HIGHLY CONFIDENTIAL

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Nivolumab for previously treated unresectable advanced oesophageal cancer [ID1249]

Appraisal Committee Meeting – 5 January 2021
2nd Committee meeting

The following documents are made available to the Committee:

1. **Appraisal Consultation Document (ACD)** as issued to consultees and commentators

2. **Comments on the Appraisal Consultation Document from BMS**

3. **Evidence Review Group critique of company comments on the ACD**

   Appraisal Committee Meeting presentation slides – to follow

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Appraisal consultation document

Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer

The Department of Health and Social Care has asked the National Institute for Health and Care Excellence (NICE) to produce guidance on using nivolumab in the NHS in England. The appraisal committee has considered the evidence submitted by the company and the views of non-company consultees and commentators, clinical experts and patient experts.

This document has been prepared for consultation with the consultees. It summarises the evidence and views that have been considered, and sets out the recommendations made by the committee. NICE invites comments from the consultees and commentators for this appraisal and the public. This document should be read along with the evidence (see the committee papers).

The appraisal committee is interested in receiving comments on the following:

- Has all of the relevant evidence been taken into account?
- Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?
- Are the recommendations sound and a suitable basis for guidance to the NHS?
- Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

Issue date: October 2020

Note that this document is not NICE's final guidance on this technology. The recommendations in section 1 may change after consultation.

After consultation:

- The appraisal committee will meet again to consider the evidence, this appraisal consultation document and comments from the consultees.
- At that meeting, the committee will also consider comments made by people who are not consultees.
- After considering these comments, the committee will prepare the final appraisal document.
- Subject to any appeal by consultees, the final appraisal document may be used as the basis for NICE's guidance on using nivolumab in the NHS in England.

For further details, see NICE's guide to the processes of technology appraisal.

The key dates for this appraisal are:

Closing date for comments: [Day month year]

Second appraisal committee meeting: [Day month year]

Details of membership of the appraisal committee are given in section 5

1 Recommendations

1.1 Nivolumab is not recommended, within its anticipated marketing authorisation, for treating unresectable advanced, recurrent or metastatic oesophageal squamous cell carcinoma in adults after fluoropyrimidine and platinum-based therapy.

1.2 This recommendation is not intended to affect treatment with nivolumab that was started in the NHS before this guidance was published. People having treatment outside this recommendation may continue without change to the funding arrangements in place for them before this guidance was published, until they and their NHS clinician consider it appropriate to stop.

Why the committee made these recommendations

Unresectable advanced, recurrent or metastatic oesophageal squamous cell carcinoma is usually first treated with fluoropyrimidine and platinum-based therapy. Then if the cancer progresses, it is treated with a taxane (docetaxel or paclitaxel).

Clinical trial evidence suggests nivolumab does not improve how well the disease responds or how long people live without their disease progressing compared with taxane treatment. In the trial, the rate of death in the first 3 months of treatment was higher with nivolumab than with taxanes, even though the trial excluded people with a life expectancy of less than 3 months. After that, evidence suggests people live for longer with nivolumab compared with taxane treatment, but clear evidence of long-term survival after 3 months is needed.

Because of the uncertainty in the clinical evidence, there is substantial uncertainty about the most appropriate estimates for costs associated with nivolumab. New data based on further follow up from the trial (up to 36 months) has just become available to the company, but the effect on cost-effectiveness estimates is unknown.

Nivolumab meets NICE's criteria to be considered a life-extending treatment at the end of life. However, the most likely cost-effectiveness estimates are above what

NICE normally considers an acceptable use of NHS resources. So, nivolumab is not recommended for routine use.

Nivolumab is not recommended for use within the Cancer Drugs Fund because it is unlikely to be cost effective at its current price (even if the uncertainty about its effectiveness is reduced).

2 Information about nivolumab

Anticipated marketing authorisation indication

2.1 On 15 October 2020, the Committee for Medicinal Products for Human Use (CHMP) adopted a positive opinion recommending a variation to the terms of the marketing authorisation for the medicinal product nivolumab. The CHMP adopted a new indication as follows: Nivolumab (Opdivo, Bristol-Myers Squibb) as monotherapy is indicated for the treatment of adult patients with unresectable advanced, recurrent or metastatic oesophageal squamous cell carcinoma after prior fluoropyrimidine and platinum-based combination chemotherapy.

Dosage in the marketing authorisation

2.2 The dosage schedule will be available in the summary of product characteristics.

Price

2.3 Nivolumab is available in 3 different sizes as a concentrate for solution for infusion vials. The cost varies according to vial size: £439 (40 mg per 4 ml), £1,097 (100 mg per 10 ml) and £2,633 (240 mg per 24 ml) (excluding VAT; BNF online, accessed October 2020). The cost for 1 dose of treatment is £2,633 (240 mg per 24 ml).

The company has a commercial arrangement. This makes nivolumab available to the NHS with a discount and it would have also applied to this indication if the technology had been recommended. The size of the

discount is commercial in confidence. It is the company's responsibility to let relevant NHS organisations know details of the discount.

3 Committee discussion

The appraisal committee (section 5) considered evidence submitted by Bristol-Myers Squibb, a review of this submission by the evidence review group (ERG), NICE's technical report, and responses from stakeholders. See the committee papers for full details of the evidence.

The appraisal committee was aware that 3 issues were resolved during the technical engagement stage, and agreed that:

- The model time horizon (issue 7, see technical report page 8) used by the company in the economic model of 40 years was sufficient to capture data for everyone having nivolumab or taxanes.
- Nivolumab is likely to improve overall survival by at least 3 months (issue 13, see technical report page 14), meeting the second criteria for end-of-life treatment.
- The approach used to calculate the cost of monitoring response to treatment (issue 12, see technical report page 13) was appropriate.

Clinical need

People would welcome a new treatment option

3.1 The clinical experts explained that people with unresectable advanced, recurrent or metastatic oesophageal squamous cell carcinoma whose disease has progressed after fluoropyrimidine and platinum-based combination therapy have a poor prognosis and no curative treatment options. It disproportionately affects people from lower socioeconomic backgrounds and smoking and alcohol consumption are risk factors. The taxanes paclitaxel and docetaxel are standard treatment for most people and weekly or 3-weekly hospital visits are needed for infusions. People often feel unwell and may experience debilitating fatigue and loss of appetite. Many people find the weekly or 3-weekly treatment regimens

difficult to tolerate because of the associated adverse events. Frequent blood tests are needed to monitor neutropenia. The NHS England clinical lead noted that taxanes have limited efficacy and people are often not well enough to have third-line treatment if taxanes do not control the disease. People who are unable to tolerate taxane chemotherapy have best supportive care, which has no effect on disease progression. Older people are less likely to tolerate chemotherapy, and about 40% of people diagnosed with squamous oesophageal cancer are over 75. The committee recognised the unmet need for a treatment with lower toxicity than chemotherapy, which provides long-term benefit and improves quality of life. The clinical expert explained that if people are not well enough to tolerate taxane therapy they are unlikely be well enough to tolerate nivolumab. Although immunotherapy is generally better tolerated, it still carries risks, notably immune-related side effects. The committee concluded that patients and clinicians would welcome an effective treatment that is better tolerated, particularly if it offers an option of further third-line treatment after disease progression.

Trial design

The ATTRACTION-3 study is appropriate for estimating clinical effectiveness

3.2 The company's clinical evidence came from ATTRACTION-3. This included people with unresectable oesophageal squamous cell carcinoma whose disease was refractory or were intolerant to combination therapy with fluoropyrimidine and platinum-based drugs, and who had a life expectancy of at least 3 months. People were monitored every 6 weeks and assessed using RECIST 1.1 criteria. They could continue treatment after first disease progression in both treatment groups, based on the investigators' judgement. The clinical expert explained that immunotherapies are associated with pseudo-progression, which is a distinct radiological pattern of apparent progression from baseline that is not confirmed with subsequent assessment. For this reason, if there is

evidence of progression but the person feels well, they usually continue having nivolumab for another cycle and then radiological progression is assessed at the next monitoring appointment. The committee concluded that ATTRACTION-3 was an appropriate source of clinical data and could be used for estimating clinical effectiveness.

Clinical evidence

The results from ATTRACTION-3 are generalisable to people in the NHS

3.3 ATTRACTION-3 was done in the US, Europe and Asia. Of the people included in the study, 96% were of Asian family origin, and two-thirds of these people were of Japanese family origin. Oesophageal squamous cell cancer is more prevalent in Asia than in Western countries. The clinical expert commented that although the trials were mainly done in Asia, there is no difference in the underlying biology of oesophageal squamous cell cancer compared with people in the UK. Also, treatment is similar because of consensus in the management of advanced oesophageal cancer. The company accepted that the population in the clinical trial was generally younger and fitter (with an Eastern Cooperative Oncology Group performance status of 0 to 1) than the population seen in NHS practice. The committee agreed with the clinical expert and concluded that the clinical trial was broadly generalisable to people with advanced oesophageal squamous cell cancer in the UK.

Nivolumab improves overall survival but disease progresses faster in the first 3 months of treatment

3.4 Nivolumab is associated with a difference in median overall survival of 2.58 months compared with the combined taxane therapy arm (median overall survival 10.91 months for nivolumab, 8.38 months in the taxane arm). However, median progression-free survival was slightly lower for nivolumab (1.68 months compared with 3.35 months), as was the overall response rate (19.3% compared with 21.5%). More people had disease progression with nivolumab than with taxanes, and most of the overall

survival benefit from nivolumab was after progression. The committee questioned why the benefit was predominantly seen after progression rather than before, which is what would be expected if nivolumab had the potential to be curative. It discussed whether this could be because of people having nivolumab after disease progression and it slowing progression, a carry-over effect after stopping nivolumab into the progression phase, or because people remained well enough for follow-on therapies at progression. The committee concluded that it was unclear why the survival benefit mainly happened after disease progression.

People are at more risk of dying having nivolumab in the first 3 months

3.5 Results up to 24 months for overall survival were provided by the company and analysed by the ERG. At 2 months and 4 months, people having nivolumab had worse overall survival than people having taxanes. However, from 6 months onwards overall survival was higher for nivolumab compared with taxanes (the data cannot be reported here because the company submitted it as academic in confidence). The clinical expert explained that this pattern in overall survival is commonly found with immunotherapies. This is because of the delay in benefit as the immune system is activated, while chemotherapy immediately acts on the cancer cells. The higher death rate in the first 3 months seen with nivolumab was particularly concerning because people in ATTRACTION-3 were expected to survive at least 3 months. The NHS England clinical lead suggested that people generally have worse performance scores in the NHS than in the trial. In clinical practice, it is possible to distinguish between people who are and are not likely to tolerate nivolumab therapy. The company stated that an additional dataset for 36 months was now available for overall survival, progression-free survival and time on treatment. NICE, the ERG and the committee have not had an opportunity to review this and so it could not be taken into account for decision making. Based on the available data, the committee concluded that nivolumab improves overall survival despite a greater death rate in the first 3 months.

Adverse events

Nivolumab is better tolerated than taxanes, but immunotherapies can cause significant side effects

3.6 Fewer patients experienced drug-related adverse events in the nivolumab group compared with taxanes in the clinical trial (the data cannot be reported here because the company submitted it as academic in confidence). The clinical experts agreed that nivolumab is better tolerated than taxanes, and that taxane therapy can be associated with long-term adverse events, such as neuropathy of the hands and feet. The NHS England clinical lead noted that nivolumab is also associated with rare but potentially life-threating gastrointestinal, renal, endocrine and hepatic adverse events. The clinical expert commented that there are standard guidelines for managing immunotoxicity associated with treatments like nivolumab, which are well managed in clinical practice. The committee concluded that nivolumab is better tolerated than taxanes, but immunotherapies can cause significant immune-related side effects.

Comparator

Taxane chemotherapy is the relevant comparator

3.7 The clinical trial compared nivolumab with a combined taxane arm (paclitaxel and docetaxel). The clinical experts and NHS England clinical lead agreed that there is a class effect for taxanes, both in efficacy and side-effect profile. Best supportive care was not considered to be a relevant comparator, because people who are not well enough to tolerate taxane therapy are unlikely to benefit from nivolumab. The committee concluded that the relevant comparator for nivolumab therapy is taxane chemotherapy.

Cost effectiveness

There is uncertainty over the method of extrapolating overall survival

3.8 The company used a semi-parametric approach to model overall survival to capture the changing risk of death over time with nivolumab treatment. Kaplan−Meier curves from the trial were used in both groups up to 2.99 months. Then parametric extrapolation was used based on a log-logistic distribution in the nivolumab arm and an exponential distribution in the taxane arm. The ERG used the Kaplan−Meier curves with a cut-point at 5.75 months and then used a generalised gamma extrapolation for both arms. It chose a later point at which to switch from the Kaplan−Meier curves to parametric extrapolation so that this was at a point after the overall survival curves crossed, and also to maximise the use of clinical data from the trial. The ERG also commented that the choice of extrapolation method should be informed by visual fit to the Kaplan−Meier curve, goodness-of-fit statistics and clinical plausibility. It considered that a generalised gamma distribution gave a better visual fit to observed data in both groups. The company's method assumed a constant risk of death for taxanes and a high initial risk of death that reduced in the long term for nivolumab. The committee considered that the company's model was not a good fit to the currently available Kaplan–Meier curves and was likely to overestimate the overall survival benefit with nivolumab. At the meeting, the company made the committee aware of a later data cut providing estimates for overall survival up to 36 months. However, this could not be taken into account because the NICE technical team, ERG and committee did not have an opportunity to review it before the meeting. The committee considered that the most recent survival data may resolve some of the uncertainty about the most appropriate methods of extrapolation. It concluded that there is uncertainty over the optimal method of extrapolating overall survival.

