with those from an unblinded, non-placebo controlled trial.	~~The ERG considers that these differences are likely, in part, to be driven by the lack of a blinding and the lack of a placebo in the COMBI-v trial and may have had an impact on the duration of patient OS~~ **Page 69: Section 4.11** ~~In conclusion, the ERG considers that results from a high quality blinded and placebo controlled trial (COMBI-d) are more likely to show the true effectiveness of a drug than results from an unblinded, non-placebo controlled trial (COMBI-v), or results generated from pooling data from a blinded and placebo controlled trial with those from an unblinded, non-placebo controlled trial. The rationale behind the company's use of pooled OS data in their model may be based on pragmatism. Analyses undertaken by the ERG show that use of pooled OS data in the company model generates cost effectiveness results that make trametinib+dabrafenib appear much more cost effective (compared to~~		company has ignored the finding, reported by Chvetzof,[1] that double blinded placebo controlled trials should be used to mitigate a range of potential causes of bias – not just the placebo effect. The ERG does not consider that the Chvetzof[1] review invalidates those analyses undertaken by the ERG that only employed data from the COMBI-d trial. In fact, the review supports the view that placebo controlled blinded trials are important and these design elements can influence trial results. The ERG has provided the AC with cost effectiveness results based on use of only COMBI-d trial data as well as results based on using pooled data.

Page 92: Section 5.5.3: Pooling of OS data from COMBI-d and COMBI-v trials	~~either dabrafenib or vemurafenib monotherapy)~~ ~~than those generated using OS data only from the COMBI-d trial.~~ **Page 71: Section 4.14** ~~The company used pooled data from the COMBI-d and the COMBI-v trials in the cost effectiveness analysis. The ERG considers that results from a high quality blinded and placebo controlled trial (COMBI-d) are more likely to show the true effectiveness of a drug than results from an unblinded, non-placebo controlled trial (COMBI-v), or results generated from pooling data from a blinded and placebo controlled trial with those from an unblinded, non-placebo controlled trial.~~ **Page 92: Section 5.5.3: Pooling of OS data from COMBI-d and COMBI-v trials** Entire section to be deleted.		

Issue 3 Relevance of the clinical trial population to UK practice

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG incorrectly state that patients with brain metastases were not included in the Combi-D and Combi-V studies: **Page 9: Section 1.1** The extent to which the submitted evidence reflects clinical practice is limited by the fact that patients included in the COMBI-d and COMBI-v trials all had an Eastern Oncology Cooperative Group (ECOG) performance status (PS) of 0 or 1 and no brain metastases. Clinical advice to the ERG is that both patients with an ECOG PS>1 and those with brain metastases are treated in UK clinical practice. **Page 17: Section 1.8.2** There is no clinical evidence to support the use of trametinib+dabrafenib in patients with a PS of 2 or higher, or for patients with brain metastases. **Page 9: Section 1.1** The findings from the COMBI-d and COMBI-v trials are only relevant to the cohort of patients in clinical practice with BRAF mutation-positive disease who have an ECOG PS 0-1 and no	Suggested re-wording **Page 9: Section 1.1** The extent to which the submitted evidence reflects clinical practice is limited by the fact that patients included in the COMBI-d and COMBI-v trials all had an Eastern Oncology Cooperative Group (ECOG) performance status (PS) of 0 or 1 ~~and no brain metastases~~. Clinical advice to the ERG is that ~~both~~ patients with an ECOG PS>1 ~~and those with brain metastases~~ are treated in UK clinical practice. **Page 17: Section 1.8.2** There is no clinical evidence to support the use of trametinib+dabrafenib in patients with a PS of 2 or higher, or for patients with <u>active</u> brain metastases. **Page 9: Section 1.1** The findings from the COMBI-d and COMBI-v trials are only relevant to the cohort of patients in clinical practice with BRAF mutation-positive disease who have an ECOG PS 0-1 and <u>stable</u> brain metastases.	Patients with brain metastases were enrolled within the Combi-D and Combi-V clinical trials. Per the exclusion criterion in the protocol: Brain lesion(s), if still present, must be confirmed stable (i.e., no increase in lesion size) for ≥ 12 weeks prior to randomization (*Ref: Comb-V and Combi-d protocols*).	The ERG report is amended (pages 9, 17, 28 and 43) to reflect the fact that patients with stable brain metastases were eligible for inclusion in the COMB-d and COMBI-v trials although fewer than 20 patients with stable brain metastases were included across the trials. Page 9: The extent to which the submitted evidence reflects clinical practice is limited by the fact that patients included in the COMBI-d and COMBI-v trials all had an Eastern Oncology Cooperative Group (ECOG) performance status (PS) of 0 or 1 and fewer than 20 patients had stable brain metastases. Clinical advice to the ERG is that both patients with an ECOG PS>1 and those with brain metastases are treated in UK clinical practice. The findings from the COMBI-d and COMBI-v trials are only relevant to the cohort of patients in clinical practice with BRAF mutation-positive disease who have an ECOG PS 0-1. Clinical advice to the ERG is that some of these patients could have been treated with an immunotherapy

brain metastases. **Page 28: Section 3.1** Clinical advice to the ERG is that patients with an ECOG PS of 2 or higher and patients with brain metastases are treated in UK clinical practice. **Page 43 Section 4.4.3** The ERG notes that the patients recruited to the COMBI-d and COMBI-v trials were limited to those with an ECOG PS of 0 or 1; an ECOG PS of 0 or 1 suggests that the patients in the trials were relatively well at the time of enrolment. The presence of brain metastases was also an exclusion factor in both trials. Clinical advice to the ERG is that patients with ECOG PS of 2, 3 and 4 are treated in the NHS and patients with brain metastases are also treated in the NHS.	**Page 28: Section 3.1** Clinical advice to the ERG is that patients with an ECOG PS of 2 or higher and patients with active brain metastases are treated in UK clinical practice. **Page 43 Section 4.4.3** The ERG notes that the patients recruited to the COMBI-d and COMBI-v trials were limited to those with an ECOG PS of 0 or 1; an ECOG PS of 0 or 1 suggests that the patients in the trials were relatively well at the time of enrolment. The presence of active brain metastases was also an exclusion factor in both trials. Clinical advice to the ERG is that patients with ECOG PS of 2, 3 and 4 are treated in the NHS and patients with active brain metastases are also treated in the NHS.		instead of a BRAF inhibitor. It is not known whether trametinib+dabrafenib is more effective than an immunotherapy in this group of patients as there are no data available to allow this comparison to be undertaken. Page 17 There is no clinical evidence to support the use of trametinib+dabrafenib in patients with a PS ≥2 and the clinical trials included fewer than 20 people with a stable brain metastases Page 28 The ERG cautions that the inclusion and exclusion criteria for the COMBI-d and COMBI-v trials meant that only patients with an Eastern Co-operative Oncology Group (ECOG) performance status (PS) of 0 or 1 and stable brain metastases (n=20) were recruited. There is, therefore, no clinical evidence to support the use of trametinib+dabrafenib in patients with an ECOG PS of 2 or higher, or for patients with brain metastases. Clinical advice to the ERG is that patients with an ECOG PS of 2 or higher and patients with brain metastases are treated in UK clinical practice. The SmPC[2,3] explicitly states that the clinical effectiveness of

			trametinib+dabrafenib in people with brain metastases is currently unknown. Page 43 The ERG notes that the patients recruited to the COMBI-d and COMBI-v trials were limited to those with an ECOG PS of 0 or 1; an ECOG PS of 0 or 1 suggests that the patients in the trials were relatively well at the time of enrolment. The presence of active brain metastases was also an exclusion factor in both trials. Clinical advice to the ERG is that patients with ECOG PS of 2, 3 and 4 are treated in the NHS and patients with brain metastases are also treated in the NHS.

Issue 4 Immunotherapies as relevant comparators

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
The ERG have made factual inaccuracies within the report when discussing the relevant comparator for dabrafenib + trametinib combination therapy: **Page 17: Section 1.8.2** In current clinical practice, ipilimumab and pembrolizumab are treatment options recommended by NICE for patients with unresectable or advanced BRAF V600 mutation-positive melanoma. The ERG considers that the absence of these comparators limits the relevance of this appraisal to the NHS. **Page 23: Section 142** The ERG considers the company's description to be accurate at the time the CS was written, with the caveat that the ERG is concerned that ipilimumab and pembrolizumab are not specified as comparators for this appraisal by NICE as they have been considered as comparators in previous appraisals and, consequently, they are not considered by the company **Page 25: Section 2.2.2** However, the ERG considers that the	Suggested re-wording: **Page 17: Section 1.8.2** In current clinical practice, ipilimumab and pembrolizumab are treatment options recommended by NICE for patients with unresectable or advanced BRAF V600 mutation-positive melanoma. However, the current standard of care for patients with BRAFV600 positive melanoma is dabrafenib or vemurafenib monotherapy.~~The ERG considers that the absence of these comparators limits the relevance of this appraisal to the NHS.~~ **Page 23: Section 142** The ERG considers the company's description to be accurate at the time the CS was written. ~~with the caveat that the ERG is concerned that ipilimumab and pembrolizumab are not specified as comparators for this appraisal by NICE as they have been considered as comparators in previous appraisals and, consequently, they are not considered by the company~~ **Page 25: Section 2.2.2**	BRAF monotherapy (dabrafenib or vemurafenib) was and is considered standard of care in this patient population. They were selected as the relevant comparators in the NICE scopes for both trametinib + dabrafenib, and cobimetinib + vemurafenib (August 2015) submissions. The ERG was fully involved in both scoping processes, and should understand the reasons why immunotherapies were omitted. To suggest inclusion of additional comparators i.e. immunotherapies at this stage undermines the NICE scoping process. We would like to draw attention to the following statement within the final NICE scope: "In clinical practice, for people with BRAF mutation positive advanced melanoma, <u>a BRAF inhibitor is the usual first line treatment</u>; ipilimumab may be considered for first line use in a subgroup of patients who are relatively well and in whom the disease is not progressing rapidly". Pembrolizumab received NICE approval in November 2015.	This is not a factual error but a matter of opinion. The ERG appreciates the company's position. However, the treatment pathway for melanoma is a rapidly changing field and treatment options have changed since the introduction of pembrolizumab and nivolumab into the NHS. The ERG considers it important to raise this as a potential issue that has arisen between the time of publication of the NICE final scope and the time of the appraisal. The possibility of an indirect treatment comparison between trametinib+dabrafenib is explored in the ERG report. However, the ERG has noted in the report that the comparison is unlikely to produce reliable results.

