APN311 for treating high-risk neuroblastoma

ERRATA

This report was commissioned by the NIHR HTA Programme as project number 15/194/02

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

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

Page No.	Change
30	First complete paragraph on page deleted. Deleted text outlined ERG's reservations around the KM data provided by the company.
33	The sentence "The ERG has severe concerns with the estimation of treatment effectiveness in the economic analysis. These, stem mainly from three overarching issues. The first one is related to the lack of face validity of the OS and EFS KM data from APN311-302. The second relates to the lack of maturity of OS data and the non-existence of EFS data in historical control R1. Finally, the third issue relates to the naïve (unadjusted) analysis of the relative treatment effectiveness of dinutuximab beta, when compared with isotretinoin." has been amended to "The ERG has severe concerns with the estimation of treatment effectiveness in the economic analysis. These, stem mainly from two overarching issues. The first one is related to the lack of maturity of OS data and the non-existence of EFS data in historical control R1. The second issue relates to the naïve (unadjusted) analysis of the relative treatment effectiveness of dinutuximab beta, when compared with isotretinoin." The following paragraph was deleted: "1) The ERG investigated the KM data provided by the company in the model and noted an inexplicable inconsistency in the proportion of patients moving out of the OS and EFS KM curves in the APN311-302 trial. The ERG produced Figure B to show the proportion of patients in cycle t minus the proportion of patients in cycle t+1 in the OS and EFS KM curves in APN311-302. As the proportion of patients in the EFS and OS curves decreases over time (because patients progress or die), the difference in the proportion of patients each cycle is always positive (Figure B). The red curve in Figure B shows the proportion of patients who leave the EFS curve over time (representing the additional number of patients who progress, relapse or die in that cycle) and the blue curve shows the proportion of patients who leave the OS curve over time (representing the additional number of patients who die that cycle). What would be expected is that the change in the EFS curve is always higher (or the same) as the change in the OS curve. This is because the OS curve only takes into account death events, while the EFS curve takes into account disease progression or relapse, second neoplasm and death events (according to the CS). Therefore, the ERG does not see any possible logical…"
34	The following paragraph was deleted: "explanation for why the proportion of deaths in in the OS curve are higher than the proportion of deaths, added to the proportion of disease, relapse and neoplasm events (captured in the EFS curve). In Figure B, this is illustrated where the blue curve is above the red curve. This might be related with the company potentially misreporting the outcomes included in the KM curves (for example, if the EFS curve censored death events), or with the time intervals not being consistent across the OS and EFS curves. Either case is worrying, and removes the validity of the KM curves in APN311-302 provided by the company. Finally, the ERG is also concerned that the company did not provide numbers at risk to accompany the unadjusted KM data for APN311-302 and R1, despite the ERG's requests for these data at the clarification stage. In conclusion, the ERG considers that the uncertainty and the lack of face validity of the KM data from APN311-302 renders the use of these data inappropriate in the analysis. Using the fitted Gompertz curves to the KM data helps adding some face validity to the OS and EFS curves for dinutuximab beta, however, the fitted and extrapolated curves are still based on the underlying KM data from APN311-302, and are therefore, flawed." Figure B was also deleted. The text "2) Equally concerning, is the fact that the company's model relies on the naïve …" was replaced with "1) The company's model relies on the naïve …"
36, 37, 38,39,40,41,42	Throughout the text, Figure C has been replaced with Figure B; Figure D has been replaced with Figure C; Figure E has been replaced with Figure D; Figure F has been replaced with

	Figure E; Figure G has been replaced with Figure F; Figure H has been replaced with Figure G and Figure I was replaced with Figure I. The title “Figure C” has been replaced with “Figure B”; The title “Figure D” has been replaced with “Figure C”; The title “Figure E” has been replaced with “Figure D”; the title “Figure F” has been replaced with “Figure E” and the title “Figure G” has been replaced with “Figure F”; the title “Figure H” has been replaced with “Figure G” and the title “Figure I” has been replaced with “Figure H”.
39	The first row of Table D was removed from the table.
43	The sentence “and also to try and minimise the structural issues found in the KM data from APN311-302” has been removed from the text.
127	Third complete paragraph on page deleted. Deleted text outlined ERG’s reservations around the KM data provided by the company.
128	Figure 9 deleted.
138	Final bullet point deleted. Deleted text outlined ERG’s reservations around the KM data provided by the company.
153,154	The following text “The ERG is extremely concerned with the lack of face validity of the KM data provided by the company. While visual inspection of the OS and EFS curves for APN311-302 might appear valid (Figure 17), the difference between the curves (which gives the proportion of patients in the failure state) and the between-curve relationship lacks face validity, as seen in Figure 18. The ERG investigated the KM data provided by the company in the model and noted an inexplicable inconsistency in the proportion of patients moving out of the OS and EFS KM curves in the APN311-302 trial. To illustrate this issue, the ERG produced Figure 19 to show the proportion of patients in cycle t minus the proportion of patients in cycle t+1 in the OS and EFS KM curves in APN311-302. As the proportion of patients in the EFS and OS curves decrease over time (because patients progress or die), the difference in the proportion of patients each cycle are always positive (Figure 19). The red curve in Figure 19 shows the proportion of patients who leave the EFS curve over time (representing the additional number of patients who progressed, relapsed or died that cycle) and the blue curve shows the proportion of patients who leave the OS curve over time (representing the additional number of patients who die that cycle). What would be expected to observe, would be that the change in the EFS curve is always higher (or the same) as the change in the OS curve. This is because the OS curve only takes into account death events, while the EFS curve takes into account disease progression or relapse, second neoplasm and death events (according to the CS). Therefore, the ERG does not see any possible logical explanation for why the proportion of deaths in in the OS curve are higher than the proportion of deaths, added to the proportion of disease, relapse and neoplasm events (captured in the EFS curve). In Figure 19, this is illustrated where the blue curve is above the red curve. This might be related with the company potentially misreporting the outcomes included in the KM curves (for example, if the EFS curve censored death events), or with the time intervals not being consistent across the OS and EFS curves. Either case is worrying, and removes the validity of the KM curves in APN311-302 provided by the company. The same issue was identified for the OS curve in R1 and the estimated EFS curve for isotretinoin (also shown in Figure 18). Finally, the ERG is also concerned with the fact that the company did not provide numbers at risk to accompany the unadjusted KM data for APN311-302 and R1, despite the ERG’s requests for these data at the clarification stage.” has been replaced with “Figure 17 presents the OS and EFS curves for APN311-302, while Figure 18 shows the FS curve, derived by estimating OS-EFS. The ERG is concerned with the fact that the company did not provide numbers at risk to accompany the unadjusted KM data for APN311-302 and R1, despite the ERG’s requests for these data at the clarification stage.” Figure 19 has been deleted.
155	The text “In conclusion, the ERG considers that the uncertainty and the lack of face validity of the KM data from APN311-302 renders the use of these data inappropriate in the analysis. Using the fitted Gompertz curves to the KM data helps adding some face validity to the OS and EFS curves for dinutuximab beta, however, the fitted and extrapolated curves are still based on the underlying KM data from APN311-302, and are therefore, flawed.” has been replaced with