No adjustment was made to efficacy or additional costs of third-line therapy

3.9 In the clinical trial, patients were able to continue initial treatment (see section 3.2) and have subsequent treatment (surgery, radiotherapy or pharmacotherapy) after disease progression. The proportion of people having subsequent therapy after progression was similar in both the nivolumab and taxane groups. However, more people in the nivolumab arm continued having their initial treatment compared with the taxane arm. The clinical expert explained that nivolumab may be continued after disease progression until the next scheduled scan confirms that the disease has progressed, but treatment would be stopped when progression was confirmed. However, because it is better tolerated than taxanes, more people would be able to have further active treatment after nivolumab than after taxanes. The committee considered the opportunity for active third-line treatment to be an important consideration for patients. It concluded that nivolumab would be more likely to be continued in the short term after progression than taxanes, as seen in the trial. It is not possible to tell whether any differences between the third-line treatments in ATTRACTION-3 and the NHS would affect the relative effectiveness of nivolumab in the NHS compared with the trial.

Utility values

Using different utilities after progression in the nivolumab and taxane arms is not adequately justified

3.10 The company estimated the utilities before and after progression using a statistical model fit to EQ-5D data from the clinical trial, with missing values imputed under the assumption that they were missing at random. Nivolumab had a higher utility before progression than taxanes because of its more favourable safety profile (the data cannot be reported here because the company submitted it as academic in confidence). The company model assumed a higher utility after progression for nivolumab

compared with taxanes because of the continued benefit of nivolumab. The committee considered it plausible for the utility before progression for nivolumab to be higher than the taxane arm, based on differences in tolerability and adverse events. But it noted that the difference was greater in the company's analysis compared with the ERG's analysis, which used values from an alternative statistical model fit by the company that did not include imputation of missing values. The clinical expert explained that it often takes people 6 months to recover from the adverse effects of chemotherapy. The NHS England clinical lead advised that if nivolumab increased the use of third-line treatments, a constant utility after progression was not plausible. The committee concluded that a differential utility before progression was reasonable, but the company had not given adequate justification for a long-term difference in utility after progression.

Costs

The company's method for estimated medical resource use costs is not adequately justified, eMIT should be the source for treatment costs

3.11 The company used the Monthly Index of Medical Specialities list price of taxanes and subsequent treatment for their economic model. Section 5.5.2 of NICE's guide to the methods of technology appraisals recommends using electronic market information tool (eMIT) prices because this is the most reflective source of average prices paid by NHS trusts. The committee concluded that eMIT should have been used to estimate the costs of treatment. This would increase the company base-case model incremental cost-effectiveness ratio (ICER) to £53,459 per quality-adjusted life year (QALY) gained.

The company's model underestimates the cost of inpatient treatment

3.12 The company estimated the cost of each episode of hospitalisation at £534.07 based on an average of 1 bed day per person. The ERG did not consider this method appropriate, instead using the cost of full length of

hospitalisation without adjusting for the length of stay. This increased the cost of hospitalisation to £3,379.73. The committee noted that this remains an uncertainty that has a substantial effect on the ICER. It concluded that the company had not given adequate justification for the estimation of hospital costs based on the duration of stay of 1 bed day.

The range of plausible ICERs is above what is considered cost effective

3.13 The committee noted that the company base-case ICER (including eMIT costs for taxanes) was £53,459 per QALY gained. There were several modelling uncertainties remaining, including the extrapolation of overall survival, progression-free survival and time on treatment. All of these could be affected by evidence from the 36-month data cut. The ERG base-case analysis included different assumptions for overall survival, time on treatment, utility values before and after progression, and medical resource use costs. This gave a cumulative ICER of £125,984 per QALY gained. Using the data available so far, the ICER may be between £53,459 (company base case with eMIT taxane prices) and £125,984 (ERG base case) per QALY gained. The committee concluded that nivolumab could not be recommended as a cost-effective use of NHS resources. It noted that the lowest ICER is also above what is considered plausibly cost effective for consideration in the Cancer Drugs Fund.

End of life

Nivolumab meets the end-of-life criteria

3.14 The committee considered the advice about life-extending treatments for people with a short life expectancy in NICE's guide to the methods of technology appraisal. The committee considered whether nivolumab meets the end-of-life criteria for people with unresectable, advanced or recurrent oesophageal cancer who have had fluoropyrimidine and platinum-based therapy. The company and ERG both agreed based on their analyses that life expectancy in this population is less than 24 months. The committee concluded that nivolumab was indicated for

people with a short life expectancy. The observed median overall survival benefit with nivolumab of 2.5 months was extrapolated. This gave an expected overall mean survival benefit of 7.8 months in the company's base-case model and 4.0 months in the ERG model. The committee considered it likely that the extension to life criterion was met but would like to see the effect of the 36-month data on modelled survival benefit.

Conclusion

Nivolumab is not recommended given the uncertainty in clinical and cost-effectiveness data

3.15 Data from the clinical trial shows that nivolumab offers improved survival benefit compared with taxanes in the long term, but not the short term. The committee has not seen the most recent results for overall survival, progression-free survival and time on treatment. Further justification and supporting evidence is needed for methods of extrapolation, differential utility after progression between treatment arms and hospitalisation costs. The most plausible ICER is currently likely to range between £53,459 (company base case with eMIT taxane prices) and £125,984 per QALY gained (ERG base case). Based on the current evidence, nivolumab is not cost effective for routine use or inclusion in the Cancer Drugs Fund.

4 Proposed date for review of guidance

4.1 NICE proposes that the guidance on this technology is considered for review by the guidance executive 3 years after publication of the guidance. NICE welcomes comment on this proposed date. The guidance executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Jane Adam

Chair, appraisal committee

October 2020

5 Appraisal committee members and NICE project team

Appraisal committee members

The 4 technology appraisal committees are standing advisory committees of NICE. This topic was considered by committee A.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each appraisal committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Farhaan Jamadar
Technical lead

Eleanor Donegan
Technical adviser

Jeremy Powell
Project manager

ISBN: [to be added at publication]

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Response to the Appraisal consultation document

Nivolumab for unresectable, advanced oesophageal cancer when standard chemotherapy has failed

ID1249

Bristol-Myers Squibb Pharmaceuticals Ltd

November 2020

Contents

Executive summary

This document provides a response to the Appraisal Consultation Document describing the use of nivolumab for the treatment of previously treated unresectable, advanced oesophageal squamous cell carcinoma.[1] In line with the Appraisal Consultation Document, this response outlines the additional clinical and economic evidence as requested by the Appraisal Committee, which can be used to support decision making. Thus, Bristol-Myers Squibb (BMS) Pharmaceuticals Ltd believes that the Committee has not yet reviewed all relevant evidence; however, this will be remedied following receipt of the evidence contained in this report.

Key points of the additional evidence presented in this response:

- Additional evidence provided from the ATTRACTION-3 August 2020 database lock demonstrate that results remain consistent with the results of the primary analysis presented in the company submission. Patients treated with nivolumab continued to demonstrate improved OS rates.

- The additional follow-up data from ATTRACTION-3 August 2020 database lock was sufficient to demonstrate that end of life criteria was met in terms of at least three months of additional survival based on restricted mean OS.

Further, an updated cost-effectiveness analysis is provided (Section 6), aligned to several of the preferred assumptions from the NICE Appraisal Committee and utilising the additional database lock from ATTRACTION-3. Mean OS is increased in both arms versus the original company base case analysis. It was predicted that the use of nivolumab will result in an additional 0.512 discounted QALYs and an additional 0.724 undiscounted life years, comprised of additional time in the pre-progression and post-progression health states. Incremental discounted costs were predicted to be £24,665 for nivolumab over taxanes, under base case assumptions. The resultant ICER estimate for nivolumab versus taxanes was £48,205 per QALY gain. Therefore, the base case ICER is below a £50,000 per QALY wiliness-to-pay threshold when the current nivolumab PAS discount is applied.

In summary, the availability of nivolumab would provide an opportunity to make a significant and substantial impact on health-related benefits and address a current unmet need. BMS believes that the Committee recommendations do not take into account all relevant evidence, do not accurately reflect clinical and cost-effectiveness conclusions, and do not provide a sound and suitable basis for guidance to the NHS. It is anticipated that further evidence presented in response to the Appraisal Consultation Document will be considered by the Appraisal Committee, and will further demonstrate that nivolumab is cost-effective and is associated with substantial clinical benefit in a population with very short survival and limited treatment options. The adoption of nivolumab for this therapeutic indication within NHS England would represent a significant advance in the management of this life-threating condition.

1. Evidence requested by NICE: ATTRACTION-3 August 2020 database lock

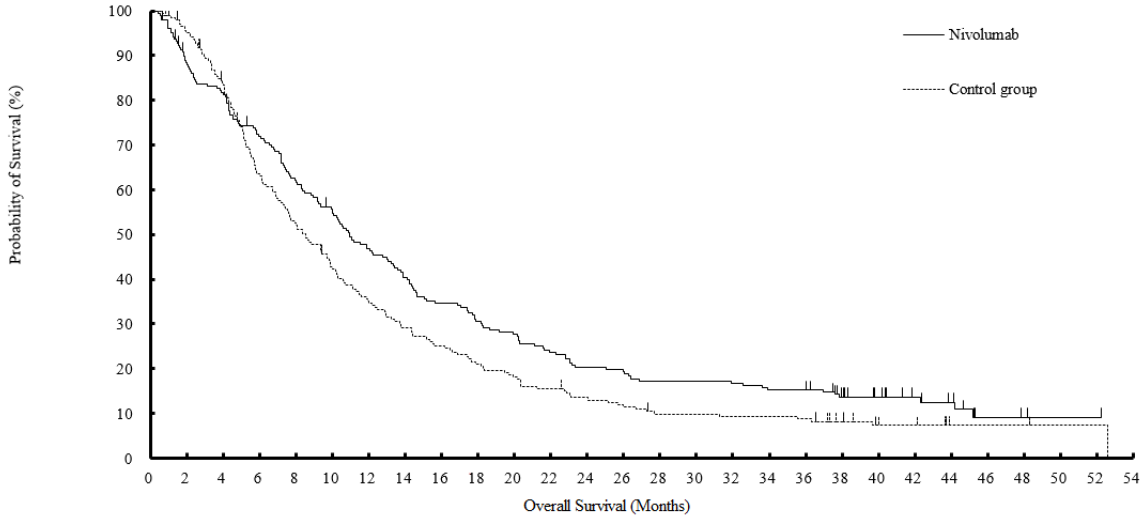
As requested in the Appraisal Consultation Document, additional evidence is provided from the ATTRACTION-3 August 2020 database lock (data cut-off: 25 May 2020), reporting a minimum follow-up of 36.04 months (Table 1).

These results remain consistent with the results of the primary analysis presented in the company submission. As can be expected, median outcomes and outcomes at one year were relatively unchanged. However, outcomes in the nivolumab arm continued to demonstrate improved OS rates at 24 months (20.2% vs 13.5%) and 36 months (15.3% vs 8.7%) compared with the chemotherapy control.

Table 1. ATTRACTION-3 updated outcomes[2, 3]

Endpoint	Nivolumab	Control
Number evaluable	210	209
Overall Survival (OS)		
Median, months (95% CI)	10.91 (9.23, 13.34)	8.51 (7.29, 9.86)
HR (95% CI)	0.79 (0.64, 0.97)	
p-value	p = 0.0264	
OS Rates (95% CI), %		
12-month	46.9	34.7
	[redacted]	[redacted]
24-month	20.2	13.5
	[redacted]	[redacted]
36-month	15.3	8.7
	[redacted]	[redacted]
Investigator-assessed Progression-free Survival (PFS)		
Median, months (95% CI)	1.68 (1.51, 2.73)	3.35 (2.99, 4.21)
HR (95% CI)	1.07 (0.87, 1.33)	
PFS Rates (95% CI), %		
12-month	11.9	7.2
	[redacted]	[redacted]
24-month	5.4	2.4
	[redacted]	[redacted]
36-month	4.3	1.6
	[redacted]	[redacted]
Investigator-assessed Objective Response Rate (ORR)		
Number evaluable	171	158
Responders, n (%)	33 (19.3)	34 (21.5)
95% CI	[redacted]	[redacted]

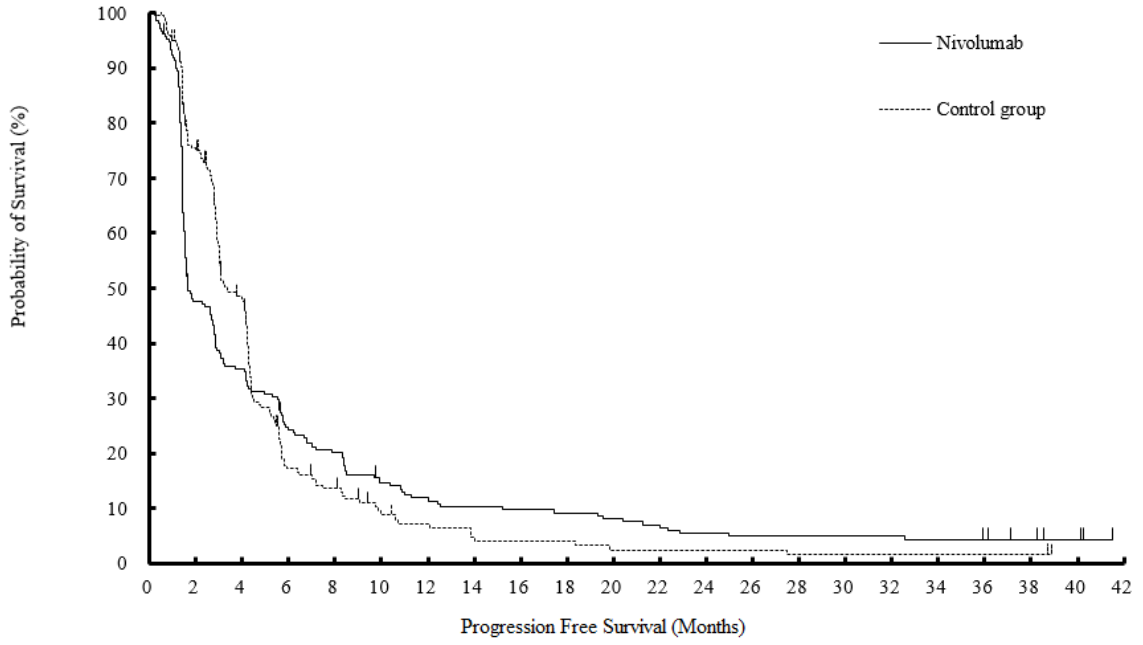
Figure 1. ATTRACTION-3: Kaplan-Meier Plot of Overall survival - 25-Aug-2020 DBL



Analysis Set : ITT

At risk	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54
Nivolumab	210	182	167	147	126	111	95	82	70	62	56	48	41	39	35	35	34	31	31	22	17	12	9	3	2	1	1	0
Control group	209	196	170	127	106	85	69	58	50	42	36	31	26	22	18	18	17	17	16	11	7	6	2	2	2	1	1	0

Figure 2. ATTRACTION-3: Kaplan-Meier Plot of progression-free survival - 25-Aug-2020 DBL



Analysis Set : ITT

At risk	0	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42
Nivolumab	210	96	71	48	40	27	22	19	18	17	15	12	10	9	9	9	9	8	7	5	3	0
Control group	209	147	89	29	22	12	9	5	5	5	3	3	3	3	2	2	2	2	2	2	0	0

2. Has all the relevant evidence been taken into account?

Since publication of the Appraisal Consultation Document, additional evidence has been sought and further economic evaluations have been undertaken in order to address the Committee's requests. This includes:

- Data describing additional follow-up in ATTRACTION-3
- Economic evaluations applying the Committee's preferred assumptions and modelling methods or addressing the Committee's stated concerns.
- Economic evaluations using the data describing additional follow-up from ATTRACTION-3.

