NICE National Institute for Health and Care Excellence

Single Technology Appraisal

Avelumab for untreated metastatic Merkel cell carcinoma

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

SINGLE TECHNOLOGY APPRAISAL

Avelumab for untreated metastatic Merkel cell carcinoma

Contents:

The following documents are made available to consultees and commentators:

The **final scope and final stakeholder list** are available on the NICE website.

1. **Company submission** from **Merck, Sharpe & Dohme**
   a. **Addendum**
2. **Clarification questions and company responses**
   a. **Additional responses**
3. **Patient group, professional group and NHS organisation submission** from:
   a. Neuroendocrine Cancer UK
4. **PHE SACT Report**
5. **Evidence Review Group report** prepared by **BMJ-TAG**
   a. Addendum
6. **Evidence Review Group – factual accuracy check**
7. **Technical engagement response** from **Merck, Sharpe & Dohme**
8. **Technical engagement responses and statments from experts:**
   a. Nickie Jervis, Patient Support & Information Nurse Specialist – patient expert, nominated by Neuroendocrine Cancer UK
9. **Technical engagement response from consultees and commentators:**
   a. Neuroendocrine Cancer UK
   b. NCRI-ACP-RCP-RCR
10. **Evidence Review Group critique of company response to technical engagement** prepared by **BMJ-TAG**

Clinical expert perspectives from:

- Dr Peter Goon, Consultant Dermatologist – clinical expert, nominated by the British Association of Dermatologists – *to follow*
- Dr Paul Nathan, Consultant Medical Oncologist – clinical expert, nominated by Merck – *to follow*

Any information supplied to NICE which has been marked as confidential, has been redacted. All personal information has also been redacted.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Cancer Drugs Fund Review of TA517

Avelumab for treating metastatic Merkel cell carcinoma

Company evidence submission for committee

3 March 2020

File name	Version	Contains confidential information	Date
ID1617 Company submission	v1-0	Yes	3 March 2020

Contents

Tables and figures

Cancer Drugs Fund review submission

A.1 Background

Merkel cell carcinoma (MCC) is an ultra-rare, aggressive, neuroendocrine skin cancer with limited treatment options. It often progresses rapidly and, as it affects the surface of the skin, is a very visible disease. Initial responses with chemotherapy can be relatively high but are short lasting, meaning prognosis with chemotherapy and best supportive care (BSC) is poor. As such, there is an unmet need for effective treatment for people with metastatic MCC.

Avelumab (Bavencio®) is currently recommended for use within the Cancer Drugs Fund (CDF) as an option for treating metastatic MCC in adults, only if they have not had chemotherapy for metastatic disease, and the conditions in the managed access agreement (MAA) for avelumab are followed.[1,2] The CDF recommendation was optimised as avelumab is also recommended as an option for treating metastatic MCC in adults, only if they have had one or more lines of chemotherapy for metastatic disease, in routine National Health Service (NHS) practice.[1]

The clinical-effectiveness evidence for avelumab for adults who have not had chemotherapy for metastatic disease was taken from one cohort (Part B) in the JAVELIN Merkel 200 (JM200) clinical trial, a single-arm non-randomised trial of metastatic MCC patients.[3*] The committee acknowledged that immature data† were modelled, and that ongoing data collection in JM200: Part B would reduce the uncertainty about the overall survival (OS) and progression-free survival (PFS) benefit, as well as time-on-treatment (ToT) with avelumab.

The committee considered incremental cost-effectiveness ratios (ICERs) based on the list price of avelumab. It concluded that, based on its preferred assumptions, the most plausible ICER for first-line treatment was £58,315 to £72,033 per quality-adjusted life-year (QALY) gained.‡ It considered there is plausible potential for first line use of avelumab to be cost effective, if further trial data prove favourable.

---

* Two cohorts of patients were studied in JM200. Part A comprises of patients with treatment experienced metastatic MCC (i.e. those who have previously received chemotherapy). Part B is comprised of patients with treatment naïve metastatic MCC (i.e. those who have not previously received chemotherapy). Accordingly, this CDF review focuses predominantly on data from Part B of JM200.

† At the time of the publication of TA517, data from Part B of JM200 were available for a total of n=39 patients. Of these, n=29 patients were followed up for 3 months or more, and n=14 were followed up for 6 months or more.

‡ The lower-bound ICER corresponds to the company's preferred base-case analysis at the time of the second appraisal committee meeting in TA517. The upper-bound ICER of £72,033 corresponds to the ERG's preferred base-case analysis prior to the second committee meeting. Please see Appendix 5 for more information.

A.2 Key committee assumptions

The key committee assumptions are summarised in Table 1. The contents of Table 1 are based on NICE's Terms of Engagement (ToE) document, with minor edits made for brevity and/or completeness.[4]

Table 1: Key committee assumptions

Topic	Assumption
Population	The CDF review will focus on the population that were recommended under the MAA; that is, adults with metastatic MCC who have not had chemotherapy for metastatic disease.
Comparators	The most appropriate comparator for first-line treatment is chemotherapy, and so the CDF review will focus on this comparison.
Generalisability	The marketing authorisation for avelumab was granted conditionally for the first-line group because of the immaturity of the data. The committee concluded that the JM200 results should be interpreted with caution as there were some unanswered questions about its generalisability, and so the CDF review will assess the generalisability of JM200.
Comparative clinical data	The committee noted that there were no direct comparative data, and concluded that, although uncertain, the 2-part observational study 100070-Obs001 provided the most appropriate comparator data. If further comparative data is available, it should be explored in the CDF review. This is discussed further in Section A.7.
Indirect comparison	The results from the naive indirect comparison are highly uncertain. In light of new evidence now available for avelumab, the CDF review will consider the naive indirect comparison.
Overall survival (OS)	OS data may be confounded by the use of subsequent treatments, and at the time of the original CS the OS data were still relatively immature. Therefore, updated OS evidence are expected to be provided as part of the CDF review.
Model structure	The company's model structure is suitable for decision making. It is anticipated that the model structure will not change for the CDF review.
Extrapolation of survival	Due to the limitations of the data for first-line treatment, estimates for survival derived from the second-line and beyond model were used in the original first-line model. The committee were concerned that the OS and PFS estimates for first line treatment were based on clinical assumptions instead of direct evidence. It was also aware that cost-effectiveness outcomes for first-line treatment were sensitive to the hazard ratio chosen for OS. The committee concluded that the survival estimates for first-line treatment are highly uncertain, and so the method used to extrapolate survival will be explored in the CDF review.
Time-on-treatment (ToT)	Two-thirds of patients were assumed to stop treatment after 2 years (and all remaining patients were assumed stop treatment after 5 years). The clinical experts explained that they expect 95% of patients having avelumab to stop treatment by 2 years. The ERG considered the ToT extrapolation without truncation at 2 years. The committee considered both the company's and the ERG's assumptions in its decision-making. It is anticipated further evidence on ToT will be available for the review and this assumption will be reconsidered.
Utilities	The committee acknowledged that the utility values were implausibly high but it noted that, because the same utilities were applied regardless of treatment group, only the difference between health states mattered.

	The committee concluded that it could accept these utility values but acknowledged they were very high. The same utility values are anticipated to be used in the CDF review. However, the 'Terms of Engagement' document stated that if further evidence is available, an exploration of the most appropriate utilities should be performed. This is discussed further in Section A.3.
Model corrections*	The following model corrections were made during the previous appraisal: (1) adding the cost of premedication; (2) added administration costs (approximately £43); and (3) corrected an error in the calculation of background mortality. It is anticipated that these corrections will be included in the CDF review.
End of life	Avelumab meets the end-of-life criteria.

Key: CDF, Cancer Drugs Fund; ERG, Evidence Review Group; JM200, JAVELIN Merkel 200; MAA, managed access agreement; OS, overall survival; PFS, progression-free survival; ToT, time-on-treatment.
Note: *The list provided does not include assumptions discussed elsewhere in this table.

A.3 Other agreed changes

The Final Appraisal Determination (FAD) relating to the original company submission (CS) expressed some uncertainty around the most appropriate utilities to use. Consequently, the ToE document stated that if further evidence were available, an exploration of the most appropriate utilities should be performed. Following confirmation with NICE and the Evidence Review Group (ERG) via teleconference in January, additional sensitivity analysis using alternative utility values have been undertaken, which are discussed further in Section A.12.

A.4 The technology

A summary of the technology being reviewed (avelumab) is provided in Table 2.

Table 2: Technology being reviewed

UK approved name and brand name	Avelumab (Bavencio®).
Mechanism of action	Avelumab is a human IgG1 lambda monoclonal antibody that specifically targets cancer cells through the inhibition of the immune checkpoint protein, PD-L1. It has a dual mechanism of action which aims to bind and block the inhibitory signalling through PD-1/PD-L1 resulting in the activation of T-cells and cell-mediated immune responses against tumour cells or pathogens.
Marketing authorisation status	Avelumab was granted a conditional European Medicines Agency marketing authorisation on September 18, 2017.[5]
Indications and any restriction(s) as described in the summary of product characteristics	Avelumab is indicated as monotherapy for the treatment of adult patients with metastatic Merkel cell carcinoma. Avelumab is also indicated in combination with axitinib for the first-line treatment of adult patients with advanced renal cell carcinoma.
Method of administration and dosage	Avelumab is administered at a flat dose of 800mg intravenously over 60 minutes every two weeks. The dosage of avelumab administered as part of the JM200 clinical trial was 10mg/kg. In November 2019, the approved dose for avelumab was changed to be a flat dose of 800mg, based on pooled pharmacokinetics and safety data from three studies in patients with

	various tumour types (studies EMR100070-001, EMR100070-002 and EMR100070-003). A summary of these data is provided in Appendix 1. It is also important to note than until this change in the licensed dose of avelumab, NHS dose banding guidance was followed for avelumab (though this was not published at the time of the original TA517 submission).[6] This guidance allowed for increased flexibility in dosing avelumab to reduce wastage without negatively affecting safety or efficacy. As such, the switch from a weight-based dose of 10mg/kg to a flat dose of 800mg is not expected to lead to a large difference in costs incurred by the NHS (i.e. most patients would require 4 x 200mg vials using either dosing approach).
Additional tests or investigations	None.
List price and average cost of a course of treatment	The list price of avelumab is £768.00 for one 200mg vial. The average time on treatment calculated via the area under the curve within the model yields a mean of approximately 13 months for patients who have not received prior chemotherapy for metastatic disease. This yields an average cost of a course of treatment of £85,062.
Commercial arrangement (if applicable)	There is no current patient access scheme which applies to this technology with regards to the routine commissioning of avelumab. A commercial access arrangement was in place during the period of data collection to inform this CDF review.
Date technology was recommended for use in the CDF	The Final Appraisal Determination for avelumab in metastatic Merkel cell carcinoma was published on the NICE website on 1 March 2018.[1]
Data collection end date	SACT follow-up ended on 31 July 2019. JM200: Part B data discussed within this submission comprise a minimum follow-up period of 15 months.
Key: CDF, Cancer Drugs Fund; IgG1, Immunoglobulin G1; JM200, JAVELIN Merkel 200; kg, kilogram(s); mg, milligram(s); PD-1, programmed-death 1; PD-L1, programmed-death ligand-1.	

A.5 Clinical effectiveness evidence

A summary of the key evidence collected during the CDF data collection period is provided in Table 3. The primary evidence is from the maturing, pivotal JM200 trial. Additional follow-up data in the treatment-naïve population (Part B of the JM200 trial) will address the key uncertainties acknowledged by the committee.

Table 3: Primary source of clinical effectiveness evidence

Study title	**JAVELIN Merkel 200 trial: Part B**
Study design	Phase II single-arm open-label study
Population	Patients with metastatic Merkel cell carcinoma with no prior systemic therapy for metastatic disease (treatment-naïve metastatic Merkel cell carcinoma)
Intervention(s)	Avelumab
Comparator(s)	-
Outcomes collected that address committee's key uncertainties	Overall survival, progression-free survival, time-on-treatment, health-related quality of life.
Reference to section in appendix	Appendix 1

Secondary evidence collected via the Systemic Anti-Cancer Therapy (SACT) database are now also available. The SACT database provides additional information concerning the OS and ToT for treatment-naïve metastatic MCC patients receiving avelumab in NHS practice. However, these data are less useful for implementation directly into the current economic model and are used rather as a reference in discussions about the generalisability of JM200: Part B to UK clinical practice (discussed further in Section A.6).

Table 4: Secondary source of clinical effectiveness evidence

Study title	**SACT data cohort study**
Study design	Systemic Anti-Cancer Therapy (SACT) data cohort study
Population	Patients with metastatic Merkel cell carcinoma with no prior systemic therapy for metastatic disease (treatment-naïve metastatic Merkel cell carcinoma)
Intervention(s)	Avelumab
Comparator(s)	-
Outcomes collected that address committee's key uncertainties	Overall survival, time-on-treatment.
Reference to section in appendix	Appendix 1

In addition to the data from JM200 and the SACT dataset, an area of uncertainty highlighted within the FAD was the lack of direct comparative data for avelumab and chemotherapy (in both treatment-experienced and treatment-naïve patients). In the original CS, a systematic literature review (SLR) found no randomised controlled trial evidence for avelumab, and limited evidence sources to inform the comparator arm of the model.

As the SLR was last updated in March 2017, Merck Serono has conducted an update to the review to establish whether any additional, relevant evidence has since been published. The SLR update did not identify any additional evidence not known to Merck Serono that is directly relevant to this appraisal. Information concerning the update to the SLR can be found in Appendix 2. Findings from the SLR are discussed in Section A.7, and the data used to inform the comparator arm of the model are discussed in Section A.8 (i.e. observational study data per the original TA517 CS).[1,7]

A.6 Key results of the data collection

A.6.1 Baseline demographics

From JM200: Part B, data for n=116 treatment-naïve metastatic MCC patients with a minimum of 15 months follow-up are now available for analysis (data cut: May 2019). In addition, data for n=52 treatment-naïve metastatic MCC patients treated during the CDF data collection period for avelumab were recorded in the SACT dataset, with a minimum of 5

months follow-up (data cut: November 2019). The baseline demographics of patients from both sources are provided in Table 5. Sex, age, and Eastern Cooperative Oncology Group Performance Status (ECOG PS) were variables provided in the SACT report.

Table 5: Baseline demographics, JM200: Part B versus SACT

Characteristic		**JM200: Part B**	**SACT**
Sample size		116	52
Follow-up		Minimum: 15 months Median: 16 months	Minimum: 5 months Median: 6 months
Sex	Male	81 (70%)	30 (58%)
	Female	35 (30%)	22 (42%)
Age	<40	0 (0%)	0 (0%)
	40-49	4 (3%)	1 (2%)
	50-59	7 (6%)	3 (6%)
	60-69	27 (23%)	8 (15%)
	70-79	46 (40%)	22 (42%)
	80+	32 (28%)	18 (35%)
	Median	74.0 years	75.5 years
ECOG PS	0	72 (62%)	7 (13%)
	1	44 (38%)	34 (65%)
	2	0 (0%)	4 (8%)
	3	0 (0%)	1 (2%)
	4	0 (0%)	0 (0%)
	Missing/ unknown	0 (0%)	6 (12%)

Key: ECOG PS, Eastern Cooperative Oncology Group Performance Status; JM200, JAVELIN Merkel 200; SACT, Systemic Anti-Cancer Therapy.

Based on Table 5, the following may be noted:

- Compared with the previous data cut from JM200: Part B, a larger number of patients are available for analysis (n=116 versus n=39). This represents nearly three times the number of patients previously available. The SACT cohort includes data for n=52 patients. These small patient numbers are atypical for an ultra-rare disease, with JM200 representing the largest registrational trial on MCC.

- For the outcome of OS (discussed further in Section A.6.2), the median follow-up time in SACT was 6 months (191 days), whereas in JM200: Part B the median follow-

up time was 16 months (483 days).[§] Combined with the larger sample size, the maturity of the JM200: Part B data permits a more robust estimate of the (particularly longer-term) outcomes of treatment

- Both cohorts comprise of mainly male patients, though the SACT dataset includes a slightly larger proportion of female patients (42% versus 30%). MCC is expected to affect more men than women, which is evident in both sources, though gender is not expected to be a treatment effect modifier

- The SACT cohort is slightly older on average; 77% of the SACT patients were aged 70 years or older, versus 68% of the JM200: Part B cohort. Median age at baseline in the SACT cohort was 75.5 years versus 74.0 years for the JM200: Part B cohort. MCC occurs most commonly in older people, and survival is expected to be poorer in older patients (due to an increased risk of other comorbidities) however the ages of both cohorts are similar

- The majority of JM200: Part B patients had an ECOG PS of 0 (62%) whereas most patients in the SACT dataset had an ECOG PS of 1 (65%). The differences in ECOG PS between the cohorts are not considered to be reflective of a substantially dissimilar cohort of patients, as in clinical practice there is known variability in ECOG PS determined by clinicians,[8] and patients with an ECOG PS or 0 or 1 are routinely grouped together in clinical trial protocols

- In JM200: Part B, patients with an ECOG PS greater than 1 were excluded. Despite similar expectations for patients treated via the CDF[**], 10% of patients in the SACT dataset had an ECOG PS of 2 or 3; and a further 12% of patients had an unknown or missing ECOG PS. Therefore, over 20% of the SACT cohort did not explicitly meet the MAA ECOG PS criterion. While it remains unclear how this may have influenced outcomes, patients with an ECOG PS of 2 or more are certainly expected to have a poorer prognosis then those with an ECOG PS of 0 or 1

[§] Median follow-up is the patients' median observed time from the start of their treatment to death or censored date. This is reported in the SACT report, and the corresponding value from JM200: Part B was calculated from patient-level data, based on the median overall survival time-to-event value recorded (in months). The value in months was converted to days using the formula $Time_{months} * 365.25 / 12$.

[**] One of the eligibility criteria in the MAA states that the patient must have an ECOG PS of *"either 0 or 1"*, and that a patient with an ECOG PS of 2 is *"not eligible for avelumab"*.[2]

The demographic differences between the cohorts may not be surprising when considering the different approaches to enrolling patients in each cohort:

- Avelumab is currently recommended by NICE and routinely commissioned in the second-line setting. In this context, some patients who would have been candidates for treatment with first-line avelumab may have instead been treated with chemotherapy by clinicians who preferred to reserve avelumab for use after chemotherapy

- Some patients included within the SACT cohort may have previously been deemed ineligible for treatment with first-line chemotherapy owing to its associated toxicities, and consequently managed with best supportive care prior to avelumab. This means that due to the lack of availability of other options, some patients may have had poorer prognosis compared to newly-diagnosed patients

- Patients need to have sufficient life expectancy to benefit from immunotherapies (such as avelumab), due to the mechanism of action of these treatments. In JM200, patients with a life expectancy of at least 12 weeks were eligible for inclusion.[9] The initial drop apparent in the OS curve for the SACT cohort indicates the inclusion of some patients who may not have sufficient life expectancy to benefit from immunotherapy

Real-world use of avelumab for metastatic MCC in NHS practice is limited. Therefore, it is important that clinicians are able to gain first-hand experience of using avelumab in order to guide their future practice and clinical decision making. This has been made possible through avelumab being made available via a combination of the CDF in the first-line setting, and routine commissioning in the second-line and beyond setting. Unavoidably, the choice of where to use avelumab may have introduced selection bias into the SACT cohort (i.e. the first-line treated population is not expected to be entirely representative of the first-line eligible population).

The differences between the two cohorts raise an important question about which dataset is the most appropriate to use as the basis for economic modelling in this resubmission. Merck Serono has chosen to model using the mature JM200: Part B data for the following reasons:

- JM200: Part B is the primary evidence source for the efficacy and safety of avelumab in patients with treatment-naïve metastatic MCC. All outcomes specified in the final NICE scope for TA517 are available from this study. The SACT dataset is, by comparison, incomplete. Data for several key model inputs are not available (e.g.