	"The ERG considers that using fitted curves for the 10-year analysis is a more robust approach". The text "the ERG notes that using fitted curves instead of the KM data reflects smoother changes in the OS and EFS curves, (Figure 22 compared to Figure 19), however, the red curve crosses the blue curve at approximately month 22, and remains that way for the rest of the short-term model. As explained previously, this reflects an impossible scenario, where the number of deaths in a specific cycle are higher than the number of deaths, summed with the number of progression and relapse events in that same cycle." has been deleted from the paragraph.
157	Figure 22 has been deleted. The sentence "Equally concerning, is the fact that the company's model relies on the naïve (unadjusted) analysis of dinutuximab beta's effectiveness, compared with isotretinoin" has been replaced with "The ERG is concerned with the fact that the company's model relies on the naïve (unadjusted) analysis of dinutuximab beta's effectiveness, compared with isotretinoin."
167	The first row of Table 38 was removed from the table. The following text has been deleted: "When the ERG replaced the OS and EFS KM dinutuximab beta curves by the Gompertz curves in the model, it became apparent that the intrinsic problematic relationship between the OS and the EFS KM curves for dinutuximab beta (Figure 29) were carried to the isotretinoin OS and EFS curves (Figure 30), as HRs were applied to the OS and EFS dinutuximab beta curves to estimate isotretinoin curves. Using the extrapolated Gompertz curves in the short-term model for OS and EFS, is an attempt to minimise the structural issues found in the KM data from APN311-302. However, given that the underlying KM data is flawed (and the Gompertz curves seems to be a considerable good fit to the shape of the KM curves), the shape of the Gompertz curves carries the same problems as the KM curves. Even though the ERG cannot anticipate the direction or the extent of the error in the shape of the curves, it is known that the OS and EFS curves should have a wider gap, as there is either an underestimation of events being captured in the EFS curve, or an overestimation of deaths captured in the OS curve."
168	The sentence "Therefore, the ERG cannot anticipate if the "real" OS curve should sit lower than the one shown in Figure 29, or if the EFS curve should sit higher (or if both curves would move)." has been deleted.
212	The sentence "and also to try and minimise the structural issues found in the KM data from APN311-302" has been deleted.
220	The sentence "The ERG has severe concerns with the estimation of treatment effectiveness in the economic analysis. These, stem mainly from three overarching issues. The first one is related to the lack of face validity of the OS and EFS KM data from APN311-302. The second relates to the lack of maturity of OS data and the non-existence of EFS data in historical control R1. Finally, the third issue relates to the naïve (unadjusted) analysis of the relative treatment effectiveness of dinutuximab beta, when compared with isotretinoin." has been amended to "The ERG has severe concerns with the estimation of treatment effectiveness in the economic analysis. These, stem mainly from two overarching issues. The first one is related to the lack of maturity of OS data and the non-existence of EFS data in historical control R1. The second issue relates to the naïve (unadjusted) analysis of the relative treatment effectiveness of dinutuximab beta, when compared with isotretinoin."

continuously over 10 days. Evidence assessing whether rate of infusion affects clinical outcomes is not available.

The ERG considers the data from APN311-302 to be immature and the length of follow-up to be insufficient to determine fully the clinical effectiveness of dinutuximab beta, particularly whether any clinical benefit is maintained in the longer term. Additionally, there is a [redacted] between treatment groups in APN311-302 in [redacted]: based on data supplied by the company during clarification, the most recently captured event occurred at [redacted]' follow-up in the group given IL-2 compared with [redacted] for the group not receiving IL-2.

As no direct evidence on dinutuximab beta-based treatment versus comparators of interest is available, all estimates of comparative clinical effectiveness are based on naïve indirect comparisons. Furthermore, comparative effect estimates are available for only OS. EFS was not captured during the R1 phase of APN311-302 or in Garaventa, and so evaluation of EFS is not feasible. In a suspended STA (GID-TAG507) evaluating dinutuximab alpha, it was noted that immunotherapy might delay rather than prevent events (EFS in Figure C, Section 1.4.2.2). Taking the previous ERG's opinion together with the relatively short length of follow-up available for APN311-302, the ERG considers that the lack of availability of EFS estimates results in an incomplete representation of the short- and long-term clinical effectiveness of dinutuximab beta-containing regimens versus isotretinoin.

In support of the ERG's reservations about the maturity of the data presented for dinutuximab beta, the ERG proposes that results on clinical effectiveness of dinutuximab alpha could aid in understanding the clinical effectiveness, particularly in the long term, of dinutuximab beta. Considering OS, as raised by the ERG assessing dinutuximab alpha, there seems to be an abrupt change in the OS curve for the immunotherapy after approximately year 7, as depicted in Figure D (Section 1.4.2.2). Importantly, longer-term follow-up available for dinutuximab alpha (12 years) indicate a marked increase in mortality in the dinutuximab alpha group between 6.5 and 9 years (Figure D) and that the observed data for the immunotherapy-containing regimen and isotretinoin seem to converge between 6.5 and 11 years. OS at 10 years is only marginally higher for those receiving dinutuximab alpha compared with those allocated to isotretinoin alone (approximately 59% with immunotherapy vs 52% with no immunotherapy), but this observation is based on sparse data and it is unclear whether the difference is clinically meaningful (as reported by the ERG assessing dinutuximab alpha). The ERG acknowledges

2) The analysis provided by the company after the clarification stage, reporting the fully adjusted HRs, produced a HR below 1 for the relapsed population (when using the APN311-202 study), suggesting that dinutuximab is less effective that isotretinoin for this population. Therefore, the results, and thus the model results lack clinical meaningfulness;

3) Clinical expert opinion sought by the ERG reported that in the UK, dinutuximab beta is always given as a first line treatment to patients and added that they would not retreat patients with dinutuximab beta unless there was evidence substantiating the effectiveness of dinutuximab as a retreatment option (given that the company decided to not carry on with studies in the relapsed or refractory population, such studies are not foreseeable);

4) The company, in their reply to the ERG's clarification questions states that, "*given the lack of data for the use of dinutuximab beta EUSA in patients that may have already failed (relapsed) or those that are refractory to dinutuximab beta EUSA, EUSA Pharma does not support re-treatment with the drug*". The company adds that there are no on-going studies that evaluate the effectiveness of dinutuximab beta in relapsed or refractory patients;

The ERG has severe concerns with the estimation of treatment effectiveness in the economic analysis. These, stem mainly from two overarching issues. The first one is related to the lack of maturity of OS data and the non-existence of EFS data in historical control R1. The second issue relates to the naïve (unadjusted) analysis of the relative treatment effectiveness of dinutuximab beta, when compared with isotretinoin. The ERG summarises the key issues surrounding these aspects of the economic evaluation below:

1) The company's model relies on the naïve (unadjusted) analysis of dinutuximab beta's relative effectiveness, compared with isotretinoin. As reported in the NICE Decision Support Unit's Technical Support Document 18, in the case of a disconnected network of evidence, a naïve indirect comparison will include sampling error plus systematic error due to the imbalance in both prognostic factors and effect modifiers. In this case, children forming the historical control R1 were randomised in the R1 phase of HR-NBL-1 (see Section 4 for more details), which was designed to compare the effectiveness of BuMel

immunotherapy works in a different way from conventional chemotherapy, by potentially altering the disease pathway, it might be inappropriate to assume a constant HR between dinutuximab beta and isotretinoin. It is uncertain if the plateau that might be observed for immunotherapy agents is likely to be present for dinutuximab beta, and how this affects the comparison to isotretinoin.