In light of the updates to the evidence base, the Committee has not yet reviewed all relevant evidence; however, this will be remedied following receipt of the evidence contained in this report.

3. Are the summaries of clinical and resource savings reasonable interpretations of the evidence?

BMS does not believe that the summaries of clinical and cost-effectiveness are reasonable interpretations of the evidence, as detailed below.

3.1. Clinical benefits

3.1.1. Survival benefit

Although there is initial crossover in the OS Kaplan-Meier data, median OS and OS rates from 6 months to end of follow up show a beneficial impact for nivolumab versus taxanes. Landmark analyses (Figure 5) demonstrate that outcomes are significantly improved for nivolumab versus taxanes in those patients alive at three months. As noted by clinical experts in the Appraisal Consultation Document, this is a common pattern of response for immuno-oncology therapies, particularly those indications where survival is short and evidence is versus an active comparator. This is because of the delay in benefit as the immune system is activated, while chemotherapy immediately acts on the cancer cells. However, it is clear that nivolumab is associated with significant survival benefits across the population of patients with oesophageal squamous cell carcinoma.

Immunotherapies such as nivolumab have a different mechanism of action than conventional anti-cancer therapies, which typically aim to reduce the tumour burden through direct disruption of tumour cell proliferation or induction of apoptosis. By contrast, immunotherapy agents such as nivolumab, often have a delayed clinical responses[4] and differences in response patterns after immunotherapy may potentially be prematurely misclassified as disease progression under the WHO or RECIST criteria.[4, 5] For the same reasons, PFS may not be an adequate endpoint in immunotherapy trials and may not be considered a surrogate for OS for the achievement of clinical efficacy.

The Appraisal Consultation Document suggests that most of the overall survival benefit from nivolumab was after progression. However, it should be noted that there is significant

survival benefit both before and after progression. As demonstrated in Figure 3 and Figure 4, nivolumab improves pre and post-progression survival versus taxanes. In the pre-progression setting, OS is [redacted] at 36 months for nivolumab versus [redacted] in the control arm, while in the post-progression setting, OS at 12 months is [redacted] for nivolumab versus [redacted] for taxanes.

Figure 3. ATTRACTION-3 pre-progression survival

Figure 4. ATTRACTION-3 post-progression survival

Figure 5. ATTRACTION-3 landmark analysis based on patients alive at three months

3.1.2. Extension to life

When considering application of end-of-life criteria, the committee concluded that nivolumab was indicated for people with a short life expectancy and considered it likely that the extension to life criterion was met but would like to see the effect of the 36-month data on modelled survival benefit. It should be noted that there was limited impact on median survival outcomes, with a median overall survival benefit of 2.4 months in the Aug 2020 database lock. However, the additional follow-up was sufficient to demonstrate that end of life criteria was met in terms of at least three months of additional survival based on restricted mean OS ([redacted] months for nivolumab for [redacted] months for taxanes).

Based on the data provided in the company submission, the observed median overall survival benefit with nivolumab of 2.5 months was extrapolated. This gave an expected overall mean survival benefit of 7.8 months in the submission base case model and 4.0 months in the ERG model. Based on additional follow-up, this mean survival benefit was extended to [redacted].

3.2. Quality of life benefits and resource use savings

3.2.1. Hospitalisation cost

Based on the clinician survey detailed in the company submission, the model assumes that disease management requires a mean of 0.095 hospitalisations per week. However, the hospitalisation cost is derived from NHS National Cost Collection based on a weighted mean of hospitalisation costs, which have a length of stay ranging from 3 days (cost: £1,907) to 19 days (cost: £8,986). This cost is applied on a weekly basis, raising an implausible scenario where the weekly cost incurred is appropriate for a period of time longer than a week.

3.2.2. Utility values

The Committee considered it plausible for the utility before progression for nivolumab to be higher than the taxane arm, based on differences in tolerability and adverse events. Further, the Committee concluded that most of the overall survival benefit from nivolumab was after progression. However, the Committee concluded that the company had not given adequate justification for a long-term difference in utility after progression. As utility in oncology is typically a function of time to death, improved OS rates are a key component in postponing quality of life decrements. The Appraisal Consultation Document suggests that most of the overall survival benefit from nivolumab was after progression. Hence, it is appropriate to reflect this benefit in the post-progression utility value.

Further, patients in the nivolumab arm frequently continued receiving nivolumab following progression, as noted in the ERG report. Hence, any beneficial impact associated with nivolumab treatment is continued into the post-progression state for those patients. Pooling post-progression quality of life data assumes that patients in the taxane arm receive benefit equivalent to patients receiving nivolumab. Additional analysis is presented in the appendix to this response, demonstrating that benefit can be stratified by treatment status, rather than by progression status. This is limited by poor data collection when off treatment. However, it demonstrates that treatment status may be a more reliable predictor of benefit than progression status.

3.2.3. Impact of subsequent therapies

As noted in the ERG report, subsequent therapy (i.e. not allocated study therapy) was received by 119 (57%) of 210 patients in the nivolumab group and 115 (55%) of 209 patients in the taxanes group.[2] However, it should be noted that this has limited impact on survival outcomes, as demonstrated in Figure 6, as censoring patients who receive subsequent therapy does not greatly impact the comparison between nivolumab and taxanes.

Figure 6. ATTRACTION-3: Overall survival censored for subsequent therapy

4. Are the recommendations sound and a suitable basis for guidance to the NHS?

BMS does not believe that the recommendations can be considered sound and a suitable basis for guidance to the NHS. A thorough discussion of the Appraisal Committee recommendations and Appraisal Consultation Document has been provided above, primarily, this response outlines additional clinical and economic evidence that can be used to support decision-making. Thus, the recommendations made within the Appraisal Consultation Document should be reviewed in the light of this evidence.

5. Are there any aspects of the recommendations that need particular consideration to ensure we avoid unlawful discrimination against any group of people on the grounds of race, gender, disability, religion or belief, sexual orientation, age, gender reassignment, pregnancy and maternity?

The Committee recognised that there is a significant unmet need in patients with unresectable advanced, recurrent or metastatic oesophageal squamous cell carcinoma whose disease has progressed after fluoropyrimidine and platinum-based combination therapy. Further, the Committee noted that it disproportionately affects people from lower socioeconomic backgrounds.

As noted in the company submission, the incidence of oesophageal cancer is strongly correlated to age, where around 41% of new cases in the UK between 2014 to 2015 were diagnosed in those over 75 years old.[6] In addition, the five-year net survival of oesophageal cancer patients aged 70 years and over is notably poorer compared with younger patients, particularly in female patients. Nivolumab provides a treatment option with proven efficacy and tolerability, with the potential to impact on symptoms, progression and survival. Ageing well and tackling premature mortality is a priority for NHS England.[7]

6. Updated cost-effectiveness analysis

6.1. Cost-effectiveness analysis methods

6.1.1. Survival extrapolation

Based on the ATTRACTION-3 August 2020 database lock, it can be observed that both the company and ERG base case analysis underestimated long-term overall survival outcomes for nivolumab and taxanes (Table 2). As these values are underestimated, it is necessary to assess the impact of using the updated ATTRACTION-3 data to inform cost-effectiveness outcomes.

Using the methodology outlined in the company submission, patient-level data from the ATTRACTION-3 August 2020 database lock were used to inform long-term extrapolations.

In line with preferences stated by the ERG, the patient-level data was assessed using a semi-parametric fit, applying Kaplan-Meier data until 5.75 months followed by parametric extrapolation.

Table 2. Comparison of previously predicted overall survival outcomes versus observed outcomes from ATTRACTION-3 August 2020 database lock

OS Rates (%)		12 months	24 months	36 months
Nivolumab	Observed	46.9	20.2	15.3
	Company submission base case	45.6	21.3	12.3
	Updated base case	44.7	22.2	13.9
	ERG base case	46.1	20.7	10.2
Taxane	Observed	34.7	13.5	8.7
	Company submission base case	36.5	11.1	3.3
	Updated base case	35.7	15.1	7.0
	ERG base case	35.4	12.2	4.4

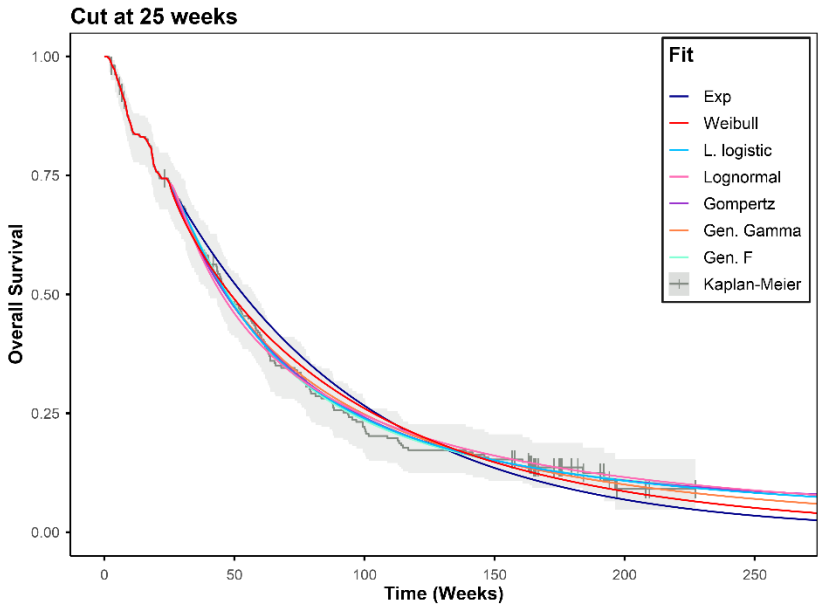
6.1.1.1. Overall survival

In order to model OS in the nivolumab arm, Kaplan-Meier data was applied until 25 weeks followed by parametric extrapolation using the log-logistic distribution to provide an appropriate fit. This approach predicted a median OS of 47.0 weeks and a mean OS of 170.4 weeks. When assessing the Akaike and Bayesian Information Criteria (AIC and BIC, respectively), the log-logistic distribution provided the best goodness-of-fit, indicating it had a strong fit to the data, whilst this was also supported by a strong visual fit to the data, capturing the hazard of the tail of the Kaplan-Meier. The Gompertz function can be excluded due to implausibly long survival and the exponential function provided a visibly poor fit to the data.

Similar to the nivolumab arm, Kaplan-Meier data was applied until 25 weeks for the taxane arm; however, a Weibull distribution followed for the extrapolation period. The Weibull distribution provided a clinically plausible estimation of the mean OS (59.0 weeks), whilst

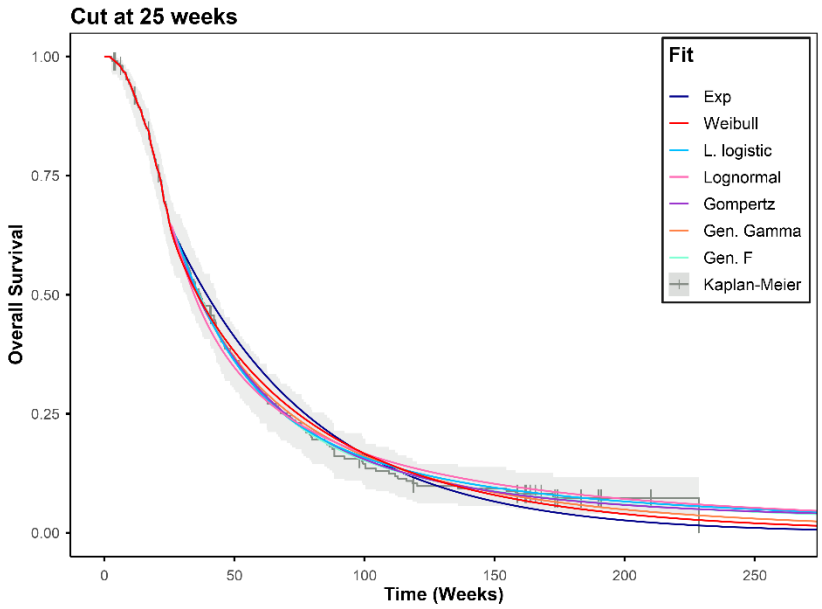
also providing a reasonable goodness-of-fit to the data. Fits predicting mean OS greater than 104 weeks were considered implausible based on clinical expert opinion.