PFS and HRQoL). In addition, data for other outcomes specified in the final NICE scope (such as response rate and adverse effects of treatment) are also not available from the SACT dataset

- In JM200, patients were considered candidates for avelumab only if they had an ECOG PS of 0 or 1, and sufficient life expectancy to benefit from cancer immunotherapy. A substantial proportion of patients in the SACT cohort either explicitly did not meet these criteria (i.e. 10% had an ECOG PS of 2 or 3), or it was unclear if the criteria were met (i.e. unknown ECOG PS in 12%), which collectively comprises over 20% of patients

- Data from JM200: Part B comprise a larger sample of patients compared with the SACT dataset (n=116 versus n=52), also with a longer follow-up (a minimum follow-up in JM200: Part B of 15 months versus 5 months in SACT). As noted in the TA517 FAD, the full benefits of avelumab are only realised in the longer term, and so any extrapolations based on limited, interim data are subject to inherent uncertainty

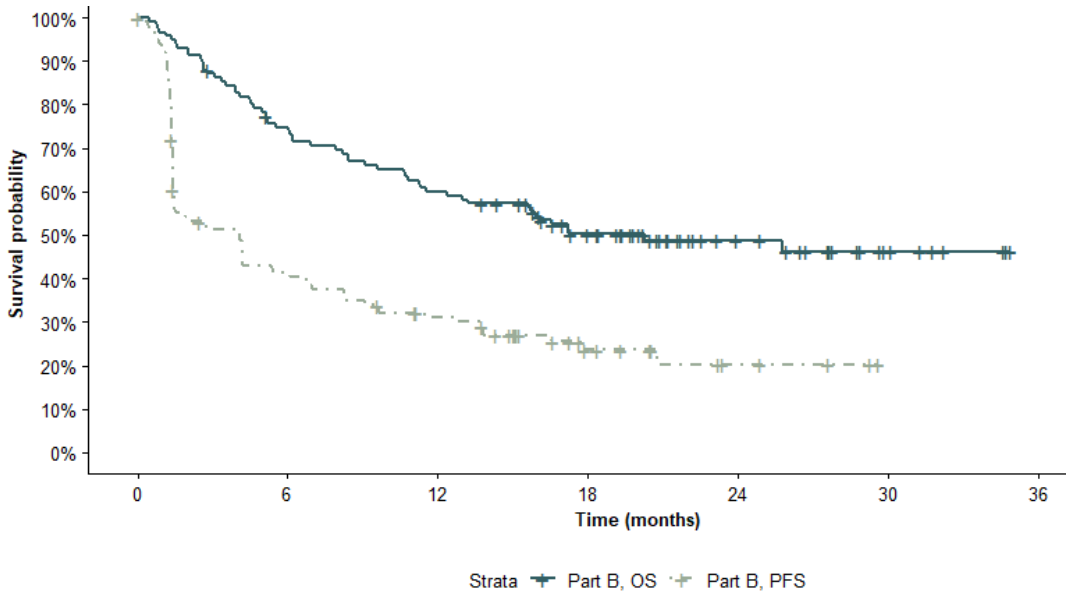
The mature JM200: Part B dataset constitutes a more reliable source to inform estimates of clinical and cost effectiveness. On balance, the population expected to be treated in practice is anticipated to more closely resemble the JM200: Part B cohort if performance status criteria are followed (i.e. those with an ECOG PS of 0 or 1). It may be no surprise to treating clinicians that when patients with more significant comorbidities are treated (i.e. those with an ECOG PS of 2 or more), the outcomes in the real world are slightly poorer than those seen under trial conditions.

A.6.2 Overall and progression-free survival

JM200: Part B

OS and PFS data for avelumab-treated patients were obtained from mature JM200: Part B data. All patients (n=116) were followed up for a minimum period of 15 months (data cut: May 2019). Corresponding Kaplan-Meier curves for OS and PFS are shown in Figure 1. An overlay of the previous and updated Kaplan-Meier curves for OS and PFS from JM200: Part B are provided in Appendix 1.

Figure 1: JM200: Part B, OS and PFS



At risk							
Part B, OS	116	85	68	45	20	7	0
Part B, PFS	116	45	31	12	4	0	0

Key: OS, overall survival; PFS, progression-free survival.

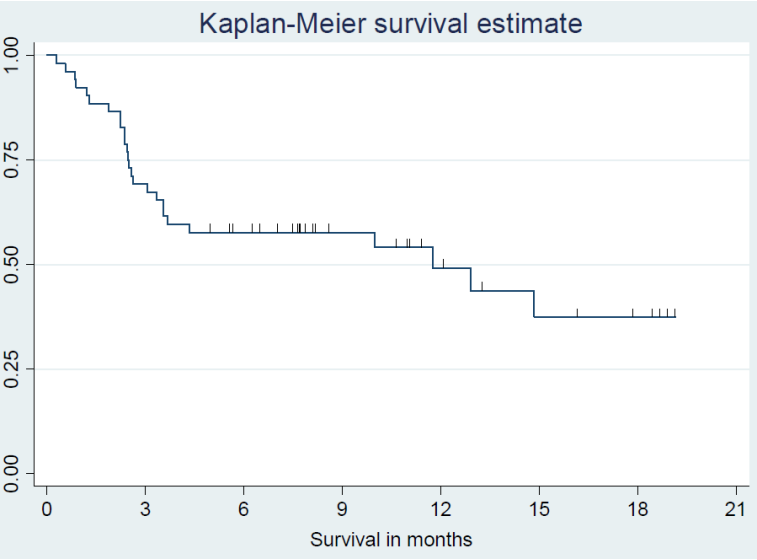
The expected 'plateau' in survival (referenced in the original TA517 CS) is now evident in mature JM200: Part B data, with n=50 patients still alive at 2 years. Median OS is approximately 20 months. OS at 6 and 12 months is 75% and 60%, respectively.

There is an initial drop in the PFS curve, followed by a levelling out (i.e. 'plateau'). In the original CS, PFS for treatment-naïve patients was assumed to be equivalent to that of treatment-experienced patients, due to a previous lack of mature data available for treatment-naïve patients. Based on the latest data from JM200: Part B, median PFS is 4.1 months, with 6- and 12-month PFS estimated at 41% and 31%, respectively.

SACT dataset

OS data from the SACT dataset are presented in Figure 2 (though notably absent from this plot are PFS data, which were not available from the SACT dataset).

Figure 2: SACT, OS



Note: Extracted from *"Avelumab for treating metastatic Merkel cell carcinoma – data review"*, Figure 4.

Median OS in the SACT dataset was 11.8 months, with estimated proportions of patients still alive at 6 and 12 months of 58% and 50%, respectively. The OS data are based on a relatively small sample of patients (n=52, versus n=116 in JM200: Part B), including some patients with an ECOG PS greater than 1. Furthermore, median follow-up time was 6.3 months (191 days), compared with 15.9 months (483 days) in JM200: Part B. Therefore, the tail-end of the Kaplan-Meier curve in particular should be considered with caution, as a true picture of the OS for the SACT cohort cannot be inferred from these interim data.

Within the first 3 months of follow-up, there were 16 deaths recorded (31% of the cohort), and an additional 6 deaths recorded up until the minimum follow-up period of five months (giving a total of 43% of the cohort having died before 5 months). Conversely, in JM200: Part B there were 25 recorded deaths up until 5 months (equivalent to 22% of the cohort), of which 15 deaths occurred in the first three months (13% of the cohort). These differences are indicative of the inclusion of patients in the SACT cohort with a life expectancy that is unlikely to be sufficient to derive benefit from an immunotherapy (such as avelumab).

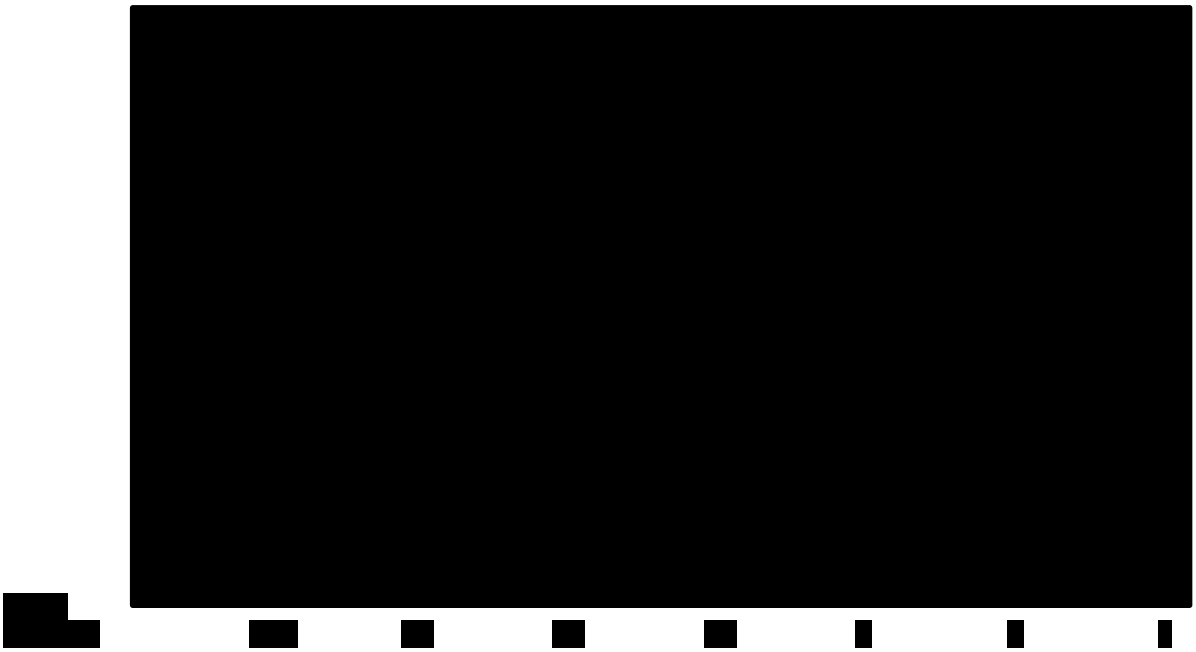
The differences in the cohorts of patients enrolled in JM200: Part B and followed up in the SACT dataset are discussed in more detail in Section A.6.1.

A.6.3 Time on treatment

JM200: Part B

ToT data for patients treated with avelumab were also obtained from JM200: Part B (n=116 patients, 15 months minimum follow-up, data cut: May 2019). The corresponding Kaplan-Meier curve for ToT is presented in Figure 3.

Figure 3: JAVELIN Merkel 200 Part B Kaplan-Meier ToT



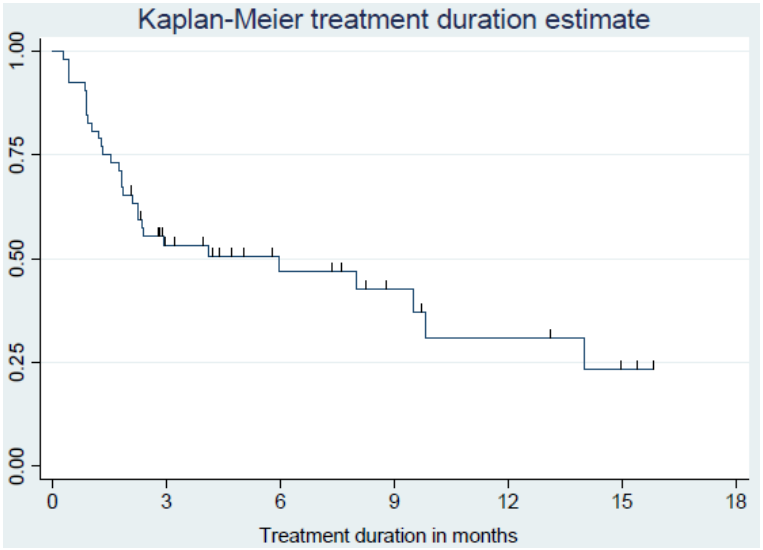
Key: ToT, time on treatment.

As described in Section A.6.2, minimum follow-up in JM200: Part B was 15 months, and so after this time there are a number of censored observations. It remains unclear when these patients are expected to discontinue treatment.

SACT dataset

Where possible to compare, the distribution of ToT from the SACT dataset is similar to that seen in JM200: Part B. The corresponding Kaplan-Meier curve from the SACT dataset is presented in Figure 2.

Figure 4: SACT database ToT



Note: Extracted from *"Avelumab for treating metastatic Merkel cell carcinoma – data review"*, Figure 3.

Median ToT was 6.0 months. 46% of patients were still receiving treatment at 6 months, and 31% of patients were still receiving treatment at 12 months. The maximum follow-up time in the SACT dataset for the outcome of ToT is approximately 16 months, and so the proportion of patients still on treatment at 24 months cannot be established.

A.7 Evidence synthesis

No new evidence was identified in Merck Serono's updated literature search considered suitable to inform the comparator arm within the economic model (see Appendix 2). In addition to the company-sponsored observational studies, three additional studies were identified as potential comparator data sources:

- Klink *et al.*, (2017): A retrospective cohort study of n=44 first-line metastatic MCC patients treated with a range of chemotherapy regimens.[10] Only an abstract was available for this study, in which no model-relevant data were presented

- Chang *et al.*, (2019): A case series of n=3 metastatic MCC patients.[11] This study population was considered too small to be deemed suitable for use within the economic model

- Zheng *et al.*, (2019): A retrospective cohort study of n=38 metastatic MCC patients treated with a range of chemotherapy regimens.[12] As per Klink *et al.*, only an abstract was available for this study, in which no model-relevant data were presented

As none of these studies were deemed suitable for use in the economic model, the indirect comparison used to inform the economic model is unchanged from the previous CS.

A.8 Incorporating collected data into the model

The mature data from JM200: Part B were used to update the estimation of OS, PFS, and ToT within the economic model. As described above, the SACT dataset was not used to directly inform the model, but rather to reflect on the generalisability of JM200: Part B to UK clinical practice (see Section A.6). Furthermore, no additional data for the comparator arm were identified (see Section A.7), and so the estimation of outcomes for the chemotherapy comparator is unchanged (recap provided below).

Recap of estimation of outcomes for chemotherapy (from previous CS)

JM200: Part B is an uncontrolled clinical trial and for the purposes of comparison with the current standard of care, data were identified from an observational study (conducted in patients similar to the JM200 cohort) that was conducted by Merck KGaA/Pfizer, as well as a number of historical controls identified in the literature. The observational study was designed with the primary purpose of serving as a comparator to the JM200 trial population. These are described in further detail in the original CS.

Regression analysis and visual inspection of individual patient data (age, ECOG PS, gender, immunosuppression status, stage at diagnosis) suggested that no patient characteristic had prognostic importance, beyond the line of therapy in which it was given. Given the low between-study variability, the treatment-naïve data from both the Merck KGaA/Pfizer observational study and studies identified in the literature were naïvely pooled and parametric curves fit to inform the model. The analyses and extrapolations are discussed in greater detail in Section 5.3.3 in the original CS (2017). There is no change to this approach in the current submission.

A.8.1 Overall and progression-free survival

As the Kaplan-Meier curves for OS and PFS in JM200: Part B are incomplete, survival extrapolation was required to inform the economic model. As per the original CS, NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14 guidance was followed for the fitting of parametric survival models.[13] For OS and PFS, simple and spline-based parametric models were fitted, which cover the range of extrapolation approaches discussed by the committee in TA517. The remainder of this section discusses the choice of models to inform the base-case analysis. Additional information concerning model selection (including statistical goodness-of-fit scores) is provided in Appendix 2.

Overall survival

The majority of the simple (i.e. non-spline) parametric survival models failed to fully realise the expected long-term OS estimates for avelumab in metastatic MCC, whereas the spline models produced more realistic longer-term extrapolations. More specifically, the majority of

the simple parametric models (e.g. lognormal and log-logistic) resulted in extrapolations wherein the curve crossed the latest Kaplan-Meier curve from JM200: Part A (see Appendix 3 for further information). The exception to this is the generalised gamma model which (due to its use of three parameters and associated flexibility) provides both a good fit to the Kaplan-Meier curves as well as plausible longer-term extrapolations. Alternative model fits are compared in Appendix 2.

In the base-case analysis, a “1-knot odds” spline-based model was used to inform the estimation of OS. This model provides a good visual fit to the Kaplan-Meier curve, has favourable statistical goodness-of-fit scores (within 1 point of the best fitting model according to AIC [Akaike’s Information Criteria], and within 4 points of the best fitting model according to BIC [Bayesian Information Criteria]), as well as clinically-plausible long-term extrapolation. The “1-knot odds” spline-based model provides a lower estimate of OS versus the generalised gamma model, but a higher estimate of OS versus the “1-knot hazard” and “1-knot normal” spline-based models. A “1-knot odds” spline-based models was also used in the original CS to extrapolate OS for treatment-experienced metastatic MCC patients receiving avelumab. Alternative models are explored within sensitivity analysis.

Progression-free survival

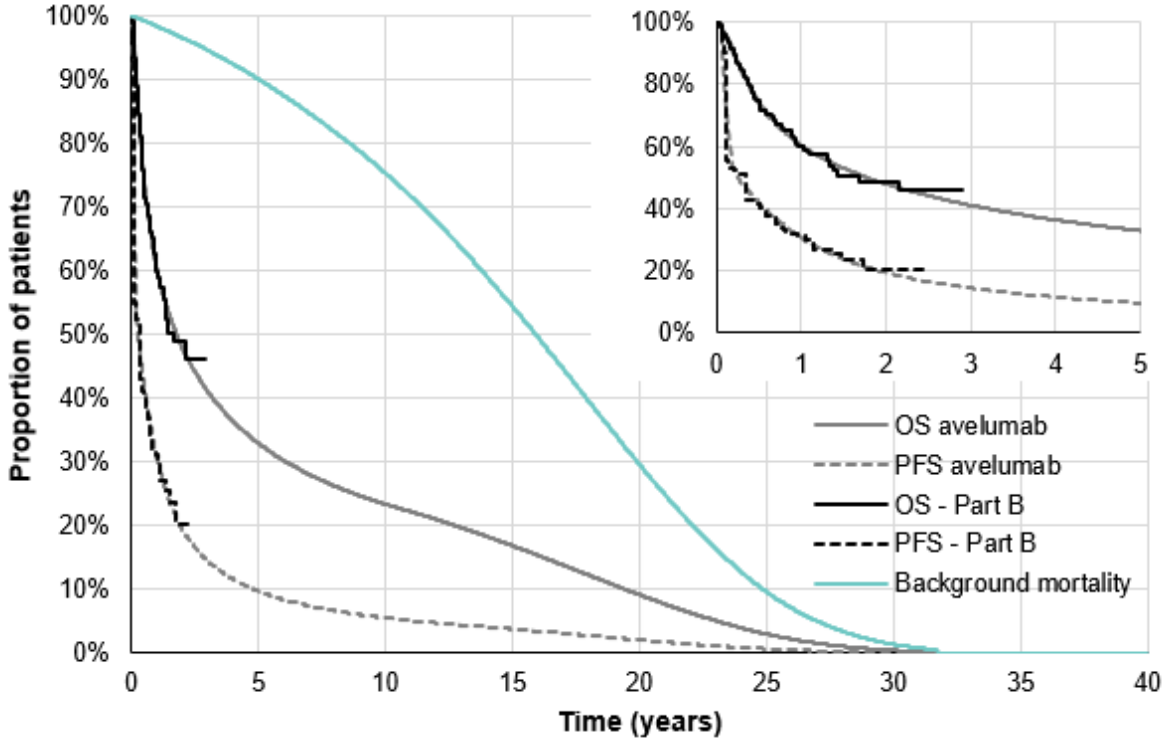
For the outcome of PFS, it was noted in TA517 that a flexible modelling approach was necessary to reflect the abrupt change in hazards that may be inferred through inspection of the Kaplan-Meier curve in JM200: Part A. In the mature JM200: Part B data, a similar change in hazards can be inferred from the Kaplan-Meier curve, and so a similarly flexible modelling approach was sought to inform the economic model.

When fitting a range of parametric models for the outcome of PFS in JM200: Part B, a better visual fit was realised for the spline-based models with 2 versus 3 or 1 internal knot(s). Non-spline-based models were not considered sufficiently flexible to capture the shape of the PFS curve. Of the 2-knot spline-based models, the ‘odds’ functional form provided the best statistical goodness-of-fit, and aligned with the functional form previously used in TA517 for treatment-experienced patients. Therefore, a “2-knot odds” spline-based model was used to inform the base-case analysis.

As with the choice of OS model, alternative models were explored. However, the choice of model for PFS is not a key driver of cost-effectiveness results, and so the results of these analyses are not presented within this submission for brevity.

The base-case OS and PFS models are provided in Figure 5.

Figure 5: Base-case, OS and PFS extrapolations, avelumab



Key: OS, overall survival; PFS, progression-free survival.

A.8.2 Time on treatment

Extrapolation was also required to estimate ToT within the model. Based on clinical expert advice provided to Merck Serono, the majority of patients are expected to have discontinued treatment within 2 years of initiation with avelumab, and all are expected to have discontinued treatment by 5 years. However, without adjustment, all models projected some patients to continue treatment longer than 5 years.

In line with the ERG's preferred approach at the time of the last appraisal committee meeting, a Weibull model was fitted to JM200: Part B data. However, two adjustments were made to the extrapolation:

- The conditional probability (i.e. rate) of treatment discontinuation beyond the minimum follow-up period (15 months) was informed by extrapolation of data from JM200: Part A (which has a notably longer minimum follow-up of 36 months) using a separate Weibull model – i.e. a piecewise modelling approach was adopted, where the estimate hazard of discontinuation switches from a Weibull model fitted to JM200: Part B data at 15 months to a Weibull model fitted to JM200: Part A data. This approach is aligned with the "stitching" analysis previously provided as part of TA517 where extrapolations based on JM200: Part B data alone were not considered appropriate, and was considered reasonable following clinical expert consultation

- At 5 years, all patients remaining on treatment are assumed to immediately discontinue. This time point is varied in sensitivity analysis

The base-case extrapolation is presented in Figure 6. Further explanation of the methods used to adjust the extrapolation of ToT is provided in Appendix 2.

Figure 6: Base-case, ToT extrapolation, avelumab

Key: ToT, time on treatment.
Note: The ToT base-case extrapolation incorporates adjustments to the base extrapolation approach to align the estimated longer-term discontinuation rate with clinical expectation. The green shaded region represents the period of the extrapolation based on data from Part B of JM200. The blue shaded region represents the period of the extrapolation based on data from Part A of JM200. Further information is provided in Appendix 2.

A.8.3 Health-related quality of life

As described in Section A.3, the mature JM200: Part B enabled an update to the utility values available to inform the economic model. In the revised base-case analysis, updated utility values using data from both cohorts studied in JM200 are used to inform the model. Alternative specifications of utility values are explored within sensitivity analysis. The analytical approach and corresponding results are discussed in Appendix 4.

A.9 Key model assumptions and inputs

Table 6 provides a summary of the key model assumptions and inputs used to inform the economic model. Further information concerning updates to key model assumptions and inputs is provided in Appendix 5.