introduction of pembrolizumab into the NHS is currently changing the treatment pathway for patients with BRAF V600 mutation-positive unresectable or metastatic melanoma and is a relevant comparator in this appraisal. **Page 9: Section 1.1** The ERG notes that there is uncertainty around whether pembrolizumab and ipilimumab should have been specified as comparators in the final NICE scope. **Page 29: Section 3.2** In summary, the ERG considers that the rationale is not clear for not including ipilimumab and pembrolizumab to be relevant comparators to trametinib+dabrafenib in this appraisal for patients with BRAF mutation positive melanoma who are relatively well and whose disease is not progressing. **Page 29: Section 3.3** In summary, the ERG considers that the rationale is not clear for not including ipilimumab and pembrolizumab to be relevant comparators to trametinib+dabrafenib in this appraisal for patients with BRAF mutationpositive melanoma who are relatively well and whose disease is not progressing.	However, the ERG considers that the introduction of pembrolizumab into the NHS is currently changing the treatment pathway for patients with BRAF V600 mutation-positive unresectable or metastatic melanoma ~~and is a relevant comparator in this appraisal.~~ Suggest to remove the following statements form the report: **Page 9: Section 1.1** ~~The ERG notes that there is uncertainty around whether pembrolizumab and ipilimumab should have been specified as comparators in the final NICE scope~~. **Page 29: Section 3.2** ~~In summary, the ERG considers that the rationale is not clear for not including ipilimumab and pembrolizumab to be relevant comparators to trametinib+dabrafenib in this appraisal for patients with BRAF mutationpositive melanoma who are relatively well and whose disease is not progressing.~~ **Page 29: Section 3.3** ~~In summary, the ERG considers that the rationale is not clear for not including ipilimumab and pembrolizumab to be relevant comparators to trametinib+dabrafenib in this appraisal for patients with BRAF~~	Pembrolizumab was not the standard of care at the time of this submission. The ERG has suggested an indirect treatment comparison via a network (ERG Page 64), with the caveat that such comparisons would have been unreliable. If a network had been provided within the submission it would have been inherently flawed with limitations, which would have been subsequently rejected by the ERG.	

Page 25: Section 2.2.2 The ERG accepts that ipilimumab was not considered as a comparator in the appraisal under discussion; however, the rationale for the decision is unclear.	~~mutationpositive melanoma who are relatively well and whose disease is not progressing~~. Suggested amendments: **Page 25: Section 2.2.2** The ERG accepts that ipilimumab was not considered as a comparator in the appraisal under discussion~~; however, the rationale for the decision is unclear.~~ .		

Issue 5 Incorrect reference to study

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 52: Section 4.7.2 At the March 2015 data cut, median PFS was 12.6 months in the trametinib+dabrafenib arm and 7.3 months in the vemurafenib monotherapy arm, a difference of 5.3 months. The risk reduction associated with trametinib+dabrafenib was 39% (HR=0.61, 95% CI: 0.51 to 0.73, p<0.001). The company (CS, p102) emphasises that the 2015 data analysis was not preplanned and that the p-value was not adjusted for multiple testing.	Suggested re-wording **Page 52: Section 4.7.2** At the March 2015 data cut, median PFS was 12.6 months in the trametinib+dabrafenib arm and 7.3 months in the vemurafenib monotherapy arm, a difference of 5.3 months. The risk reduction associated with trametinib+dabrafenib was 39% (HR=0.61, 95% CI: 0.51 to 0.73, p<0.001). ~~The company (CS, p102) emphasises that the 2015 data analysis was not preplanned and that the p-value was not adjusted for multiple testing~~	CS Page 102 refers to the Combi-D study, not Combi-V. The statement by the ERG: “The company (CS, p102) emphasises that the 2015 data analysis was not pre-planned and that the p-value was not adjusted for multiple testing” is incorrect.	This is a factual error and the statement has been removed from the ERG report.

Issue 6 Incorrect assertion that the combination patient population could have received an immunotherapy

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Page 17: Section 1.8.2 The treatment sequence of BRAF inhibitors and immunotherapy in patients with BRAF mutation-positive melanoma is uncertain. Patients recruited to the COMBI-d and COMBI-v trials appear to be fitter patients who, in clinical practice, could have received an immunotherapy (pembrolizumab or ipilimumab) as a first-line treatment	**Page 17: Section 1.8.2** The treatment sequence of BRAF inhibitors and immunotherapy in patients with BRAF mutation-positive melanoma is uncertain. Patients recruited to the COMBI-d and COMBI-v trials appear to be fitter patients who, in clinical practice, some of whom could have received an immunotherapy (pembrolizumab or ipilimumab) as a first-line treatment.	We would like to draw particular attention to the following statement within the final NICE scope: "In clinical practice, for people with BRAF mutation positive advanced melanoma, a BRAF inhibitor is the usual first line treatment; ipilimumab may be considered for first line use in a subgroup of patients who are relatively well and in whom the disease is not progressing rapidly". The patient population in the Combi studies included patients with high LDH, stage IV M1c and rapidly progressing disease. Most of this patient population would not be considered for ipilimumab and some would not be considered appropriate for pembrolizumab, particularly those with high tumour burden, heavy organ involvement and rapidly-progressing / life-threatening disease.	This is not a factual error. As noted previously, the treatment of melanoma is a rapidly changing field and treatment options have changed. It is not clear what proportion of patients from the COMBI-d and COMBI-v trials would be considered for treatment with recently recommended immunotherapies such as pembrolizumab or nivolumab.

Issue 7 End of life

<table>
<tr><th>Description of problem</th><th>Description of proposed amendment</th><th>Justification for amendment</th><th>ERG Response</th></tr>
<tr>
<td>At page 11, it is stated that: “[…] however, when mean OS is considered, treatment with a BRAF inhibitor monotherapy may increase patient life expectancy to about 3 years”.

At page 118 of the ERG report, it is stated that “[…] However, the ERG notes that in the company’s economic analyses mean OS (life expectancy) for patients receiving BRAF inhibitor monotherapy is estimated to be about 3 years. Using the ERG’s preferred approach to OS modelling, the ERG calculates a mean</td>
<td>Delete these statements from this section and only report median OS for dabrafenib and vemurafenib to assess whether Taf+Mek does meet the EoL criteria, that is the first row in table 38 of the submission, reproduced below:

Table 38 Company End of Life criteria
<table><tr><th>Criterion</th><th>Data available</th></tr><tr><td>The treatment is indicated for patients with a short life expectancy, normally less than 24 months</td><td>Median OS data from BRAF inhibitor monotherapy arms of RCTs
Dabrafenib
BREAK-3=20.1 months
COMBI-d=18.7 months
Vemurafenib
BRIM-3=13.6 months
COMBI-v=18 months</td></tr></table></td>
<td>When assessing the first criterion to evaluate whether the combination of dabrafenib+trametinib meets the end-of-life criteria (that is, “The treatment is indicated for patients with a short life expectancy, normally less than 24 months”), the ERG accepts that, on the basis of the <u>median</u> OS data presented by Novartis in the submission, the life expectancy on dabrafenib and vemurafenib is 20.1 months (from BREAK-3) and 18.7 months (from combi-d). For vemurafenib, the date presented by Novartis report a median OS of 13.6 months (from BRIM-3) and 18 months (from combi-v). However, the ERG then seems to questions the approach based on the consideration of median OS, and they refer to the mean OS for dabrafenib and vemurafenib generated in the CE model, thus suggesting to use another element on the basis of which to assess whether EoL should to be granted. This is in clear contradiction with the approach that the same ERG has</td>
<td>This is matter of opinion, not a factual error.

The supplementary advice,[4] issued to ACs, which should be taken into account when appraising ‘End of Life treatments’ does not specify whether a mean value or a median value should be used when determining life expectancy. In the past, the practice has been to use the median value. The ERG, however, considers that as the difference between median and mean survival for the patient population considered in this appraisal is so large, the AC should be informed of the magnitude of the difference. Whether the AC has been provided with such information in the past is not relevant.</td>
</tr>
</table>

OS (life expectancy) to be about 4 years."		adopted in TA366, where only median OS was considered to assess the life expectancy for patients on dabrafenib and vemurafenib, despite the fact that the mean OS for these two treatments were above 2 years. For consistency, the same approach needs to be used for the purpose of this submission and therefore accept that Taf+Mek does meet the end-of-life criteria.	

Issue 8 Patient pathway

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 90: This failure to incorporate second-line treatment with pembrolizumab in the company model casts doubt on the OS estimates presented by the company (and those produced by the ERG). Given the information now available on the efficacy and tolerability of pembrolizumab, as well as the NICE recommendations for its use, the ERG considers that it is likely that if patients in the COMBI-d and COMBI-v trials had been presented with the option of receiving pembrolizumab at the time of disease progression then more patients would have chosen to switch from their study treatment to receive pembrolizumab. The potential impact of pembrolizumab on patients' choice to switch treatment at the time of	Page 90: ~~This failure to incorporate second-line treatment with pembrolizumab in the company model casts doubt on the OS estimates presented by the company (and those produced by the ERG).~~ Given the information now available on the efficacy and tolerability of pembrolizumab, as well as the NICE recommendations for its use, the ERG considers that it might be possible ~~is likely~~ that if patients in the COMBI-d and COMBI-v trials had been presented with the option of receiving pembrolizumab at the time of disease progression then more patients would have chosen to switch from their study treatment to receive pembrolizumab. ~~The potential impact of pembrolizumab on patients' choice to switch treatment at the time of disease progression and the resultant impact on survival should be considered when interpreting the model survival~~	Novartis did not "fail" to incorporate second-line treatment with pembrolizumab after disease progression and treatment discontinuation on the combination, for the simple fact that <u>there are no clinical data</u> that could have possibly informed this analysis in any meaningful way. At the time of the submission, pembrolizumab had just been approved and cannot be considered SOC for BRAF+ patients. The 7% of patients who went on to receive pembro as 2nd line treatment have been taken into account - duly so, reflecting the evidence from the trials. Accounting for the use of pembro as 2nd line treatment would	This is a matter of opinion, not a factual error. One of the ERG's roles is to highlight areas of uncertainty. The ERG considers that whilst modelling the effect of treatment with pembrolizumab would not have been straightforward, its use should have been discussed by the company in their submission as a potential area of uncertainty.

disease progression and the resultant impact on survival should be considered when interpreting the model survival and economic results.	~~and economic results.~~	have required a number of untested, unproven assumptions on the efficacy of that treatment after the combination is used that the ERG for first would have certainly critized and not accepted. As a consequence, the modelling of OS by the company should be seen as conservative. In addition, pembro has only been approved by NICE (TA366) in December 2015; its use post BRAF-inhibitors is not known; there is no clinical evidence available to substantiate the ERG's claim at this point in time.	