As the ERG did not have any other available source of comparator data for EFS, it turned to the previous STA for dinutuximab alpha vs isotretinoin (GID-TAG507). Figure B and Figure C show the difference in OS and EFS KM curves when the latest data cut-off point became available for dinutuximab alpha and isotretinoin. The results show that the observed data for immunotherapy and standard therapy appear to converge between 4.5 and 11 years in the longer follow-up analysis. This could suggest that, had a longer follow-up period been allowed in APN311-302, the EFS and OS curves for dinutuximab beta would eventually drop to be closer to the EFS curve for isotretinoin. However, the unadjusted analysis of dinutuximab beta (Figure D and Figure E) shows a substantial separation of EFS and OS curves at around year 7. With regards to EFS, the ERG considers this separation to be unsubstantiated as it is not evidence-based (as R1 did not provide EFS data) and is very likely to represent an overestimation of the effect of dinutuximab beta in terms of preventing disease progression. Based on visual inspection of Figure B, long term EFS is only slightly better by 7% among immunotherapy patients (approximately 52% vs 45%) at 10 years. Despite the apparent difference between the two curves, this was not found to be statistically significant (p-value for log rank test: 0.153 as stated in the dinutuximab alpha ERG report).

Figure B. Observed EFS data for updated 4-year (March 2014) and primary 2-year (June 2009) data analysis (Figure 19 in ERG report for dinutuximab alpha STA [GID-TAG507], page 86)

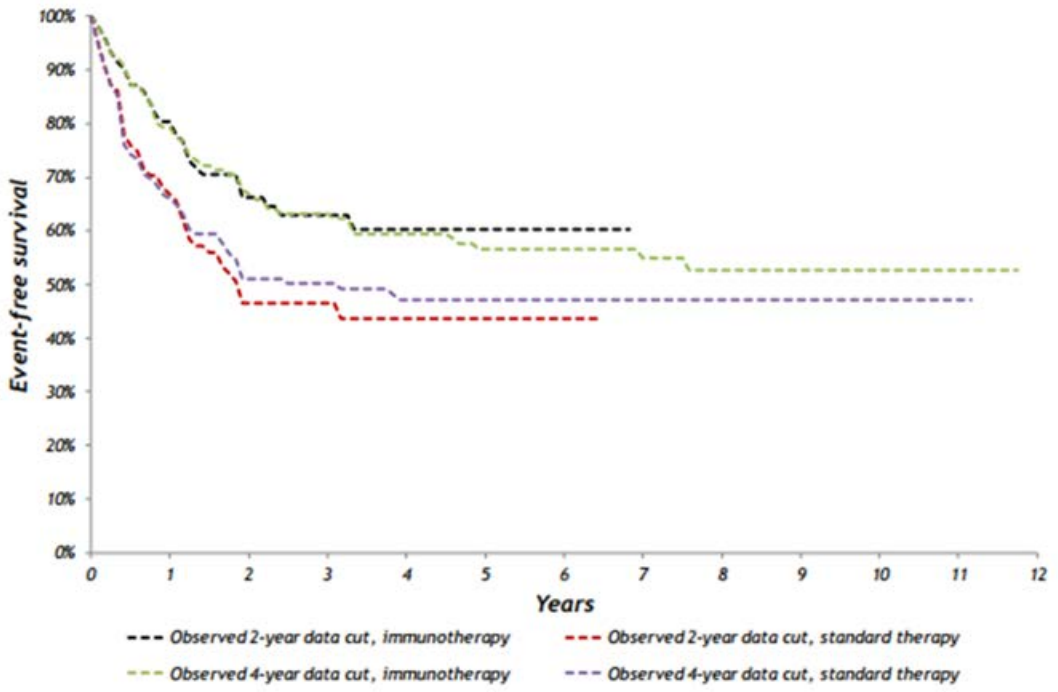


Figure C. Observed OS data for updated 4-year (March 2014) and primary 2-year (June 2009) data analysis (Figure 20 in ERG report for dinutuximab alpha STA [GID-TAG507], page 87)

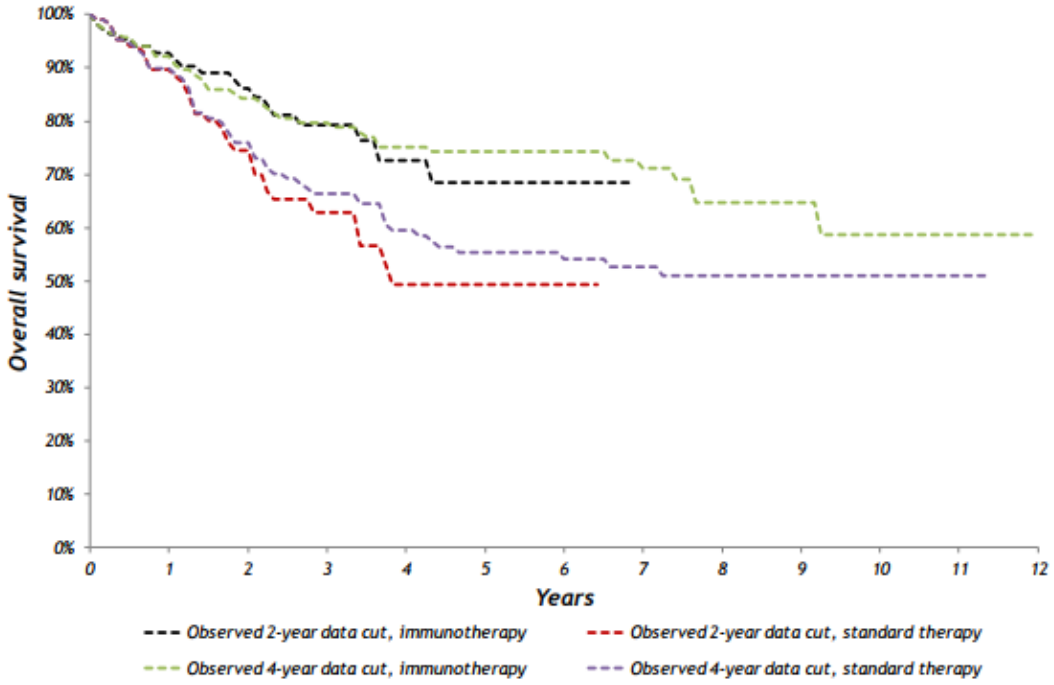


Figure D. Unadjusted EFS curve for dinutuximab beta and estimated unadjusted EFS curve for isotretinoin