Table 6: Key model assumptions and inputs

Model input	Original parameter/ assumption	Updated parameter/ assumption	Source/ justification
OS model	CS: HR of 0.80 applied to base curve for TE patients (1-knot odds-based spline model) FAD: 1-knot normal-based spline model fitted to interim JM200: Part B data	1-knot odds-based spline model, fitted to JM200: Part B data	Updated JM200: Part B data allows for a more robust estimation of parametric curves, as opposed to an assumed benefit or difference versus treatment-experienced patients. A range of possible extrapolations were explored for all outcomes. • For OS, a 1-knot odds model provides a good visual and statistical fit, while also providing realistic long-term projections • For PFS, similar inferences may be noted, yet a 1- or 3-knot model did not provide as good of a fit versus a 2-knot model, hence this was instead used • For ToT, the extrapolation was adjusted to make use of the longer-term data from Part A of JM200 to better reflect the expected pattern (rate) of treatment discontinuation
PFS model	CS: HR of 1.00 applied to base curve for TE patients (3-knot odds-based spline model) FAD: 1-knot normal-based spline model fitted to interim JM200: Part B data	2-knot odds-based spline model, fitted to JM200: Part B data	
ToT model	CS: HR of 1.00 applied to base curve for TE patients (log-logistic model fitted to 2 years, with adjustment at 2 years to account for one-third of patients continuing up until a maximum of 5 years) FAD: Weibull model fitted to interim JM200: Part B data	Weibull model fitted to JM200: Part B data, adjusted in two aspects: the estimated hazard of discontinuation was based on JM200: Part A data after 15 months (minimum follow-up in JM200: Part B), and all patients were assumed to discontinue treatment by 5 years.	
Dosing	Weight-based dose (10mg/kg)	Flat-dose (800mg)	Changed to align with updated product label. This is not expected to affect NHS practice markedly given that dose banding guidance meant that the majority of patients would be treated with 4 vials prior to the label change.
HRQoL	Time-to-death utility analysis, using JM200: Part A data	Time-to-death utility analysis, using JM200: Part A and B data	Updated analysis of patient HRQoL making use of all data from JM200 (with a covariate for line). See Appendix 4 for further information.

Key: CS, Company Submission; FAD, Final Appraisal Determination; HR, hazard ratio; OS, overall survival; PFS, progression-free survival; TE, treatment-experienced; ToT, time on treatment.

The clinical validity of the survival model projections was independently tested with a UK clinician with experience using avelumab in metastatic MCC yet was not one of the three clinicians consulted by Merck Serono as part of the original TA517 submission. Projected outcomes (e.g. proportions projected to be alive, progression free, and on treatment) were felt to be reasonable, although in some cases pessimistic (in relation to projected survival). The clinician considered that 100% of patients would have discontinued treatment by 5 years, which is consistent with clinical opinion at the time of the original TA517 submission.

A.10 Cost-effectiveness results (deterministic)

A summary of the deterministic cost-effectiveness analysis (CEA) results is provided in Table 7. This table contains results in three categories, as per the ToE document:

- **CEA 1a and b:** Replication of the key results at entry to the CDF (1a = upper bound, 1b = lower bound)
- **CEA 2a and b:** Results using mature JM200 data, other parameters unchanged
- **CEA 3:** Revised company base-case results

CEA 2a and 2b fix all model settings and assumptions except for the choice of survival models for OS, PFS, and ToT (which use mature JM200: Part B data). The revised base case (CEA 3) includes updated survival models, the fixed dose of avelumab, and utility values based on updated JM200 data. Further details are provided in Appendix 5.

Table 7: Comparison of cost-effectiveness results (deterministic)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
CEA 1a*							
Chemotherapy	10,608	2.02	1.37				
Avelumab	102,812	4.16	2.65	92,204	2.14	1.28	72,033
CEA 1b*							
Chemotherapy	11,116	1.94	1.34				
Avelumab	99,610	4.58	2.86	88,494	2.64	1.52	58,315
CEA 2a**							
Chemotherapy	10,611	1.94	1.34				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CEA 2b**							
Chemotherapy	11,116	1.94	1.34				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
CEA 3***							
Chemotherapy	11,116	1.94	1.32				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; CDF, Cancer Drugs Fund; CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Notes: *CEA 1a and 1b are based on a replication of analysis that demonstrated plausible potential for cost-effectiveness at CDF entry (a is based on the upper bound, and b is based on the lower bound). **CEA 2a and 2b are analyses based on 1a and 1b, respectively; but incorporating updated clinical evidence (hence the same estimated QALYs and LYs). CEA 2a includes the 'fix' for background mortality which affects the estimated total costs, QALYs, and LYs for the chemotherapy arm. ***CEA 3 represents the new company base-case analysis. Only one set of results is presented, based on the totality of the evidence base now available, and incorporation of additional information (as discussed in Section A.9).

A.11 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was conducted (based on CEA 3) to explore the impact of parameter uncertainty within the cost-effectiveness model when all parameters were varied simultaneously. Model parameters were sampled within their respective distribution and bounds of uncertainty for 1,000 iterations. The results of each iteration were recorded and the average of the results are presented in Table 8. A summary of the model parameters varied within the PSA, including bounds of uncertainty and distributions used, is provided in Appendix 6.

Table 8: Updated base-case results (probabilistic)

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Chemotherapy	[redacted]	1.95	1.33				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

Figure 7 presents the PSA scatterplot for avelumab versus chemotherapy (for CEA 3), where the mean results correspond to Table 8. At a willingness-to-pay threshold for and and-of-life treatment of £50,000 per QALY gained, avelumab is associated with an [redacted] probability of being cost effective.

Figure 7: Scatterplot of probabilistic results

Key: QALY, quality-adjusted life year; WTP, willingness-to-pay.

A.12 Key sensitivity and scenario analyses

The results of a deterministic one-way sensitivity analysis (OWSA) are presented as a tornado diagram in Figure 8 (centred on CEA 3). In this analysis, individual model parameters were varied at their respective lower and upper bounds, with the impact on model results recorded (tabulated model parameters provided in Appendix 6).

As per the original TA517 CS, one of the largest drivers of cost-effectiveness results in the OWSA was medical costs. In addition, both age and sex were shown to lead to variations in the ICER (as each of these parameters influence background mortality within the model). [redacted].

Notably, survival and utility-related values are excluded from this analysis owing to the specification of variance-covariance matrices in the updated utility regression and parametric survival models. This means that each modelled utility or curve parameter cannot be robustly varied in isolation of the other values. The uncertainties concerning utility values and survival models are instead captured within the PSA and explored in scenario analysis.

Figure 8: Tornado diagram

Key: CT, computed tomography; freq, frequency; GP, General Practitioner; EoL, end of life; ICER, incremental cost-effectiveness ratio; MRU, medical resource use; PF, progression-free.

A series of deterministic scenario analyses were undertaken to explore alternative model settings and assumptions. The scenarios expected to be of greatest relevance for decision

making (based on the contents of the ToE document) are presented in Table 9. The scenarios explored cause the ICER to increase by up to [redacted] or decrease by as much as [redacted]; [redacted].

Table 9: Key scenario analyses

Scenario and cross reference	Scenario detail	Brief rationale	Impact on base-case ICER (£)
Base case			[redacted]
OS for avelumab (Section A.8.1)	Use generalised gamma model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (generalised gamma) which projects higher OS versus the base-case analysis	[redacted]
	Use 1-knot normal spline-based model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (1-knot normal spline) which projects lower OS versus the base-case analysis	[redacted]
Clinical expectation of ToT (Section A.8.2)	Assume all patients discontinued by 10 years, with no interim capping at 2 years	Approach intended to serve as an upper bound of potential long-term treatment with avelumab. In practice, discontinuation with avelumab is expected to occur before 5 years	[redacted]
	Assume one-third of patients continue treatment after 2 years, and all discontinue by 5 years	Approach aligned with clinical expert opinion at the time of the original TA517 CS, and is still expected to be broadly representative of clinical practice	[redacted]
Utility values (Section A.8.3)	Use original utility values from TA517	Allows for assessment of impact on cost-effectiveness results through updating utility values	[redacted]
	Use only data from JM200: Part B to inform utility values	Allows exploration of using utility values derived only from a treatment-naïve metastatic MCC population	[redacted]

Key: BSC, best supportive care; CS, company submission; ICER, incremental cost-effectiveness ratio; JM200, JAVELIN Merkel 200; OS, overall survival.

A.13 Key issues and conclusions based on CDF data

This submission presents a summary of the additional evidence collected concerning the use of avelumab for people with previously-untreated metastatic MCC during the period of CDF data collection. Two key evidence sources inform this submission – updated data from Part B of the JM200 clinical trial, as well as the SACT dataset. The updated data from JM200: Part B facilitated an update to the economic model, with reduced uncertainty concerning the clinical effectiveness of avelumab used in the first line setting.

The updated base-case analysis demonstrates that avelumab [redacted]. The economic analysis is based on the previously-submitted modelling approach, with updates to clinical data and assumptions based on data from JM200: Part B. Sensitivity analysis results provided similar findings to the base-case analysis, with the key drivers of cost effectiveness being assumptions relating to OS, ToT, and HRQoL.

The lack of direct comparative evidence for avelumab and chemotherapy remains a key uncertainty within the context of the clinical and cost effectiveness of avelumab. As has been previously discussed, this – in part - is a consequence of metastatic MCC being an ultra-rare disease, and the difficulties of conducting comparative trials in these cases. In lieu of a direct comparison, the model utilises observational study data to facilitate a comparison with the current standard of care. This comparison demonstrates a clear survival advantage for patients treated with avelumab.

Data collected from the SACT cohort provide additional information relevant to this appraisal, yet are subject to a number of important limitations. However, it is unclear how generalisable the SACT cohort is to an avelumab-eligible metastatic MCC population in NHS practice owing to the availability of avelumab in multiple treatment lines.

In conclusion, this submission [redacted] option for people with previously-untreated, metastatic MCC. Avelumab represents a step change in therapy, allowing patients to forgo toxic chemotherapy and derive the full benefits provided by avelumab when used in the first-line setting.

Appendices

Appendix 1. Supporting information from JM200 and SACT

Information relating to JM200

A conference poster presenting the primary analysis of JM200: Part B (the data cut considered within this submission) may be accessed by double-clicking the icon below:

D'Angelo (2019).pdf

Figure 9 presents an overlay of the original and updated OS and PFS Kaplan-Meier curves from JM200: Part B.

Figure 9: Overlay of original OS and PFS Kaplan-Meier curves (JM200: Part B)

Key: OS, overall survival; PFS, progression-free survival.

Data for the outcomes of OS and PFS from the latest available data cut from Part A of JM200 (treatment-experienced patients) were presented at the ISMCC conference in October 2019.

Figure 10 presents a summary of these data, alongside the latest ToT data from Part A and data for all three outcomes (OS, PFS, and ToT) from JM200: Part B.

Figure 10: OS, PFS, and ToT PFS Kaplan-Meier curves (JM200: Parts A and B)

Key: OS, overall survival; PFS, progression-free survival; ToT, time on treatment.

Information relating to SACT

The SACT report for avelumab in metastatic MCC may be accessed by double-clicking the icon below:

SACT report for avelumab in mMCC.

Information relating to pharmacokinetics

Background

As described in the main document, avelumab was studied in patients with metastatic MCC in the pivotal JM200 trial. Within this trial, avelumab was administered at a target dose of

10mg/kg once every two weeks (q2w), through 1-hour intravenous infusions. However, administration of a flat dose is expected to have several practical advantages over the use of a weight-based dose. These advantages include ease of dose preparation, reduced chance of dosing errors, and minimised drug wastage.

In November 2019, the licensed dose of avelumab for metastatic MCC changed from the weight-based dose studied in JM200 to a flat dose of 800mg for all patients. The 800 mg dose was selected as median body weight for adults with various tumour types has been shown to be approximately 80 kg in previous studies, and a dose of 800 mg would directly correspond to the same dose with 10 mg/kg weight-based dosing.

This appendix provides a summary of the information provided to the European Medicines Agency (EMA) to inform its decision to amend the licensed dose of avelumab in metastatic MCC. For further information, a recent publication by Novakovic *et al.*, (2020) details the analytical approach to assessing the alternative doses of avelumab, which includes further description of the results.[14]

Methods

Pharmacokinetic (PK) analyses were performed for avelumab, comparing the 10mg/kg weight-based dosing and the 800 mg flat dosing regimen. Data for 1,827 patients enrolled in 3 clinical trials (NCT01772004, NCT01943461, and NCT02155647) were used to inform the analysis:

- JAVELIN Solid Tumour (Phase I, NCT01772004/EMR100070-001)
  - Phase 1a: patients with various tumours received avelumab at 1, 3, 10, or 20 mg/kg every 2 weeks via a 1-hour intravenous infusion
  - Phase 1b: patients were enrolled into tumour-specific cohorts and received avelumab at 10 mg/kg every 2 weeks
- JAVELIN Solid Tumour JPN (Phase I, NCT01943461/EMR100070-002)
  - Initial dose-escalation: patients received avelumab at 3, 10, or 20 mg/kg every 2 weeks
  - Dose-expansion: patients with advanced/ metastatic gastric or gastroesophageal cancer received avelumab at 10 mg/kg every 2 weeks
- JAVELIN Merkel 200: Part A (Phase II, NCT02155647/EMR100070-003)
  - Patients received 10mg/kg every 2 weeks (analysis included only treatment-experienced metastatic MCC patients)

The analysis had three distinct objectives:

- To compare avelumab exposure between weight-based and flat-dosing regimens using simulations based on previously developed population PK models

- To compare the simulated probability of experiencing an adverse event (AE) of special interest with weight-based and flat-dosing regimens, specifically immune-related AEs (irAEs) and infusion-related reactions (IRRs)

- To compare the simulated probability of objective response (OR) between weight-based and flat-dosing regimens in patients with mMCC or advanced/metastatic urothelial carcinoma (UC)

Results

Exposure

Across the three clinical trials the median body weight among participants was 70.6 kg (range 30.4–204 kg). Avelumab was administered at doses of 1 mg/kg (n = 4), 3 mg/kg (n = 18), 10 mg/kg (n = 1,778), or 20 mg/kg (n = 27) q2w. The flat dose resulted in slightly higher exposures than weight-based dosing, with the median area under the curve (AUC) during the first dosing interval increasing by approximately 12%. This was expected as the median weight (70.6 kg) was lower than the weight used to determine the flat dose (80.0 kg).

For weight-based dosing, exposures were lowest in the lightest participants, whereas for flat dosing exposures were lowest in the heaviest participants. However, exposures in all weight groups showed considerable overlap. Overall, population PK modelling and simulations suggested that exposure to avelumab was similar with flat and weight-based dosing regimens.

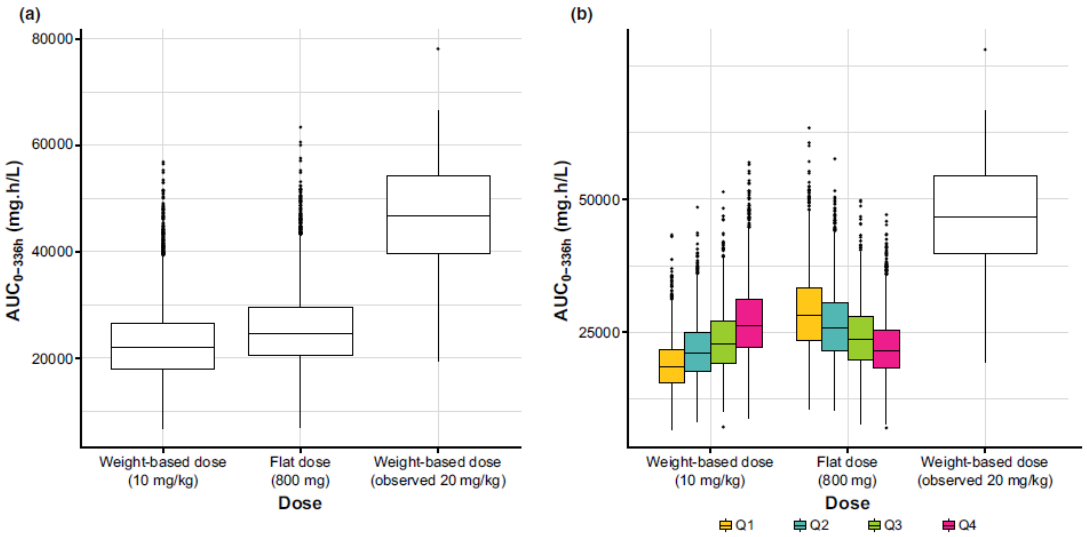


Figure 1 Simulated $AUC_{0-336\,h}$ values for weight-based (10 mg/kg q2w) and flat (800 mg q2w) dosing of avelumab using the first-cycle population pharmacokinetic model. Box and whisker plots for (**a**) the entire population and (**b**) the population split by quartiles of weight; observed data with avelumab 20 mg/kg dosing are included for comparison purposes (n = 27). $AUC_{0-336\,h}$, area under the curve during the first dosing interval.

Adverse events

The simulated probability of irAEs across all patients with weight-based dosing was 11.9%, which was the same as the observed rate with avelumab 10 mg/kg dosing across the clinical trials included. The simulated probability of experiencing an irAE based on the AUC during the first dosing interval showed a similar and overlapping distribution between flat and weight-based dosing. The probability of irAEs was slightly higher for flat dosing (12.6%) compared with 10mg/kg weight-based dosing (11.4%), which can be attributed to the higher exposure of the flat-dosing regimen resulting from the median weight of the sampled patients being <80 kg.

Simulated probabilities of IRRs based on exposure were strongly concordant between weight-based and flat dosing, with no trends seen in weight quartiles. This finding was expected because a previously developed exposure-IRR model concluded that the probability of IRRs does not change with exposure. Overall, the exposure-based analyses suggested that the safety profile of flat dosing for avelumab is similar to that of weight-based dosing.

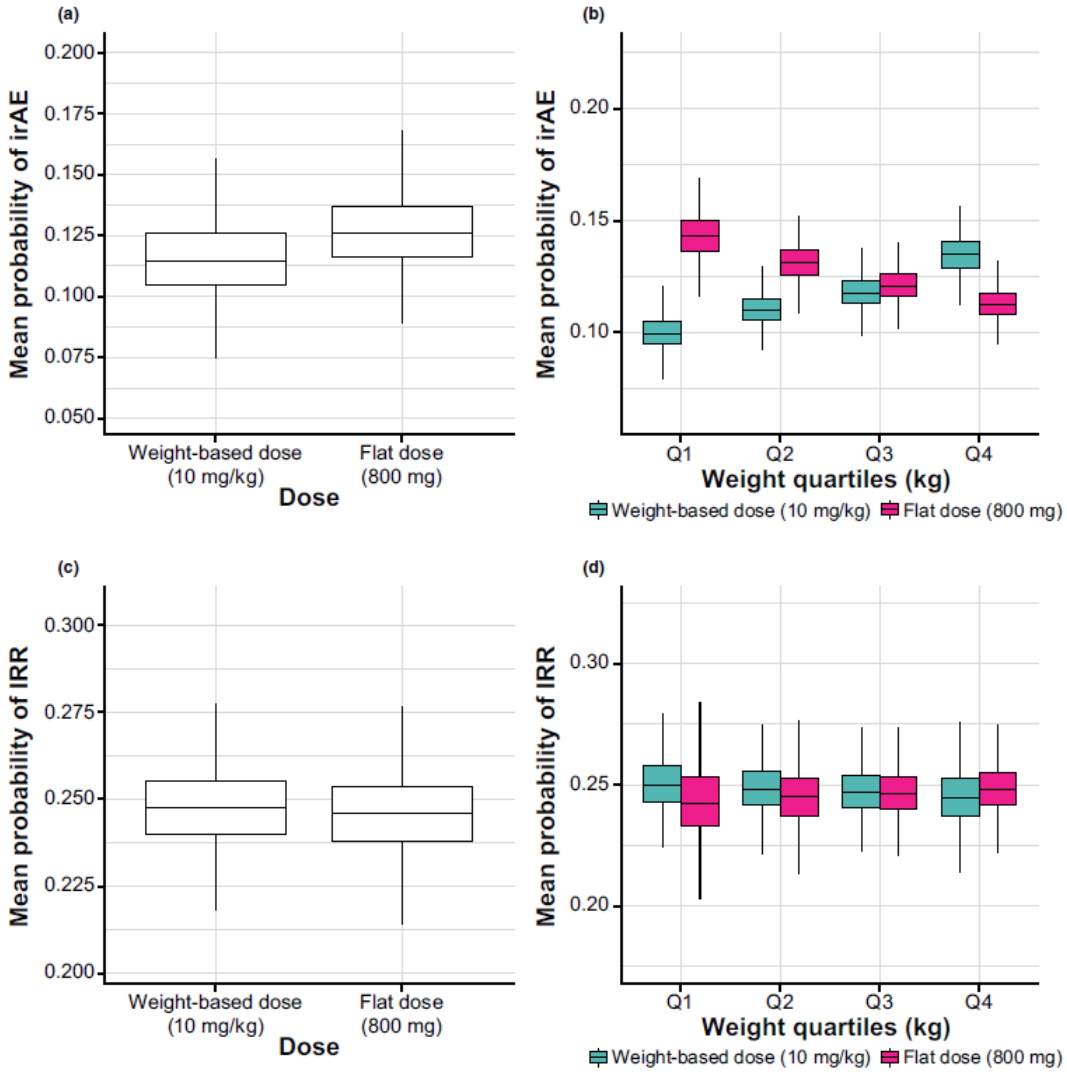


Figure 2 Mean probability of experiencing an irAE (upper panels) or IRR (lower panels) for weight-based (10 mg/kg q2w) and flat (800 mg q2w) dosing with avelumab based on the first-cycle population pharmacokinetic model. Box and whisker plots for (**a**) probability of irAEs based on $AUC_{0-336\ h}$ in all patients; (**b**) probability of irAEs based on $AUC_{0-336\ h}$ stratified by quartiles of weight; (**c**) probability of IRRs based on C_{max} in all patients; and (**d**) probability of IRRs based on C_{max} stratified by quartiles of weight. $AUC_{0-336\ h}$, area under the concentration curve during the first dosing interval; C_{max}, maximum concentration; irAE, immune-related adverse event; IRR, infusion-related reaction.