Figure 7. ATTRACTION-3 August 2020 database lock: nivolumab OS



Fit Name	AIC	BIC	Parameters	Median (weeks)	Mean (weeks)
Exp.	1317.06	1320.07	lambda: 0.0135	53.36 (45.67, 62.24)	75.64 (64.64, 89.72)
Weibull	1314.52	1320.54	shape: 0.8580; scale: 71.9442	48.56 (40.92, 57.83)	78.53 (66.18, 96.19)
L. logistic	1307.62	1313.64	shape: 1.2137; scale: 41.2077	47.04 (40.48, 55.07)	170.40 (101.92, 546.86
Lognormal	1311.86	1317.88	mu: 3.6938; sigma: 1.4663	45.14 (38.62, 53.14)	107.85 (81.60, 158.58)
Gompertz	1308.37	1314.39	shape: -0.0072; rate: 0.0192	46.58 (40.48, 54.18)	N/A (67.87, N/A)
Gen. Gamma	1311.36	1320.40	mu: 3.9363; sigma: 1.3518; Q: 0.3956	46.82 (39.95, 54.98)	87.58 (70.00, 137.85)
Gen. F	1310.92	1322.96	mu: 3.6906; sigma: 1.0000; Q: -0.0610; P: 1.9951	47.59 (40.63, 54.90)	N/A (82.57, N/A)

Figure 8. ATTRACTION-3 August 2020 database lock: taxane OS



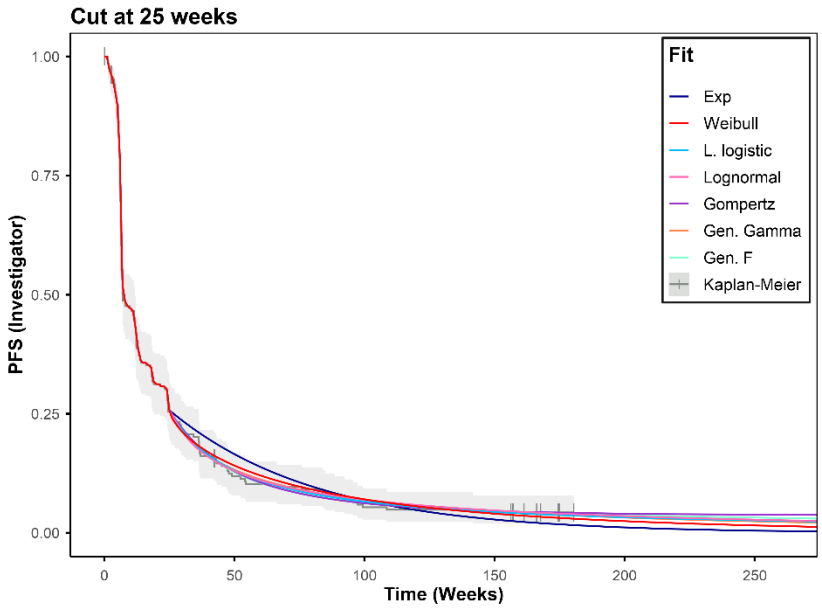
Fit Name	AIC	BIC	Parameters	Median (weeks)	Mean (weeks)
Exp.	1152.09	1154.97	lambda: 0.0183	39.40 (33.02, 45.46)	57.60 (49.60, 65.80)
Weibull	1148.97	1154.72	shape: 0.8484; scale: 52.1163	35.83 (30.21, 41.55)	59.00 (50.13, 69.94)
L. logistic	1145.12	1150.87	shape: 1.2330; scale: 29.8737	36.30 (31.40, 41.39)	110.62 (70.74, 357.90)
Lognormal	1152.31	1158.07	mu: 3.3453; sigma: 1.4821	34.57 (30.18, 39.43)	77.43 (59.64, 108.66)
Gompertz	1144.03	1149.78	shape: -0.0080; rate: 0.0255	35.81 (30.72, 41.02)	N/A (51.44, N/A)
Gen. Gamma	1148.08	1156.71	mu: 3.7098; sigma: 1.2998; Q: 0.5774	35.75 (30.74, 41.19)	61.59 (51.62, 81.00)
Gen. F	1147.09	1158.59	mu: 3.5121; sigma: 1.0000; Q: 0.2220; P: 1.7325	36.52 (31.23, 41.48)	106.25 (56.71, 501.64)

6.1.1.2. Progression-free survival

In order to model PFS in the nivolumab arm, Kaplan-Meier data was applied until 25 weeks followed by parametric extrapolation using the log-normal distribution to provide an appropriate fit. This approach predicted a median PFS of 7.3 weeks and a mean PFS of 44.0 weeks. When assessing the AIC and BIC, the log-normal distribution provided the best goodness-of-fit from the plausible distributions (log-logistic and gompertz distributions are deemed implausible), indicating it had a strong fit to the data, whilst this was also supported by a strong visual fit to the data.

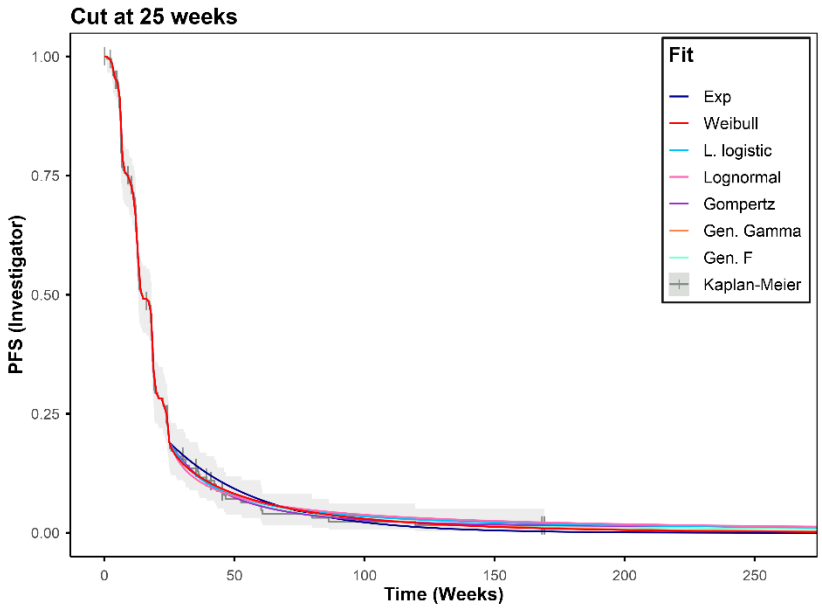
Similar to the nivolumab arm, Kaplan-Meier data was applied until 25 weeks for the taxane arm, however, a Weibull distribution followed for the extrapolation period. The Weibull distribution provided a clinically plausible estimation of the mean PFS (22.9 weeks), whilst also providing a strong fit to the data via the goodness-of-fit statistics.

Figure 9. ATTRACTION-3 August 2020 database lock: nivolumab PFS



Fit Name	AIC	BIC	Parameters	Median (weeks)	Mean (weeks)
Exp.	414.77	416.71	lambda: 0.0177	7.29 (6.71, 12.14)	27.56 (20.70, 35.94)
Weibull	407.67	411.54	shape: 0.7005; scale: 51.9358	7.29 (6.71, 12.14)	29.94 (21.46, 44.08)
L. logistic	404.17	408.04	shape: 1.0062; scale: 25.9476	7.29 (6.71, 12.14)	N/A (28.91, N/A)
Lognormal	405.16	409.02	mu: 3.2438; sigma: 1.7582	7.29 (6.71, 12.14)	43.98 (24.82, 107.03)
Gompertz	403.37	407.24	shape: -0.0175; rate: 0.0340	7.29 (6.71, 12.14)	N/A (20.29, N/A)
Gen. Gamma	406.93	412.73	mu: 3.4080; sigma: 1.6936; Q: 0.2198	7.29 (6.71, 12.14)	36.95 (22.59, 197.07)
Gen. F	407.49	415.22	mu: 3.0198; sigma: 1.0000; Q: -0.4023; P: 3.5606	7.29 (6.71, 12.14)	N/A (21.87, N/A)

Figure 10. ATTRACTION-3 August 2020 database lock: taxane PFS

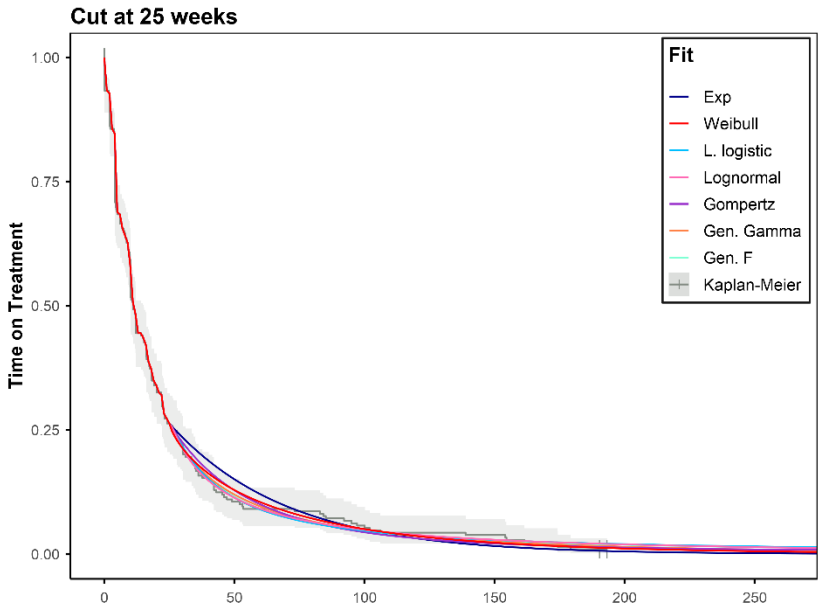


Fit Name	AIC	BIC	Parameters	Median (weeks)	Mean (weeks)
Exp.	229.70	231.17	lambda: 0.0286	14.57 (13.00, 18.14)	22.12 (18.33, 27.34)
Weibull	226.87	229.80	shape: 0.7280; scale: 32.0559	14.57 (13.00, 18.14)	22.92 (18.34, 31.92)
L. logistic	227.09	230.03	shape: 1.0363; scale: 17.6088	14.57 (13.00, 18.14)	107.87 (20.32, 302.18)
Lognormal	229.68	232.61	mu: 2.7725; sigma: 1.8471	14.57 (13.00, 18.14)	32.22 (20.23, 95.08)
Gompertz	225.69	228.63	shape: -0.0171; rate: 0.0460	14.57 (13.00, 18.14)	N/A (17.58, N/A)
Gen. Gamma	228.72	233.11	mu: 3.3490; sigma: 1.4421; Q: 0.8164	14.57 (13.00, 18.14)	23.21 (18.40, 54.23)
Gen. F	228.69	234.55	mu: 3.0918; sigma: 1.0000; Q: 0.4714; P: 2.6577	14.57 (13.00, 18.14)	54.98 (18.78, 130.69)

6.1.1.3. Time on treatment

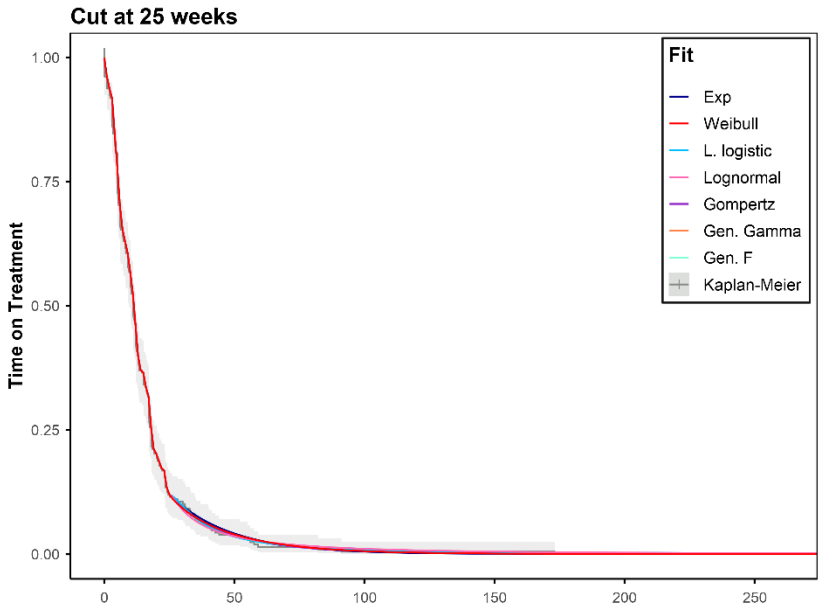
In order to model time on treatment in the nivolumab arm, Kaplan-Meier data was applied until 25 weeks followed by parametric extrapolation using the Weibull distribution to provide an appropriate fit. This approach predicted a median time on treatment of 11.1 weeks and a mean time on treatment of 25.3 weeks. When assessing the AIC and BIC statistics, the Weibull distribution a reasonable goodness-of-fit, whilst this was also supported by a strong visual fit to the data. Similar to the nivolumab arm, Kaplan-Meier data was applied until 25 weeks for the taxane arm, however, a log-logistic distribution followed for the extrapolation period. The log-logistic distribution provided a clinically plausible estimation of the mean time on treatment (16.3 weeks), whilst also providing a reasonable goodness-of-fit to the data.