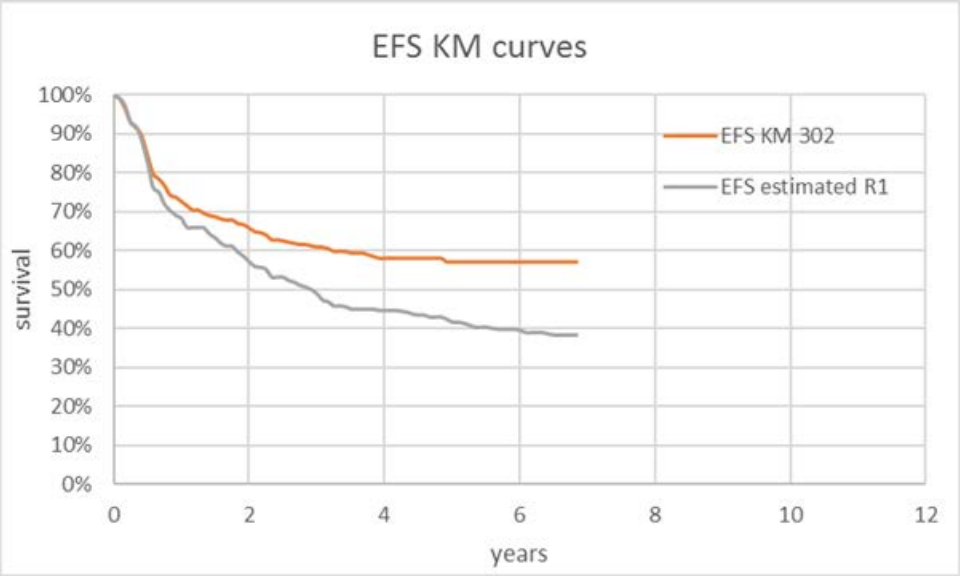
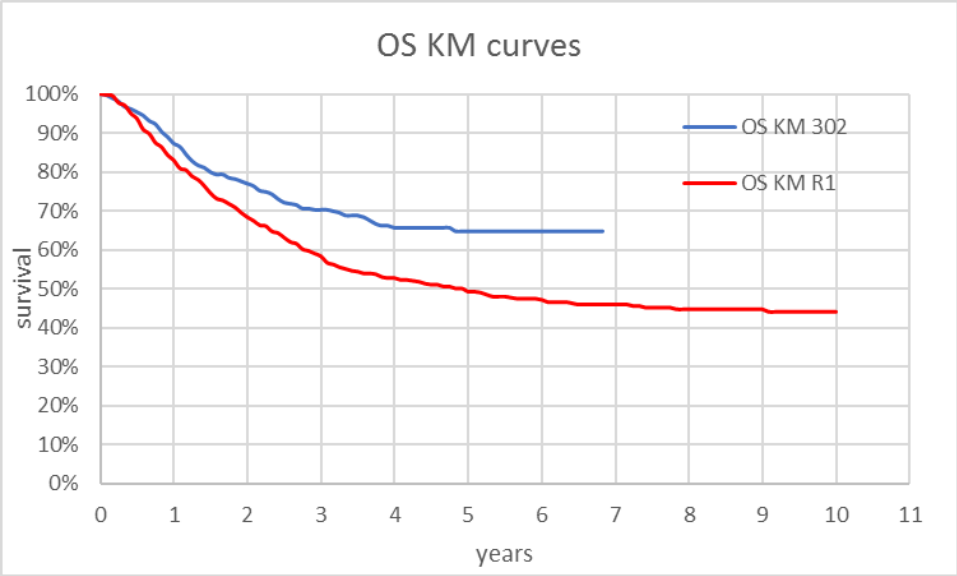


Figure E. Unadjusted OS KM curves

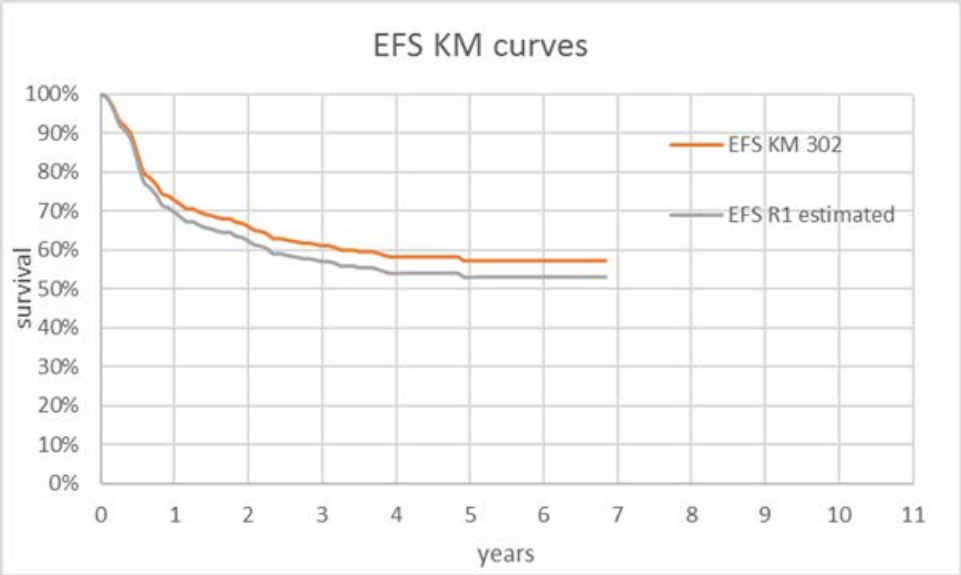


The ERG took the relative difference between the OS HR and the EFS HR in the dinutuximab alpha submission and applied it to the adjusted OS HR estimated for dinutuximab beta. The ERG estimated EFS HR for dinutuximab beta compared with isotretinoin is 1.656/1.319*[redacted]=[redacted].

The ERG acknowledges that the underlying assumption in the ERG's approach is that there is a constant relative risk between EFS and OS for dinutuximab alpha, and furthermore, that the latter relationship is also only observed for dinutuximab beta vs isotretinoin. This is a caveat to the ERG's approach as not only are these assumptions strong, but also the ERG has no evidence to corroborate these. However, the ERG notes that these were the best available data to overcome undertaking a naïve analysis of treatment effectiveness in the model.

After applying the HR of [redacted] to estimate the EFS curve for isotretinoin, the ERG arrived at the curves shown in Figure F. At year 7, the EFS curves seem to be separated by approximately 4% (57% vs 53%). This separation, albeit smaller than the 7% shown in Figure B, is likely to be a better approximation of the relative effectiveness of dinutuximab beta compared with isotretinoin than the 20%, shown in Figure D (resulting from non-evidence based assumptions made by the company, as R1 did not provide EFS data). Finally, the separation of the curves is also linked to the use of a HR to estimate the EFS curve for isotretinoin. As previously mentioned, the ERG cannot be certain if this is a correct methodological approach in this case.

Figure F. Unadjusted EFS curve for dinutuximab beta and estimated isotretinoin EFS curve with adjusted HR



The ERG also notes that about 50% of patients in Figure C were event-free at year 11, regardless of having received dinutuximab alpha or not. With regards to the other 50% of patients, who have progressed, it could be hypothesised that dinutuximab alpha delays, rather than prevents a further event. While it would appear that patients receiving isotretinoin experience the majority of their events over the first two years, a considerable number of events experienced by patients receiving dinutuximab alpha occur between year 2 and year 7. The ERG sought clinical expert opinion with regards to the role

of dinutuximab beta in preventing or delaying events. The clinical experts advising the ERG confirmed that dinutuximab beta was expected to delay events, rather than prevent them.

The ERG's proposed alternatives to overcome the several methodological shortcomings of the company's analysis are, to some degree, flawed, when considered in isolation (for example an assumption of proportional hazards in order to use HRs). However, when combined and incorporated in the final analysis, the synergies resulting from the individual changes made by the ERG, contribute to an increase in the level of uncertainty in the analysis. The ERG summarises the main methodological changes undertaken in Table D.