Issue 9 Pooling of combi-d and combi-v

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
In terms of trametinib+dabrafenib therapy, the K-M OS curves are almost identical in both the COMBI-d and COMBI-v trials (Log-rank statistic: -0.0004, p=0.984). This means that there is likely to be very little difference in the results derived from an assessment of AUC irrespective of whether the data are pooled or whether data are used from either the COMBI-d trial or the COMBI-v trial. The ERG, therefore, completely rejects the argument that pooling the data increases the accuracy of the company's AUC analysis.	In terms of trametinib+dabrafenib therapy, the K-M OS curves are almost identical in both the COMBI-d and COMBI-v trials (Log-rank statistic: -0.0004, p=0.984). This means that there is likely to be very little difference in the results derived from an assessment of AUC irrespective of whether the data are pooled or whether data are used from either the COMBI-d trial or the COMBI-v trial. ~~The ERG, therefore, completely rejects the argument that pooling the data increases the accuracy of the company's AUC analysis.~~	In providing a rationale for pooling, the manufacturer stated that pooling the data from the two trials would improve the *precision* of the estimates. Contrary to what is suggested by the ERG, the manufacturer did not state that pooling would improve the *accuracy* of the estimates. Accuracy refers to the degree to which there is any systematic bias in the estimate. Precision refers to the extent of statistical uncertainty in the estimates and can be improved by increasing the sample size upon which the estimates are derived.	This is not a factual error The ERG recognises that the word 'precision' has been correctly used within the CS. However, accuracy is the key issue of concern to the ERG - introducing potentially biased data does not improve the accuracy of a data-set.

Issue 10 Model structure (I)

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 87: The ERG identified a flaw in the model design that casts doubt on the size of any of the ICERs per QALY gained generated by the model, both in the company base case and after ERG amendments. Rectifying the flaw would require extensive reconstruction, a task which is beyond the remit of the ERG.	Page 87: ~~The ERG identified a flaw in the model design that casts doubt on the size of any of the ICERs per QALY gained generated by the model, both in the company base case and after ERG amendments. Rectifying the flaw would require extensive reconstruction, a task which is beyond the remit of the ERG.~~	We dispute the statement that there is a "flaw in the model design that cast doubts on the model results". The model uses a partitioned survival analysis approach. The flaw in the model cited by the ERG is a limitation that is inherent in any partitioned survival model. While there are limitations with this modelling approach, it has been used in numerous prior evaluations and has been considered acceptable for evaluating ICERs in these evaluations. Additionally, the ERG provides no evidence to support the statement that this structure "that casts doubt on the size of any of the ICERs per QALY gained generated by the model".	This is a matter of opinion and not a factual error. In the ERG report that formed part of the NICE appraisal of pembrolizumab for treating advanced melanoma in people previously untreated with ipilimumab (TA366[5]) the ERG made it clear that they had had concerns about the reliability of the submitted three state model, namely that the structure of the model meant that the mortality rate was not influenced by whether a patient was in the pre-progression or post-progression state. Even if these issues had not been raised in that report, past failure to identify flaws is not grounds to ignore flaws once they have been identified. Where increased effectiveness results in increased life expectancy an increase in costs is not necessarily an issue. However, in this case, the increase in effectiveness (i.e. PFS) does not lead to an increase in total life expectancy. The

			detailed description of the ERG concerns presented in the ERG report does not need repeating here, but the concerns are sufficient for the ERG to consider that an alternative approach to modelling rather than the three state model used by the company (and other companies in the past) should be considered in future submissions. The ERG has, however, changed the paragraph on page 89 to the following: Effectively, the model structure makes it impossible to directly explore how changes in the assumption of the effectiveness of treatment with trametinib+dabrafenib at delaying or stopping progression impact on the size of the ICERs. The ERG has partially addressed this issue by using time to treatment discontinuation (TTD) data rather than PFS as the basis for estimating costs and utilities, but the problem still essentially remains.

Issue 11 Model structure (II)

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 88: The ERG considers that an appropriate model structure would generate cost effectiveness results that show that a reduction in the efficacy of the intervention leads to an increase in the size of the ICER in respect of the comparator.	Page 88: ~~The ERG considers that an appropriate model structure would generate cost effectiveness results that show that a reduction in the efficacy of the intervention leads to an increase in the size of the ICER in respect of the comparator~~.	We dispute the ERG statement that "an appropriate model would generate cost effectiveness results that show that a reduction in the efficacy of the intervention leads to an increase in the size of the ICER in respect of the comparator". If increasing effectiveness is also associated with increasing costs, the impact on the ICER should depend on the magnitude of the increase in costs relative to the magnitude in the increase in effectiveness. As the ERG has explained, in the model, increases in PFS are projected to increase both QALYs and the costs of treatment. This result is reasonable given that treatment duration and hence treatment costs are largely determined by time to progression. Because the proportionate increase in costs is greater than the proportionate increase in QALYs, the ICER becomes less favourable with increasing PFS. While the ERG may find these results counterintuitive, they can be easily explained, as the ERG has done in their report. It should be noted that one could equally argue that this "counterintuitive" finding is a consequence structure of the partitioned survival analysis model, in which changes in PFS do not result in changes in OS. If the model were structured such that changes in PFS were associated with corresponding increases in OS (rather than reductions in post-progression survival), the finding that the ERG expects would be obtained. However, in partitioned survival models, PFS and OS are not linked. While this must be recognized as a limitation of partitioned survival models, it is not unique to the model used in this evaluation.	This is a matter of opinion and not a factual error. Please see ERG response to issue 10

Issue 12 Model structure (III)

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 89: Effectively, the model structure makes it impossible to explore how changes in the assumption of the effectiveness of treatment with trametinib+dabrafenib at delaying or stopping progression impact on the size of the ICERs. The ERG has partially addressed this issue by using time to treatment discontinuation (TTD) data rather than PFS as the basis for estimating costs and utilities, but the problem still essentially remains.	Page 89: ~~Effectively, the model structure makes it impossible to explore how changes in the assumption of the effectiveness of treatment with trametinib+dabrafenib at delaying or stopping progression impact on the size of the ICERs.~~ The ERG has ~~partially~~ explored a scenario that uses ~~addressed this issue by using~~ time to treatment discontinuation (TTD) data rather than PFS as the basis for estimating costs and utilities. ~~, but the problem still essentially remains.~~	This statement is incorrect. It *is* possible to explore how changes in the assumption of the effectiveness of treatment with trametinib+dabrafenib at delaying or stopping progression impact on the size of the ICERs. It is correct to state that the model is not structured to *directly* explore how changes in the assumption of the effectiveness of treatment with trametinib+dabrafenib at delaying or stopping progression impact on the size of the ICERs *under the assumption that changes in PFS are not associated with changes in treatment costs.* However, it is feasible to use the model to explore such a scenario. For example, one could run the model using two different sets of PFS assumptions. One could then use the estimated QALYs from this analysis with the estimated costs of treatment from the base case.	This is a matter of opinion and not a factual error. Please see ERG response to issue 10

Issue 13 AJCC curve and time since diagnosis

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 90: The survival curves generated from the AJCC data cover the period from the time of initial melanoma diagnosis and reflect staging at time of diagnosis. To use AJCC data to generate a survival curve for a specific	Page 90: The survival curves generated from the AJCC data cover the period from the time of initial melanoma diagnosis and reflect staging at time of diagnosis. To use AJCC data	The 1.1% monthly hazard in year 5 of the AJCC registry corresponds to a 12% annual mortality rate. In the DTIC+placebo arm of the CA184-024 trial of ipilimumab in treatment-naïve patients, which should be more comparable to the COMBI-d and COMBI-v trials in terms of time since diagnosis, the	This is a matter of opinion, not a factual error. As the company states in issue 13, you cannot use alternative trial data to estimate survival at 5 years without adjusting for case mix and also taking

population, the data must be adjusted by the time since diagnosis and melanoma stage. Whilst the company adjusted for the case mix (age, sex and stage) of patients included in the COMBI-d and COMBI-v trials, no adjustment was made for the time since diagnosis.	to generate a survival curve for a specific population, the data must be adjusted by the time since diagnosis and melanoma stage. ~~Whilst the~~ company adjusted for the case mix (age, sex and stage) of patients included in the COMBI-d and COMBI-v trials. ~~, no adjustment was made for the time since diagnosis.~~	annual mortality rate in year 5 post-randomization was 9%. This data provide support to the ERG's contention that that the use of the AJCC data without adjustment may lead to overestimating of the mortality rate in both treatment arms. However, such overestimation is likely to result in a conservative bias to estimates of cost-effectiveness. The ERG's failure to adjust their estimates for this known bias that would favor the experimental treatment raises the possibility of selectivity in parameter estimation.	into account the other treatments available to patients at the time of the trials. As such, the simple comparison made by the company in issues 13 and 18 to results from other trials has to be treated with caution. Essentially, the company seems to be saying that there is no issue with BRAF inhibitor therapy having lower OS at 5 years than having only BSC as is the case in the AJCC registry data. The ERG considers that this is a significant issue and casts doubt on the OS projection for the BRAF inhibitors. The ERG accepts that in ERG report that formed part of the NICE appraisal of vemurafenib for the treatment of locally advanced or metastatic BRAF V600 mutation positive malignant melanoma (TA26[96]), the ERG projection of OS did effectively meet the AJCC data at 5 years. However, at that time the data available did not take into account the second-line treatments that were available (notably immunotherapy) to patients participating in the COMBI-d and COMBI-v trials.