Figure 11. ATTRACTION-3 August 2020 database lock: nivolumab time on treatment



Fit Name	AIC	BIC	Parameters	Median (weeks)	Mean (weeks)
Exp.	511.14	513.15	lambda: 0.0223	11.14 (10.14, 15.14)	25.09 (19.96, 31.02)
Weibull	506.17	510.19	shape: 0.7646; scale: 38.9120	11.14 (10.14, 15.14)	25.29 (20.03, 31.75)
L. logistic	506.12	510.14	shape: 1.1388; scale: 19.5941	11.14 (10.14, 15.14)	51.37 (24.85, 255.02)
Lognormal	505.79	509.81	mu: 2.9524; sigma: 1.5281	11.14 (10.14, 15.14)	29.49 (21.52, 44.27)
Gompertz	508.29	512.30	shape: -0.0079; rate: 0.0316	11.14 (10.14, 15.14)	N/A (20.83, N/A)
Gen. Gamma	506.36	512.38	mu: 3.2695; sigma: 1.4432; Q: 0.4357	11.14 (10.14, 15.14)	26.20 (20.69, 44.51)
Gen. F	510.04	518.07	mu: 3.0635; sigma: 1.2379; Q: 0.1320; P: 1.0000	11.14 (10.14, 15.14)	38.22 (22.28, 157.55)

Figure 12. ATTRACTION-3 August 2020 database lock: taxane time on treatment



Fit Name	AIC	BIC	Parameters	Median (weeks)	Mean (weeks)
Exp.	209.95	211.17	lambda: 0.0425	11.14 (10.13, 12.14)	14.88 (12.71, 17.33)
Weibull	211.46	213.90	shape: 0.9038; scale: 22.3378	11.14 (10.13, 12.14)	14.87 (12.70, 17.31)
L. logistic	211.72	214.16	shape: 1.4585; scale: 13.6067	11.14 (10.13, 12.14)	16.27 (13.05, 33.21)
Lognormal	215.08	217.52	mu: 2.5006; sigma: 1.3527	11.14 (10.13, 12.14)	15.71 (12.99, 19.89)
Gompertz	210.69	213.13	shape: -0.0073; rate: 0.0518	11.14 (10.13, 12.14)	N/A (12.57, N/A)
Gen. Gamma	213.06	216.72	mu: 2.9643; sigma: 1.1394; Q: 0.7464	11.14 (10.13, 12.14)	14.86 (12.78, 17.63)
Gen. F	214.28	219.15	mu: 2.8365; sigma: 1.0000; Q: 0.5564; P: 1.0000	11.14 (10.13, 12.14)	15.16 (13.15, 25.89)

6.1.2. Drug costs

In line with stated preferences in the Appraisal Consultation Document, drug costs in the model have been updated to use eMIT.

6.1.3. Hospitalisation cost

Based on the clinician survey detailed in the company submission, the model assumes that disease management requires a mean of 0.095 hospitalisations per week. However, the hospitalisation cost is derived from NHS National Cost Collection based on a weighted mean of hospitalisation costs, which have a length of stay ranging from 3 days (cost: £1,907) to 19 days (cost: £8,986). This cost is applied on a weekly basis, raising an implausible scenario where the weekly cost incurred is appropriate for a period of time of time longer than a week.

6.1.4. Utility values

The Committee considered it plausible for the utility before progression for nivolumab to be higher than the taxane arm, based on differences in tolerability and adverse events. Further, the Committee concluded that most of the overall survival benefit from nivolumab was after progression. However, the Committee concluded that the company had not given adequate justification for a long-term difference in utility after progression. As utility in oncology is typically a function of time to death, improved OS rates are a key component in postponing quality of life decrements.

Further, patients in the nivolumab arm frequently continued receiving nivolumab following progression, as noted in the ERG report. Hence, any beneficial impact associated with nivolumab treatment is continued into the post-progression state for those patients. Pooling post-progression quality of life data assumes that patients in the taxane arm receive benefit equivalent to patients receiving nivolumab.

Additional analysis was undertaken to assess the impact of treatment status on quality of life. Using a mixed effects model, as per ERG preference, data were stratified by treatment status. Collection of data was notably poorer in the off-treatment setting, leading to increased missing values. Hence, these values should be considered as supportive evidence. However, this clearly demonstrates the impact of treatment status may be greater than the impact of progression status, as demonstrated in Table 6.

Table 3. ATTRACTION-3 utility values by treatment status

	On initial treatment		Off initial treatment	
	Mean value	Standard error	Mean value	Standard error
Nivolumab	[redacted]	[redacted]	[redacted]	[redacted]
Taxane	[redacted]	[redacted]	[redacted]	[redacted]

Scenario analyses have been explored assessing treatment-independent utilities and utilities stratified by treatment status.

6.1.5. Impact of subsequent therapies

As noted in the ERG report, subsequent therapy (i.e. not allocated study therapy) was received by 119 (57%) of 210 patients in the nivolumab group and 115 (55%) of 209 patients in the taxanes group.[2] However, it should be noted that this has limited impact on survival outcomes, as demonstrated in Figure 6, as censoring patients who receive subsequent therapy does not greatly impact the comparison between nivolumab and taxanes. Hence, this is unlikely to impact on outcomes in the economic model.

Figure 13. ATTRACTION-3: Overall survival censored for subsequent therapy

[redacted]

6.2. Impact of alternative assumptions on cost-effectiveness outcomes

6.2.1. Updated base case analysis

The results of the base case analysis are summarised in Table 4.

In terms of comparator treatments (taxanes), the model predicts a median OS of 0.690 years, with an accrual of [redacted] discounted QALYs over the modelled time horizon. By comparison, it was predicted that the use of nivolumab will result in an additional 0.512 discounted QALYs (total: [redacted] discounted QALYs) and an additional 0.724 undiscounted life years (total: [redacted] undiscounted life years), respectively. It was estimated that patients receiving nivolumab would spend [redacted] years in the pre-progression health state (versus [redacted] for taxanes), with a

subsequent [redacted] years in the post-progression health state (versus [redacted] for taxanes), indicating that nivolumab is associated with incremental benefit across all health states.

Total discounted costs associated with nivolumab (with PAS), accrued over the modelled time horizon, were predicted to be £[redacted]. By comparison, total discounted costs associated with taxanes were notably lower, predicted to be £[redacted]. Incremental discounted costs were predicted to be £24,665 over taxanes, under base case assumptions. The resultant ICER estimate for nivolumab versus taxanes was £48,205 per QALY gain. Therefore, the base case ICER is below a £50,000 per QALY wiliness-to-pay threshold when the current nivolumab PAS discount is applied.

Table 4. Base case analysis results (with PAS, lifetime horizon)

	Nivolumab	Taxanes
Patient-level survival (undiscounted)		
Median ToT (years)	0.230	0.230
Mean ToT (years)	0.498	0.318
Median PFS (years)	0.153	0.287
Mean PFS (years)	0.703	0.447
Median OS (years)	0.901	0.690
Mean OS (years)	1.848	1.125
Patient-level progression		
Time in pre-progression (years)	[redacted]	[redacted]
- Time initial therapy (years)	[redacted]	[redacted]
- Time in subsequent therapy (years)	[redacted]	[redacted]
Time in post-progression (years)	[redacted]	[redacted]
Costs (with PAS)		
HS costs	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]
BSC costs	[redacted]	[redacted]
Average AE costs per patient	[redacted]	[redacted]
Total costs	[redacted]	[redacted]
Health benefits		
Total QALYs	[redacted]	[redacted]
Total life years (undiscounted)	[redacted]	[redacted]
Incremental results		
Incremental total costs	-	£24,665
Incremental QALYs	-	0.512
Incremental life years (undiscounted)	-	0.724
ICER		
Cost/QALY	**-**	**£48,205**

AE: adverse event; BSC: best supportive care; HS: health state; ICER: incremental cost-effectiveness ratio; OS: overall survival; PAS: patient access scheme; PFS: progression-free survival; QALY: quality-adjusted life year; ToT: time on treatment

6.2.2. Alternative survival extrapolations

In order to assess the impact of alternative parametric fittings on the cost-effectiveness of nivolumab, alternative survival curves based on the updated ATTRACTION-3 data have been applied within the model as scenario analyses.

All extrapolations have been assessed for completeness. However, it should be noted that several of these extrapolations are not considered appropriate. Clinically implausible fits are presented in grey italics and are defined as extrapolations that exceed the 95% confidence intervals of the Kaplan-Meier data or provides mean survival that cannot be considered plausible.

The impact of applying alternative survival extrapolations for the nivolumab and taxane arms (OS, PFS and time on treatment) is shown in Table 5. Predicted discounted incremental QALYs ranged from 0.409 to 0.512; while PFS extrapolations did not greatly impact on the QALY gains, OS extrapolations had a large impact, with shorter extrapolations reducing survival benefit; conversely, longer extrapolations increasing QALY accrual. There was a similar variation in discounted incremental costs ranging from £22,826 to £28,289. This had an associated impact on ICERs versus taxanes, which ranged between £48,205 per QALY and £56,959 per QALY.

Table 5. Scenario analysis: impact of alternative extrapolations using updated ATTRACTION-3 database lock

Scenario			Incremental QALYs	Incremental Costs (£)	ICER (£/QALY)
PFS: Semi-parametric with Kaplan-Meier to 5.75 months	Nivolumab	*Exponential*	*0.491*	*£24,665*	***£50,257***
		Generalised Gamma	0.506	£24,665	**£48,727**
		Gompertz	*0.537*	*£24,665*	***£45,900***
		Log-logistic	*0.512*	*£24,665*	***£48,218***
		Log-normal	0.512	£24,665	**£48,205**
		Weibull	0.496	£24,665	**£49,713**
	Taxanes	*Exponential*	*0.514*	*£24,665*	***£47,991***
		Generalised Gamma	0.511	£24,665	**£48,280**
		Gompertz	*0.507*	*£24,665*	***£48,681***
		Log-logistic	*0.504*	*£24,665*	***£48,914***
		Log-normal	0.503	£24,665	**£49,057**
		Weibull	0.512	£24,665	**£48,205**
OS: Semi-parametric with Kaplan-Meier to 5.75 months	Nivolumab	*Exponential*	*0.316*	*£22,563*	***£71,365***
		Generalised Gamma	0.413	£23,551	**£56,959**
		Gompertz	*0.666*	*£26,435*	***£39,690***
		Log-logistic	0.512	£24,665	**£48,205**
		Log-normal	0.506	£24,605	**£48,583**
		Weibull	0.344	£22,826	**£66,451**
	Taxanes	*Exponential*	*0.521*	*£24,795*	***£47,557***
		Generalised Gamma	0.494	£24,418	**£49,462**
		Gompertz	*0.364*	*£22,650*	***£62,246***
		Log-logistic	0.409	£23,263	**£56,875**
		Log-normal	0.414	£23,336	**£56,347**
		Weibull	0.512	£24,665	**£48,205**
Time on treatment: Semi-parametric with Kaplan-Meier to 5.75 months	Nivolumab	*Exponential*	*0.512*	*£24,595*	***£48,069***
		Generalised Gamma	0.512	£25,232	**£49,314**
		Gompertz	*0.512*	*£28,261*	***£55,235***
		Log-logistic	0.512	£28,289	**£55,289**
		Log-normal	0.512	£26,967	**£52,705**
		Weibull	0.512	£24,665	**£48,205**
	Taxanes	Exponential	0.512	£24,810	**£48,490**
		Generalised Gamma	0.512	£24,814	**£48,497**
		Gompertz	*0.512*	*£24,797*	***£48,464***
		Log-logistic	0.512	£24,665	**£48,205**
		Log-normal	0.512	£24,682	**£48,240**
		Weibull	0.512	£24,813	**£48,495**

ICER: incremental cost-effectiveness ratio; OS: overall survival; PFS: progression-free survival; QALY: quality-adjusted life year

Grey italics denotes highly implausible extrapolations, defined as extrapolations that exceed the 95% confidence intervals of the Kaplan-Meier data or provides mean survival that cannot be considered plausible.

6.2.3. Alternative utility values

In order to assess the impact of utility values on the cost-effectiveness of nivolumab, scenario analyses have been undertaken using alternative utility values. Results from the analysis is detailed in Table 6, where application of alternative utilities resulted in ICER estimates ranging between £50,580 per QALY to £59,995 per QALY.

Table 6. Impact of alternative utilities on base case analysis

Scenario	Pre-progression	Post-progression	Incremental QALYs	Incremental costs (£)	ICER (£/QALY)
Base case analysis CS values					
Nivolumab	[redacted]	[redacted]	0.512	£24,665	**£48,205**
Taxanes	[redacted]	[redacted]			
Base case analysis ERG values					
Nivolumab	[redacted]	[redacted]	0.411	£24,665	**£59,955**
Taxanes	[redacted]				
ERG values with non-pooled post-progression values					
Nivolumab	[redacted]	[redacted]	0.485	£24,665	**£50,850**
Taxanes	[redacted]	[redacted]			

Additionally, scenario analyses were undertaken assessing the impact of utility values stratified by initial treatment status. Results from the analysis is detailed in Table 7, where application of alternative utilities resulted in ICER estimates ranging between £46,448 per QALY and £50,042 per QALY.