Table D. Summary of fundamental problems in CS and ERG's ammendmants

	Problem in CS	**ERG's amendment**	**Level of mitigation**	**Proposed approach**
	Naïve comparison of OS data	Use of adjusted HR for OS	Problem partially mitigated. Some level of adjustment for patients' characteristics and previous treatments was applied in the analysis. However, the HR estimation method is flawed and it is unlikely that the use of HRs is an appropriate method of analysis.	An indirect comparison of dinutuximab beta versus isotretinoin and versus dinutuximab alpha should be undertaken. The major methods outlined in the DSU TSD18 applicable in this case are an MAIC and/or an STC. The ERG considers that, depending on what assumptions are made on the nature of the data being compared (e.g. whether proportional hazards hold), an MAIC or an STC will be the most appropriate method to use (please see Section 4 for more details)
	Naïve comparison of EFS data + lack of EFS data for isotretinoin in historical control R1	Taking the relative difference between the OS HR and the EFS HR in the dinutuximab alpha submission and applying it to the adjusted OS HR estimated for dinutuximab beta.	Problem partially mitigated. Some level of adjustment for patients' characteristics and previous treatments was applied in the analysis, through the adjusted OS HR. However, the EFS HR carries the same flaws as the OS HR. Furthermore, it relies on the naïve comparison of the relative treatment effectiveness of dinutuximab alpha vs isotretinoin and isotretinoin beta vs isotretinoin.	
Robustness of the final analysis	Economic analysis unfit for purpose. Resulting ICERs are meaningless	Economic analysis unfit for purpose	Problem partially mitigated	As above

When applying the OS and EFS HRs to the dinutuximab beta curves, the ERG obtained the curves shown in Figure H. The fact that the relative positioning of the dinutuximab beta curves (Figure G) was maintained, allied to the fact that the OS HR and the EFS HR used in the ERG's analysis come from different data sources (thus different populations), leads to the fact that the final relationship between the isotretinoin OS and EFS curves has different and cumulative layers of embedded uncertainty. This is illustrated by the EFS curve crossing the OS curve at approximately 70 months. The ERG had to subsequently cap the EFS curve by the OS curve in the isotretinoin arm of the model.

In conclusion, the ERG does not consider that the changes made to the company's model are robust enough to provide results suitable for robust decision making. The economic analysis needs reconsideration before a meaningful ICER can be produced.

Figure G. Gompertz OS and EFS curves for dinutuximab beta

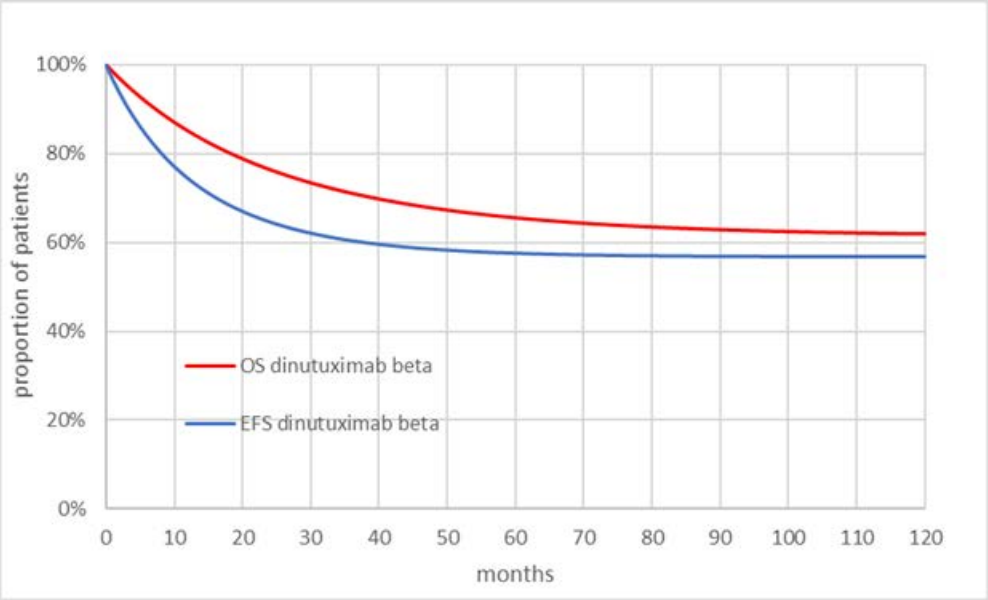
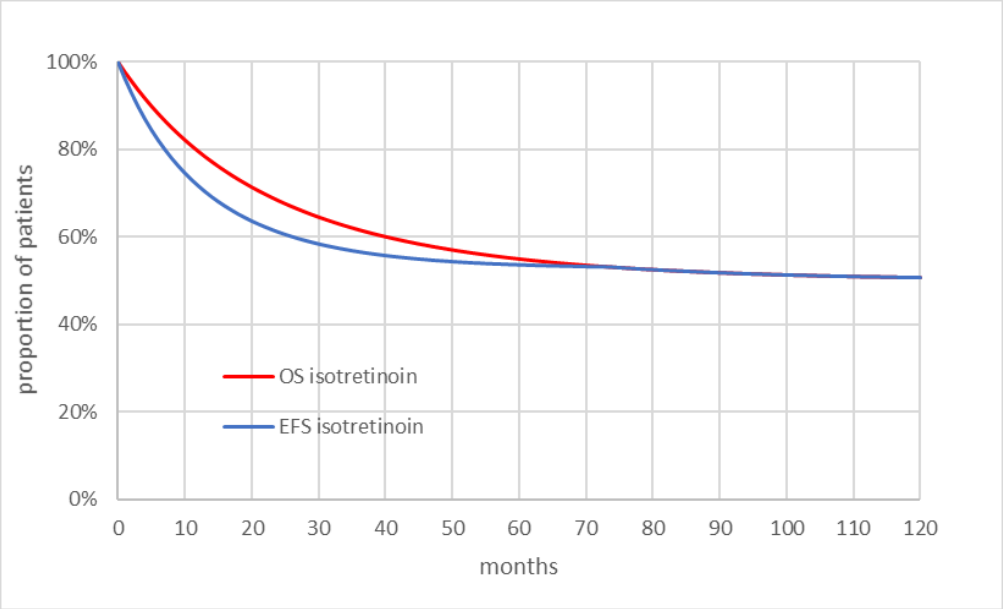


Figure H. Gompertz OS and EFS curves for isotretinoin



The ERG identified issues relating to the estimation of costs and utility values in the economic analysis. These, however, only become relevant once the aforementioned fundamental issues are addressed.

1.5 Summary of exploratory and sensitivity analyses undertaken by the ERG

1.5.1 Economic

The ERG describes the errors found in the company's analysis throughout Section 5 of the report. The company's base case ICER rose from £22,338 to £31,366 per QALY gained, when the ERG corrections were applied.

As the ERG disagrees with carrying out a naïve analysis of treatment effectiveness, two additional corrections were implemented in terms of relative treatment effectiveness in the model:

1. Restructuring the high-risk economic model to incorporate the use of the OS HR ([redacted]) to estimate OS for isotretinoin.

2. Using the relative difference between the OS HR and the EFS HR (for dinutuximab alpha compared with isotretinoin) in the dinutuximab alpha submission and applying it to the adjusted OS HR estimated for dinutuximab beta of [redacted]. To note is that the EFS HR for dinutuximab alpha vs isotretinoin was found to be not statistically significant in the dinutuximab alpha STA. The ERG's estimated EFS HR for dinutuximab beta compared with isotretinoin is 1.656/1.319*[redacted]=[redacted];

Furthermore, the ERG replaced the dinutuximab beta KM curves for OS and EFS by the fitted and extrapolated Gompertz curves in the short-term model, in order to estimate OS after the 7-year KM OS curve. In doing so, the ERG had to subsequently cap the EFS curve by the OS curve in the isotretinoin arm of the model as the curves cross in the model at approximately 70 months.