Issue 14 K-M curves in CE model

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 90: The ERG considers that, unless there are compelling reasons not to use all of the data available from a trial, it is preferable to incorporate the trial evidence as it is rather than extrapolate over the period for which data are available. As such, the ERG amended the model by using the K-M data directly (but separately) from the two arms of the COMBI-d trial over the first 27 months. After 27 months there were no further deaths recorded in either arm of both trials, the only events being censoring events.	Page 90: ~~The ERG considers that, unless there are compelling reasons not to use all of the data available from a trial, it is preferable to incorporate the trial evidence as it is rather than extrapolate over the period for which data are available. As such, the ERG amended the model by using the K-M data directly (but separately) from the two arms of the COMBI-d trial over the first 27 months. After 27 months there were no further deaths recorded in either arm of both trials, the only events being censoring events.~~	The approach used in the manufacturer's model is consistent with that suggested by Bagust and Beale (2013) in their recommendations for estimating OS in economic evaluations of cancer therapies which states "*If a long-term constant hazard trend is indicated, then an exponential model may be fitted to the remaining data after left truncation of the data set at the point when the trend is clearly established.*" The resulting survival distribution using this approach very closely approximates the Kaplan Meier curve beyond the cut point. We believe there is a compelling reason to use this approach. Specifically, this approach employs an objective criteria for selecting the point at which to transition from the empirical data to the projection rather than an arbitrarily chosen cut-point as employed by the ERG. As the choice of the cut point can have a large impact on the projection, the use of an objective rather than a subjective method is essential to avoid "cherry picking" the cut-point to influence the findings. While the ERG makes the decision to terminate the Kaplan Meier curve at 27 months because "there are no additional events after this time point", the ERG provides no theoretical or empirical basis to support this criteria.	This is a matter of opinion, not a factual error. The ERG did not 'arbitrarily' choose a cut-point, rather the ERG used the Kaplan-Meier (K-M) data for the period up to the point at which the only events were due censoring, i.e. no further deaths occurred. The ERG notes that, within the company model, hazard rates for the intervention and comparator arms were merged from 30 months and no justification for this decision was provided in the CS. In general, the ERG's preferred approach to modelling survival is to use individual patient data (IPD) rather than a projection. As stated by the company, over the first 27 months of modelled OS there was little difference between trial results and modelled estimates. However, the mortality rate used beyond the period for which trial data are available is important, especially as, at this point, monthly mortality rates were declining in the BRAF inhibitor monotherapy arms of both the COMBI-d and COMBI-v trials.

Issue 15 Placebo effect in combi-v trial

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 92: The COMBI-d trial was a blinded and placebo controlled trial, whilst the COMBI-v trial was an open-label trial that did not involve a placebo. The methodology employed in the COMBI-d trial is, therefore, considered to be “gold standard”, whilst the method used in the COMBI-v trial is recognised as having limitations and being prone to potential biases. For example, the placebo effect from trametinib+dabrafenib therapy was not accounted for in the COMBI-v trial, whereas it was accounted for in the COMBI-d trial	Page 92: The COMBI-d trial was a blinded and placebo controlled trial, whilst the COMBI-v trial was an open-label trial that did not involve a placebo. The methodology employed in the COMBI-d trial is, therefore, considered to be “gold standard”. ~~whilst the method used in the COMBI-v trial is recognised as having limitations and being prone to potential biases. For example, the placebo effect from trametinib+dabrafenib therapy was not accounted for in the COMBI-v trial, whereas it was accounted for in the COMBI-d trial~~	While it is true that the COMBI-v was not placebo-controlled, the ERG has provided no evidence that, in trials of cancer therapies, outcomes such as OS, PFS, TTD, incidence of AEs, or health-related quality of life based on the EQ-5D index value are influenced by the placebo-effect. We know of no data supporting the hypothesis that the placebo effect has any material impact on progression free or overall survival for cancer therapies. Hróbjartsson and Gøtzsche (2005) conducted a systematic review of 156 randomized clinical trials (representing 11 737 patients) that had compared placebo-treated with untreated patients and concluded that there was no evidence of a generally large effect of placebo interventions.” Chvetzoff and Tannock (2003) reviewed the responses of patients on placebo or best supportive care in 47 randomized trials in a variety of tumors to assess placebo effects in oncology and concluded that “placebos are sometimes associated with improved control of symptoms such as pain and appetite but rarely with positive tumor response. Substantial improvements in symptoms and quality of life are unlikely to be due to placebo effects.” Although the COMBI-v trial also was unblinded and was therefore subject to potential observer bias, this bias would not have impacted outcomes such as overall survival, which is objective and not subject to investigator interpretation, or measures of health-related quality of life, which were assessed by the study subjects and not the investigators. While PFS was not assessed by blinded independent review committee (BIRC) in COMBI-v,	This is a matter of opinion, not a factual error. The ERG has not argued that the placebo effect is large. The ERG’s focus is on how the open-label nature of the COMBI-v trial may have influenced trial results. The ERG’s focus is not solely (or even largely) on the placebo effect. The ERG does not dispute that the difference in OS between the BRAF inhibitor monotherapy arms of the two trials is small, but the difference in terms of the K-M data trajectory is real and this makes a difference to how OS may be projected over the model time horizon. Analyses carried out by the ERG show that if only data from the COMBI-d trial are used as the basis for modelling OS, then the OS gains associated with trametinib+dabrafenib are lower than if pooled IPD from the COMBI-d and COMBI-v trials had been used. One of the roles of the ERG is to highlight areas of uncertainty and the influence these unknowns may have on the magnitude of cost effectiveness results. The ERG has presented concerns about the impact of using pooled IPD from the COMBI-d and COMBI-v trials as the basis of modelling patient survival. Comparability of trials does

		analyses of hazard ratios for PFS from controlled trials have found no evidence of bias with investigator assessed versus BIRC-assessed PFS (U.S. Food and Drug Administration. ODAC briefing document. Evaluation of radiologic review of progression-free survival in non-hematologic malignancies. July 24, 2012. Available at: http://www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/drugs/oncologicdrugsadvisorycommittee/ucm311141.pdf. Accessed August 19, 2015). It is important to not that BIRC-assessed PFS was not evaluated in the final analyses of COMB-d. In the primary analyses of PFS for COMBI-d based on the 26 August 2013 data cut-off, PFS results based on BIRC assessment were consistent with the investigator-assessed results. Any differences between IRC and investigator assessed PFS must be interpreted cautiously, as IRC assessed PFS may be subject to bias due to informative censoring in cases of un-confirmed investigator-assessed progression (Dodd *et all*, 2008). While it has been recognised that in randomized double-blinded, placebo-controlled trials, presumably with minimum sources of bias, placebos are sometimes associated with improved control of symptoms, such as pain and appetite, this is rarely with positive tumor response. Substantial improvements in symptoms and quality of life are unlikely to be due to placebo effects (see http://jnci.oxfordjournals.org/content/95/1/19.full) In light of these factors, which the ERG failed to consider, the likelihood of significant bias due to the lack of placebo-control and blinding in COMBI-v on the outcomes of PFS and OS is small and the decision of the ERG to disregard the evidence from COMBI-v relating to these outcomes is unwarranted.	not mean equality, and with gold standard RCT evidence from the COMBI-d trial available the ERG considers that company needs to present a compelling reason to omit cost effectiveness results generated by using only data from that trial.

Issue 16 Pooling

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 92: The ERG has concerns that the method used by the company to pool OS data fails to take any account of potential trial design bias. The PSA carried out by the company used bootstrapping; however, this was in recognition of potential differences in baseline patient characteristics, rather than concern about any potential trial design bias. The ERG is unaware of any methods that could be employed to account for unknown bias when pooling a dataset to produce unbiased K-M curves, and the company did not suggest a method.	Page 92: ~~The ERG has concerns that the method used by the company to pool OS data fails to take any account of potential trial design bias. The PSA carried out by the company used bootstrapping; however, this was in recognition of potential differences in baseline patient characteristics, rather than concern about any potential trial design bias. The ERG is unaware of any methods that could be employed to account for unknown bias when pooling a dataset to produce unbiased K-M curves, and the company did not suggest a method~~.	The ERG has not defined what it means by "potential trial design bias". Nor has the ERG defined what it means by "unknown bias when pooling a dataset to produce unbiased curves". Unless the ERG can clearly identify or explain the bias which is the source of their concern it is unreasonable to expect the manufacturer to attempt to account for such bias.	This is a matter of opinion, not a factual error. Please see ERG response to issue 16

Issue 17 2nd projection phase for OS

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 95: On checking the company model, the ERG identified that the second projection phase was not as described in the CS. The constant	Page 95: ~~On checking the company model, the ERG identified that the second projection phase was not as~~	The Manufacturer's submissions states "Consistent with patterns of survival from long-term data (suggesting a rapid drop in survival initially and a plateau from approximately 5 years),	This is a point of accuracy. The following statement, included in the CS is correct: "Consistent with patterns of survival from long-term data (suggesting a rapid drop in

mortality hazard identified at the breakpoint was not applied until year 5 as stated by the company, but only until month 30 for trametinib+dabrafenib therapy and month 31 for BRAF inhibitor monotherapy (these time points approximately match the duration of available follow up data). The monthly mortality hazard in both arms then decline at a constant rate to converge with the mortality hazard at 5 years, which was derived from the AJCC registry data. From 5 years, the monthly mortality hazard is assumed to match the mortality hazard from year 5 in the AJCC registry data each month to year 20 regardless of treatment arm.	~~described in the CS. The constant mortality hazard identified at the breakpoint was not applied until year 5 as stated by the company, but only until month 30 for trametinib+dabrafenib therapy and month 31 for BRAF inhibitor monotherapy (these time points approximately match the duration of available follow up data). The monthly mortality hazard in both arms then decline at a constant rate to converge with the mortality hazard at 5 years, which was derived from the AJCC registry data. From 5 years, the monthly mortality hazard is assumed to match the mortality hazard from year 5 in the AJCC registry data each month to year 20 regardless of treatment arm.~~	it was assumed that *mortality rates during the projection period would converge with case-mix adjusted estimates of mortality* (emphasis added) from the AJCC Melanoma Registry from year 5." This is consistent with the approach employed.	survival initially and a plateau from approximately 5 years), it was assumed that *mortality rates during the projection period would converge with case-mix adjusted estimates of mortality* (emphasis added) from the AJCC Melanoma Registry from year 5" The ERG does not dispute that in the company model the mortality hazard rates for both the intervention and the comparator merge with the AJCC registry data hazard at 5 years. The issue is the point at which convergence starts and how it happens. In the CS, the company states that "*mortality rates during the projection period would converge".* This implies that convergence would occur at some point during the projection period. However, in the company model, the projection period for the BRAF inhibitor monotherapy starts at 3 months but the hazard rate does not start to converge with the AJCC hazard until 30 months. It would have been more accurate for the company to state that mortality rates start to converge at month 30 in a linear fashion to meet the AJCC mortality rate at 5 years. The ERG notes that there is no mention in the CS of a 30 month point that hazard rates start to change and no justification as to why they start to converge from 30 months rather than earlier or later.