Response

To compare the simulated probability of OR, data were analysed in 88 patients with mMCC from part A of JAVELIN Merkel 200 and 249 patients with advanced/metastatic UC from JAVELIN Solid Tumour. In both mMCC and advanced/metastatic UC populations, the simulated probability of OR based on exposure was slightly higher with the 800 mg flat dose than weight-based dosing with 10 mg/kg.

Across all weight quartiles for both tumour types, there was substantial overlap between the weight-based and flat-dose regimens but with opposing trends. More specifically, the probability of OR was highest for the heaviest weight quartile with weight-based dosing and for the lowest weight quartile with flat dosing. Overall, the exposure-efficacy simulations

indicated that the probability of OR is likely to be similar with flat or weight-based dosing in the populations examined.

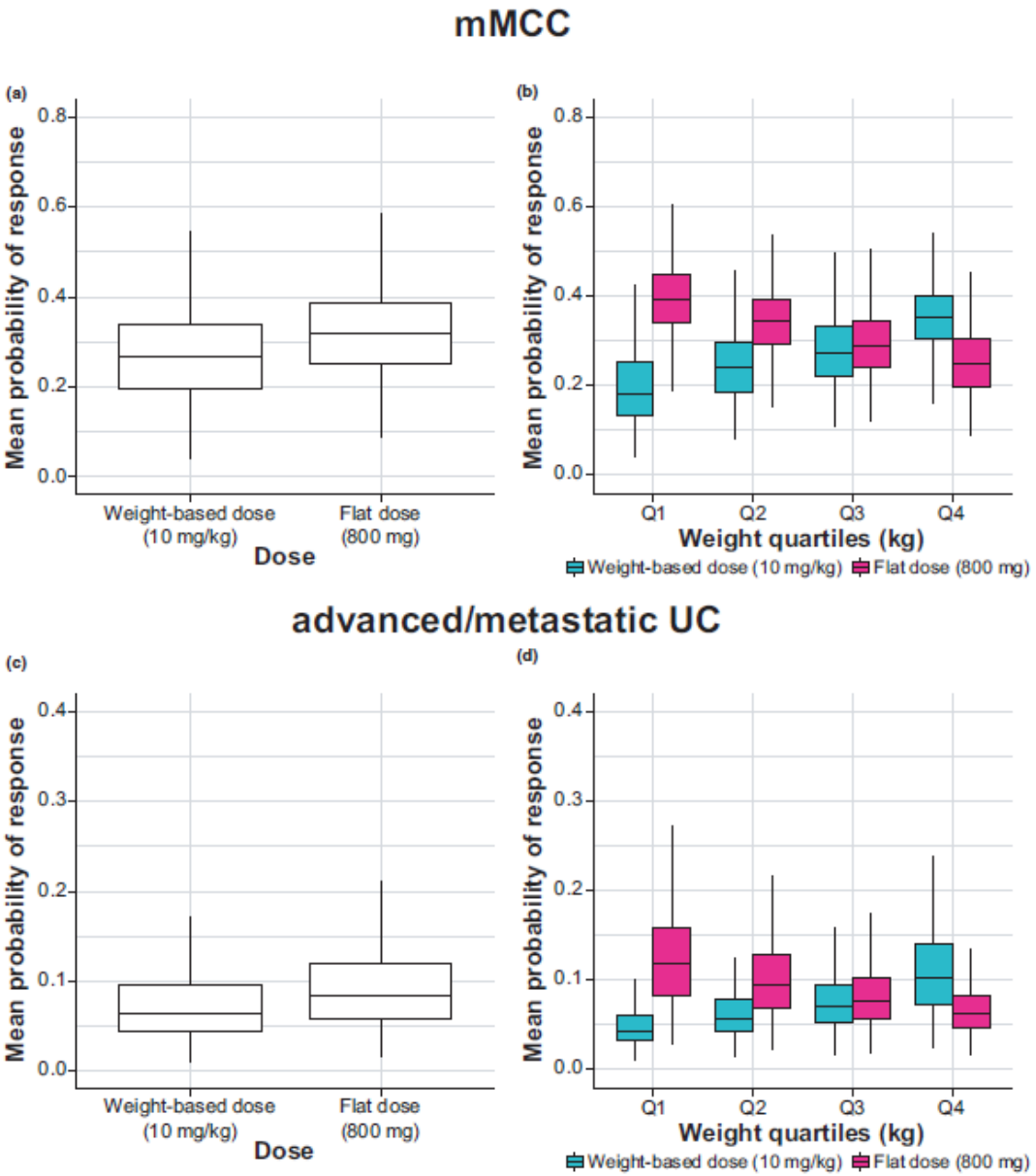


Figure 3 Mean probability of objective response in patients with mMCC (upper panels) or advanced/metastatic UC (lower panels) for weight-based (10 mg/kg q2w) and flat (800 mg q2w) dosing with avelumab based on $AUC_{0-336\,h}$ (first-cycle population pharmacokinetic model). Box and whisker plots in (**a**) all patients with mMCC, (**b**) patients with mMCC stratified by quartiles of weight, (**c**) all patients with advanced/metastatic UC, and (**d**) patients with advanced/metastatic UC stratified by quartiles of weight. $AUC_{0-336\,h}$, area under the curve during the first dosing interval; C_{trough}, minimum serum concentrations; mMCC, metastatic Merkel cell carcinoma; UC, urothelial carcinoma.

Conclusion

Population PK modelling and simulation, based on a large patient data set, support the use of a flat avelumab dose of 800 mg q2w instead of the weight-based 10 mg/kg q2w dose that was approved initially.

Appendix 2. Update to the systematic literature review

Clinical SLR

The update to the clinical SLR can be accessed via the icon below:

Clinical SLR_v1-0.docx

Cost-effectiveness, healthcare resource use, and utilities SLR

In addition to the update to the clinical SLR, for completeness updates have also been undertaken to identify any additional cost-effectiveness analyses, healthcare resource use, and utility analysis. These update search reports can be accessed using the icons below:

Cost-effectiveness SLR_v1-0.docx

HCRU and costs SLR_v1-0.docx

Utilities SLR_v1-0.docx

Appendix 3. Supporting survival modelling information

Statistical goodness-of-fit scores

Statistical goodness-of-fit scores are provided in Table 10. The lowest score for each outcome and statistic are presented in **bold print**. Models that provide an AIC or BIC within 2 points of the lowest score are highlighted in black. Models that provide and AIC or BIC within 4 points of the lowest score are highlighted in grey. The selected base-case curves are filled in green.

Table 10: Statistical goodness-of-fit scores (OS, PFS, and ToT)

Model	OS		PFS		ToT*	
	AIC	BIC	AIC	BIC	AIC	BIC
1-knot, odds	501.76	510.02	483.25	491.51	1,217.40	1,225.66
1-knot, normal	**501.46**	509.72	481.43	489.69	1,217.00	1,225.26
1-knot, hazard	501.90	510.16	488.24	496.50	1,216.10	1,224.36
2-knots, odds	503.77	514.78	462.71	473.72	1,219.27	1,230.28
2-knots, normal	503.30	514.32	473.65	484.66	1,218.05	1,229.06
2-knots, hazard	503.90	514.92	463.94	474.95	1,218.18	1,229.20
3-knots, odds	505.68	519.45	**455.04**	**468.81**	1,210.65	1,224.41
3-knots, normal	505.25	519.01	**	**	1,210.65	1,224.42
3-knots, hazard	505.77	519.54	461.70	475.46	**1,209.55**	1,223.31
Exponential	510.52	513.27	552.28	555.03	1,242.81	1,245.56
Weibull	509.90	515.41	536.83	542.34	1,216.39	1,221.90
Gompertz	503.08	508.59	514.33	519.83	1,225.60	1,231.11
Log-logistic	505.37	510.87	517.22	522.73	1,215.42	**1,220.93**
Lognormal	502.04	**507.55**	512.03	517.54	1,217.38	1,222.89
Generalised gamma	501.05	509.31	486.91	495.17	1,216.77	1,225.04

Key: AIC, Akaike's information criterion; BIC, Bayesian information criterion; OS, overall survival; PFS, progression-free survival; ToT, time on treatment.
Note: *For the outcome of ToT, models were fitted based on time in the unit of "days", whereas for OS and PFS models were fitted based on time in the unit of "months". For this reason, the AIC and BIC scores for ToT are much larger (in absolute terms) compared to the values for OS and PFS. Please note that these figures refer to the fits without adjustment, and so for the outcome of ToT these values should be interpreted with this in mind. **For the outcome of PFS, the standard optimisation algorithm resulted in a model error when fitting the 3-knots normal spline model. Consequently, this model was fitted using the alternative BFGS optimisation algorithm. As such, the AIC and BIC scores based on this model should not be compared to the other AIC and BIC scores.

Based on the statistical goodness-of-fit scores, the following may be inferred:

- For OS, 1-knot spline-based models are preferred based on the AIC score; whereas simpler parametric models are preferred according to the BIC score. The 1-knot spline-based models provide AIC and BIC scores within 4 points of the statistically best-fitting model

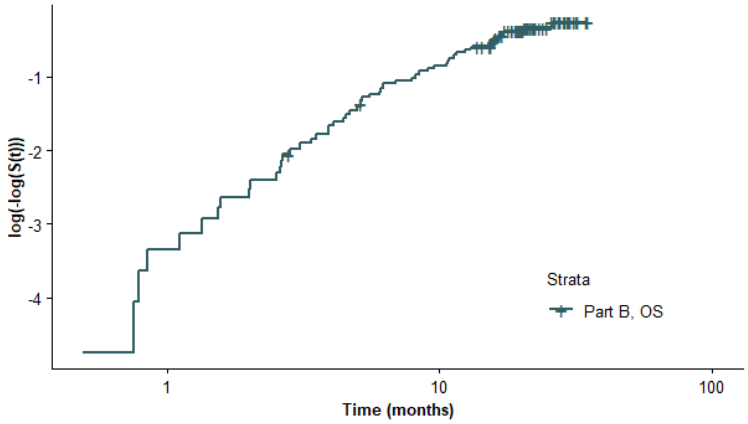
- For PFS, a spline-based model with multiple knots is preferred to reflect the complex patterns of hazards. However, of the simpler parametric models, the generalised gamma provides a notably lower score compared with the other alternatives

- For ToT, AIC scores support the use of a complex 3-knot based spline model, whereas the BIC scores indicate a preference for a simpler parametric model. Both the exponential and Gompertz parameterisations provide relatively poor statistical goodness-of-fit scores

Log-cumulative hazard plots

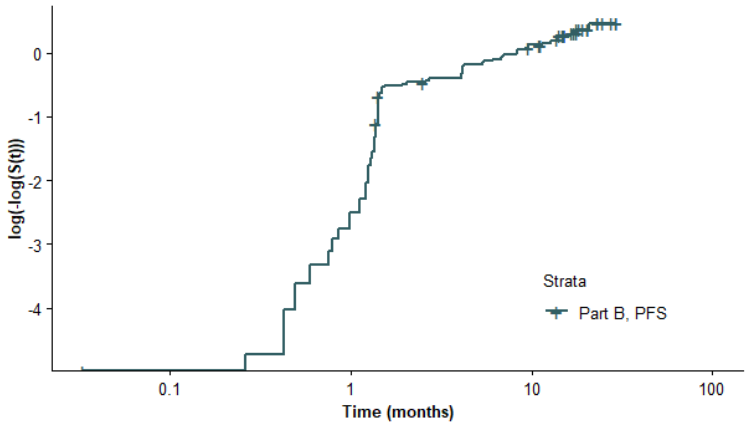
Log-cumulative hazard plots for the outcomes of OS, PFS, and ToT are presented in Figure 11, Figure 12, and Figure 13, respectively. The plot for OS exhibits a shape indicative of a reducing hazard over time (i.e. a flattening of the curve as time goes on). This shape is much more pronounced in the PFS curve, which shows a sharp change in the hazard at approximately 1 month. No clear patterns of hazards for the outcome of ToT is evident from inspection of the log-cumulative hazard plot.

Figure 11: Log-cumulative hazard plot – JM200: Part B, OS



Key: OS, overall survival.

Figure 12: Log-cumulative hazard plot – JM200: Part B, PFS



Key: PFS, progression-free survival.

Figure 13: Log-cumulative hazard plot – JM200: Part B, ToT

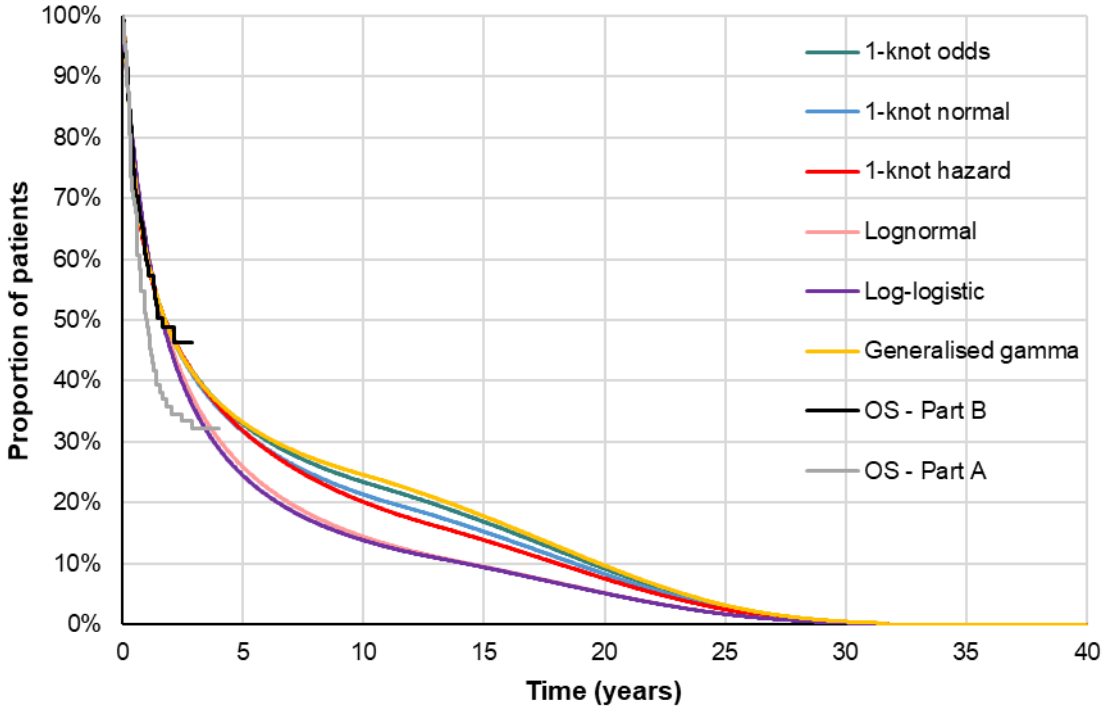
[redacted]

Key: ToT, time on treatment.

Alternative extrapolations

Alternative extrapolations for the outcome of OS are provided in Figure 14.

Figure 14: Alternative parametric models - OS



Key: OS, overall survival.

Time-on-treatment extrapolation

As described in Section A.8.2, the base extrapolation for ToT was adjusted to align with clinical expectation. The base extrapolation without adjustment is presented alongside the adjusted extrapolation in

Figure 15 (dashed green versus solid green line, respectively). This diagram shows that up until the minimum follow-up period in JM200: Part B (15 months), the adjusted (solid) and unadjusted (dashed) extrapolations are identical.

The unadjusted extrapolation, while providing a reasonably good fit to the Kaplan-Meier curve, was considered to over-estimate the duration of treatment exposure, both in terms of discontinuation in the medium-term after the minimum follow-up period until the end of the Kaplan-Meier curve, and in the longer-term (where some patients were predicted to remain on treatment for more than 5 years).

Figure 15: ToT extrapolation, avelumab – presentation of adjustments

Key: ToT, time on treatment.

Two adjustments were made to the base extrapolation:

- Firstly, the probability of discontinuing treatment after 15 months was assumed to be based on the extrapolation for treatment-experienced patients. This approach allows for the use of the more mature treatment-experienced data from Part A of JM200 while maintaining the initial curve fit based on data from Part B

- Secondly, to avoid the longer-term projections of ToT estimating some patients to be treated beyond 5 years, the model assumes all patients still receiving treatment at 5 years would immediately discontinue. In reality, nearly all patients are expected to have discontinued prior to this point in time, though it remains unclear exactly when all patients would have discontinued treatment. Based on clinical advice provided to Merck Serono, nearly all patients would be expected to have discontinued treatment by 5 years. Therefore, Merck Serono selected this time point to represent a definitive cap for the ToT curve

Appendix 4. Updated utility analysis

Background

In JM200: Part A and Part B, data were collected via completion of the EQ-5D-5L questionnaire. EQ-5D-5L data were scored using the recommended 'crosswalk' algorithm between the EQ-5D-5L and the EQ-5D-3L, to produce utility values suitable for inclusion within the model.[15] The EQ-5D-3L were subsequently analysed using generalised estimating equation (GEE) regression to account for multiple observations per patient.

In the original CS, data were available from Part A of JM200 to inform model utility values, but no data were available from Part B. Since this time, sufficient data are now available from both cohorts enrolled within the JM200 trial. This appendix provides a summary of the updated utility analysis available to inform the model.

Methods

Since the model makes use of utility values based on the time between the recorded utility and death, a number of approaches were taken to estimate possible utility values:

- Use of the original utility values based solely on data from Part A of JM200
    - This analysis allows for the production of results using the original utility values, as requested in the ToE document
- Analysis of data from either Part A or Part B of JM200
    - This analysis provides utility values in treatment-experienced or treatment-naïve only populations (the latter being the population of primary relevance to this appraisal)
- Analysis of data from both Part A and Part B of JM200
    - This analysis aims to make use of utility data from both populations, accounting for differences via the inclusion of a covariate for treatment history (i.e. naïve or experienced)

The approach taken to analyse the utility data from Part B (either alone or in combination with data from Part A) of JM200 is similar to the approach taken in the original CS of using an algorithm to select the best fitting groupings given a set of criteria. The original CS did this using the *optim* function in R, whereas in this resubmission a more sophisticated

approach was used with additional criteria. In this updated approach other criteria were first applied including:

- Health states must be in multiples of 7 days, to align with the cycle length of the economic model
- The shortest clinically plausible health state duration was deemed to be 14 days
- Health states were required to be the same or increasing in duration as the time from death increased e.g. if the final 21 days of life were the first health state, the second must be at least 21 days in duration
- All potential health states were required to have at least 10 utility values included to ensure reliable estimation of coefficients and confidence intervals

To identify optimal groupings (or cut points), the full JM200 data set was used. Alternative model specifications were ranked in a 'league table' able to be sorted by different goodness of fit statistic – Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or Quasi-Information Criterion (QIC). A fourth option was added, which was the mean rank of the regression across all three metrics.

A criticism of the original utility analysis was that the values appeared "implausibly high". This was noted as a potential limitation of the approach to select relevant cut points for the analysis, as the largest drop in utility used to inform the model was in the region of 0.06-0.07 (i.e. from 0.7744 to 0.7082). To address this, an additional assumption was imposed within the time-to-death analysis to require the minimum difference between distinct health state modelled to be at least 0.08. This assumption ensures that modelled changes in utility are categorised according to meaningful changes. The value of 0.08 was selected based on the minimal clinically important difference in cancer for the EQ-5D.[16]

Results

When looking at the fitted models preferred by each of the 4 criteria, all used a 35-day cut point to define the group closest to death. In terms of the preferred model there is little to choose between them, except for the QIC-preferred model which performed poorly on other metrics, and was thus discarded from further considerations. The candidate models included by MAE ranking 35/266 days, by RMSE ranking 35/105/294 days, and by joint ranking 35/70/245 days. If one has to be selected, that by MAE ranking (35/266) was considered the

preferred model, simply as the most parsimonious – in requiring one fewer coefficient, yet ranking first in MAE, and well on other criteria.

As a sensitivity analysis, cut points were selected to resemble the original analysis presented in TA517. The lower cut-point of 30 days was changed to 28 days (to align with the model cycle length, where 28 days = 4 weeks), and the upper cut-point of 100 days was changed to 84 days (to align with the model cycle length, where 84 days = 12 weeks).

The resultant utilities considered for use within the economic model are presented in Table 11. Although all the values are similar, the updated analysis using data from Part B only yielded similar results for the state furthest from death, but with lower values closer to death. The analysis incorporating data from both Part A and B suggested different cut points – with the group furthest from death defined as those with approximately 9 months until death, and the group closest to death defined as those with approximately 1 month until death.