Using the Gompertz survival curves and the OS and EFS HRs to estimate relative treatment effectiveness in the model leads to an ICER of £111,858 per QALY gained (with all the ERG's corrections incorporated in the analysis).

The ERG considers that while some of the amendments made to the model provide step changes in the right direction, when combined in the final analysis these produce inconsistent outcomes and introduce a paramount level of uncertainty in the analysis. Therefore, the ERG does not consider that the changes made to the company's model are robust enough to produce an ICER fit for purpose and emphasises that the final ICER of £111,858 is provided for illustrative purposes only.

Given the ERG's assessment that the departing ICER of £111,858 is fundamentally flawed, the ERG did not proceed to implement further scenario analyses as all the resulting ICERs. The ERG lists below the analyses that would be required to explore further uncertainty in the economic model, once the base case ICER is robust enough to be used to carry sensitivity analysis:

1. Changing the assumption that patients entering the failure state of the economic model receive chemotherapy for the rest of their lives. In the base case model, some patients receive chemotherapy for more than 20 years, which is not clinically plausible. Therefore, the partitioned survival model should be changed to estimate newly progressed patients in both the dinutuximab beta and isotretinoin arms of the model. Once newly progressed patients are estimated, an assumption needs to be made for treatment duration. For example, it could be assumed that relapsed patients would stay on treatment for a maximum of one year. An assumption should also be made for the resource use required to manage relapsed patients who have gone off chemotherapy treatment, but are still alive and in the failure state;

2. The cost estimations regarding the chemotherapy regimens used in the failure state should include wastage;

3. The cost of treatment administration in the failure state should use the cost of an inpatient stay (£4,670 for five days), instead of procurement cost for chemotherapy drugs, which is used in the base case model (£2,620.54);

4. Concomitant medication costs in the stable state should include wastage for gabapentin;

4.4.2 Methods

The company evaluated the difference in OS between dinutuximab beta and no dinutuximab beta using the log rank test. Estimates of effect and accompanying 95% CIs were not reported. As part of the clarification process, the ERG requested that, for high-risk neuroblastoma, the company carry out an MAIC using the RCT by Yu *et al*.[80] to inform the comparator group of isotretinoin alone. In case the company considered an MAIC infeasible, as an alternative, the ERG requested HRs and 95% CIs for the indirect comparisons of the relevant APN311 study versus historical control and asked that the HR be adjusted for prior treatment (BuMel vs CEM), MYCN status, and age at diagnosis and INSS stage. As discussed in the paragraph introducing Section 4.4, the company did not carry out the MAIC, instead reporting adjusted HRs, initially adjusted for each individual factor and, after further clarification, adjusted simultaneously for all factors. The company presents p values for chi squared tests for potential association between each prognostic factor and treatment effect. Minimal details on the methods and tools used to generate the HRs are available in the clarification response. Cox proportional hazards regression methods have been implemented to generate multivariate adjusted estimates of effect.

4.4.3 Results

The ERG notes that effect estimates for the indirect comparisons are available for only OS. EFS was not captured during the R1 phase of APN311-302 or in Garaventa, and so evaluation of EFS is not feasible. Given that the ERG evaluating dinutuximab alpha raised the point that the immunotherapy might be delaying rather than preventing events, together with the relatively short length of follow-up available for APN311-302, the ERG considers that the lack of availability of EFS estimates results in an incomplete representation of the short- and long-term clinical effectiveness of dinutuximab beta-containing regimens versus isotretinoin.

Figure 9. Deleted by ERG

4.4.3.1 High-risk neuroblastoma

As the company highlights in the CS, mean OS was substantially longer in those receiving isotretinoin alone (2,447.1 days) compared with those receiving dinutuximab beta plus isotretinoin with or without IL-2 (1,359.4 days; Table 30). Similarly, there was variation between groups in median OS, with a median OS of 1,869 days for those receiving isotretinoin and median OS yet to be reached in the group receiving the dinutuximab beta-containing regimen: estimation of the median OS time was not possible in the group receiving dinutuximab beta-containing regimen as <50% of patients had died at the time of analysis. The company proposes that the large difference in mean OS between the groups is likely due to those in the isotretinoin group being followed for longer. The ERG considers that data from the combined analysis for APN311-302 is immature and has concerns about the disparity in length of follow-up between the two studies.

The company reports that the difference in OS between the two groups was statistically significant when evaluated using the log rank test ($p < 0.0001$; unadjusted HR not available; Table 30) and favoured treatment including dinutuximab beta: unadjusted KM curves for OS are presented in Figure 10. The company reported that Cox regression models had been investigated and that INSS stage at initial diagnosis (combined stage 2 vs stage 4S, stage 3 vs stage 4S and stage 4 vs stage 4S) and prior myeloablative consolidation therapy (BuMel vs CEM) were identified as having statistically significant associations with all-cause mortality ($p = 0.0011$ for INSS stage and $p = 0.001$ for prior myeloablative

- [redacted] compared with isotretinoin alone [redacted]): the reported HR is adjusted for age, INSS stage at initial diagnosis, MYCN status, and prior myeloablative therapy.

- In relapsed neuroblastoma, [redacted]. For example, OS estimate for [redacted] with dinutuximab beta-based treatment [redacted].

- Data on the adverse effect profile of dinutuximab beta are primarily derived from a safety database comprising 514 people who have undergone treatment with the immunotherapy, with a focus on 98 people who received dinutuximab beta as a continuous infusion over 10 days. Administration of dinutuximab beta is known to be associated with pain, hypersensitivity reactions, and capillary leak syndrome. Each person in APN311-202 and APN311-303 experienced a TEAE. The company reported that, although the number of TEAEs decreased substantially with each treatment cycle, the proportion of people experiencing a TEAE remained high throughout the study (data not presented).

4.5.1 Clinical issues

- Methods implemented to search and appraise the literature for clinical effectiveness undermine the robustness of the company's systematic review process, including omission of index terms for neuroblastoma from the search strategies, review of abstract and full text publications by one reviewer, potential non-validation of data extraction.

- Potential sources of bias associated with design and conduct of APN311-302 include uncertainty around concealment of allocation, open label design of the study and lack of masked independent assessment of EFS, and the possible disparity within the study in timing of follow-up and recording of clinical effectiveness outcomes.

(instead of using just the tail) but decided to use KM data (instead of the fitted curve) for the period of time where KM data were available.

Despite these technical shortcomings, the ERG notes that estimated survival data are only used for a maximum of 3 years in the company's base case model, for the dinutuximab beta arm of the model, when the 10-year cure threshold is used. Nonetheless, the ERG disagrees with the approach of using OS and EFS KM data for dinutuximab beta for seven years, and then using estimated survival data for three years. To note is that this approach was not justified by the company. The ERG discusses the issues related with the KM data for OS and EFS in APN311-302 in the next section.

5.4.5.2.1 Kaplan–Meier data from APN311-302

Figure 17 presents the OS and EFS curves for APN311-302, while Figure 18 shows the FS curve, derived by estimating OS-EFS. The ERG is concerned with the fact that the company did not provide numbers at risk to accompany the unadjusted KM data for APN311-302 and R1, despite the ERG's requests for these data at the clarification stage.