Table 7. Impact of using on-treatment and off-treatment utilities on base case analysis

Scenario	On-treatment/ Pre-progression	Off-treatment/ Post-progression	Incremental QALYs	Incremental costs (£)	ICER (£/QALY)
On and off-treatment utilities in both treatment arms					
Nivolumab	On treatment: [redacted]	Off-treatment [redacted]	0.493	£24,665	£50,042
Taxanes	On treatment: [redacted]	Off-treatment: [redacted]			
On and off-treatment utilities in nivolumab arm only					
Nivolumab	On treatment: [redacted]	Off-treatment [redacted]	0.531	£24,665	£46,448
Taxanes	Pre-progression: [redacted]	Off-treatment: [redacted]			

7. References

1. National Institute for Health and Care Excellence. Appraisal consultation document - Nivolumab for previously treated unresectable advanced or recurrent oesophageal cancer. 2020.
2. Kato K, Cho BC, Takahashi M, Okada M, Lin CY, Chin K, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. The Lancet Oncology. 2019;20(11):1506-17.
3. Chin K, editor Three-year follow-up of ATTRACTION-3: A phase III study of nivolumab (Nivo) in patients with advanced esophageal squamous cell carcinoma (ESCC) that is refractory or intolerant to previous chemotherapy2021; Gastrointestinal Cancers Symposium: American Society of Clinical Oncology.
4. Borcoman E, Nandikolla A, Long G, Goel S, Le Tourneau C. Patterns of Response and Progression to Immunotherapy. American Society of Clinical Oncology Educational Book. 2018(38):169-78.
5. Chiou VL, Burotto M. Pseudoprogression and immune-related response in solid tumors. Journal of Clinical Oncology. 2015;33(31):3541-3.
6. Cancer Research UK. Oesophageal cancer incidence statistics. 2019 [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence.
7. NHS England. Improving care for older people 2019 [Available from: https://www.england.nhs.uk/ourwork/clinical-policy/older-people/improving-care-for-older-people/.

NIVOLUMAB FOR UNRESECTABLE, ADVANCED OESOPHAGEAL CANCER WHEN STANDARD CHEMOTHERAPY HAS FAILED [ID1249]

A Single Technology Appraisal

Addendum #1

ERG response to ACD consultation

December, 2020

Produced by

Peninsula Technology Assessment Group (PenTAG)
University of Exeter Medical School
South Cloisters
St Luke's Campus
Heavitree Road
Exeter
EX1 2LU

Authors

Mr Ash Bullement, Associate[1] and Analyst[2]
Dr Linda Long, Research Fellow[1]
Dr Kevin Deighton, Associate[1] and Analyst[2]
Ms Naomi Shaw, Information Specialist[1]
Dr Stephen Falk, Consultant Clinical Oncologist[3]
Dr Nicole Dorey, Consultant Clinical Oncologist[3]
Ms Louise Crathorne, Senior Research Fellow[1]
Prof G.J. Melendez-Torres, Professor[1]
Dr Maxwell S. Barnish, Research Fellow[1]

[1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter
[2] Delta Hat Limited, Nottingham

[3] Bristol Haematology and Oncology Centre, Bristol

[4] Royal Devon and Exeter NHS Foundation Trust, Exeter

Correspondence to Dr Maxwell S. Barnish

3.09f South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; m.s.barnish@exeter.ac.uk

Source of funding This report was commissioned by the NIHR Systematic Reviews Programme as project number 18/54/01.

Declared competing interests of the authors Dr Dorey declares educational support from Boehringer-Ingelheim; Roche; Lilly and Astra-Zeneca.

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

This addendum is linked to ERG report Bullement A, Long L, Deighton K, Shaw N, Falk S, Dorey N, Crathorne L, Melendez-Torres GJ, Barnish MS. Nivolumab for unresectable, advanced oesophageal cancer when standard chemotherapy has failed [ID1249]. Peninsula Technology Assessment Group (PenTAG), 2020.

Abbreviations

ACD	Appraisal Consultation Document
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
DBL	database lock
eMIT	electronic market information tool
ERG	Evidence Review Group
gen	Generalized
HTA	health technology assessment
ICER	incremental cost-effectiveness ratio
LYs	life years
MIMS	Monthly Index of Medical Specialities
NHS	National Health Service
OS	overall survival
PD	progressed disease
PF	progression free
PFS	progression free survival
QALYs	quality-adjusted life years

1. INTRODUCTION

The purpose of this addendum is to review and critique additional clinical- and cost-effectiveness data provided by the company from the updated August 2020 database lock (DBL).

2. UPDATED CLINICAL EFFECTIVENESS ANALYSIS

In the company's response to the ACD, updated clinical effectiveness analysis results are provided from the August 2020 DBL. This section of the ERG's response discusses the updated clinical effectiveness evidence provided by the company.

2.1. Summary of updated clinical effectiveness data

The company has presented updated clinical effectiveness results for the efficacy outcomes. These are described below in Table 1.

Table 1. ATTRACTION-3 updated outcomes

Endpoint	Nivolumab	Taxanes
Number evaluable	210	209
Overall survival (OS)		
Median, months (95% CI)	10.91 (9.23, 13.34)	8.51 (7.29, 9.86)
HR (95% CI)	0.79 (0.64, 0.97)	
p-value	p = 0.0264	
OS Rates (95% CI), %		
12-month	46.9	34.7
	[redacted]	[redacted]
24-month	20.2	13.5
	[redacted]	[redacted]
36-month	15.3	8.7
	[redacted]	[redacted]
Investigator-assessed progression-free survival (PFS)		
Median, months (95% CI)	1.68 (1.51, 2.73)	3.35 (2.99, 4.21)
HR (95% CI)	1.07 (0.87, 1.33)	
p-value	[redacted]	
PFS Rates (95% CI), %		
12-month	11.9	7.2
	[redacted]	[redacted]
24-month	5.4	2.4
	[redacted]	[redacted]
36-month	4.3	1.6
	[redacted]	[redacted]

Endpoint	Nivolumab	Taxanes
Investigator-assessed objective response rate (ORR)		
Number evaluable	171	158
Responders, n (%)	33 (19.3)	34 (21.5)
95% CI	[redacted]	[redacted]

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

Source: Company ACD Response, Table 1, p.4.

2.2. Review of updated clinical effectiveness data

The ERG was satisfied with the company's claim in the ACD response that the results from the updated August 2020 DBL were consistent with the results of the primary analysis from the original company submission, and that therefore the conclusions drawn from the clinical effectiveness evidence in the original ERG report still hold.

The estimated hazard ratio (HR) for overall survival (OS) is largely consistent with the previous DBL, though as the curves cross this measure should be interpreted with caution. In the previous DBL, 24-month OS was estimated to be 19.1% versus 15.1% (CS, Table 13), which has changed to 20.2% versus 13.5% (Table 1). The HR for progression-free survival (PFS) should also be interpreted with caution; however, it is consistent with the previous DBL. The objective response rate is unchanged from the previous analysis.

The concerns expressed by the ERG in its original report, for example around the generalisability of the ATTRACTION-3 trial to UK clinical practice still hold. The ERG noted that the company did not address this matter in its ACD response and has not provided updated subgroup results for Japan versus the rest of the world. The ERG also noted that the company did not provide updated adverse event data, including relating to early deaths, which have been noted to be a key issue in this appraisal.

3. UPDATED COST-EFFECTIVENESS ANALYSIS

In the company's response to the ACD, updated cost-effectiveness analysis results are provided. This section of the ERG's response discusses the changes made to the model and the impact on results.

3.1. Summary of changes

In the company's revised base-case analysis, several changes have been made to the original base-case analysis. These are described in Table 2, alongside the ERG's previous preferred base-case analysis settings.

Table 2: Summary of base-case analysis changes

Model feature	Original base case		Revised base case
	Company	**ERG**	**Company**
OS (nivolumab)	Semi-parametric model Cut at 2.99 months Log-logistic extrapolation November 2018 DBL	Semi-parametric model Cut at 5.75 months Gen gamma extrapolation November 2018 DBL	Semi-parametric model Cut at 5.75 months Log-logistic extrapolation August 2020 DBL
OS (taxanes)	Semi-parametric model Cut at 2.99 months Exponential extrapolation November 2018 DBL	Semi-parametric model Cut at 5.75 months Gen gamma extrapolation November 2018 DBL	Semi-parametric model Cut at 5.75 months Weibull extrapolation August 2020 DBL
PFS (nivolumab)	Semi-parametric model Cut at 2.99 months Weibull extrapolation November 2018 DBL	Per company's original base-case analysis	Semi-parametric model Cut at 5.75 months Log-normal extrapolation August 2020 DBL
PFS (taxanes)	Semi-parametric model Cut at 2.99 months Weibull extrapolation November 2018 DBL	Per company's original base-case analysis	Semi-parametric model Cut at 5.75 months Weibull extrapolation August 2020 DBL
ToT (nivolumab)	Fully-parametric model Gen gamma extrapolation November 2018 DBL	Semi-parametric model Cut at 5.75 months Weibull extrapolation November 2018 DBL	Semi-parametric model Cut at 5.75 months Weibull extrapolation August 2020 DBL
ToT (taxanes)	Fully-parametric model Gen gamma extrapolation	Semi-parametric model Cut at 5.75 months	Semi-parametric model Cut at 5.75 months

Model feature	Original base case		Revised base case
	Company	**ERG**	**Company**
	November 2018 DBL	Weibull extrapolation November 2018 DBL	Log-logistic extrapolation August 2020 DBL
Taxanes acquisition costs	MIMS	eMIT	eMIT
Hospitalisation cost	£534.07	£3,379.73	£534.07 (per company's original base-case analysis)
Utility values	Nivolumab, PF: [redacted] Nivolumab, PD: [redacted] Taxanes, PF: [redacted] Taxanes, PD: [redacted]	Nivolumab, PF: [redacted] Nivolumab, PD: [redacted] Taxanes, PF: [redacted] Taxanes, PD: [redacted]	Per company's original base-case analysis (some additional scenarios provided)

Key: DBL = database lock; eMIT = electronic market information tool; gen = generalised; MIMS = Monthly Index of Medical Specialities; OS = overall survival; PD = progressed disease; PF = progression-free; PFS = progression-free survival; ToT = time on treatment.

3.2. Review of the updated model provided

The updated model included the specification of a new function on the patient flow sheet (_xlfn.SWITCH) which caused the model to produce errors when run on the ERG's machines. The switch was introduced to allow for an alternative utility analysis provided by the company where utility values were based on treatment status as opposed to progression status.

It is the ERG's understanding that this function is a relatively recent addition to Excel, and is therefore not fully compatible with all versions of Excel. To address this, the ERG simply edited the necessary functions to instead use an IF statement. After making this change to the model, the ERG was able to successfully replicate the company's preferred base-case analysis.

As the model has been updated to replace prior functionality, the ERG could not replicate both the original company-preferred base-case analysis and the updated company-preferred base-case analysis within the same model file. However, after replacing the cost of taxanes with the company's original values, and pasting the outputs from the survival curves from the original model into the revised model patient flow sheets, the ERG obtained the same results as per the previously submitted model. Therefore, while a comparison to the previous base-case analysis would have been preferred within the same model file, the ERG is confident that the updated model is consistent with the original model file provided.

The model provided to the ERG allowed for the exploration of some, but not all, scenarios discussed at some stage either in the CS, the ERG report, or the company's response to the ACD. Further information concerning the reproducibility of specific scenario analyses is provided within this document, along with a description of which model edits the ERG was able to incorporate within its updated, preferred base-case analysis.

3.3. Updated survival analysis

In the company's response to the ACD, updated data from the ATTRACTION-3 August 2020 DBL are provided. Using these updated data, the company re-fitted the parametric survival models for the outcomes of OS) PFS, and time-on-treatment (ToT) used within the cost-effectiveness analysis (for both treatment arms). These three outcomes are discussed in turn in the following sub-sections.

3.3.1. Overall survival

The ERG agrees with the company that, in light of the updated data from the August 2020 DBL, the company's and the ERG's previously preferred models for both the nivolumab and taxanes arm provide OS estimates which are lower than the updated Kaplan-Meier curve evaluated at 36 months (Company ACD response, Table 2).

In the previous November 2018 DBL, the maximum follow-up time was approximately 34 months (based on CS Figure 10), and so it was previously not possible to compare extrapolations at the landmark of 36 months. The August 2020 DBL includes a minimum follow-up of 36.04 months which means the Kaplan-Meier curve can be considered reasonably stable up until this time point (within the context of the trial). The number of patients still at risk for the outcome of OS at 36 months is, however, relatively small: n=31 (nivolumab) versus n=16 (taxanes). Based on the ERG's previous concerns regarding the generalisability of the ATTRACTION-3 trial population to the UK NHS population, it should be noted that the Kaplan-Meier curve is not equivalent to the "true" survival curve (given that it is based on a finite sample of patients, potential issues with generalisability etc.).

The ERG acknowledges the company's decision to adhere to the ERG's previously preferred cut-point at 5.75 months, though within the context of updated data, it is not necessarily the case that the ERG would continue to prefer this choice of cut-point specifically. Alternative cut-points were not provided in the company's ACD response (or fully parametric models), and so the ERG cannot rule out the possibility that alternative models may provide a relatively better fit.

In spite of this, it is the ERG's view that a semi-parametric approach is likely to provide at least a similar (if not better) fit than a fully-parametric approach, as the semi-parametric approach was shown to provide a better fit to the previous DBL.

The choice of the most optimal cut-point for the August 2020 DBL would ideally be based on an exploration of different options. The more mature data (affecting numbers at risk particularly at the tail-end of the Kaplan-Meier curve) may allow for models fitted from different cut-points to provide better or worse overall fits. The ERG highlights that its preference for the 5.75-month cut-point was guided predominantly by overall visual fit to the November 2018 DBL Kaplan-Meier curve. Given that the Kaplan-Meier curve has now changed, a revised assessment of the most appropriate cut-point would ideally have been performed (but was not presented in the company's ACD response).

The company's revised base-case analysis involves the specification of the following preferred models:

- **Nivolumab:** Semi-parametric approach, with Kaplan-Meier curve used until 5.75 months (25 weeks), followed by a *log-logistic* model
- **Taxanes:** Semi-parametric approach, with Kaplan-Meier curve used until 5.75 months (25 weeks), followed by a *Weibull* model

For ease of comparison, the ERG has superimposed the company's updated preferred models for overall survival on top of the corresponding August 2020 DBL Kaplan-Meier curves, presented in Figure 1. The ERG notes that this is a relatively crude approach to comparing the Kaplan-Meier curves to the fitted models, but was necessary due to the limited time available for the ERG to perform its critique, and the fact that the company did not provide the Kaplan-Meier curve data points within the update model.