Table 11: Options for utility values within the updated model

Label	Health state	Utility value/ coefficient		
		Part B only	Part A only	Part A and B
Original TA517 values	> 100 days to death		0.7744	
	30-100 days to death		0.7540	
	< 30 days to death		0.7082	
Optimal cut-points	> 266 days to death	0.8128	0.7561	0.8019
	35-266 days to death	0.6893	0.6943	0.7096
	< 35 days to death	0.4206	0.4174	0.4411
	Coefficient: Treatment-experienced			-0.0348
Sensitivity analysis cut-points	> 84 days to death	0.7837	0.7494	0.7839
	28-84 days to death	0.6487	0.6208	0.6525
	< 28 days to death	0.3951	0.2804	0.3513
	Coefficient: Treatment-experienced			-0.0349

For context, it was also considered important to estimate the average utility of avelumab-treated patients over the course of the modelled time horizon. To do this, the annual discount rate for QALYs and life-years were both set to 0% and the quotient of the total values accrued over the model time horizon was taken. The results of this analysis are presented in Table 12.

Table 12: Modelled average utility

Description	Average utility value
Original TA517 values	0.7744
Optimal cut-points (35, 266), Parts A and B	0.7753
Optimal cut-points (35, 266), Part B only	0.7453
Optimal cut-points (35, 266), Part A only	0.7401
Sensitivity analysis cut-points (28, 84), Parts A and B	0.7881
Sensitivity analysis cut-points (28, 84), Part B only	0.7750
Sensitivity analysis cut-points (28, 84), Part A only	0.7958

Discussion

This analysis provides utility values to inform the economic model using three similar, yet distinct, approaches. Each of the analyses exhibit a fall in utility as observations are taken closer to death whether using the two sets of trial data separately, or combining them. Based on the analysis using data from both cohorts (Table 11), it can be seen that average utility in patients with treatment-naïve metastatic MCC is slightly higher than average utility in patients with treatment-experienced metastatic MCC as may be expected.

The average utility using each approach (Table 12) ranged between 0.7401 to 0.7958. The original TA517 analysis was based on data from Part A of the JM200 trial only, and importantly using the original data cut of the study in a treatment-experienced population. In this data cut, relatively few observations were taken close to death (acknowledging that 565 utility values were available for analysis, provided by 79 patients). As previously highlighted, this approach allowed for the difference between modelled health states to be any value, which is why a small difference between each of the utility values is observed (Table 11).

In Part B of JM200, there are an increased number of patients available to inform the utility analysis (725 utility values provided by 103 patients), and in this approach a minimum difference in utility of 0.08 between health states was imposed. The cut points broadly align with the original Part A analysis, though by moving the cut-point for the state furthest from death to be closer to death, a larger decline in utility is modelled. When setting the cut points similar to the original TA517 analysis, a similar, yet slightly higher average utility is noted when analysing only data from Part B of JM200.

Due to increasing maturity of the data from Part A of JM200, the number of observations available to inform an updated analysis from this dataset increased from 362 observations provided by 71 patients, to 565 observations provided by 79 patients. Combined with the data from Part B, a more in-depth assessment of the relationship between the time to death, treatment line, and utility can be established (with 1,290 observations available). By

considering the totality of the available utility data, the optimal cut point for the group furthest from death shifted from approximately 3 months to approximately 9 months. By moving the upper limit further from death, the utility value furthest from death increases, yet the utility value closest to death decreases markedly (Table 11).

Conclusion

The analyses undertaken demonstrate similar patterns of changes in utility as observations are taken closer to death. Cost-effectiveness analyses using all three of these approaches to capture the HRQoL for patients with treatment-naïve metastatic MCC are expected to be relevant for decision making. However, given that the third analysis makes use of data from both cohorts in JM200 (accounting for differences in utility by treatment line through the use of a covariate), these utility values were considered the most appropriate for informing decision making.

Appendix 5. Key model assumptions and inputs

A description of the different model assumptions, inputs, and settings that correspond to each of the modelled CEAs is provided in Table 13. The information provided in Table 13 is intended to serve as a reference point for comparing the differences between the modelled CEAs. CEA 1a serves as the reference scenario, from which all other CEAs are described in relative terms.

Table 13: Description of differences in model settings and assumptions

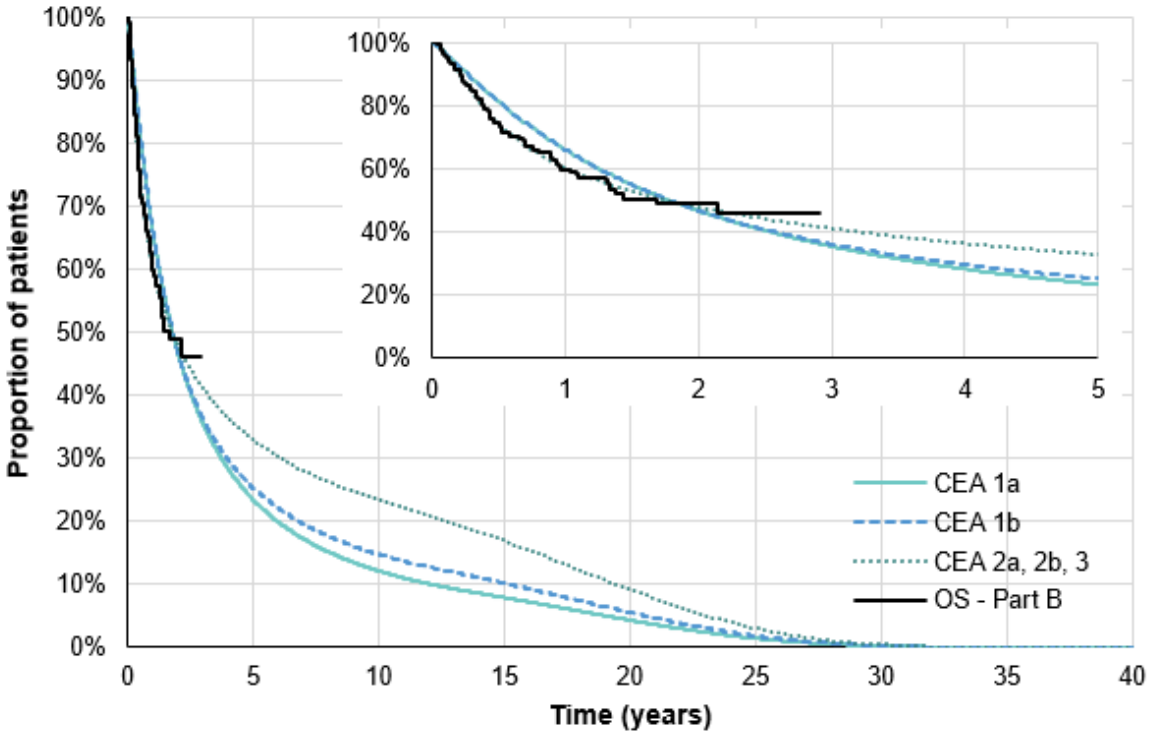
CEA	Summary of changes in modelled analyses
1a*	Reference scenario
1b**	In addition to settings specified for 1a: • Includes revised administration cost (£253 instead of £199) • Caps proportion of patients receiving avelumab at 2 years at 5% • Informs OS extrapolation with JM200: Part A hazards after 21 months • Fixes error in background mortality
2a	In addition to settings specified for 1a: • Uses parametric survival models for OS, PFS, and ToT, based on JM200: Part B data • Includes adjustment in ToT extrapolation based on JM200: Part A hazards • Fixes error in background mortality***
2b	In addition to settings specified for 1a: • Uses parametric survival models for OS, PFS, and ToT, based on JM200: Part B data • Includes adjustment in ToT extrapolation based on JM200: Part A hazards • Includes revised administration cost (£253 instead of £199) • Fixes error in background mortality
3	In addition to settings specified for 2b: • Uses parametric survival models for OS, PFS, and ToT, based on JM200: Part B data • Includes adjustment in ToT extrapolation based on JM200: Part A hazards • Includes revised administration cost (£253 instead of £199) • Fixes error in background mortality • Includes flat dose of avelumab • Includes updated utility values (based on analysis of JM200: Part A and B data)

Key: CDF, Cancer Drugs Fund; CEA, cost-effectiveness analysis; ERG, Evidence Review Group; FAD, Final Appraisal Determination; JM200: ICER, incremental cost-effectiveness ratio; JM200, JAVELIN Merkel 200; OS, overall survival; PFS, progression-free survival; ToE, Terms of Engagement; ToT, time on treatment.
Note: * The ICER for this CEA is £72,033. In the TA517 FAD, this is referred to as the ERG's revised base-case ICER. ** The ICER for this CEA is £58,315. In the TA517 FAD, this is referred to as the company's revised base-case ICER. ***As discussed within the ToE document, one of the expected corrections to be made in the CDF review is an error that was found in the calculation of background mortality. For completeness, this error is fixed in this analysis, which affects the total costs, QALYs, and LYs for the chemotherapy arm. Should the equivalent results without fixing this error be of interest, please compare the results for chemotherapy in CEA 1a with the results for avelumab in CEA 2a.

In the economic model provided alongside this submission, macros have been included to automatically generate the results associated with each scenario. Therefore, Merck Serono encourages reference to the economic model file itself for an exhaustive set of model assumptions and inputs.

A key driver of model results is the extrapolation of OS for patients receiving avelumab. Thus, for comparison purposes, Figure 16 provides a comparison of the original extrapolations that inform each of the modelled CEAs. The same extrapolation is used in CEAs 2a, 2b, and 3.

Figure 16: Comparison of OS extrapolations



Key: OS, overall survival.
Notes: CEA 1a and 1b are based on a replication of analysis that demonstrated plausible potential for cost-effectiveness at CDF entry (a is based on the upper bound, and b is based on the lower bound). CEA 2a, 2b, and 3 are analyses that incorporate updated clinical evidence.

The original approaches estimated a larger proportion of patients to be alive up until approximately 18 months, after which the curves cross the Kaplan-Meier and continue to project long-term OS estimates that fall below the updated extrapolation. This is unsurprising, as noted in Merck Serono's previous response to the Appraisal Consultation Document (ACD):

> *"The ERG's* [preferred] *model projections do not account for the expected immuneoncology (IO) plateau which has been demonstrated with longer follow up in all IOs including avelumab's treatment experienced data."* – Merck Serono's response to the ACD, page 3.[17]

Appendix 6. Economic model parameters

Table 14: Tabulated economic model parameters

Parameter	Value	SE	Distribution
Annual discount rate - costs	3.5%		Fixed
Annual discount rate - QALYs	3.5%		Fixed
Annual discount rate - LYs	0.0%		Fixed
Cycle length (weeks)	1.00		Fixed
Time horizon (years)	40.0		Fixed
Age (years)	69.3	1.14	Normal
Weight (kg)	78.50		Fixed*
Proportion male	79.3%	0.08	Beta
RDI: Chemotherapy	0.67	0.07	Normal
Utility: >266 days to death	0.78		Varied using variance-covariance matrix
Utility: 35-266 days to death	0.65		
Utility: <35 days to death	0.40		
Admin cost: All drugs	£253.32	25.33	Normal
Cost: GP visit	£36.00	3.60	Normal
Cost: CT scan	£120.99	12.10	Normal
Cost: FBC	£3.00	0.30	Normal
Cost: LFT	£1.00	0.10	Normal
Cost: RFT	£1.00	0.10	Normal
Cost: TFT	£1.00	0.10	Normal
Cost: Radiotherapy	£126.60	12.66	Normal
Cost: EoL, Health care	£4,868	486.75	Normal
Cost: EoL, Social care	£2,152	215.16	Normal
MRU freq: GP visit, avelumab, PF	0.25	0.03	Normal
MRU freq: CT scan, avelumab, PF	0.08	0.01	Normal
MRU freq: FBC, avelumab, PF	0.50	0.05	Normal
MRU freq: LFT, avelumab, PF	0.50	0.05	Normal
MRU freq: RFT, avelumab, PF	0.50	0.05	Normal
MRU freq: TFT, avelumab, PF	0.50	0.05	Normal
MRU freq: GP visit, chemo, PF	0.33	0.03	Normal
MRU freq: CT scan, chemo, PF	0.12	0.01	Normal
MRU freq: FBC, chemo, PF	0.33	0.03	Normal
MRU freq: LFT, chemo, PF	0.33	0.03	Normal
MRU freq: RFT, chemo, PF	0.33	0.03	Normal
MRU freq: TFT, chemo, PF	0.00	0.00	Normal
MRU freq: Duration of radiotherapy	3.75	0.38	Normal
Cost: Avelumab	£768.00		Fixed**
Cost: Carboplatin	£25.25	0.26	Normal
Cost: Etoposide (oral)	£87.23		Fixed**
Cost: Anaemia	£799.39	79.94	Normal
Cost: Dyspnoea	£256.62	25.66	Normal
Cost: Fatigue	£66.45	6.65	Normal
Cost: Febrile neutropenia	£4,543.44	454.34	Normal
Cost: Low haemoglobin	£66.45	6.65	Normal
Cost: Hyponatremia	£66.45	6.65	Normal
Cost: Infections	£256.62	25.66	Normal
Cost: Leukopenia	£281.67	28.17	Normal

Cost: Lymphopenia	£281.67	28.17	Normal
Cost: Muscle pain	£153.49	15.35	Normal
Cost: Nausea/vomiting	£218.27	21.83	Normal
Cost: Neutropenia	£281.67	28.17	Normal
Cost: Low platelets	£281.67	28.17	Normal
Cost: Sensory neuropathy	£446.59	44.66	Normal
Cost: Thrombocytopenia	£286.12	28.61	Normal
Cost: Hair loss	£0.00	0.00	Normal
Disutility: Anaemia	-0.09	0.02	Beta
Disutility: Dyspnoea	-0.05	0.01	Beta
Disutility: Fatigue	-0.07	0.02	Beta
Disutility: Febrile neutropenia	-0.09	0.02	Beta
Disutility: Low haemoglobin	-0.08	0.02	Beta
Disutility: Hyponatremia	-0.09	0.02	Beta
Disutility: Infections	-0.12	0.01	Beta
Disutility: Leukopenia	-0.09	0.02	Beta
Disutility: Lymphopenia	-0.09	0.02	Beta
Disutility: Muscle pain	-0.05	0.02	Beta
Disutility: Nausea/vomiting	-0.05	0.02	Beta
Disutility: Neutropenia	-0.09	0.02	Beta
Disutility: Low platelets	-0.09	0.02	Beta
Disutility: Sensory neuropathy	-0.23	0.02	Beta
Disutility: Thrombocytopenia	-0.11	0.01	Beta
Disutility: Hair loss	-0.04	0.01	Beta
Duration: Anaemia	21.00	2.10	Normal
Duration: Dyspnoea	21.00	2.10	Normal
Duration: Fatigue	21.00	2.10	Normal
Duration: Febrile neutropenia	4.00	0.40	Normal
Duration: Low haemoglobin	21.00	2.10	Normal
Duration: Hyponatremia	1.90	0.19	Normal
Duration: Infections	14.90	1.49	Normal
Duration: Leukopenia	1.90	0.19	Normal
Duration: Lymphopenia	1.90	0.19	Normal
Duration: Muscle pain	7.20	0.72	Normal
Duration: Nausea/vomiting	3.00	0.30	Normal
Duration: Neutropenia	1.90	0.19	Normal
Duration: Low platelets	1.90	0.19	Normal
Duration: Sensory neuropathy	35.30	3.53	Normal
Duration: Thrombocytopenia	23.80	2.38	Normal
Duration: Hair loss	21.00	2.10	Normal

Key: CT, computed tomography; EoL, end of life; FBC, full blood count; freq, frequency; GP, General Practitioner; kg, kilogram; LFT, liver function test; LY, life-year; PF, progression-free; QALY, quality-adjusted life year; RFT, renal function test; SE, standard error; TFT, thyroid function test.

Notes: *Fixed for simplicity, as weight only affects comparator costs in revised base-case analysis (due to flat dosing of avelumab). **Fixed as branded drug cost.

References

1. National Institute for Health and Care Excellence (NICE). TA517: Final Appraisal Determination [Internet]. 2018 [cited 2020 Feb 21]. Available from: https://www.nice.org.uk/guidance/ta517/documents/final-appraisal-determination-document

2. National Institute for Health and Care Excellence (NICE). TA517: Managed Access Agreement [Internet]. 2018 [cited 2020 Feb 21]. Available from: https://www.nice.org.uk/guidance/ta517/documents/final-appraisal-determination-document-2

3. D'Angelo SP, Lebbé C, Mortier L, Brohl AS, Fazio N, Grob J-J, et al. First-line avelumab treatment in patients with metastatic Merkel cell carcinoma: primary analysis after ≥15 months of follow-up from JAVELIN Merkel 200, a registrational phase 2 trial. SITC 2019; 2019 Nov 6; National Harbor, Maryland, USA.

4. National Institute for Health and Care Excellence (NICE). TA517: Terms of Engagement. 2019 Dec.

5. European Medicines Agency (EMA). Avelumab Summary of Product Characteristics [Internet]. 2019 [cited 2020 Feb 21]. Available from: https://www.ema.europa.eu/en/documents/product-information/bavencio-epar-product-information_en.pdf

6. National Health Service (NHS) England. National Dose Banding Table – Avelumab [Internet]. [cited 2020 Feb 21]. Available from: https://www.england.nhs.uk/wp-content/uploads/2018/01/national-tables-avelumab-20mgmL-v2.pdf

7. Cowey CL, Mahnke L, Espirito J, Helwig C, Oksen D, Bharmal M. Real-world treatment outcomes in patients with metastatic Merkel cell carcinoma treated with chemotherapy in the USA. Future Oncol Lond Engl. 2017 Aug;13(19):1699–710.

8. Datta SS, Ghosal N, Daruvala R, Chakraborty S, Shrimali RK, van Zanten C, et al. How do clinicians rate patient's performance status using the ECOG performance scale? A mixed-methods exploration of variability in decision-making in oncology. ecancermedicalscience [Internet]. 2019 Mar 28 [cited 2020 Mar 2];13. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6467460/

9. Clinicaltrials.gov. Avelumab in Subjects With Merkel Cell Carcinoma (JAVELIN Merkel 200) [Internet]. 2019 [cited 2020 Feb 21]. Available from: https://clinicaltrials.gov/ct2/show/NCT02155647

10. Klink AJ, Phatak H, Bharmal M, Kaufman J, Feinberg B. Merkel Cell Cancer: Poor Response To Chemotherapy Exposes Significant Unmet Need. Value Health. 2017 Oct 1;20(9):A415.

11. Chang JW-C, Chang Y-Y, Huang Y-L, Lo Y-F, Ho T-Y, Huang Y-T, et al. Merkel cell carcinoma in Taiwan: A series of 24 cases and literature review. Medicine (Baltimore). 2019 Oct;98(42):e17538.

12. Zheng Y, Kim R, Yu T, Dreyfus J, Gayle JA, Wassel CL, et al. Real-World Study of Metastatic Merkel Cell Carcinoma Patients Receiving Checkpoint Inhibitors (CPIs) vs. Chemotherapy Treatments. Value Health. 2019 May 1;22:S117.

13. Latimer NR. NICE DSU Technical Support Document 14: Survival Analysis for Economic Evaluations Alongside Clinical Trials - Extrapolation with Patient-Level Data [Internet]. 2011 [cited 2019 Jan 18]. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

14. Novakovic AM, Wilkins JJ, Dai H, Wade JR, Neuteboom B, Brar S, et al. Changing Body Weight–Based Dosing to a Flat Dose for Avelumab in Metastatic Merkel Cell and Advanced Urothelial Carcinoma. Clin Pharmacol Ther. 2020;107(3):588–96.

15. van Hout B, Janssen MF, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health J Int Soc Pharmacoeconomics Outcomes Res. 2012 Aug;15(5):708–15.

16. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. Health Qual Life Outcomes. 2007 Dec 21;5:70.

17. National Institute for Health and Care Excellence (NICE). TA517: Company response to Appraisal Consultation Document [Internet]. 2018 [cited 2020 Feb 21]. Available from: https://www.nice.org.uk/guidance/ta517/documents/committee-papers-2

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Cancer Drugs Fund Review of TA517

Avelumab for treating metastatic Merkel cell carcinoma

RESULTS ADDENDUM (with PAS) to Company evidence submission for committee

August 2020

File name	Version	Contains confidential information	Date
ID1617 RESULTS ADDENDUM to Company submission	v1-0	Yes	04 September 2020

A confidential simple discount of [redacted] off the list price of avelumab has recently been approved by NHS England in an avelumab indication (RCC, TA645). In this addendum, Merck apply this cross-indication patient access scheme to the cost-effectiveness analysis for avelumab in metastatic Merkel Cell Carcinoma, which is currently under CDF review by NICE.

A.1 Cost-effectiveness results (deterministic), including PAS

A summary of the deterministic cost-effectiveness analysis (CEA) results is provided in Table 1. This table contains results in three categories, as per the ToE document:

- **CEA 1a and b:** Replication of the key results at entry to the CDF (1a = upper bound, 1b = lower bound), now including avelumab PAS
- **CEA 2a and b:** Results using mature JM200 data, other parameters unchanged
- **CEA 3:** Revised company base-case results, with PAS

CEA 2a and 2b fix all model settings and assumptions except for the choice of survival models for OS, PFS, and ToT (which use mature JM200: Part B data). The revised base case (CEA 3) includes updated survival models, the fixed dose of avelumab, and utility values based on updated JM200 data. Further details are provided in **Error! Reference source not found.**.