Figure 17. Kaplan-Meier curves for OS and PFS in APN311-302

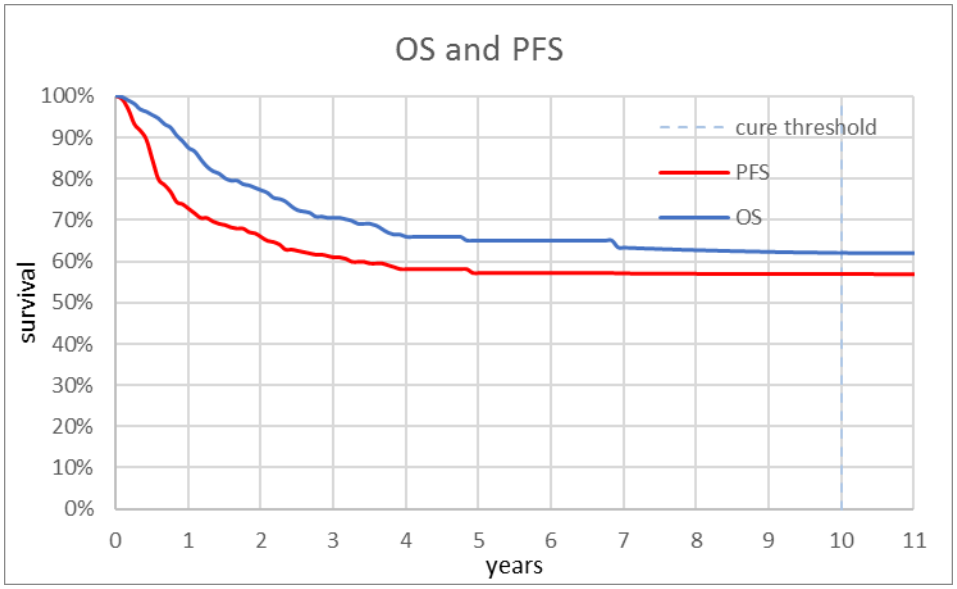
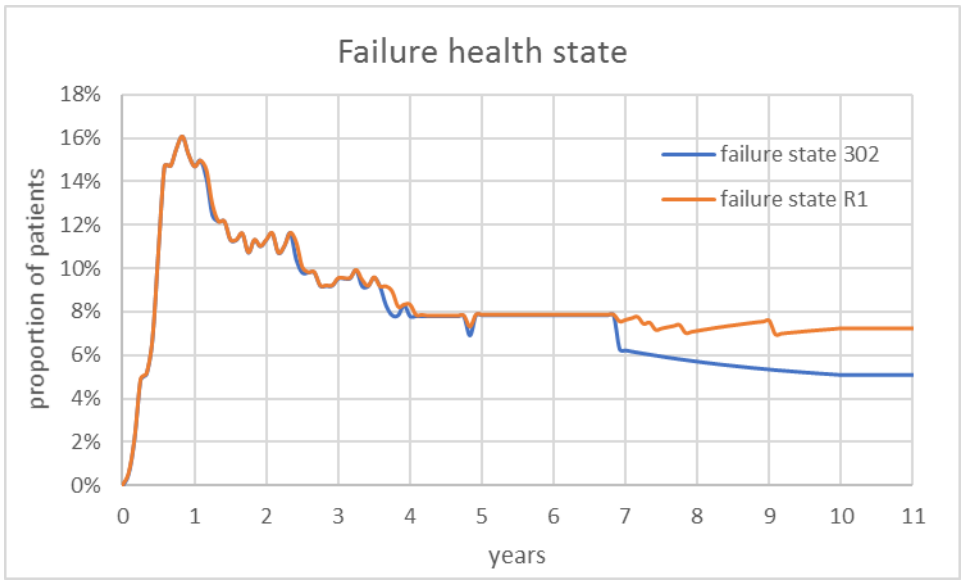


Figure 18. Failure state KM data



The ERG considers that using fitted curves for the 10-year analysis is a more robust approach. Figure 20 shows the unadjusted OS and EFS KM curves for dinutuximab beta, along with the fitted Gompertz curves, and Figure 21 shows the OS KM curves for isotretinoin taken from R1 and the estimated EFS data for R1 (using APN311-302 data), along with the fitted Gompertz curves.

In terms of assessment of fit, the ERG can only rely of visual fit and the measure of variance provided by the company. Both seem to suggest that the Gompertz, lognormal and log-logistic models are the more suitable models to fit the KM data for APN311-302. The same is true for the Gompertz curves fitted to the OS data from R1 and the estimated EFS data for isotretinoin.

The ERG is concerned with the fact that the company's model relies on the naïve (unadjusted) analysis of dinutuximab beta's effectiveness, compared with isotretinoin. As reported in NICE DSU TSD 18, in the case of a disconnected network of evidence, a naïve (unadjusted) indirect comparison will include sampling error plus systematic error due to the imbalance in both prognostic factors and effect modifiers. The guidance adds that the size of this systematic error can be reduced, and probably substantially, by appropriate use of a matching-adjusted indirect comparison (MAIC).[79]

As part of the clarification process, the ERG requested that the company carry out an MAIC. Furthermore, the ERG proposed that an MAIC of the full trial population in APN311-302 versus the group receiving isotretinoin alone in the RCT published by Yu *et al*.[29] (with the updated follow-up data from the dinutuximab alpha submission) would have constitute a better comparison than using R1 (and would have provided a source EFS data for the comparator arm). The company decided against carrying out an MAIC, and instead provided adjusted HRs. The ERG disagrees with the company's arguments for deciding against an MAIC and considers this to have been a most robust method of analysis in this case (details on the company's justification and ERG's views on the latter can be found in Section 4 of the ERG report).

As an alternative, the company provided HRs and 95% confidence intervals (CIs) for the indirect comparisons of OS in the APN311-302 study versus historical control R1, adjusting for prior treatment (BuMel vs CEM), MYCN status, and age and INSS stage at diagnosis. Hazard ratios were initially adjusted for each individual factor and, after further clarification, adjusted simultaneously for all factors. The company presented p-values for chi-squared tests for potential association between each prognostic factor and treatment effect. Cox proportional hazards regression methods have been implemented to generate multivariate adjusted estimates of effect. These are reported in Table 31 below. However, the

Table 38. Summary of fundamental problems in CS and ERG's ammendments

	Problem in CS	ERG's amendment	Level of mitigation	Proposed approach
	Naïve comparison of OS data	Use of adjusted HR for OS	Problem partially mitigated. Some level of adjustment for patients' characteristics and previous treatments was applied in the analysis. However, the HR estimation method is flawed and it is unlikely that the use of HRs is an appropriate method of analysis.	An indirect comparison of dinutuximab beta versus isotretinoin and versus dinutuximab alpha should be undertaken. The major methods outlined in the DSU applicable in this case are an MAIC and/or an STC. The ERG considers that, depending on what assumptions are made on the nature of the data being compared (e.g. whether proportional hazards hold), an MAIC or an STC will be the most appropriate method to use (please see Section 4 for more details).
	Naïve comparison of EFS data + lack of EFS data for isotretinoin in historical control R1	Taking the relative difference between the OS HR and the EFS HR in the dinutuximab alpha submission and applying it to the adjusted OS HR estimated for dinutuximab beta.	Problem partially mitigated. Some level of adjustment for patients' characteristics and previous treatments was applied in the analysis, through the adjusted OS HR. However, the EFS HR carries the same flaws as the OS HR. Furthermore, it relies on the naïve comparison of the relative treatment effectiveness of dinutuximab alpha vs isotretinoin and isotretinoin beta vs isotretinoin.	
Robustness of the final analysis	Economic analysis unfit for purpose. Resulting ICERs are meaningless	Economic analysis unfit for purpose	Problem partially mitigated	As above

When the ERG replaced the OS and EFS KM dinutuximab beta curves by the Gompertz curves in the model, it became apparent that the intrinsic problematic relationship between the OS and the EFS KM curves for dinutuximab beta (Figure 29) were carried to the isotretinoin OS and EFS curves (Figure 30), as HRs were applied to the OS and EFS dinutuximab beta curves to estimate isotretinoin curves.