Issue 18 Case-mix adjustment

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 95: Examination of the survival curves produced by the company, compared to the survival data from the COMBI-d and COMBI-v case mixed adjusted AJCC registry data (Figure 7), shows that fewer patients are alive at year 5 in the BRAF inhibitor monotherapy arm (14.0%) than would be predicted by the AJCC data (14.2%). The ERG considers this implausible, especially given the issues raised earlier in this report about the AJCC registry data probably overestimating the underlying mortality for the patient groups in the COMBI-d and COMBI-v trials at all time points during the trial.	Page 95: Examination of the survival curves produced by the company, compared to the survival data from the COMBI-d and COMBI-v case mixed adjusted AJCC registry data (Figure 7), shows that the survival for patients in the BRAF monotherapy arm is similar to that predicted by the AJCC data.~~that fewer patients are alive at year 5 in the BRAF inhibitor monotherapy arm (14.0%) than would be predicted by the AJCC data (14.2%).~~ ~~The ERG considers this implausible, especially given the issues raised earlier in this report about the AJCC registry data probably overestimating the underlying mortality for the patient groups in the COMBI-d and COMBI-v trials at all time points during the trial.~~	The statement suggests that projected survival for BRAF inhibitors is projected to be materially less than that predicted by the AJCC data. However, a difference of 0.2% in the survival probabilities. It is not reasonable to describe this as “fewer patients”. In deeming that 14% survival at 5 years is implausible, the ERG has failed to compare this projection with estimates from other external sources to assess the validity of this projection. Although long-term data on OS for dabrafenib and vemurafenib from BREAK-3 and BRIM-3 out to 5 years are unavailable, data from other sources can be used to inform the plausibility of this projection. Specifically, the 14% projected survival at 5 years for BRAF monotherapy is intermediate between 5-year survival reported for the DTIC+placebo and ipilimumab+DTIC arms of the CA184-024 trial of ipilimumab in treatment-naïve patients (8.8% and 18.2% respectively) (Maio 2015). Additionally, the 14% survival at 5-years projected by the model is *virtually identical* to the 5-year survival projected by the LiRG for both vemurafenib and DTIC in their evaluation of vemurafenib as first line treatment of BRAF mutation positive melanoma (Beale 2012). The failure of the ERG to consider these alternative sources for validation of the projection, including projections made by the some of the same investigators, raises questions regarding the thoroughness and objectivity of their evaluation of the manufacturer’s submission.	This is a matter of opinion, not a factual error. As the company states in issue 13, you cannot use alternative trial data to estimate survival at 5 years without adjusting for case mix and also taking into account the other treatments available to patients at the time of the trials. As such, the simple comparison made by the company in issues 13 and 18 to results from other trials has to be treated with caution. Essentially, the company seems to be saying that there is no issue with BRAF inhibitor therapy having lower OS at 5 years than having only BSC as is the case in the AJCC registry data. The ERG considers that this is a significant issue and casts doubt on the OS projection for the BRAF inhibitors. The ERG accepts that in ERG report that formed part of the NICE appraisal of vemurafenib for the treatment of locally advanced or metastatic BRAF V600 mutation positive malignant melanoma (TA26[96]), the ERG projection of OS did effectively meet the AJCC data at 5 years. However, at that time the data available did not take into account the second-line treatments that were available (notably immunotherapy) to patients participating in the COMBI-d and COMBI-v trials.

Issue 19 Mortality hazard

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 95: From 3 months the company applied a mortality hazard rate for BRAF inhibitor monotherapy of 4.4%. Whilst the mortality hazard rate between months 3 and 4 suggested by the AJCC registry data is 7.0%, by month 14 the monthly mortality hazard has fallen to 4.3% and falls to 1.1% by month 60 (year 5 – when the mortality hazard rate for BRAF therapy converges with the AJCC registry mortality hazard in the company model). So for almost 4 years the monthly mortality hazard rate for BRAF inhibitor monotherapy is higher in the company model than for patients in the AJCC registry over the same time period. For patients receiving trametinib+dabrafenib, the monthly constant mortality hazard is 2.3% from 18	See justification: the ERG should consider alternative explanations of the pattern of hazards as given in the justification section Page 95: From 3 months the company applied a mortality hazard rate for BRAF inhibitor monotherapy of 4.4%. Whilst the mortality hazard rate between months 3 and 4 suggested by the AJCC registry data is 7.0%, by month 14 the monthly mortality hazard has fallen to 4.3% and falls to 1.1% by month 60 (year 5 – when the mortality hazard rate for BRAF therapy converges with the AJCC registry mortality hazard in the company model). So for ~~almost 4~~ 3.6 years the monthly mortality hazard rate for BRAF inhibitor monotherapy is higher in the company model than for	The actual hazard rates used in the model are shown in the figure below. 26-month moving averages are provided for dabrafenib-trametinib to facilitate interpretation of the hazards from the Kaplan Meier curve. **Figure 1. Monthly hazards of death from AJCC survival distribution (normal RCS, adjusted to case-mix for COMBI-d and COMBI-v) and for dabrafenib-trametinib and BRAF inhibitors used in model** [redacted] It should be noted that the hazard rate is greater for BRAF monotherapy than the AJCC from week 70 or month 16.1 until week 259. This represents 3.6 years, not “almost 4 years” as stated by the	This includes a point of accuracy and also a matter of opinion. In terms of accuracy, the ERG considers that the statement in their report is not misleading but agrees with the company that it would be more accurate to display the figure to one decimal place so that ‘4 years’ is changed to ‘3.6 years’. This change has been implemented in the erratum to the ERG report. The paragraph on page 96 of the ERG report now reads: For 3.6 years the monthly mortality hazard rate for BRAF inhibitor monotherapy is higher in the company model than for patients in the AJCC registry over the same time period. In terms of opinion, whilst the ERG accepts that the proposed ‘shift’ in the OS curve as shown in the company’s justification is one explanation for the difference in hazard at different time points, using OS curves which do not

months at which point the monthly mortality hazard in the AJCC registry is 3.4%. The monthly mortality hazard rate for the AJCC data at 30 months – the point at which the company has started to converge mortality hazard rates to the AJCC data – is 2.4%. The monthly mortality hazard for trametinib+dabrafenib, therefore, essentially converges with the AJCC registry data at 30 months rather than at 60 months and is always lower until the hazards converge at year 5.	patients in the AJCC registry over the same time period. For patients receiving trametinib+dabrafenib, the monthly constant mortality hazard is 2.3% from 18 months at which point the monthly mortality hazard in the AJCC registry is 3.4%. The monthly mortality hazard rate for the AJCC data at 30 months – the point at which the company has started to converge mortality hazard rates to the AJCC data – is 2.4%. The monthly mortality hazard for trametinib+dabrafenib, therefore, essentially converges with the AJCC registry data at 30 months rather than at 60 months and is always lower until the hazards converge at year 5.	ERG. While not stated explicitly, the ERG seems to be suggesting that it is implausible that the mortality rate for BRAF inhibitor monotherapy would be greater than that in the AJCC registry. We believe that such a result is not only plausible but is also entirely consistent with a "shifting" of the AJCC distribution such as that suggested by the LiRG in their assessment of vemurafenib (Beale 2012). That is, if the hazards for death for BRAF inhibitors are similar to that for AJCC, but are "shifted" to the right several months, then given the observed pattern of hazards for the AJCC—i.e., increasing for the first four months, then diminishing—then the hazard rate for BRAF inhibitor monotherapy would be expected to be greater than that for the AJCC for all time points after the point at which the hazards crossed. A stylized example of such a shift in the hazards is shown in the Figure 2. below **Figure 2. Monthly hazards of death from AJCC survival distribution (normal RCS, adjusted to case-mix for COMBI-d and COMBI-v) and shifted AJCC survival distribution (97 day shift)** [redacted]	reflect the impact of treatment with an immunotherapy on survival is not particularly informative. Moreover, the ERG notes that the difference in the monthly hazard rate shown is not the same as in the company model, being much larger in the model and happening much later. What matters is whether there are grounds to: i) project a mortality hazard for patients treated with BRAF inhibitor monotherapy that is higher than that recorded in the AJCC data during the period beyond the availability of trial data ii) assume that there would be more people dead at 5 years than would have been the case if the only treatment option had been BSC. Analyses carried out by the ERG suggest that there are no grounds to support either premise. In the last 6 months of available COMBI-d trial BRAF inhibitor monotherapy data the monthly hazard had almost converged with that in the AJCC data set at the same time point.

		This analysis suggests that the hazards used in the model are not only plausible but consistent with the suggested pattern of hazards by the LiRG in their assessment of vemurafenib. It should be noted that it is possible that there is both a shift in the hazard function and reduction in the maximum hazard. This pattern would yield a result such as that with the combination arm in which the hazard function remains below the AJCC for the entirety of the survival distribution. The failure to consider these alternative explanations for the pattern of hazards, including a potential explanation that the same investigators used in their analyses of BRAF inhibitor monotherapy, raises questions regarding the objectivity of their evaluation of the manufacturer's submission.	