Figure 1: Comparison of updated ATTRACTION-3 database lock versus revised company extrapolations of overall survival

Figure redacted

It is the ERG's view that the choice of model for the nivolumab arm seems broadly appropriate, but the corresponding choice of model for the control group (taxanes) does not appear to provide a particularly good fit to the Kaplan-Meier curve. Unfortunately, an in-depth assessment of the updated Kaplan-Meier curves was not provided as part of the company's response to the ACD (e.g. including provision of hazard plots), and so a comprehensive assessment of the updated data from ATTRACTION-3 by the ERG was not possible. However, visual inspection of the Kaplan-Meier curves for both treatment arms after the cut point does not suggest a substantially different pattern of hazards that would necessitate the selection of two fundamentally very different parametric models.

A plot comparing the ERG's original preferred extrapolation arm and the company's revised extrapolation for the nivolumab arm is provided in Figure 2. The ERG acknowledges that its previously-preferred extrapolation does not fit the tail-end of the Kaplan-Meier curve for the nivolumab arm well, and even crosses the Kaplan-Meier curve for the taxanes arm (though the tails of both curves are palpably uncertain).

The ERG considers that the company's revised choice of model for the nivolumab arm provides a *relatively* better fit versus the ERG's original preferred extrapolation. In spite of this, as plots of alternative models were provided only as an image (where seven models overlapped), and

other model specifications (e.g. different cut-points) were not provided, the ERG cannot rule out the potential for other choices of model to be a more appropriate selection.

Figure 2: Comparison of ERG's original preferred extrapolation for overall survival (nivolumab arm) versus revised company's base-case extrapolation

Figure redacted

Key: ERG = Evidence Review Group.

The corresponding plot for the taxanes arm is presented in

Figure 3, from which it may be argued that in places, the ERG's original preferred model provided a better fit than the company's updated model. For example, it can be seen that OS is overestimated by the company's revised model for the taxanes arm between approximately [redacted] and [redacted] months, after which the extrapolated tail relatively quickly (compared with the nivolumab arm) approaches zero. While the ERG's previously preferred extrapolation for the taxanes arm was shown to under-estimate the latter portion of the Kaplan-Meier curve, it provided a notably much better fit to the earlier portion of the Kaplan-Meier curve, which can be seen in

Figure 3.

Based predominantly on

Figure 3, it is the ERG's view that the parametric model selected for the taxanes arm is not sufficiently flexible to reflect the underlying pattern of hazards, and that an alternative choice of model may provide a better fit. However, as noted with respect to the nivolumab arm, the ERG was unable to explore alternative models due to the lack of options provided within the company's updated model.

Figure 3: Comparison of ERG's original preferred extrapolation for overall survival (taxanes arm) versus revised company's base-case extrapolation

Figure redacted

Key: ERG = Evidence Review Group.

In selecting its preferred base-case models, the company stated that *"fits predicting mean OS* [for the taxanes arm] *greater than 104 weeks were considered implausible based on clinical expert opinion"* (Company ACD response, page 13). The ERG highlights that the criticality of the mean survival of 2 years has not been previously highlighted by the company, other than with respect to the expectation that most patients have a survival that is less than 2 years (in relation to NICE's end-of-life criteria).

The ERG questions whether or not it is truly *"clinically implausible"* for the taxanes group to achieve a mean survival of greater than 104 weeks, given that the company's preferred extrapolation for the nivolumab arm estimated mean survival of [redacted] weeks (Company ACD response, Figure 8), excluding adjustment for background mortality (which is a separate consideration within the cost-effectiveness model). No corresponding upper limit is specified in relation to the nivolumab group.

In the company's base-case analysis, the baseline survival (for the taxanes group) is estimated to be [redacted] weeks, meaning that nivolumab is estimated to add an additional [redacted] weeks of survival (Company ACD response, Figure 9), equivalent to a relative improvement of [redacted] ([redacted]). The ERG stresses that these values are independent of adjustment according to background mortality, but

nevertheless imply substantially different long-term projections of survival in the choice of parametric model.

To further illustrate this, the ERG highlights the difference in extrapolated tail between the two treatment arms, shown in Figure 4 (which also accounts for background mortality). This projection illustrates that at 5 years, [redacted] of the nivolumab group and [redacted] of the taxanes group are expected to still be alive. At 10 years, these percentages fall to [redacted] and [redacted], respectively.

Figure 4: Comparison of extrapolated tails for overall survival in revised company's base-case extrapolation

Figure redacted

The ERG notes the provision of statistical goodness-of-fit scores for the different models, based on Akaike's and Bayesian information criteria (AIC and BIC, respectively – see company ACD response Figures 8 and 9). For the nivolumab arm, the company explains *"*[based on the AIC and BIC] *the log-logistic distribution provided the best goodness-of-fit, indicating it had a strong fit to the data, whilst this was also supported by a strong visual fit to the data, capturing the hazard of the tail of the Kaplan-Meier."*

However, the AIC and BIC scores for the taxanes arm are not discussed. Instead, it may be inferred from the company's response to the ACD that three models (log-logistic, Gompertz, and generalised F) were considered clinically implausible (due to their prediction of a mean survival beyond 104 weeks), although it may be noted that the log-logistic model has the lowest AIC and BIC score (after excluding the Gompertz model based on this projecting a proportion of patients

to survive indefinitely). This implies that long-term plausibility aside, the log-logistic model provides the best representation of the survival for patients treated with taxanes.

Of the four models deemed clinically plausible by the company (exponential, Weibull, lognormal, and generalised gamma), the Weibull model has the lowest AIC and BIC score. However, the remaining models each have similar scores (e.g. generalised gamma has AIC and BIC scores each within two points of the Weibull model).

In summary, it is the ERG's view that the choice of parametric model for the taxanes arm in particular is likely sub-optimal, and therefore the ERG considers further exploration of models for this arm in particular is necessary. The ERG cannot determine the most appropriate choice of model based on the information provided in the company's ACD response.

3.3.2. Progression-free survival

The ERG previously noted that the choice of PFS model has a limited impact on cost-effectiveness results. Based on Table 5 of the company's ACD response, the choice of model for PFS continues to have a relatively limited impact on cost-effectiveness results. However, as with OS, the cost-effectiveness model provided by the company does not allow for alternative models to be selected, and so only a limited commentary on the PFS models is provided here.

In the original base-case analysis, the company used a semi-parametric model (cut at 2.99 months) with a Weibull extrapolation for both treatment arms. In the revised base-case analysis, the company has used the cut-point at 5.75 months, with a Weibull model for the taxanes arm, but a log-normal extrapolation for the nivolumab arm. The ERG highlights that if a Weibull model were selected for both arms, the ICER would increase from £48,205 to £49,713. However, based on the information provided by the company, the ERG cannot appropriately comment on which model appears to provide the best overall fit.

3.3.3. Time on treatment

As noted in the ERG's report, ToT has a greater influence on cost-effectiveness results compared to PFS. In the company's original base-case analysis, a fully-parametric generalised gamma model was selected for both treatment arms. In the updated analysis, a semi-parametric model (cut at 5.75 months) was selected for both arms, with a Weibull extrapolation for the nivolumab arm, and a log-logistic extrapolation for the taxanes arm.

The ERG considers this selection of models to be inappropriate, especially when considered alongside the choice of models for the outcome of OS. This is because the company has selected a model with a "heavy tail" for the outcome of OS for the nivolumab group, but a model without this feature for the outcome of ToT. Conversely, the company has selected the opposite models for the taxanes group (i.e. a "heavy tail" for ToT, but not for OS). The rationale for this particular combination of models is not provided by the company within its response to the ACD.

For the ToT model fitted to the taxanes arm, the company's response to the ACD states: *"The log-logistic distribution provided a clinically plausible estimation of the mean time on treatment (16.3 weeks), whilst also providing a reasonable goodness-of-fit to the data."* It is unclear to the ERG why 16.3 weeks is considered to *"clinically plausible"*, yet the range for the other models ([redacted] to [redacted] weeks) is considered less clinically plausible, or perhaps even clinically implausible (though, again, this is unclear based on the information presented in the company's ACD response).

Based on Figure 13 from the company's response to the ACD, there is almost no difference between the visual fit of the models, but the projected mean ToT differs based on the extrapolated tail. The log-logistic model predicts the largest mean duration of treatment out of all the models considered (except the Gompertz model which was considered implausible due to it extrapolating indefinitely). The Weibull model has a better AIC and BIC score than the log-logistic, and predicts a mean duration of treatment ([redacted] weeks) very similar to both the generalised gamma ([redacted] weeks) and exponential ([redacted] weeks) models. When selecting the Weibull model for both arms, a small increase in the ICER is noted (from £48,205 to £48,495).

For the nivolumab arm, a selection of any of the other three models deemed to be "plausible" causes the ICER to increase. The base-case ICER (£48,205) increases to £49,314 if the generalised gamma model is selected, and increases further to £52,705 or £55,289 if the log-normal or log-logistic models are selected, respectively. The generalised gamma predicts a mean ToT of [redacted] weeks, versus [redacted] weeks for the Weibull model – a difference in the mean of less than [redacted] which casues the ICER to increase by over £1,000. It is the ERG's view that given the influence on the ICER, both of these models (as well as the other choices of model deemed plausible) may be important to consider in decision making.

In summary, as with the ERG's perspective on the outcome of OS, it is the ERG's view that the choice of parametric models for ToT may be sub-optimal, and further exploration of models is

necessary. The ERG cannot determine the most appropriate models based on the information provided in the company's ACD response.

3.4. Drug and administration costs

The ERG agrees with the company's update to the unit costs of taxanes per eMIT.

The ERG highlights that another change made by the ERG within its preferred base-case analysis was the specification of different costs for administration. These changes were made for two reasons:

1. Administration of nivolumab is expected to take place over 30 minutes, versus at least 60 minutes for taxanes. Therefore, the ERG's preferred base-case analysis included the specification of a higher administration cost for taxanes versus nivolumab (to reflect this difference in chair time).
2. The company's base-case analysis assumed administration would take place in an outpatient setting. However, clinical advice to the ERG was that administration would predominantly take place in a day case setting. Therefore, the ERG's preferred costs were based on administration taking place in a day case setting.

In the company's revised base-case analysis, the administration costs for both nivolumab and taxanes are as per its original base-case analysis. The ERG has seen no evidence to change its preference for the administration costs per its base-case analysis.

3.5. Hospitalisation cost

As noted in the company's ACD response, the submitted model assumes a mean of 0.095 hospitalisations per week. The company comments that the cost taken from NHS National Cost Collection is based on an *average* hospital stay, which have a length ranging from 3 days (at a total cost of £1,907; equivalent to £635.67 per day) to 19 days (at a total cost of £8,986; equivalent to £472.95 per day). However, it should be noted that the method and source used for calculating these length of stay values have not been provided by the company, and therefore cannot be reproduced by the ERG. Within the context of the model, the cost is applied on a weekly basis, and so the company explains that because of this, the cost for hospitalisations could extend beyond the model cycle length, meaning that costs incurred within a given week could actually apply beyond that week itself.

The ERG refers to CS Table 69. The surrounding text explains that a clinician survey was administered, part of which asked respondents to provide estimates of resource use associated with disease management. An excerpt of this table is re-produced in Table 3, focusing on hospitalisations – one of the medical resource use items respondents were asked about.

Table 3: Hospitalisation based on company's clinical expert survey

Frequency	**N**	**%**
Every 3 months	21	53%
Monthly	9	23%
Biweekly	3	8%
Weekly	2	5%
Never	5	13%
Mean frequency per week	**0.095**	

Based on the ERG's understanding of the survey, the results show (for example) that over half of clinicians expect patients to be hospitalised once every three months (53%), whereas only 5% expect patients to be hospitalised every week. However, when averaging over the full set of responses, the mean number of hospitalisations per week is calculated to be 0.095. This value is independent of the length of stay, as based on the information presented in the CS, clinicians were not asked how many days patients spend in hospital. Instead, respondents were asked how often patients would be admitted.

The ERG therefore understands the value of 0.095 to mean the average patient incurs the cost of 9.5% of a hospital stay per week. It should be noted however, that the ERG is not able to re-produce the value of 0.095 in Table 3 using the information presented in the CS.

Again, based on the ERG's understanding, the implausibility point raised by the company is that some of the hospitalisation costs incurred within a given model cycle may include some cost that extends into the next cycle. However, rather than being justification for considering a smaller hospitalisation cost, this is instead a problem either with the company's choice of model cycle length or the elicitation of a given rate for hospitalisation (i.e. an additional question could have been how long the average length of stay is, in order to appropriately estimate the mean number of bed days per model cycle).

The ERG still does not believe sufficient justification has been provided by the company for applying a hospitalisation cost equivalent to a length of stay of only one day (which is calculated by the company as £534.07, based on a weighted average of the cost codes used to inform the model). The ERG is unable to reproduce the value of £534.07 using the information presented in the CS, with the reference[1] provided in the company response to clarification questions detailing that: *"The unit cost of day case, elective inpatient and non-elective inpatient is per finished consultant episode (FCE)"*. This does not provide the length of stay as required to calculate the cost per day.

It also remains unclear to the ERG why the company has chosen to convert long-stay costs to a cost per day (effectively converting these costs into an equivalent short-stay admission), rather than simply using the short-stay costs available within the cited National Cost Collection for the NHS 2018/19.[2] This would allow for a more appropriate costing of a 1-day length of stay, if it is expected that all hospital admissions are for exactly one-day. However, it is the ERG's understanding that not all patients would be admitted for only one day, and that some patients would have a length of stay longer than one day.