Table 1: Comparison of cost-effectiveness results (deterministic), including avelumab PAS

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
CEA 1a*							
Chemotherapy	10,608	2.02	1.37				
Avelumab	[redacted]	4.16	2.65	[redacted]	2.14	1.28	[redacted]
CEA 1b*							
Chemotherapy	11,116	1.94	1.34				
Avelumab	[redacted]	4.58	2.86	[redacted]	2.64	1.52	[redacted]
CEA 2a**							
Chemotherapy	10,611	1.94	1.34				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,258
CEA 2b**							
Chemotherapy	11,116	1.94	1.34				

Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	19,686
CEA 3* **							
Chemotherapy	11,116	1.94	1.32				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947

Key: BSC, best supportive care; CDF, Cancer Drugs Fund; CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Notes: *CEA 1a and 1b are based on a replication of analysis that demonstrated plausible potential for cost-effectiveness at CDF entry (a is based on the upper bound, and b is based on the lower bound). **CEA 2a and 2b are analyses based on 1a and 1b, respectively; but incorporating updated clinical evidence (hence the same estimated QALYs and LYs). CEA 2a includes the 'fix' for background mortality which affects the estimated total costs, QALYs, and LYs for the chemotherapy arm. ***CEA 3 represents the new company base-case analysis. Only one set of results is presented, based on the totality of the evidence base now available, and incorporation of additional information (as discussed in Section **Error! Reference source not found.**).

A.2 Probabilistic sensitivity analysis

Probabilistic sensitivity analysis (PSA) was conducted (based on CEA 3) to explore the impact of parameter uncertainty within the cost-effectiveness model when all parameters were varied simultaneously. Model parameters were sampled within their respective distribution and bounds of uncertainty for 1,000 iterations. The results of each iteration were recorded and the average of the results are presented in Table 2. A summary of the model parameters varied within the PSA, including bounds of uncertainty and distributions used, is provided in **Error! Reference source not found.**.

Table 2: Updated base-case results (probabilistic), including avelumab PAS

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Chemotherapy	[redacted]	1.95	1.33				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,939

Key: BSC, best supportive care; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

Figure 1 presents the PSA scatterplot for avelumab versus chemotherapy (for CEA 3), where the mean results correspond to Table 2. At a willingness-to-pay threshold for an end-of-life treatment of £50,000 per QALY gained, avelumab is associated with a 99.6% probability of being cost effective.

Figure 1: Scatterplot of probabilistic results, with PAS

Key: QALY, quality-adjusted life year; WTP, willingness-to-pay.

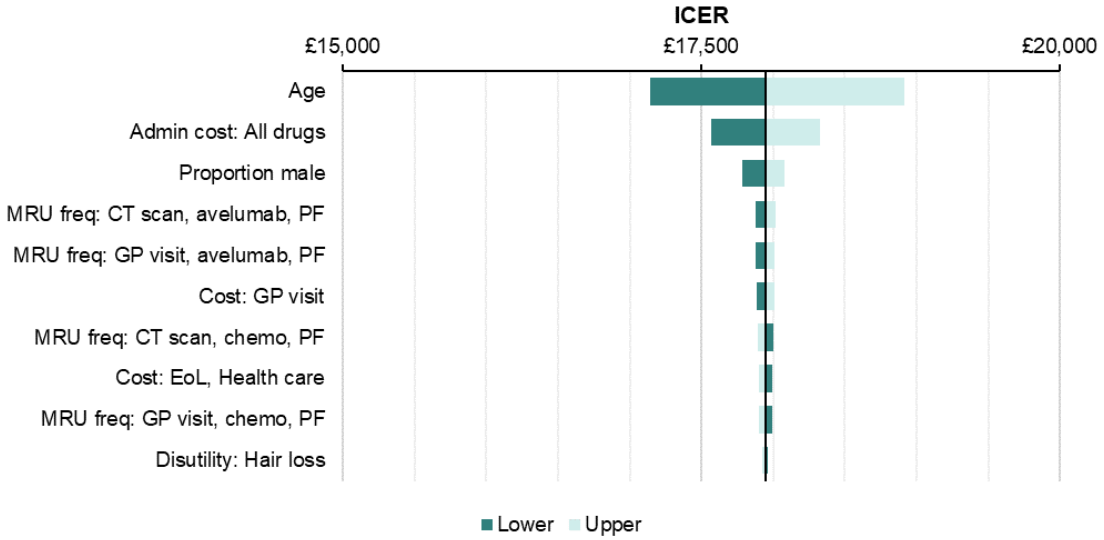
A.3 Key sensitivity and scenario analyses, including avelumab PAS

The results of a deterministic one-way sensitivity analysis (OWSA) are presented as a tornado diagram in Figure 2 (centred on CEA 3). In this analysis, individual model parameters were varied at their respective lower and upper bounds, with the impact on model results recorded (tabulated model parameters provided in **Error! Reference source not found.**).

As per the original TA517 CS, one of the largest drivers of cost-effectiveness results in the OWSA was medical costs. In addition, both age and sex were shown to lead to variations in the ICER (as each of these parameters influence background mortality within the model). However, no single parameter (when varied in isolation of all other model parameters) caused the ICER to exceed £50,000.

Notably, survival and utility-related values are excluded from this analysis owing to the specification of variance-covariance matrices in the updated utility regression and parametric survival models. This means that each modelled utility or curve parameter cannot be robustly varied in isolation of the other values. The uncertainties concerning utility values and survival models are instead captured within the PSA and explored in scenario analysis.

Figure 2: Tornado diagram, including avelumab PAS



Key: CT, computed tomography; freq, frequency; GP, General Practitioner; EoL, end of life; ICER, incremental cost-effectiveness ratio; MRU, medical resource use; PF, progression-free.

A series of deterministic scenario analyses were undertaken to explore alternative model settings and assumptions. The scenarios expected to be of greatest relevance for decision making (based on the contents of the ToE document) are presented in Table 3. The scenarios explored cause the ICER to increase by up to £1,329 or decrease by as much as £2,671; with no scenario yielding an ICER greater than £50,000.

Table 3: Key scenario analyses

Scenario and cross reference	Scenario detail	Brief rationale	Impact on base-case ICER (£)
Base case			**17,947**
OS for avelumab (Section **Error! Reference source not found.**)	Use generalised gamma model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (generalised gamma) which projects higher OS versus the base-case analysis	17,363 (-584)
	Use 1-knot normal spline-based model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (1-knot normal spline) which projects lower OS versus the base-case analysis	19,276 (+1,329)
Clinical expectation of ToT (Section **Error! Reference source not found.**)	Assume all patients discontinued by 10 years, with no interim capping at 2 years	Approach intended to serve as an upper bound of potential long-term treatment with avelumab. In practice, discontinuation with avelumab is expected to occur before 5 years	18,895 (+948)
	Assume one-third of patients continue treatment after 2 years, and all	Approach aligned with clinical expert opinion at the time of the original TA517 CS, and is still	15,278 (-2,671)

	discontinue by 5 years	expected to be broadly representative of clinical practice	
Utility values (Section **Error! Reference source not found.**)	Use original utility values from TA517	Allows for assessment of impact on cost-effectiveness results through updating utility values	18,655 (+708)
	Use only data from JM200: Part B to inform utility values	Allows exploration of using utility values derived only from a treatment-naïve metastatic MCC population	18,395 (+448)
Key: BSC, best supportive care; CS, company submission; ICER, incremental cost-effectiveness ratio; JM200, JAVELIN Merkel 200; OS, overall survival.			

A.4 Key issues and conclusions based on CDF data

This submission presents a summary of the additional evidence collected concerning the use of avelumab for people with previously-untreated metastatic MCC during the period of CDF data collection. Two key evidence sources inform this submission – updated data from Part B of the JM200 clinical trial, as well as the SACT dataset. The updated data from JM200: Part B facilitated an update to the economic model, with reduced uncertainty concerning the clinical effectiveness of avelumab used in the first line setting.

The updated base-case analysis demonstrates that avelumab provides a cost-effective end-of-life treatment, with an ICER of £17,947 per QALY gained. The economic analysis is based on the previously-submitted modelling approach, with updates to clinical data and assumptions based on data from JM200: Part B. Sensitivity analysis results provided similar findings to the base-case analysis, with the key drivers of cost effectiveness being assumptions relating to OS, ToT, and HRQoL.

The lack of direct comparative evidence for avelumab and chemotherapy remains a key uncertainty within the context of the clinical and cost effectiveness of avelumab. As has been previously discussed, this – in part - is a consequence of metastatic MCC being an ultra-rare disease, and the difficulties of conducting comparative trials in these cases. In lieu of a direct comparison, the model utilises observational study data to facilitate a comparison with the current standard of care. This comparison demonstrates a clear survival advantage for patients treated with avelumab.

Data collected from the SACT cohort provide additional information relevant to this appraisal, yet are subject to a number of important limitations. However, it is unclear how generalisable the SACT cohort is to an avelumab-eligible metastatic MCC population in NHS practice owing to the availability of avelumab in multiple treatment lines.

In conclusion, this submission supports the expectation that avelumab is a cost-effective treatment option for people with previously-untreated, metastatic MCC. With the incorporation of avelumab's patient access scheme, the base case, sensitivity analyses and scenario analyses all result in ICERs well below the willingness-to-pay threshold. Avelumab represents a step change in therapy, allowing patients to forgo toxic chemotherapy and derive the full benefits provided by avelumab when used in the first-line setting.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Avelumab for treating metastatic Merkel cell carcinoma

(CDF review TA517) – [ID1617]

Clarification question responses

March 2020

File name	Version	Contains confidential information	Date
2020-03-27_ID1617 Avelumab MCC ClarQ responses_FINAL_ UPDATED	v0-2	Yes	27-March-2020

Tables

Figures

Section A: Clarification on effectiveness data

Propensity score matching

A1. Priority question: Please provide a propensity score matching analysis of JAVELIN Merkel 200: part B (n=116) and Study 100070-Obs001 (n=67) for the outcomes of PFS and OS in the first line metastatic MCC population following the advice presented in the NICE DSU technical Support Document 17. Best practice guidance is to adjust for as many variables as possible, regardless of effect modifier status or level of imbalance. As a minimum, please ensure that the adjustments made in the analysis include the following (please note this list is not exhaustive):

- **Immunocompetency;**
- **Tumour PD-L1 expression status;**
- **Age;**
- **Merkel cell polyomavirus (MCPyV) status;**
- **Tumour burden;**
- **Gender;**
- **ECOG status at baseline.**

Merck Serono appreciates the flexibility provided by NICE and the ERG in allowing additional time to respond to this priority question.

In order to address any potential bias in the estimation of treatment effects that can arise with imbalances in baseline patient characteristics, propensity scores were calculated for each patient in the pooled analysis set. A propensity score is the probability of receiving treatment given an observed set of covariates, which can be used to balance covariate values between treated (avelumab) and control (chemotherapy) patients to obtain an unbiased estimate of treatment effect. The pooled analysis set comprised JAVELIN Merkel 200: Part B (n=116 patients whom received avelumab) and Study 100070-Obs001 (n=67 patients whom received chemotherapy).

Propensity scores were used to produce a balanced analysis via the implementation of two different methods: a) propensity score matching (PSM), and b) propensity score weighting (PSW). Sensitivity analyses were performed using alternative specifications of the propensity score (i.e., using a different combination of baseline patient characteristics), different patient population subgroups, and alternative methods for comparison (i.e., inverse probability of treatment weighting [IPTW], also known as general weights, vs. stabilised weights [SW]).

The propensity score in this analysis represents the probability of being in the treatment group (i.e., avelumab) given an observed set of baseline patient characteristics. This was calculated using a logistic regression model predicting treatment assignment according to the following baseline characteristics:

1. Age
    - Coded as a binary 1/0 variable according to a threshold age of 75 years (i.e., 1 = 75+ years / 0 = <75 years).
2. ECOG PS
    - Coded as a binary 1/0 variable according to a threshold ECOG score of one (i.e., 1 = ECOG score of 1+ / 0 = ECOG score of 0).
3. Sex
    - Coded as a binary 1/0 variable (i.e., 1 = woman / 0 = man).
4. Immune status
    - Coded as a binary 1/0 variable (i.e., 1 = immunosuppressed / 0 = immunocompetent)

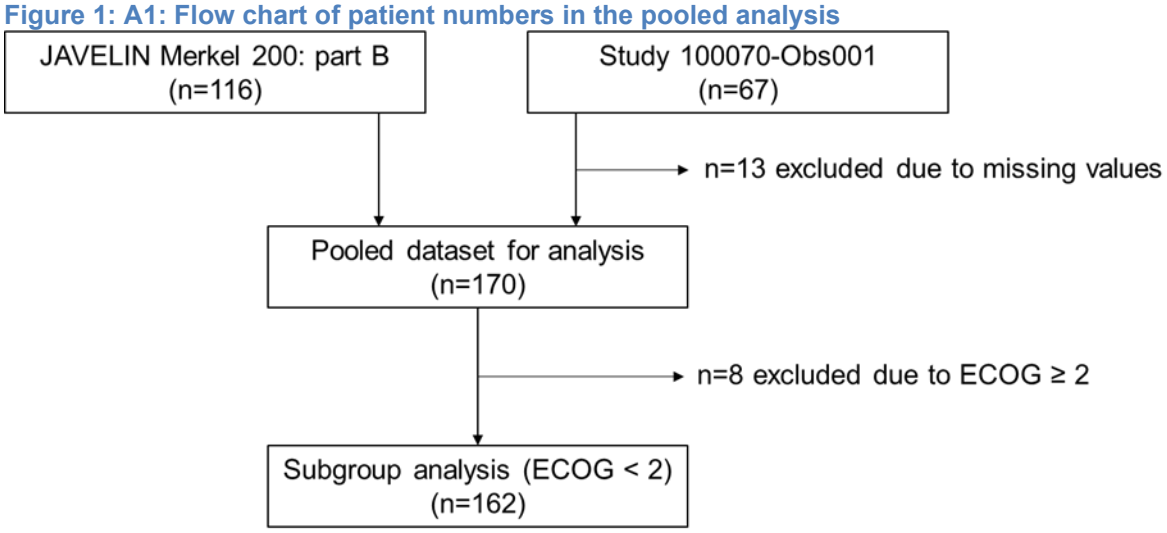
The other three variables requested (tumour PD-L1 expression status, MCPyV status, and tumour burden) are not available in the Study 100070-Obs001 dataset and were therefore not possible to consider. After considering patients with reported age, sex, ECOG PS, and immune status, n=13 patients from Study 100070-Obs001 were removed due to missing ECOG PS, leaving n=54 patients for analysis.

Merck Serono acknowledges that the list provided by the ERG is not exhaustive, and that other variables could have also been considered within the matching analysis. However, given the limited sample size to match on (especially after removing the n=13 chemotherapy patients with no reported ECOG PS at baseline), the additional

complexity introduced through specifying a more complex logistic regression model was not considered justified as for propensity scoring variables should either be prognostic, or at least correlated with unobserved important variables – here we do not believe that to be the case.

In order to perform sensitivity analysis, propensity scores were also calculated with the omission of immune status as a covariate. This was chosen due to the low prevalence of immunocompromised patients (n=13, all of whom were in Study 100070-Obs001) compared with immunocompetent patients (n=157) within the pooled dataset. Further sensitivity analysis was performed using the subpopulation of patients with baseline ECOG PS <2. This removed eight patients from the analysis who displayed baseline ECOG scores of 2 (n=6) and 3 (n=2). The total number of patients included at different stages of the analysis is shown in Figure 1.

Figure 1: A1: Flow chart of patient numbers in the pooled analysis



PSM was performed using the package 'MatchIt' (Ho et al., 2015), within R version 3.6.3 (R Core Team, 2020).* The procedure of 1:1 matching was employed, which matches each patient in the intervention group with the control patient exhibiting the nearest propensity score (this is also known colloquially as 'greedy' matching). Consequently, if all patients from the smallest group (i.e., intervention or control) are matched then the sample size for subsequent analysis becomes double the sample size of the smallest group. If an appropriate match is not available, for example due

* R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Daniel E. Ho, Kosuke Imai, Gary King, Elizabeth A. Stuart (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. Journal of Statistical Software, Vol. 42, No. 8, pp. 1-28. URL http://www.jstatsoft.org/v42/i08/.

to a lack of overlap in propensity score values between groups, then cases are discarded and the matched sample size for analysis is reduced accordingly.

PSW was implemented within R version 3.6.3 (R Core Team, 2020). Weighting was initially performed using the IPTW method, in order to create a baseline reference of propensity score weights for comparison. Using this approach, each patient is assigned a weight representing the inverse probability of being assigned to their respective group. Thus, IPTW aims at giving more importance (i.e., more "weight") to those patients that have unexpected propensity score values. Although this approach uses the totality of the data, patients with unexpected propensity score values are counted more than once in the pseudo-population, resulting in an inflated sample size. For the treatment group (i.e., avelumab), the weight, W, assigned in the IPTW method for each individual, i, based on propensity score, π, is:

$$W_i = \frac{1}{\pi_i}$$

For the control group (i.e., chemotherapy), patients receive weights of:

$$W_i = \frac{1}{(1-\pi_i)}$$

SW address some of the limitations of IPTW by reducing the weights of either those treated patients with low propensity scores or those control patients with high propensity scores. This preserves the sample sizes in the pseudo dataset and allows for appropriate estimates of the variance to be given without requiring the use of more sophisticated methods (Xu et al., 2010). SW use the term p to denote the proportion of treated (avelumab) patients. SW is therefore given as:

For the treatment group (i.e., avelumab):

$$SW_i = \frac{p}{\pi_i}$$

For the control group (i.e., chemotherapy):

$$SW_i = \frac{(1-p)}{(1-\pi_i)}$$

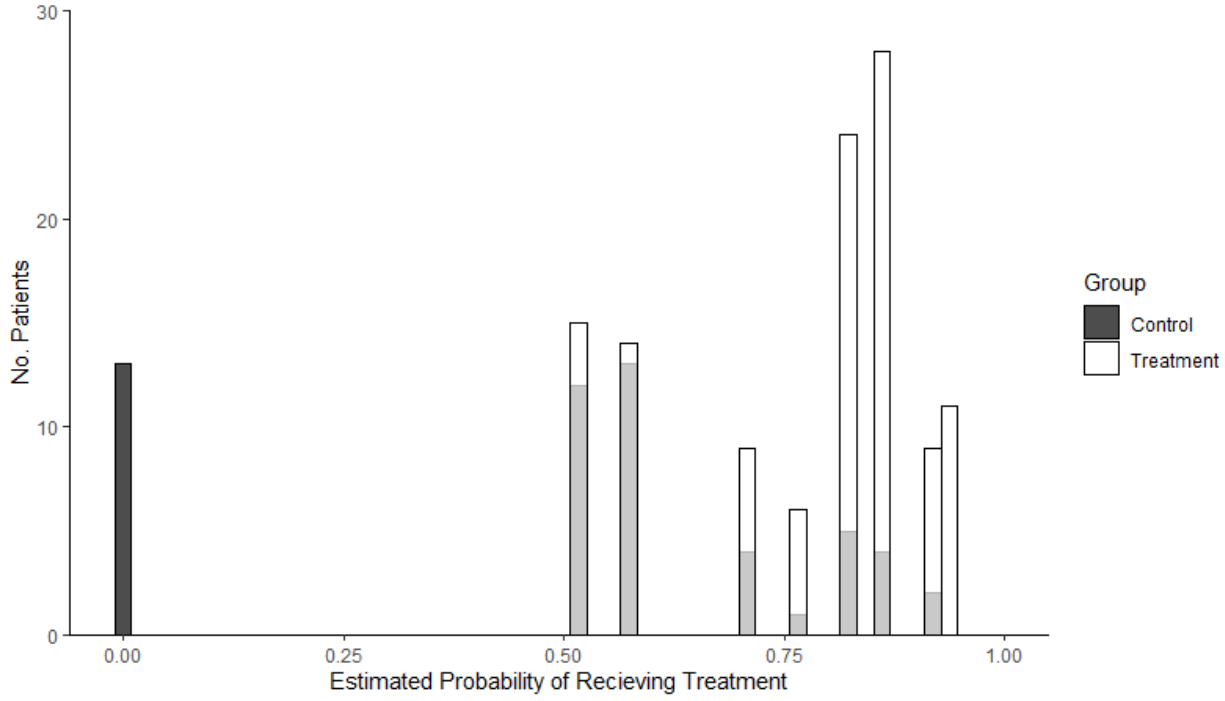
Three sets of propensity scores were therefore estimated:

1. Estimated based on age, ECOG status, sex and immune status for complete cases (i.e. patients with reported values for each variable)
2. Estimated based on age, ECOG status and sex for complete cases (i.e. patients with reported values for each variable)
3. Estimated based on age, ECOG status and sex for patients with an ECOG PS of 0 or 1

Analysis 1: Estimated based on age, ECOG status, sex and immune status for complete cases (i.e. patients with reported values for each variable)

A histogram demonstrating the distribution of propensity score values across the treatment and control groups is presented in Figure 2.

Figure 2: A1: Histogram of propensity score values – age, ECOG, sex and immune status.



NB: Grey sections represent overlap between groups. Control = chemotherapy; treatment = avelumab.