Using the extrapolated Gompertz curves in the short-term model for OS and EFS, is an attempt to minimise the structural issues found in the KM data from APN311-302. However, given that the underlying KM data is flawed (and the Gompertz curves seems to be a considerable good fit to the shape of the KM curves), the shape of the Gompertz curves carries the same problems as the KM curves. Even though the ERG cannot anticipate the direction or the extent of the error in the shape of the curves, it is known that the OS and EFS curves should have a wider gap, as there is either an underestimation

of events being captured in the EFS curve, or an overestimation of deaths captured in the OS curve.

When applying the OS and EFS HRs to the dinutuximab beta curves (Figure 29), the ERG obtained the curves shown in Figure 30. The fact that the relative positioning of the dinutuximab beta curves was maintained, allied to the fact that the OS HR and the EFS HR used in the ERG's analysis come from different data sources (thus different populations), leads to the fact that the final relationship between the isotretinoin OS and EFS curves has different and cumulative lawyers of embedded uncertainty. This is illustrated by the EFS curve crossing the OS curve at approximately 70 months. The ERG had to subsequently cap the EFS curve by the OS curve in the isotretinoin arm of the model.

Furthermore, given the possibility that immunotherapy might be delaying rather than preventing events, or simply that immunotherapy works in a different way from isotretinoin, therefore altering the disease pathway, it might be inappropriate to assume a constant HR between immunotherapy and conventional chemotherapy. It is uncertain if the plateau typically observed for immunotherapy agents is likely to be observed for dinutuximab beta, and how this compares to isotretinoin.

Consequently, the ERG considers that while some of the amendments made to the model provided step changes in the right direction, when combined in the final analysis these produce inconsistency and introduce a paramount level of uncertainty in the analysis. In conclusion, the ERG does not consider that the changes made to the company's model are robust enough to produce an economic model fit for robust decision making. Nonetheless, and for inclusiveness, the ERG provides the results of implementing the changes listed in Table 38 in the final ICER in Section 6. However, the ERG emphasises that these results are provided purely for illustrative purposes.

9. The discounting factor being applied in the model was estimated on a monthly basis instead of an annual basis. For example, at 1.5 years in the model, instead of using an annual discount factor of 1, the company used a discount factor of 1.5. The ERG corrected this to reflect annual discounting in the analysis.

The company's base case results with the implemented ERG's corrections are presented in Table 56 below. The company's base case ICER rose from £22,338 to £31,366 per QALY gained, when the corrections were applied.

Table 56. Company's corrected base case results – high-risk population

Therapy	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER
Isotretinoin	£172,236	13.61	—	—	£31,366
Dinutuximab beta + isotretinoin	£36,172	18.83	£163,808	5.22	
Abbreviations in table: ICER, incremental cost-effectiveness ratio; QALYs, quality-adjusted life-years.					

As discussed in Section 5.4.5, the ERG does not consider that a naïve comparison of APN311-302 and R1 data is a reliable method for estimating treatment effectiveness. Therefore, the ERG used the only available evidence providing an alternative to the company's analysis. This consisted on the following:

1. Restructuring the high-risk economic model to incorporate the use of the OS HR ([redacted]) to estimate OS for isotretinoin.

2. Using the relative difference between the OS HR and the EFS HR (for dinutuximab alpha compared with isotretinoin) in the dinutuximab alpha submission and applying it to the adjusted OS HR estimated for dinutuximab beta of [redacted]. The ERG notes that the EFS HR for dinutuximab alpha vs isotretinoin was found to be not statistically significant in the dinutuximab alpha STA (GID-TAG507). The ERG's estimated EFS HR for dinutuximab beta compared with isotretinoin is 1.656/1.319*[redacted]=[redacted];

As discussed in Section 5.4.5, the ERG replaced the dinutuximab beta KM curves for OS and EFS by the fitted and extrapolated Gompertz curves in the short-term model, in order to estimate OS after the 7-year KM OS curve. In doing so, the ERG had to subsequently cap the EFS curve by the OS curve in the isotretinoin arm of the model as the curves cross in the model at approximately 70 months.

The company's base case results with the implemented ERG's corrections and the applied HRs to estimate isotretinoin curves are presented in Table 57 below. Using HRs to estimate relative treatment effectiveness in the model leads to an ICER of £111,858 per QALY gained (with all the ERG's corrections incorporated in the analysis).

the comparison, the ERG considers the results of the naïve indirect comparisons in OS to be unreliable and advises that the results are interpreted with extreme caution.

The ERG has serious concerns with the robustness of the economic analysis undertaken by the company. The second (updated) version of the company's model provided to the ERG incorporated paramount changes, which were only accompanied by a brief document as a reply to the ERG's clarification questions. Thus, most of the ERG's critique is based on the inspection of the economic model and not on written evidence submitted by the company. The ERG notes that several calculations and assumptions were changed in the updated model, without being reported or justified by the company (or requested by the ERG during the clarification stage). The consequences of this are twofold: the ERG cannot guarantee that some aspects of the economic analysis and/or economic model were not missed; and there were several instances where the ERG had to make assumptions with regards to what was the company's approach. The ERG identified implementation and formulae errors in the updated economic model (described throughout the report). The ERG is concerned that this reflects a poor level of internal quality assessment of the model by the company.

Overall, the company's modelling approach and model structure is unnecessarily burdensome and removes transparency from the formulae and calculations within the model. It is the ERG's view that the use of a decision-tree to estimate short-term outcomes was unnecessary, especially when the cohort data populating the decision-tree structure is taken from the cohort-based partitioned survival model. The decision-tree model is extremely difficult to navigate and has several circular references in its data implementation. All this makes the ERG's review unnecessarily complex. This also leads to a higher probability of errors in formulae, and a lower probability of all errors being identified during the ERG's review process. In total, the company's model was structured in three different model engines, the decision-tree model, the short-term partitioned survival model and the long-term partitioned survival model. The company could have simplified the model structure, and have a single cohort-based partitioned survival model, which would have been more efficient and transparent, and potentially avoided formulae, and calculation errors.

The ERG has severe concerns with the estimation of treatment effectiveness in the economic analysis. These, stem mainly from two overarching issues. The first one is related to the lack of maturity of OS data and the non-existence of EFS data in historical control R1. The second issue relates to the naïve (unadjusted) analysis of the relative treatment effectiveness of dinutuximab beta, when compared with isotretinoin

The ERG's proposed alternatives to overcome the methodological shortcomings of the company's analysis are, to some degree, flawed, when considered in isolation. However, when combined and