As a simple calculation using the values provided in Table 3, the average yearly hospitalisation costs can be estimated using the following formulae, with an average unit hospitalisation cost of £3,379.73 (see ERG report, Section 4.2.8.3):

- Every 3 months: 4 per year x 52.5% x £3,379.73 = £7,097.44
- Every month: 12 per year x 22.5% x £3,379.73 = £9,125.28
- Biweekly: 26.09 per year x 7.5% x £3,379.73 = £6,613.11
- Weekly: 52.18 per year x 5.0% x £3,379.73 = £8,817.48
- Never: 0 per year x 12.5% x £3,379.73 = £0.00
- Total: *£31,653.32*

This can be compared to the equivalent value using the company's preferred calculation:

- Total: £534.07 x 0.095 per week x 52.18 weeks per year = *£2,647.44*

If we applied the unit cost preferred by the ERG, the following alternative total is obtained:

- Total: £3,379.73 x 0.095 per week x 52.18 weeks per year = *£16,753.22*

Each of these three approaches yields substantially different values. The ERG is therefore still not clear exactly what is being assumed in the company's base-case analysis, and how the value of £534 was calculated. However, it can be seen from the ERG's calculations above that if the company's calculated cost of a single bed day is applied (which the ERG does not considered to be appropriate), the total estimated hospitalisation costs are much lower than the ERG's analysis.

Based on the company's preference to include long-stay hospitalisation costs, the ERG would expect that these costs should, by definition, apply for more than a single day. The length of stay is an inherent property of the unit cost for hospitalisation; consequently, any deviation from the unit cost requires greater justification and detail than has been provided by the company.

3.6. Utility values

The ERG is unaware if further data concerning health-related quality of life were available with the August 2020 DBL. [redacted].
However, for clarity, the utility values used in the company's revised base-case analysis are unchanged from those used in its original base-case analysis.

The ERG notes the company's provision of an additional analysis where utility values were calculated based on treatment status (as opposed to progression status). The rationale for this new approach in estimating utilities relates to the assertion that *"utility in oncology is typically a function of time to death"*; however, it is important to note that time-to-death utilities were not considered in the analysis. To support the inclusion of a new approach to utility analysis, the ERG would expect model summary and model fit comparisons (i.e., AIC/BIC) to be provided for the progression-based and treatment-based models. The absence of this information prevents assessment of which approach provides the best fit to the underlying data when estimating utility values.

The ERG notes with particular concern that the values for on treatment for both arms are higher than the progression-free utility values used in the company's base-case analysis; and that the values for off treatment for both arms are higher than the progressed utility values used in the company's base-case analysis (see Table 4). This suggests, for example, that the minimum utility experienced by the nivolumab group is approximately [redacted] greater if a treatment-based

approach was taken versus a progression-based approach (equivalent difference for the taxanes group: [redacted]).

Table 4: Utility values by treatment versus progression status

Frequency	Treatment based	Progression based
Nivolumab on treatment or progression-free	[redacted]	[redacted]
Taxanes on treatment or progression-free	[redacted]	[redacted]
Nivolumab off treatment or progressed	[redacted]	[redacted]
Taxanes off treatment or progressed	[redacted]	[redacted]

It is the ERG's view that while the treatment-based analysis demonstrates a mean utility value after discontinuation that is greater for the nivolumab arm versus the taxanes arm, this still does not address the concerns raised about the utility analysis in general within the ERG's report. As no additional evidence has been presented in the company's ACD response regarding the methodological approach taken, these issues are still unresolved and remain within the new utility values provided as a scenario analysis.

For example, the control arm mean baseline utility (taken at screening) was significantly lower than of the nivolumab arm (ERG report Section 4.2.7.3), yet no adjustment to utility values appears to have been performed to account for this. Furthermore, the mean utility value for the nivolumab arm in the progression-free (or on-treatment) state is significantly higher than the mean utility of the UK general population aged 65-70 years, based on Ara and Brazier (2011)[3] of 0.8041 (95% CI: 0.790, 0.817).

The ERG's preferred utility values to inform its base-case analysis remain unchanged from its original base-case analysis. The ERG emphasizes that these values are still subject to substantial uncertainty.

3.7. Subsequent therapies

The ACD explains that no adjustment was made to efficacy or additional costs of third-line therapy within the company's model (see paragraph 3.9 of the ACD). In ATTRACTION-3, subsequent therapy (defined as therapy different to the allocated study therapy) was received by 119 (57%) of 210 patients in the nivolumab group and 115 (55%) of 209 patients in the taxanes group (Company ACD response section 6.1.5). This means that more patients on the nivolumab arm received a subsequent line of therapy versus the taxanes group (though the

proportions are similar). The ERG highlights that the difference in the *duration* of subsequent therapy (which may or may not be similar between treatment arms) is not reported.

In the company's response to the ACD, a plot of adjusted OS is presented, in which patients that receive a subsequent therapy are censored at the date of initiation. This plot is re-produced below in Figure 5. The company explains that this plot shows that subsequent therapy does not greatly impact the comparison between nivolumab and taxanes; and that consequently, subsequent therapy is unlikely to impact on outcomes in the economic model.

Figure 5: ATTRACTION-3: Overall survival censored for subsequent therapy (Company ACD response Figure 14)

Figure redacted

Key: OS = overall survival.

The ERG highlights that in Figure 5, the summary statistics have been mislabelled and contain some apparent errors (e.g. 95% confidence interval [CI] around the medians, and incorrect ordering of rows). The ERG therefore has not considered these values further; however, some

of the values that do not appear to contain errors are identical to those provided in Table 1 of the company's ACD response, which is for *unadjusted* OS. For example, 12-month OS for the nivolumab arm matches the point estimate and 95% CI shown in Table 1 of the Company's ACD response, though the values for the taxanes arm are not the same.

No specific overlay of the unadjusted and adjusted Kaplan-Meier curves was provided in the company's response, and so the ERG attempted to compare the curves visually (though the ERG urges caution when interpreting these plots given that they are based on a crude overlay of images). When the ERG attempted to compare Figure 5 to the original (uncensored) OS plot (through superimposing the curves), very little difference was noted, as shown in Figure 6. Moreover, the ERG could not tell if any difference was made to the nivolumab arm (especially given that 12-month OS was identical, including the 95% CI).

Figure 6: ATTRACTION-3: Comparison of adjusted and unadjusted overall survival

Figure redacted

Key: OS = overall survival.

The ERG suspects that most patients that initiated a subsequent therapy were likely censored for the main OS analysis at this time point (or a time point shortly thereafter), given that no major difference is seen in the Kaplan-Meier curves. In addition, the ERG would have ideally

been informed exactly how many patients were censored, and other potentially-relevant information to help contextualise the findings (e.g. the average difference in adjusted versus unadjusted survival times). Further information from the company is needed to understand why these curves are near identical.

The company's response to the ACD does not comment on the cost aspect of subsequent therapy. To explore the potential impact of subsequent therapy costs on the ICER, the ERG has considered two exploratory scenarios in the company's model. In the model base-case analysis, a cost of £33.73 is applied per model cycle in the third-line setting to reflect best supportive care costs. The following scenarios were considered:

- Double cost for both arms
- Double cost for nivolumab arm only

It should be noted that these scenarios involve the specification of a relatively simplistic adjustment to the per-cycle cost applied for patients in the third-line setting. In reality, it is expected that additional costs for subsequent therapy are unlikely to continue indefinitely, but the costs themselves are likely to exceed £33.73 per week (as this would include all costs related to treatment acquisition, administration, routine blood tests, resolution of adverse events, etc.). However, this analysis was undertaken as a pragmatic means of understanding the directional effect on the ICER were additional costs for the third-line setting taken into consideration, and if these were disproportionately incurred by nivolumab-treated patients.

The results of these scenarios are provided in Table 5. These results illustrate that any additional cost factored into the model to account for subsequent therapy causes an increase in the ICER. This was true both when the same increase in cost was considered for both arms, or just for the nivolumab arm. The latter scenario is of particular interest if, in practice, subsequent taxane therapy would only be considered for nivolumab-treated patients (in other words, if nivolumab treatment would displace the use of taxanes in the second-line setting, but these would still be considered in the third-line setting).

Table 5: Scenarios exploring additional costs for subsequent therapy

Technology	Total			Incremental			ICER £/QALY
	Costs (£)	QALYs	LYs	Costs (£)	QALYs	Lys	
Base-case analysis							
Nivolumab	[redacted]	[redacted]	[redacted]				
Taxanes	[redacted]	[redacted]	[redacted]	24,665	0.567	0.512	48,205
Double cost per cycle in third-line for both arms							
Nivolumab	[redacted]	[redacted]	[redacted]				
Taxanes	[redacted]	[redacted]	[redacted]	25,359	0.567	0.512	49,563
Double cost per cycle in third-line for nivolumab only							
Nivolumab	[redacted]	[redacted]	[redacted]				
Taxanes	[redacted]	[redacted]	[redacted]	26,716	0.567	0.512	52,215

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost effectiveness ratio; LYs, life-years; QALYs, quality adjusted life years

Note: LYs presented here are discounted at 3.5% per annum. Company's response to the ACD presented undiscounted LYs. The undiscounted LYs are identical to those presented in the Company's response to the ACD (as no change is made to the estimation of overall survival).

3.8. Company's revised preferred base-case analysis results

The company's revised base-case analysis results are provided in Table 6. Incremental costs have increased from £20,842 to £24,665, and the incremental QALY gain has also increased from 0.458 to 0.512. The change in incremental costs is driven by a combination of the switch to eMIT costs for taxanes, and the updated OS, PFS, and ToT models. The change in QALYs gained is driven entirely by the specification of different OS and PFS models, as the utility values used in the company's revised analysis are identical to those per its original submission.

Table 6: Company's revised base-case analysis

Technology	Total			Incremental			ICER £/QALY
	Costs (£)	QALYs	Lys	Costs (£)	QALYs	Lys	
Nivolumab	[redacted]	[redacted]	[redacted]				
Taxanes	[redacted]	[redacted]	[redacted]	24,665	0.512	0.567	48,205

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost effectiveness ratio; LYs, life-years; QALYs, quality adjusted life years

Note: LYs presented here are discounted at 3.5% per annum. Company's response to the ACD presented undiscounted LYs. The undiscounted LYs are identical to those presented in the Company's response to the ACD (as no change is made to the estimation of overall survival).

3.9. ERG's revised preferred base-case analysis

As the ERG was not able to change any of the revised survival models within the model provided, the ERG has not made any changes to the estimation of OS, PFS, or ToT versus the company's revised base-case analysis. However, the ERG highlights that it does not consider the company's selected base-case models to be the most appropriate for decision making. The ERG has provided an updated base-case analysis reflecting the ERG's other preferred base-case analysis settings for which its view is unchanged based on the Company's ACD response.

The changes made in the company's revised model are as follows:

- The ERG's preferred administration costs (see ERG report Section 6.3.4)
- The ERG's preferred utility values (see ERG report Section 6.3.5)
- The ERG's preferred hospitalisation cost (see ERG report Section 6.3.6)

The corresponding results are presented in Table 7.

Table 7: Company's revised base-case analysis with ERG preferred assumptions (excluding choice of survival models)

Technology	Total			Incremental			ICER £/QALY
	Costs (£)	QALYs	Lys	Costs (£)	QALYs	LYs	
Nivolumab	[redacted]	[redacted]	[redacted]				
Taxanes	[redacted]	[redacted]	[redacted]	31,554	0.411	0.567	76,701

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost effectiveness ratio; LYs, life-years; QALYs, quality adjusted life years

Note: LYs presented here are discounted at 3.5% per annum. Company's response to the ACD presented undiscounted LYs. The undiscounted LYs are identical to those presented in the Company's response to the ACD (as no change is made to the estimation of overall survival).

The results in Table 7 show that the three edits made (administration costs, hospitalisation costs, utility values) cause the incremental costs to increase from £24,665 to £31,554, and the incremental QALYs to decrease from 0.512 to 0.411. This causes the ICER to increase from £48,205 to £76,701.

While these results do not constitute the ERG's preferred base-case analysis (as it was not possible to change the survival models to reflect the ERG's preferences), these edits alone cause the ICER to increase above £50,000 per QALY gained. It is the ERG's expectation that

edits to the models for OS, PFS, and/or ToT may further increase the ICER, though the extent to which the ICER would increase remains unclear.

4. END OF LIFE

The company has commented further on the evidence available to support its expectation that NICE's end-of-life criteria are met. The ERG notes that ultimately, whether nivolumab fulfils the end-of-life criteria is a decision for the appraisal committee to make. However, a short commentary is provided below concerning the updated evidence provided in the company's ACD response.

NICE's end-of-life criteria are said to be met if both of the following apply:

1. The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
2. There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional three months, compared to current NHS treatment.

The uncertainty regarding whether or not nivolumab fulfils the end-of-life criteria has, in the view of the ERG, centred around the second criterion relating to the magnitude of extension to life offered by nivolumab.

In its response to the ACD, the company makes the case that the criterion of at least three months extension to life is now fulfilled based on restricted mean OS (i.e., the calculated area between the Kaplan-Meier curves).

The ERG considers that the additional follow-up data presented in the company's updated analysis helps to resolve some of the uncertainty related to OS benefit. Nevertheless, the ERG notes that the choice of survival extrapolation has the potential to reduce the survival benefit associated with nivolumab markedly. The updated data demonstrate relatively greater survival outcomes for both the nivolumab and taxanes arms, versus the company's and the ERG's original base-case analyses.

5. REFERENCES

1. NHS Improvement. Reference costs 2017/18: highlights, analysis and introduction to the data, 2018. Available from: https://improvement.nhs.uk/documents/1972/1_-_Reference_costs_201718.pdf.
2. NHS Improvement. 2018/19 National Cost Collection data, 2020. Available from: https://improvement.nhs.uk/resources/national-cost-collection/#ncc1819.
3. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. Value in Health. 2011;14(4):539-45.