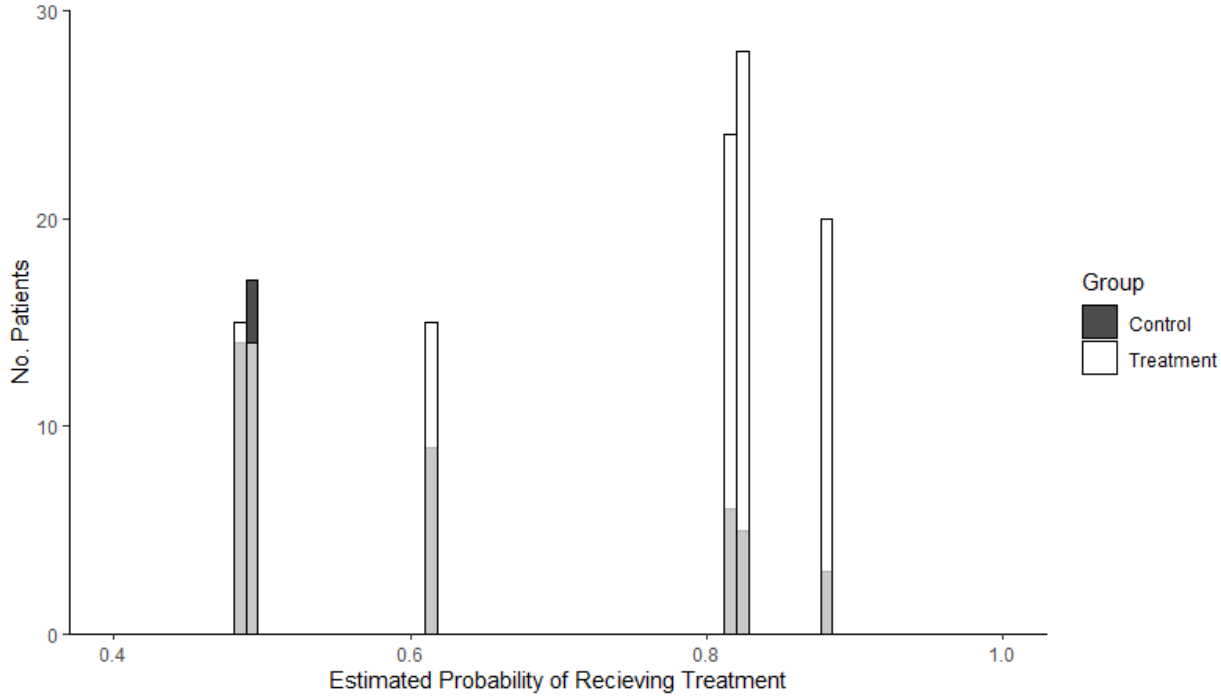
There is some overlap with the distribution of propensity score values between groups. Higher values are more prevalent within the avelumab group, though differences in the spread of propensity scores may be exacerbated due to the difference in sample sizes. A total of n=13 immunocompromised patients from the control group received a propensity score of zero due to the absence of any

immunocompromised patients within the avelumab group; therefore, these patients could not be matched. Consequently, the mean propensity score was lower for the chemotherapy group (0.49) than the avelumab group (0.77).

Analysis 2: Estimated based on age, ECOG status and sex for complete cases (i.e. patients with reported values for each variable)

A histogram demonstrating the distribution of propensity score values across the avelumab and chemotherapy groups after the removal of immune status as a covariate is presented in Figure 3.

Figure 3: A1: Histogram of propensity score values – age, ECOG and sex.



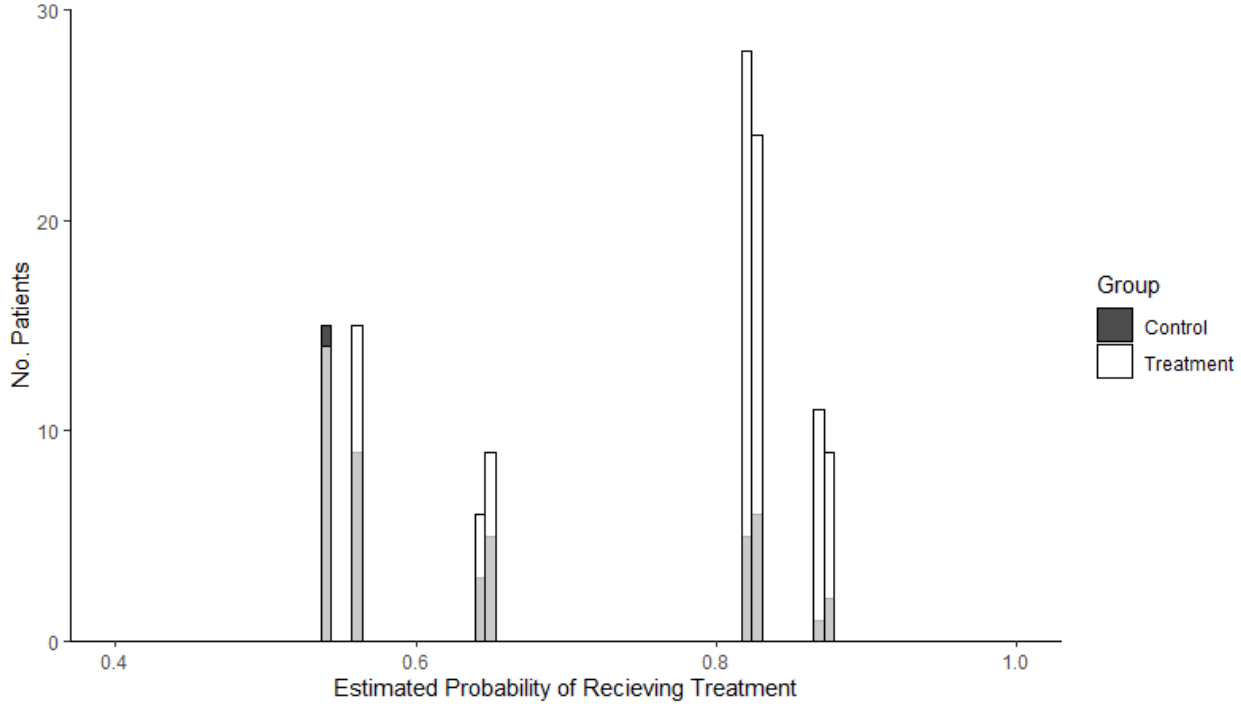
NB: Grey sections represent overlap between groups. Control = chemotherapy; treatment = avelumab.

As with Analysis 1, there is some overlap with the distribution of propensity score values between groups, which is improved though the removal of immunocompromised patients. However, values are still higher within the avelumab group; 0.72 vs 0.60.

Analysis 3: Estimated based on age, ECOG status and sex for patients with an ECOG PS of 0 or 1

A histogram demonstrating the distribution of propensity score values across the avelumab and chemotherapy groups is presented in Figure 4.

Figure 4: A1: Histogram of propensity score values – age, ECOG and sex (ECOG >1 excluded).



NB: Grey sections represent overlap between groups. Control = chemotherapy; treatment = avelumab.

The overlap in the distribution of propensity score values between groups is further improved when removing patients with an ECOG PS of 2 or more; yet higher values are still noted within the avelumab group (0.74 vs 0.65).

Following production of the three sets of propensity scores, PSM and PSW analyses were conducted. In total, seven analyses were performed, as described below. Four of these (shown in darker text) were taken forward to inform scenarios in the economic model (please see response to clarification question B1 for further details). The model also includes the ability to use the other three analyses (shown in lighter text), but these results are not reported within this response.

- PSM
  - PSM 1. Age (aged ≥75 vs. <75 years), sex (female vs. male), ECOG PS (0 vs. 1+), and immune status
  - PSM 2. Age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1+), excluding immunosuppression as a variable

PSM 3. Age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1+), excluding immunosuppression as a variable and excluding patients with an ECOG PS of 2 or more

- PSW

PSW 1. IPTW, based on all patients with available data for age (aged ≥75 vs. <75 years), sex (female vs. male), ECOG PS (0 vs. 1+), and immune status

PSW 2. SW, based on all patients with available data for age (aged ≥75 vs. <75 years), sex (female vs. male), ECOG PS (0 vs. 1+), and immune status

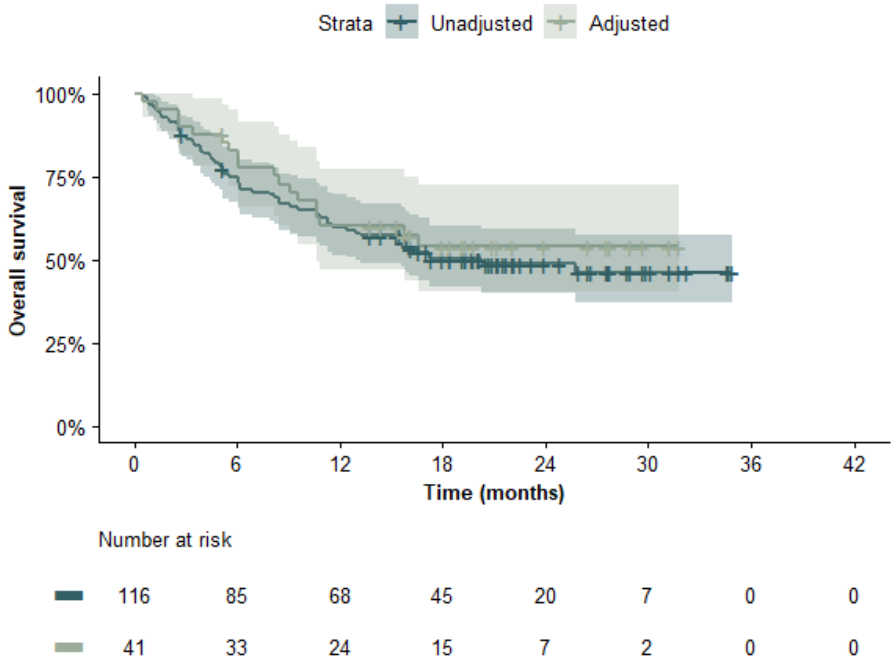
PSW 3. SW, based on all patients with available data for age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1+), excluding immunosuppression as a variable

PSW 4. SW, based on all patients with available data for age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1), excluding immunosuppression as a variable and excluding patients with an ECOG PS of 2 or more

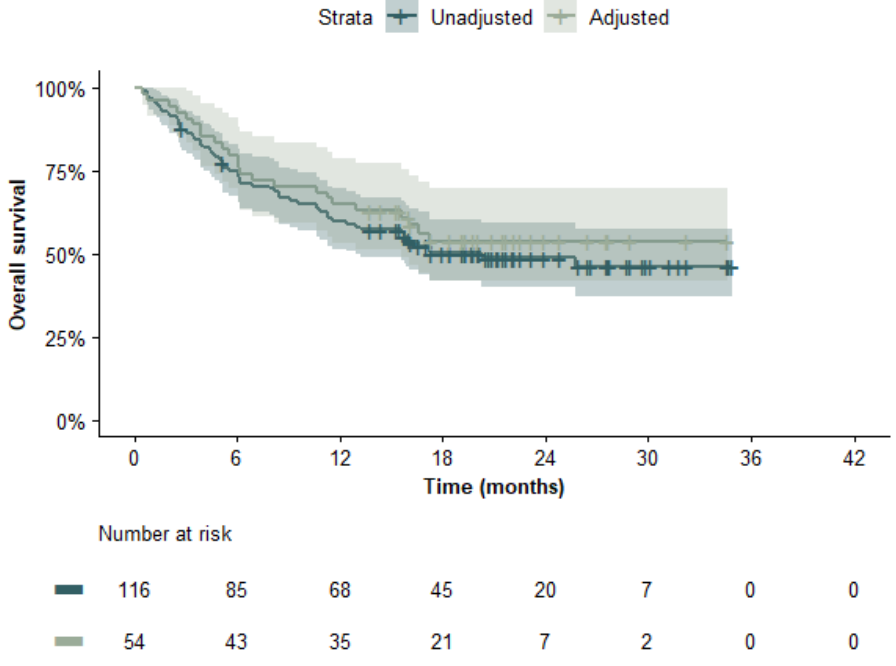
For brevity, the impact of using the different approaches to weighting/matching patients is provided below in the form of Kaplan-Meier plots for the outcome of OS only. Plots are presented comparing the unadjusted full cohort of patients for each of the seven analyses for avelumab patients (Figure 5) and chemotherapy patients (Figure 6).

Figure 5: A1: Adjusted OS plots – JM200: Part B

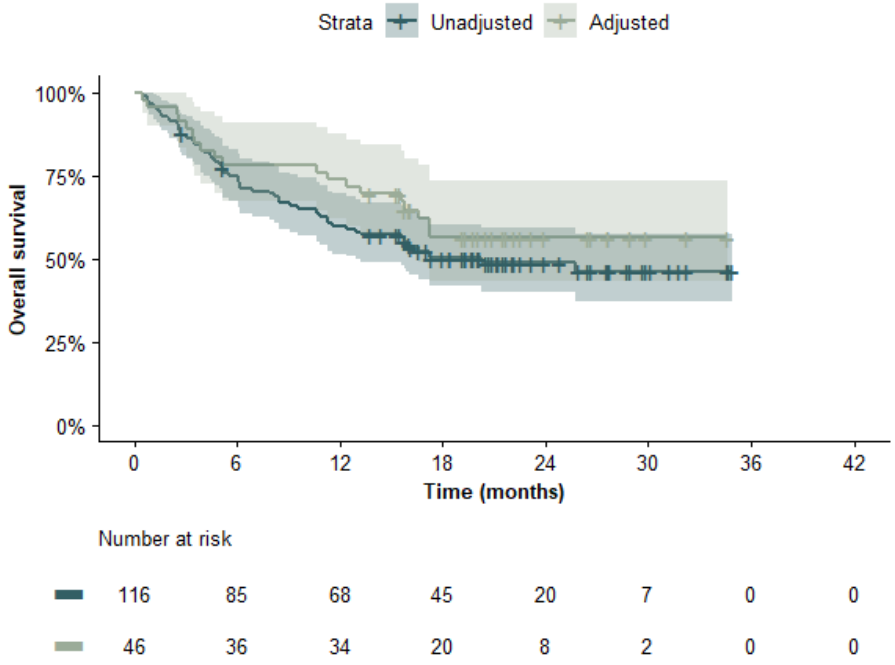
PSM 1



PSM 2



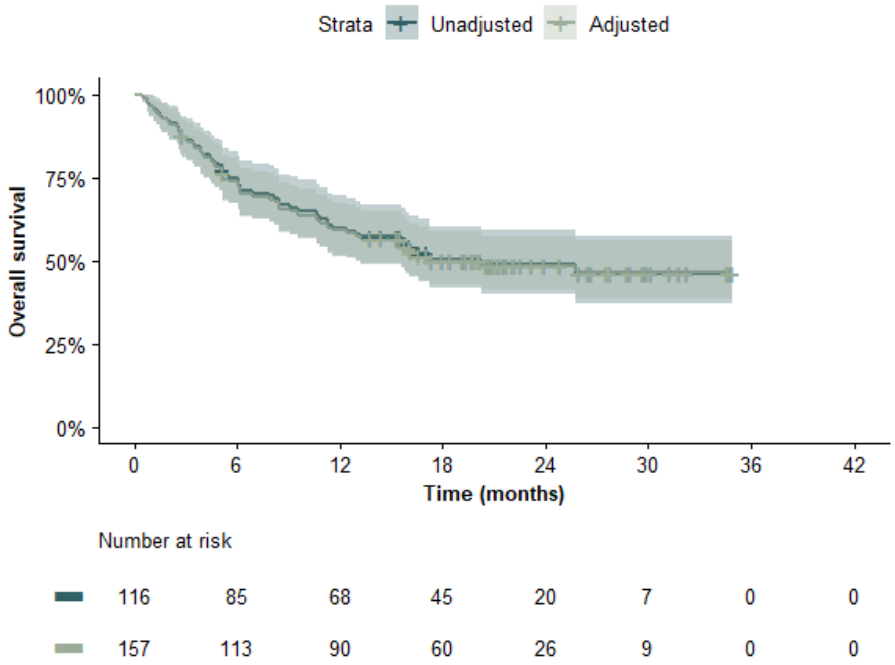
PSM 3



Number at risk

Unadjusted	116	85	68	45	20	7	0	0
Adjusted	46	36	34	20	8	2	0	0

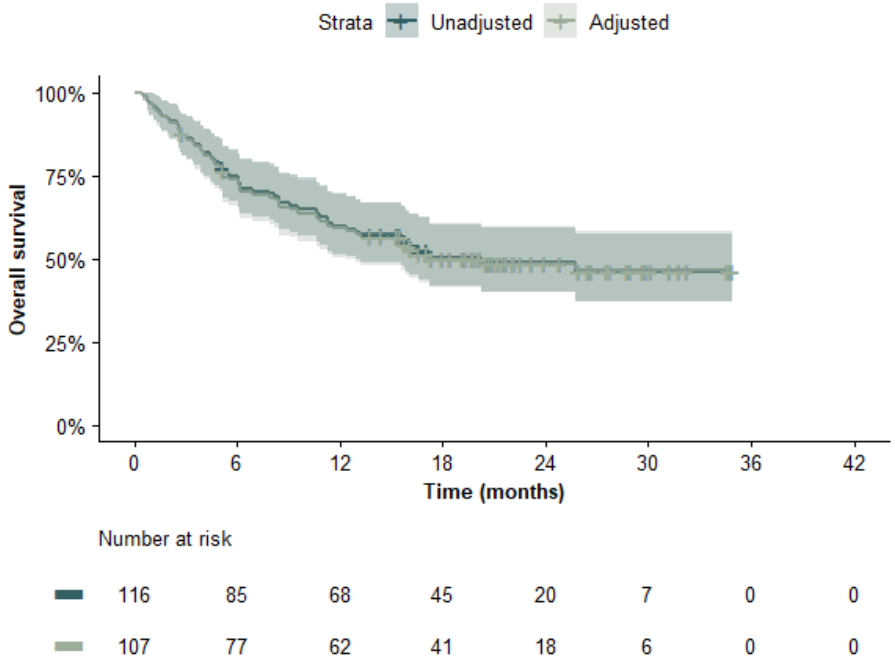
PSW 1



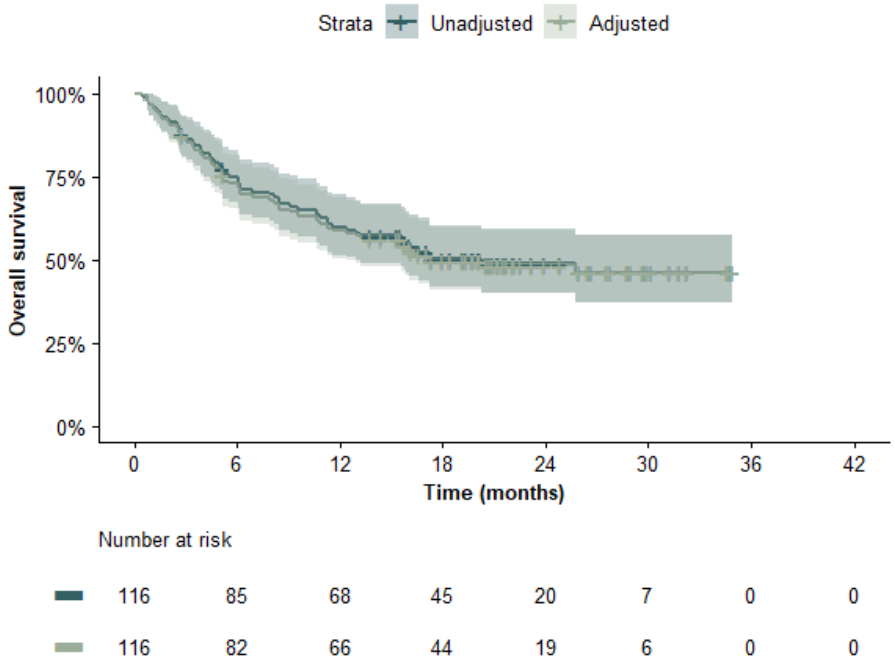
Number at risk

Unadjusted	116	85	68	45	20	7	0	0
Adjusted	157	113	90	60	26	9	0	0

PSW 2


Strata
Unadjusted
Adjusted
100%
75%
50%
25%
0%
Overall survival
0
6
12
18
24
30
36
42
Time (months)
Number at risk
116 85 68 45 20 7 0 0
107 77 62 41 18 6 0 0

PSW 3


Strata
Unadjusted
Adjusted
100%
75%
50%
25%
0%
Overall survival
0
6
12
18
24
30
36
42
Time (months)
Number at risk
116 85 68 45 20 7 0 0
116 82 66 44 19 6 0 0