Issue 20 Post-treatment discontinuation K-M analysis

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 95: Further evidence on the implausibility of such a large difference between monthly mortality hazard rates over 5 years is found in the post treatment discontinuation survival curves. If treatment with trametinib+dabrafenib really does lead to a lower monthly mortality hazard over 5 years, given more than half the patients have stopped treatment in any arm of both trials by 12 months, it would be expected that treatment with trametinib+dabrafenib would	Please refer to justification section and amend text Page 95: Further evidence on the implausibility of such a large difference between monthly mortality hazard rates over 5 years is found in the post treatment discontinuation survival curves. If treatment with trametinib+dabrafenib really does lead to a lower monthly mortality hazard over 5 years, given more than half the patients have stopped treatment in any arm of both trials by 12 months, it would be	We believe that due to potential biases inherent in the comparison of post-discontinuation survival by treatment group, that this analysis provides no support to the ERG's contention that the OS projections from the model are implausible. The limitations of such analyses have been described by García-Albéniz and Hernán (2015). In particular, when not all patients in the trial have experienced the qualifying event (e.g., progression or treatment discontinuation), the patients included in the analysis are selected for poor prognosis. Moreover, because patients in the experiment group are less likely to have discontinued than those in the control group, it follows that patients who discontinued in the experimental group have, on average, risk factors for discontinuation sufficiently strong to overcome the beneficial effects experimental therapy. By limiting the	This is a matter of opinion, not a factual error. The ERG considers that the unblinded nature of the COMBI-v trial may have had an impact on clinical behaviour and patient choice with regard to the timing of study treatment discontinuation and as such it could have biased an analysis of post discontinuation survival. The points made by the company, referenced from García-Albéniz[7] relate to post-progression survival for patients receiving second-line treatment, not survival after discontinuation of first-line treatment. The ERG considers that

confer a survival advantage over BRAF inhibitor monotherapy once treatment stopped. To explore this, the ERG asked for the company to provide K-M data from the COMBI-d trial on post treatment discontinuation survival. Only COMBI-d data were requested to avoid any bias in the decision to stop treatment due to knowledge of treatment arm by clinicians and patients that may have existed in the COMBI-v trial, due to the open-label nature of this trial. The post treatment discontinuation K-M curves from the COMBI-d trial are shown in Figure 8 and provide evidence that post treatment survival is the same regardless of treatment with trametinib+dabrafenib or dabrafenib. If anything, patients in the dabrafenib arm appear to have improved post treatment survival compared to patients in the trametinib+dabrafenib arm, at least up until about 7 months.	expected that treatment with trametinib+dabrafenib would confer a survival advantage over BRAF inhibitor monotherapy once treatment stopped. To explore this, the ERG asked for the company to provide K-M data from the COMBI-d trial on post treatment discontinuation survival. ~~Only COMBI-d data were requested to avoid any bias in the decision to stop treatment due to knowledge of treatment arm by clinicians and patients that may have existed in the COMBI-v trial, due to the open-label nature of this trial.~~ The post treatment discontinuation K-M curves from the COMBI-d trial are shown in Figure 8 and provide evidence that post treatment survival is the same regardless of treatment with trametinib+dabrafenib or dabrafenib. If anything, patients in the dabrafenib arm appear to have improved post treatment survival compared to patients in the trametinib+dabrafenib arm, at least up until about 7 months.	analysis to those who experience the qualifying event, only patients with the very worst prognosis will be selected for the experimental group, whereas those with more typical prognosis will be selected for the control group. While there are methods for controlling for this bias, they all rely on the untestable assumption that information on all prognostic factors related to time to discontinuation and post-discontinuation survival are available. Given the potential bias in this comparison, we believe that the comparison of post-discontinuation survival cited by the ERG provides no support to the ERG's contention that the OS projections from the model are implausible. Overall Survival (OS) is the gold standard measure of efficacy (U.S Department of health and humans services. Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics 2017) OS is an objective measure which is not affected by a study being placebo-controlled or open label (Gisele Chvetzof et.al 2003). There will have been no bias in the Combi-V study with respect to the overall survival results (its primary endpoint) and thus we suggest that the text suggesting that Combi V trial was biased be removed.	analysis of post discontinuation survival is informative in this instance, and helps to determine whether the company OS projections are plausible.

Issue 21 Open-label trial

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 100: The differences, shown by the data presented in Table 32, highlight how differential clinical and patient behaviour may lead to biased outcomes in open-labelled trials.	Page 100: ~~The differences, shown by the data presented in Table 32, highlight how differential clinical and patient behaviour may lead to biased outcomes in open-labelled trials~~.	The data in Table 32 show that for patients who continued treatment for more than 15 days after progression, the mean duration of treatment post-progression was greater in COMBI-v, the open label trial, than in COMBI-d, the placebo controlled trial. These differences may be due to a variety of potential factors including but not limited to difference in the use of a placebo control in COMBI-d vs COMBI-v. The ERG provides *no* evidence that establishes that these differences are due to the absence of the placebo control in the COMBI-v trial. Nor has the ERG established how the differences in post-progression treatment shown in this table contribute to biased outcomes in the open-label COMBI-v trial Even if one makes the (as yet unsupported) assumption that the differences in duration of post-progression treatment are due to the differences in the use of a placebo control, then these data suggest that when patients and clinicians are blinded to treatment, they may continue longer on such treatment than if they were not. The impact on OS of the differences in utilization of post-progression study therapy is unknown. To the extent that the differences in utilization are due to differences in the use of a placebo control, one can state that the pattern observed in COMBI-v is more likely to be representative of typical clinical practice than that in COMBI-d, as placebo controls are not used in typical clinical practice. If anything, this table demonstrates the potential impact of placebo controls on patient and clinician behavior which may make trial results less generalizable to typical clinical practice. There are several other limitations with the data in Table 32. For one, as with the analysis of post-discontinuation survival, because some patients were censored for discontinuation, the patients in this analysis may not be representative of all patients in the trial and differences between treatment groups may bias any comparisons across groups. Furthermore, data in this table provide a biased assessment of the degree to which TTD differs from PFS as it is limited to patients with TTD > 15 day more than PFS. It does not include patients who discontinued before PFS or after PFS but <=15 days from PFS. Nor does it include patients who were censored for PFS. Also the distributions are highly skewed and the means are influenced by a small number of subjects who continued for a relatively long period after PFS. Furthermore, it does not account for censoring of patients who progressed but did not discontinue. The Kaplan Meier curves below provide a more complete view of the post-progression time to	This is a matter of opinion, not a factual error. The company seems to have misunderstood the ERG's point that it is the entirety of the open label, unblinded nature of the COMBI-v trial that is of concern. In their justification for proposing an amendment to the ERG report the company states: *These differences may be due to a variety of potential factors including but not limited to difference in the use of a placebo control in COMBI-d vs COMBI-v. The ERG provides no evidence that establishes that these differences are due to the absence of the placebo control in the COMBI-v trial.* The company, therefore, acknowledges that the differences shown may be due to trial design. The ERG has no way of establishing whether this difference is due to trial design. The company does not dispute that there is a difference between trials in the meantime on treatment after progression, just the scale of the

		discontinuation for patients who discontinued after progression. These figures show that median post-progression TTD was less than 30 days for both arms in both trials, that approximately 90% of the patients in the monotherapy arm and 80% of those in the combination arms had discontinued treatment within 6 months. Also, restricted mean survival time (RMST) for time to discontinuation post-progression was slightly more than 3 months for patients receiving combination therapy and less than 2 months for those in the monotherapy arms (see table). **Figure 3. Kaplan Meier estimates of post progression TTD: COMBI-d** [redacted] **Figure 4. Kaplan Meier estimates of post progression TTD: COMBI-v** [redacted]	difference and the reason(s) why this difference occurs. As a point of accuracy, in the company FAC it is stated that mean time on treatment post progression is higher in the COMBI-v trial than in the COMBI-d trial, however, the ERG notes that it was actually higher in the COMBI-d trial. The company states, in its justification for their proposed amendment, that: *The impact on OS of the differences in utilization of post-progression study therapy is unknown. To the extent that the differences in utilization are due to differences in the use of a placebo control, one can state that the pattern observed in COMBI-v is more likely to be representative of typical clinical practice than that in COMBI-d, as placebo controls are not used in typical clinical practice. If anything, this table demonstrates the potential impact of placebo controls on patient and clinician behavior which may make trial results less generalizable to typical clinical practice.* To make the amendment the ERG would first have to accept that there is no difference in clinical practice between the trials, something the company acknowledges does exist. The ERG considers that this

Figure 5. Kaplan Meier estimates of post progression TTD: COMBI-d and COMBI-v

[redacted]

Table 1. Time to discontinuation post progression by treatment group –COMBI-d, COMB-v, and COMBI-d and COMBI-v (Days)

	RMST		Median	
	D+T	D/V/B	D+T	D/V/B
COMBI-d	[redacted]	[redacted]	28	21
COMBI-v	[redacted]	[redacted]	29	18.5
COMBI-d COMBI-v	[redacted]	[redacted]	28	19

RMST=Restricted mean survival time.

An even better comparison for assessing the degree to which it is appropriate to model treatment costs based on PFS can be obtained by comparing Kaplan Meier estimates of PFS with TTD among all patients enrolled in the two trials. As shown in the figures below, PFS and TTD are similar in both COMBI-v and COMBI-d, although TTD is somewhat greater than PFS for both arms of COMBI-d.

difference could have been driven by the fact that the COMBI-v trial was open label. The company also agree this could have been the case, stating that clinical practice during an open label trial would be more reflective of clinical practice than a blinded trial. Whether a trial being open label makes the trial more (or less) like clinical practice is outwith the consideration of the ERG. In terms of pooling data, what matters is that there is a difference between the trials. The company acknowledges this and correctly states that there is no way of knowing the impact of the difference in treatment duration on OS.

		Figure 5. Kaplan Meier estimates of TTD and PFS: COMBI-d **Figure 6. Kaplan Meier estimates of TTD and PFS: COMBI-v** 	

Figure 7. Kaplan Meier estimates of TTD and PFS: COMBI-d and COMBI-v

[redacted]

Consistent with these relatively small differences in the Kaplan Meier curves for PFS and TTD, the absolute differences in Kaplan Meier estimated median TTD and PFS are less than 30 days for all arms of both trials however (see Table 2). The restricted mean survival times (RMST) for TTD and PFS (measured up to the point at which TTD and PFS were available for both arms) also were similar for both arms for both trials. For the combination arm in COMBI-d, the RMST for TTD was 28 days greater than that for PFS. For the monotherapy arm in COMBI-d, the RMST for TTD was 17 days greater than that for PFS. For COMBI-v, the RMST for TTD is somewhat less than PFS for both the combination and the monotherapy arms. The latter finding or the monotherapy arm may be a consequence of cross-over pre-progression.

Table 2. PFS and TTTD by treatment group: COMBI-d, COMB-v, and COMBI-d and COMBI-v combined

	Combination			Monotherapy		
	PFS	TTD	Difference	PFS	TTD	Difference
Median						
COMBI-d	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
COMBI-v	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
COMBI-d COMBI-v	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

		RMST	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		COMBI-d	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		COMBI-v	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		COMBI-d COMBI-v	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
		Taken as a whole these data suggest that PFS does provide a reasonable basis for estimating treatment costs with a scalar adjustments for dose adjustment, treatment interruption and the difference between PFS and TTD and that the ERGs concern regarding the approach used in the model is unfounded.						

Issue 22 TTD and RDI

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 102: The ERG considers that application of TTD only partially negates the need for the RDI (relative dose intensity) adjustment applied by the company. In Table 51 of the CS dose interruptions were shown for almost 50% of patients across both arms of the COMBI-d and COMBI-v trials and dose reductions due to AEs were noted for between 14% and 36% of patients. As such, some estimate of RDI would be considered reasonable. The ERG did not request patient level data to recalculate the RDI based on TTD and so has applied the RDI used in the company base case.	Page 102: ~~The ERG considers that application of TTD only partially negates the need for the RDI (relative dose intensity) adjustment applied by the company.~~ In Table 51 of the CS dose interruptions were shown for almost 50% of patients across both arms of the COMBI-d and COMBI-v trials and dose reductions due to AEs were noted for between 14% and 36% of patients. As such, some estimate of RDI would be considered reasonable. The ERG did not request patient level data to recalculate the RDI based on TTD and so has applied the RDI used in the company base case.	The ERG is correct to that the use of TTD rather than PFS to model treatment costs only partially negates the need for the RDI adjustment. However, the RDI values used in the model account not only for dose adjustments and treatment interruptions, but also the difference between PFS and TTD. Accordingly, the application of the RDIs provided by the manufacturer to the TTD curves leads to overestimation of the costs of treatment. The failure to account for this casts doubt on the ERG's estimates of costs and their conclusions that the manufacturer's model has underestimated the cost of treatment.	This is a matter of opinion, not a factual error. The ERG considers that TTD data provide a better basis for estimating drug costs and patient utility than PFS data. Analysis of the TTD K-M data does not suggest that the difference in cost estimates resulting from using TTD data as opposed to using PFS data is insignificant. The ERG accepts that their use of the company's RDI multiplier slightly overestimate costs and has therefore amended the wording on p102 of their report as follows: The ERG considers that application of TTD only partially negates the need for the RDI (relative dose intensity) adjustment applied by the company. In Table 51 of the CS dose interruptions were shown for almost 50% of patients across both arms of the COMBI-d and COMBI-v trials and dose reductions due to AEs were noted for between 14% and 36% of patients. As such, some estimate of RDI would be considered reasonable. The ERG did not request patient level data to recalculate the RDI based on TTD and so has applied the RDI used in the company base case. As the RDI already adjusts in part for time on treatment, this may mean that the cost using TTD is a slight overestimate for all treatments.

Issue 23 Utility values (I)

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 105: For vemurafenib, the company data used data collected during the open-label COMBI-v trial. Whilst this evidence would appear to support the view that treatment with vemurafenib should be associated with a lower utility value than dabrafenib, it is noted that in previous NICE submissions, notably for dabrafenib first-line treatment, TA321	Modify the text to note that utility value for dabrafenib would be the same as that for vemurafenib because data on EQ-5D utility values for patients receiving vemurafenib were unavailable at the time of the submission	The ERG is correct that in the prior submission, it was assumed that the utility value for dabrafenib would be the same as that for vemurafenib. However, this is because data on EQ-5D utility values were not collected in the BRIM-3 study and therefore were unavailable for patients receiving vemurafenib at the time of the submission.	This is not a factual error. However, the ERG has changed the wording on p105 of their report to reflect the fact that at the time of the NICE appraisal of dabrafenib for the treatment of patients with unresectable/metastatic BRAF V600 mutation-positive melanoma (TA321[8]) utility values associated with treatment with vemurafenib utilities were not available. The revised wording in the ERG report is as follows: For vemurafenib, the company data used data collected during the open-label COMBI-v trial. Whilst this evidence would appear to support the view that treatment with vemurafenib should be associated with a lower utility value than dabrafenib, it is noted that in previous NICE submissions, notably for dabrafenib first-line treatment, TA321[8] the same utility value was applied for dabrafenib and vemurafenib in the absence of utility evidence on vemurafenib.

Issue 24 Utility values (II)

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 107: Whilst utility on treatment is not just driven by AEs, the ERG considers that the AE data from the COMBI-d trial for both arms does not equivocally demonstrate that treatment with trametinib+dabrafenib is more tolerable than treatment with dabrafenib. Results from the COMBI-d trial show that there were greater incidences of any AE, AEs leading to permanent discontinuation of study treatment, AEs leading to dose reduction, AEs leading to dose interruption and fatal SAEs in the trametinib+dabrafenib arm than in the dabrafenib arm (CS, Table 51). These could all be expected to impact on quality of life, even if the symptoms were relatively mild and only classed as Grade1 or 2. For some AEs, notably rash and squamous cell carcinoma, patients in the trametinib+dabrafenib arm experienced more events than those receiving dabrafenib. The ERG considers there is no AE evidence data to support a higher utility for patients receiving trametinib+dabrafenib than for those receiving dabrafenib. Furthermore, the evidence, notably on the higher likelihood of dose reduction and ceasing treatment for patients receiving trametinib+dabrafenib compared with those	We suggest deleting this paragraph and not implementing the relevant change to the CE model Page 107: ~~Whilst utility on treatment is not just driven by AEs, the ERG considers that the AE data from the COMBI-d trial for both arms does not equivocally demonstrate that treatment with trametinib+dabrafenib is more tolerable than treatment with dabrafenib. Results from the COMBI-d trial show that there were greater incidences of any AE, AEs leading to permanent discontinuation of study treatment, AEs leading to dose reduction, AEs leading to dose interruption and fatal SAEs in the trametinib+dabrafenib arm than in the dabrafenib arm (CS, Table 51). These could all be expected to impact on quality of life, even if the symptoms were relatively mild and only classed as Grade1 or 2. For some AEs, notably rash and squamous cell carcinoma, patients in the trametinib+dabrafenib arm experienced more events than those receiving dabrafenib.~~ ~~The ERG considers there is no AE evidence data to support a higher utility for patients receiving~~	The ERGs has used the differences in the percent of patients experiencing AEs to discount the possibility that utilities might be improved with the combination vs. monotherapy despite results from two randomized controlled trials demonstrating such improvements. There are several problems with the ERG's analysis. First, the state that "AE data from the COMBI-d trial for both arms does not equivocally demonstrate that treatment with trametinib+dabrafenib is more tolerable than treatment with dabrafenib" and use this to support an assumption of no difference in utilities. We know of no guidance by NICE or other bodies that require that such a threshold be met in order to assume differences in utility values between treatments. The ERG also failed to note that the utility values that were derived from both combi D and V were a reflection of the patients' quality of life as they received the treatments. The impact of AEs on QOL is intrinsically captured in the quality of life outcomes measured in the trials and therefore it does not make sense to speculate that the combination treatment is not more tolerable than with the monotherapy.	This is a matter of opinion, not a factual error. The ERG does not consider the evidence provided by the company relating to the impact of adverse events (AEs) on quality of life to be compelling. The ERG also has serious reservations about the reliability of the quality of life data collected during the COMBI-v trial, due to the open-label nature of this trial. The key point is that there is no statistical evidence to justify applying differential utilities. This is the primary reason behind the ERG's decision to apply one 'on treatment' and one 'off treatment' utility. The ERG's preferred 'on treatment' utility is higher than the baseline utility values associated with any of the COMBI-d trial or COMBI-v trial treatment arms and therefore provides an indication of treatment effect.

receiving dabrafenib, would seem to support the view that treatment with trametinib+dabrafenib is, on average, less tolerable and associated with a lower utility than treatment with dabrafenib. However, the ERG considers that the improved levels of response seen with trametinib+dabrafenib compared to dabrafenib may mitigate any reduction from increased AEs. On balance, the ERG considers that there is no robust statistical or clinical evidence to support applying different utilities for treatment with trametinib+dabrafenib compared with treatment with dabrafenib.	~~trametinib+dabrafenib than for those receiving dabrafenib. Furthermore, the evidence, notably on the higher likelihood of dose reduction and ceasing treatment for patients receiving trametinib+dabrafenib compared with those receiving dabrafenib, would seem to support the view that treatment with trametinib+dabrafenib is, on average, less tolerable and associated with a lower utility than treatment with dabrafenib. However, the ERG considers that the improved levels of response seen with trametinib+dabrafenib compared to dabrafenib may mitigate any reduction from increased AEs.~~ ~~On balance, the ERG considers that there is no robust statistical or clinical evidence to support applying different utilities for treatment with trametinib+dabrafenib compared with treatment with dabrafenib.~~	Also, the ERG has failed to consider that patients receiving the combination therapy were on treatment for longer than those receiving the monotherapies. If the incidence of AEs increases with time on therapy, then this could explain the higher incidence of some events for patients receiving combination therapy. To the extent that the differences between arms in the incidence of AEs are due to differences in duration of therapy, then the ERG's rationale for not assuming differences in utility values is flawed. In assessing the plausibility of differences in utility values, the ERG failed to discuss the observed differences between groups in other measures of quality of life. For example, in COMBI-d, Overall quality of life (QOL), as measured by the global health status from the EORTC QLQ-C30 was consistently better (significantly at Weeks 8, 16, and 80) for subjects receiving the combination therapy compared with those receiving dabrafenib monotherapy during treatment and at progressive disease. In COMBI-v, similar results were obtained. Overall QOL as measured by the global health status score was significantly better and clinically meaningful ("small" MIDs with 5.2 to 7.9 point improvement) for subjects receiving the combination therapy compared with those receiving vemurafenib at Weeks 8, 16, 24, 32, 40, 48, and at PD.	

		Additionally, the ERG has provided no theoretical basis or empirical data which would explain why *both* the COMBI-d and COMBI-v randomized controlled trials would demonstrate improvements in health related quality of life if no such improvement actually existed as they suggest. The onus should be on the ERG to explain the observed findings given their presumption of no actual benefit. In summary, the ERG's decision to disregard the results of analyses of utility values from two large randomized controlled trials using the preferred method of assessment by NICE is based on unreasonable criteria, uses a potentially biased analysis of AEs that fails to account for the duration of therapy, and which provides no alternative explanation for the results observed in the trials is flawed and should be discarded. As with other changes to the model implemented by the ERG, this change raises questions regarding the thoroughness and objectivity of their evaluation of the manufacturer's submission.	

Issue 25 Discounting

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 112: The company applied a discount rate at the end of every week after one year rather than annually as requested by NICE. The company model does not allow for the application of annual discounting without substantial reworking of the model. The ERG considers that the impact of using the NICE method on the size of the ICER, rather than that employed by the company, would be minimal.	We suggest deleting this text and all other references to this "error" Page 112: ~~The company applied a discount rate at the end of every week after one year rather than annually as requested by NICE. The company model does not allow for the application of annual discounting without substantial reworking of the model. The ERG considers that the impact of using the NICE method on the size of the ICER, rather than that employed by the company, would be minimal.~~	The statement by the ERG is incorrect. The model has a weekly cycle. For all cycles within a year, the same annual discount rate is applied. This is shown in the figure below which plots the discount factors used in the model. **Figure 8. Discount factors used in the model.** [redacted]	The company is incorrect to assert that they have correctly discounted costs and utilities. It is practice in the NHS to apply discounts annually. In the model discounting has been carried out weekly and the resulted figures summed. The order in which multiplication and addition are carried out is important, so weekly costs (and utilities) should be summed to obtain the annual figure before the discount is applied.

Issue 26 AJCC case mix adjustment

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 112: The ERG notes that the company adjusted the AJCC registry data for the BRAF inhibitor arms using only the case mix from the COMBI-d trial rather than the case mix from both the COMBI-d and COMBI-v trials. As the ERG remodelled OS data using only data from the COMBI-d trial, this error does not affect the ERG results	We suggest deleting this text and all other references to this "error" Page 112: ~~The ERG notes that the company adjusted the AJCC registry data for the BRAF inhibitor arms using only the case mix from the COMBI-d trial rather than the case mix from both the COMBI-d and COMBI-v trials. As the ERG remodelled OS data using only data from the COMBI-d trial, this error does not affect the ERG results~~	The statement by the ERG is incorrect. The model uses the survival based on AJCC melanoma registry data adjusted to the M stage distribution all arms of the COMBI-v and COMBI-d combined.	It appears that either the issue 26, or the information presented in the CS, is incorrect. Information, namely details of sample size (n=211), presented in the CS (Box 1, p199) suggests that the case-mix adjustment was based on the characteristics of patients in the monotherapy arm of the COMBI-d trial.

Issue 27 Adverse events

Description of problem	Description of proposed amendment	Justification for amendment	ERG Response
Page 112: However, the ERG considers that the cost of treating the common AEs would have been small relative to the costs of drug treatment included in the model. The model does not allow for easy addition of new AEs, so whilst not amending the model to include additional AEs, the	Page 112: However, the ERG considers that the cost of treating the common AEs would have been small relative to the costs of drug treatment included in the model. ~~The model does not allow for easy addition of new AEs, so whilst not amending the model to include additional AEs,~~ The ERG notes that as AEs that were more commonly associated with trametinib+dabrafenib treatment than with BRAF inhibitor monotherapy were not included in the model, the ICERs presented by the company and the ERG for	The statement by the ERG is incorrect. It is a simple procedure to add additional AEs to the model. It only requires estimation of the incidence of the events for each treatment arm and entering these values on the Safety Profiles sheet.	As requested by the company, the ERG has deleted the identified sentence from their report. However, the ERG remains unconvinced that the model was well designed. The new text on p111 of the ERG report is as follows: **Adverse event costs** The ERG considers that, from a cost perspective, the company model should have included more AEs than

ERG notes that as AEs that were more commonly associated with trametinib+dabrafenib treatment than with BRAF inhibitor monotherapy were not included in the model, the ICERs presented by the company and the ERG for trametinib+dabrafenib compared to BRAF inhibitor monotherapy may be slight underestimates.	trametinib+dabrafenib compared to BRAF inhibitor monotherapy may be slight underestimates.		were actually included, notably those that occurred more often in patients treated with trametinib+dabrafenib rather than a BRAF inhibitor, such as diarrhoea. However, the ERG considers that the cost of treating the common AEs would have been small relative to the costs of drug treatment included in the model. The ERG notes that as AEs that were more commonly associated with trametinib+dabrafenib treatment than with BRAF inhibitor monotherapy were not included in the model, the ICERs presented by the company and the ERG for trametinib+dabrafenib compared to BRAF inhibitor monotherapy may be slight underestimates.

References

Bagust, et al. “Survival Analysis and Extrapolation Modeling of Time-to-Event Clinical Trial Data for Economic Evaluation: An Alternative Approach” Med Decis Making 2014 34: 343

Chvetzoff, G. et al. “Placebo Effects in Oncology” J Natl Cancer Inst 2003; 95:19–29

Dodd, L.E, et al. “Blinded Independent Central Review of Progression-Free Survival in Phase III Clinical Trials: Important Design Element or Unnecessary Expense?” Journal of Clinical Oncology 2008; 26: 22

FDA Briefing Document Oncologic Drugs Advisory Committee Meeting “Evaluation of Radiologic Review of Progression-free Survival in Non-hematologic Malignancies”, 2012, July 24,

Hrobjartsson ,. et al. “IS THE PLACEBO POWERLESS? An Analysis of Clinical Trials Comparing Placebo with No Treatment”, The New England Journal of Medicine 2001; 344:1594-602

Hrobjartsson, A et al “Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment” Journal of Internal Medicine 2004; 256: 91–100

Long G, et al. Combined BRAF and MEK inhibition versus BRAF inhibition alone in melanoma. N.Engl.J.Med. 2014; 371 (20): 1877-1888 (and supplementary appendix).

Maio, M., Jean-Jacques Grab, Steinar Aamdal, Igor Bondarenko, Caroline Robert, Luc Thomas, Claus Garbe, Vanna Chiarion-Sileni, Alessandro Testori, Tai-Tsang Chen, Marina Tschaika, and Jedd D. Wolchok “Five-Year Survival Rates for Treatment-Naive Patients With Advanced Melanoma Who Received Ipilimumab Plus Dacarbazine in a Phase III Trial”, Journal of Clinical Oncology, Volume 33 · number 10 · April 1 2015

Michielin O, et al. Gaining momentum: New options and opportunities for the treatment of advanced melanoma. Cancer Treatment Reviews. 2015; 41:660-70.18. Jang S, Atkins MB. Which drug, and when, for patients with BRAF-mutant melanoma? The Lancet Oncology. 2013; 14:e60-e9.

“Pembrolizumab for advanced melanoma not previously treated with ipilimumab”, NICE TA366, ERG report available at: https://www.nice.org.uk/guidance/TA366/documents/appraisal-consultation-document

Robert C, et al. Improved overall survival in melanoma with combined dabrafenib and trametinib. New England Journal of Medicine 2015; 372(1): 30-39.

Robert C, et al. Pembrolizumab versus ipilimumab in advanced melanoma. New England Journal of Medicine 2015;372:2521-32. DOI: 10.1056/NEJMoa1503093. Supplementary materials

Spiegel, D, et al ”Correspondence”, New England Journal of Medicine, 2001; Vol. 345, No. 17, October 25,

Temple, R. J. “Implications of Effects in Placebo Groups” Journal of the National Cancer Institute, 2003; Vol. 95, No. 1, January

U.S. Department of health and humans services. Guidance for industry clinical trial endpoints for the approval of cancer drugs and biologics.2017

Xabier Garcıa-Albeniz, et al. “Why post-progression survival and post-relapse survival are not appropriate measures of efficacy in cancer randomized clinical trials International Journal of Cancer: 2015; 136, 2444–2447

Vemurafenib for the treatment of locally advanced or metastatic BRAF V600 mutation positive malignant melanoma, ERG report, available at: NICE: https://www.nice.org.uk/guidance/TA269/documents/melanoma-braf-v600-mutation-positive-unresectable-metastatic-vemurafenib-evidence-review-group-report3

Wang, L. “In Clinical Trials and In the Clinic, What Is the Placebo’s Effect?” Journal of the National Cancer Institute, 2003; 95 (1): 67.

1. Chvetzoff G, Tannock IF. Placebo Effects in Oncology. Journal of the National Cancer Institute. 2003; 95:19-29.
2. Electronic Medicines Compendium (eMC). Tafinlar SPC. 2015; Available from: https://www.medicines.org.uk/emc/medicine/28432 [accessed January 2016].
3. Electronic Medicines Compendium (eMC). Mekinist SPC. 2015; Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/002643/WC500169657.pdf.
4. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. 2013; Available from: http://www.nice.org.uk/article/pmg9/resources/non-guidance-guide-to-the-methods-of-technology-appraisal-2013-pdf (Accessed 01/06/2015).
5. National Institute for Health and Care Excellence. TA366: Pembrolizumab for advanced melanoma not previously treated with ipilimumab: NICE 2015.
6. National Institute for Health and Care Excellence. TA269: Vemurafenib for treating locally advanced or metastatic BRAF V600 mutation‑positive malignant melanoma: NICE 2012.
7. García-Albéniz X, Maurel J, Hernán MA. Why post-progression survival and post-relapse survival are not appropriate measures of efficacy in cancer randomized clinical trials. International Journal of Cancer. 2015; 136:2444-7.
8. National Institute for Health and Care Excellence. TA321:Dabrafenib for treating unresectable or metastatic BRAF V600 mutation‑positive melanoma: NICE 2014.