PSW 4

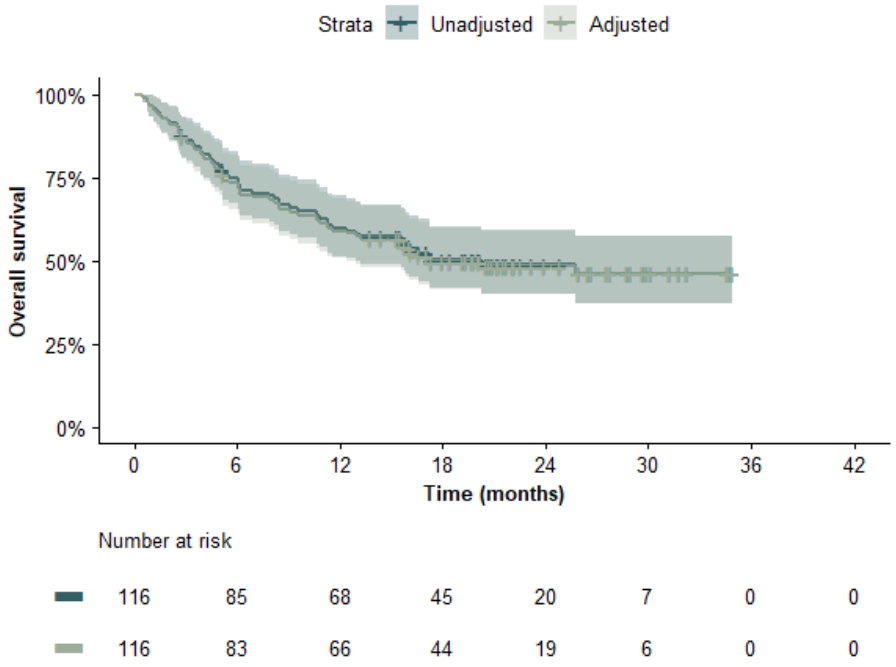
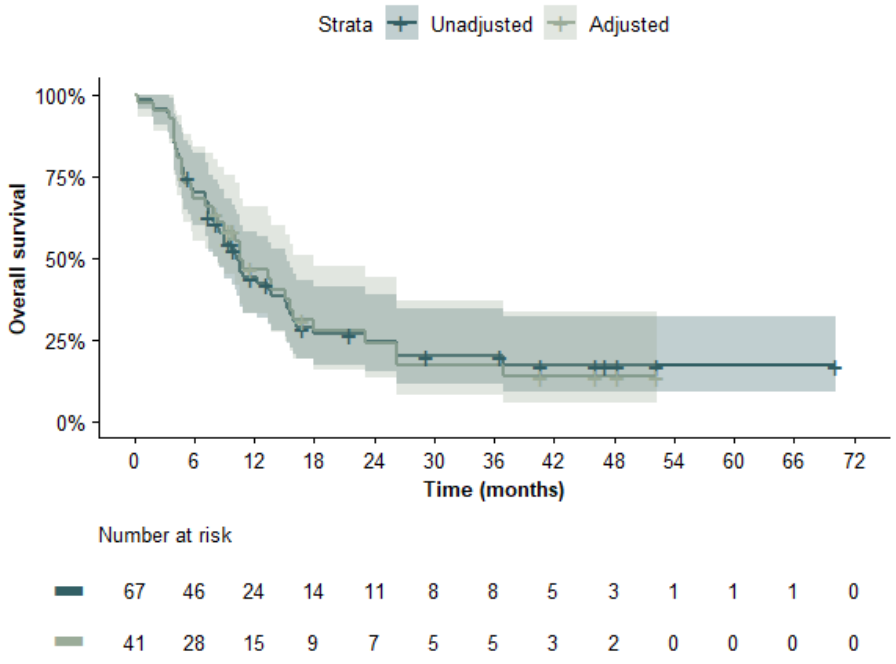
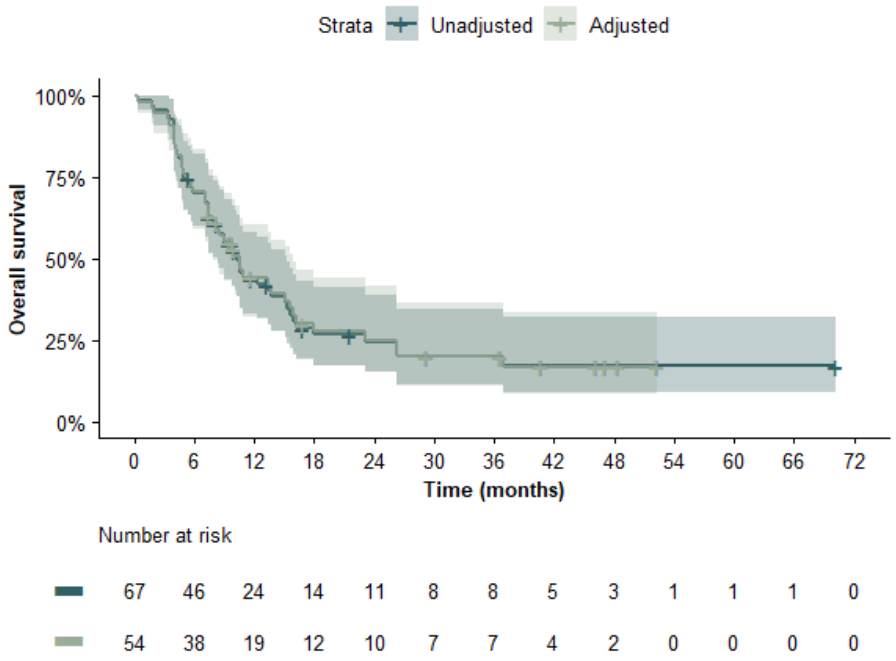


Figure 6: A1: Adjusted OS plots – Study 100070-Obs001

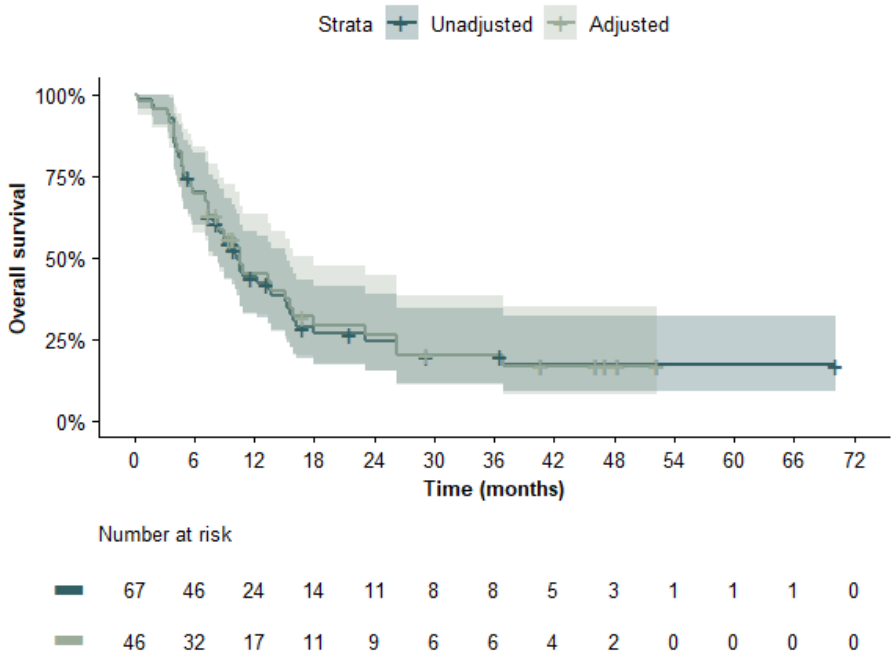
PSM 1



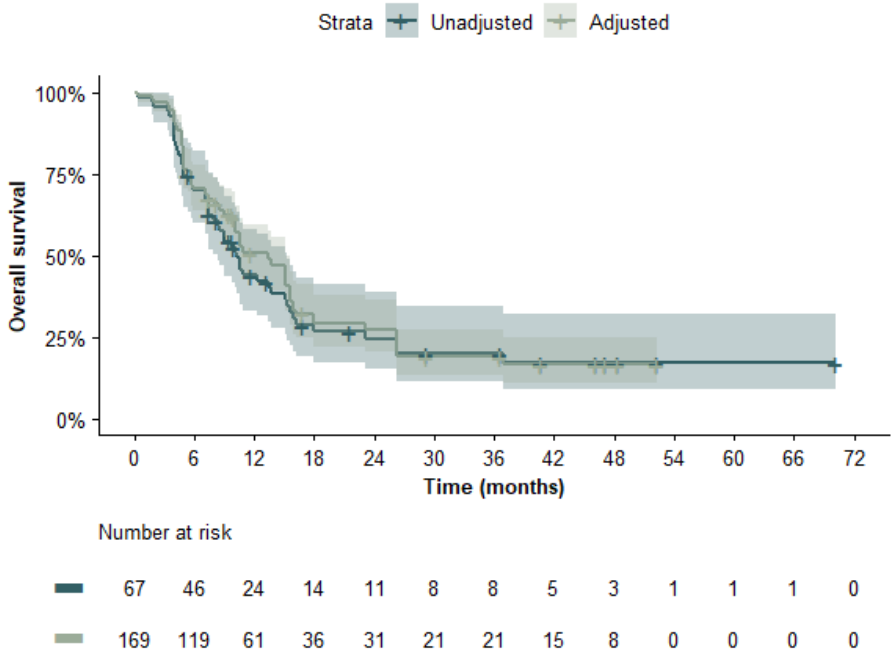
PSM 2


Strata
Unadjusted
Adjusted
100%
75%
50%
25%
0%
Overall survival
0
6
12
18
24
30
36
42
48
54
60
66
72
Time (months)
Number at risk
67 46 24 14 11 8 8 5 3 1 1 1 0
54 38 19 12 10 7 7 4 2 0 0 0 0

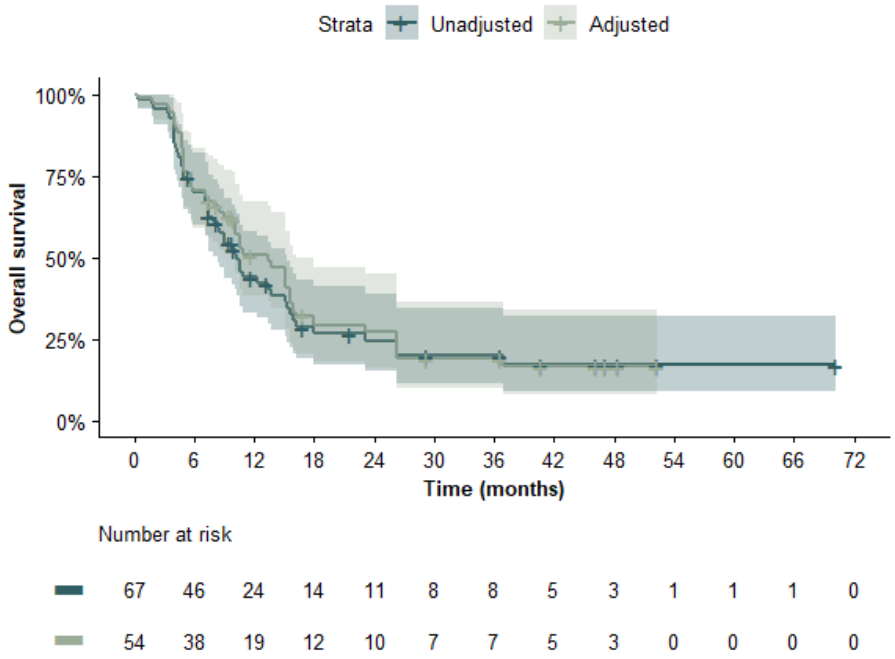
PSM 3


Strata
Unadjusted
Adjusted
100%
75%
50%
25%
0%
Overall survival
0
6
12
18
24
30
36
42
48
54
60
66
72
Time (months)
Number at risk
67 46 24 14 11 8 8 5 3 1 1 1 0
46 32 17 11 9 6 6 4 2 0 0 0 0

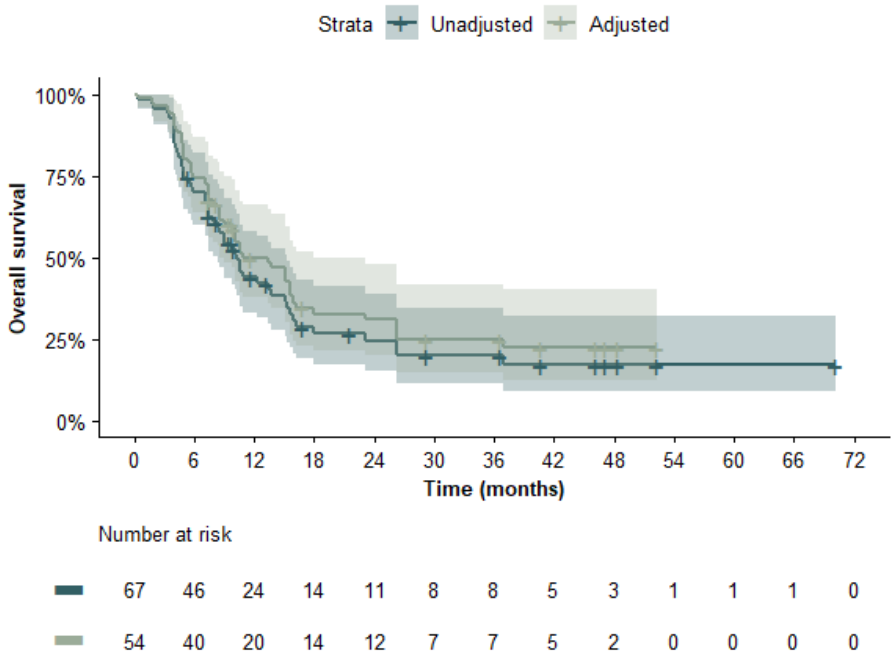
PSW 1


Strata
Unadjusted
Adjusted
100%
75%
50%
25%
0%
Overall survival
0
6
12
18
24
30
36
42
48
54
60
66
72
Time (months)
Number at risk
67 46 24 14 11 8 8 5 3 1 1 1 0
169 119 61 36 31 21 21 15 8 0 0 0 0

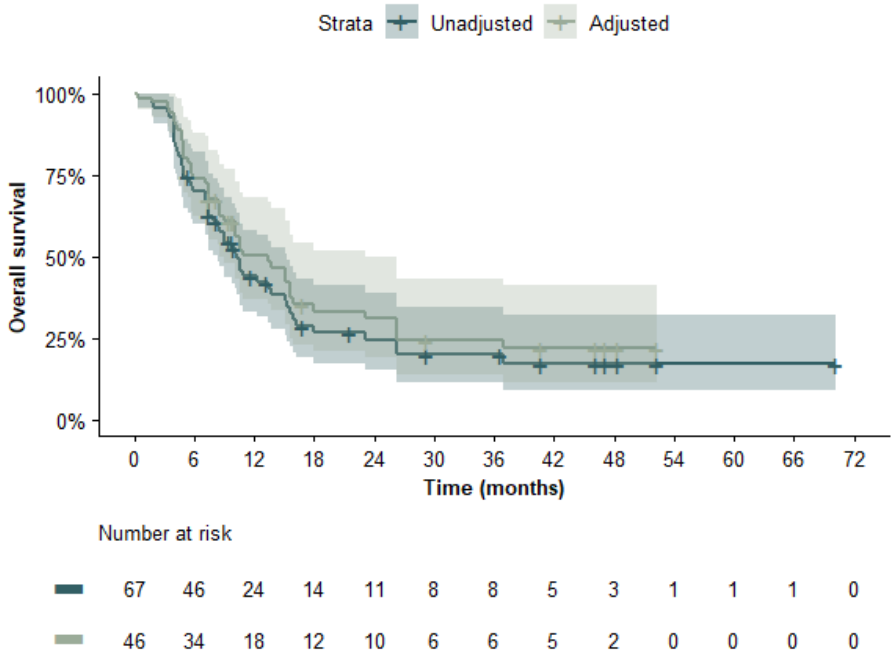
PSW 2


Strata
Unadjusted
Adjusted
100%
75%
50%
25%
0%
Overall survival
0
6
12
18
24
30
36
42
48
54
60
66
72
Time (months)
Number at risk
67 46 24 14 11 8 8 5 3 1 1 1 0
54 38 19 12 10 7 7 5 3 0 0 0 0

PSW 3



PSW 4



The largest impact on the avelumab curve is seen when using a matching analysis (given that this approach leads to the removal of more than half of the cohort), whereas the largest impact on the chemotherapy curve is seen when using a weighting analysis.

Baseline characteristics

A2. Priority question: Please provide the following additional baseline characteristics for Javelin Merkel 200: part B to supplement the data provided in table 5 of the company submission:

- **Region (e.g. Europe);**
- **Number of patients from UK sites;**
- **Site of primary tumour;**
- **Tumour size (cm [median and range]);**
- **Time from initial diagnosis to study entry (months [median and range]);**
- **Presence of distal metastases;**
- **Presence of lymph node metastases;**
- **Tumour PD-L1 expression;**
- **Tumour MCPyV status;**

Please see the requested information in Table 1. Data for the presence of distal metastases are not available, however data for visceral metastases are available and reported instead.

Table 1: A2: Additional baseline characteristics from JAVELIN Merkel 200: Part B

Baseline characteristic	Value(s)
Region (e.g. Europe), n (%)	North America: 29 (25.0) Western Europe: 75 (64.7) Australia: 9 (7.8) Asia: 3 (2.6)
Number of patients from UK sites, n (%)	0 (0%)
Site of primary tumour, n (%)	Skin: 104 (89.7) Lymph node: 1 (0.9) Not reported: 11 (9.5)
Tumour size (cm [median and range]), n (%)	Median: 3.2 Range: (0.6, 25.0) n: 57 (49.1%)

Time from initial diagnosis to study entry in months, median (range)	10.6 (0.7, 120.9)
Presence of distal metastases, n (%)	Not available
Presence of visceral metastases, n (%)	Yes: 79 (68.1) No: 35 (30.2) Not evaluable: 2 (1.7)
Presence of lymph node metastases, n (%)	Yes: 25 (21.6) No: 89 (76.7) Not evaluable: 2 (1.7)
Tumour PD-L1 expression, n (%)	Positive: 21 (18.1) Negative: 87 (75.0) Not evaluable: 8 (6.9)
Tumour MCPyV status, n (%)	Positive: 70 (60.3) Negative: 37 (31.9) Not evaluable: 9 (7.8)

Time on treatment

A3. Priority question: Please provide the number of patients censored and the numbers of patients who have ended treatment (events) for treatment duration with avelumab in JM200: Part B using the May 2019 data-cut that corresponds with the numbers at risk in Figure 3 of the company submission to complete the table below.

Please see the requested information in Table 2.

Table 2: A3: ToT numbers (JAVELIN Merkel 200: Part B)

Time intervals (months)	0	6	12	18	24	30	36
ToT number at risk	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ToT censored	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ToT events (ended treatment)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

A4. Priority question: Please complete the table below to provide data for time on treatment (ToT) from Javelin Merkel 200: part B along with associated measures of uncertainty (e.g. 95% confidence intervals) using the latest (May 2019) data-cut.

Please see the requested information in Table 3. 95% confidence interval limits were estimated using the default settings in the statistical software *R*. For the 'Number in

analysis' column, we have populated this assuming the total sample size for the median value, and the number at risk for the values at a specific time point.

Table 3: A4: ToT results (JAVELIN Merkel 200: Part B)

Outcome	Result	95 % CI	Number in analysis
Median ToT, months	[redacted]	[redacted]	[redacted]
6-month ToT rate, %	[redacted]	[redacted]	[redacted]
12-month ToT rate, %	[redacted]	[redacted]	[redacted]
15-month ToT rate, %	[redacted]	[redacted]	[redacted]

Javelin Merkel 200: part B results

A5. Priority question. Please provide the number of patients censored and the numbers of events for PFS and OS in JM200: Part B using the May 2019 data-cut that corresponds with the numbers at risk in Figure 1 of the company submission to complete the table below.

Please see the requested information in Table 4.

Table 4: A5: OS and PFS numbers (JAVELIN Merkel 200: Part B)

Time intervals (months)	0	6	12	18	24	30	36
OS number at risk	116	85	68	45	20	7	0
OS censored	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS number at risk	116	45	31	12	4	0	0
PFS censored	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS events	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

A6. Please provide details of the reasons for censoring for the PFS and OS analyses in Javelin Merkel 200: part B.

Please see the requested information in Table 5.

Table 5: A6: OS and PFS reasons for censoring (JAVELIN Merkel 200: Part B)

Outcome	Reasons for censoring
OS	[redacted]

	[redacted]
PFS	[redacted]

A7. Priority question. Please complete the table below to provide results for PFS and OS from Javelin Merkel 200: part B along with associated measures of uncertainty (e.g. 95% confidence intervals) using the latest (May 2019) data-cut.

Please see the requested information in Table 6. 95% confidence interval limits were estimated using the default settings in the statistical software R. For the 'Number in analysis' column, we have populated this assuming the total sample size for the median value, and the number at risk for the values at a specific time point. Please note for the outcome of median OS, the upper limit of the 95% confidence interval is not estimable.

Table 6: A7: OS and PFS outcomes (JAVELIN Merkel 200: Part B)

Outcome	Result	95% CI	Number in analysis
Median PFS, months (95% CI)	4.11	(1.48, 6.74)	116
6-month PFS rate, % (95% CI)	41.3%	(33.1%, 51.5%)	[redacted]
12-month PFS rate, % (95% CI)	31.0%	(23.5%, 41.0%)	[redacted]
15-month PFS rate, % (95% CI)	27.0%	(19.8%, 36.8%)	[redacted]
Median OS, months (95% CI)	20.3	(13.0, NE)	116
6-month OS rate, % (95% CI)	74.8%	(67.3%, 83.2%)	[redacted]
12-month OS rate, % (95% CI)	59.9%	(51.5%, 69.6%)	[redacted]
15-month OS rate, % (95% CI)	57.2%	(48.8%, 67.1%)	[redacted]

Key: NE, not evaluable

A8. Priority question. Please provide the results from the latest (May 2019) data-cut for Javelin Merkel 200: part B along with associated measures of uncertainty (e.g. 95% confidence intervals) for the following outcomes:

- **response rate (see table below);**

- **adverse effects of treatment;**
- **immune-related adverse events;**
- **health related quality of life.**

Response rate

Please see Table 7 for the requested information.

Table 7: A8: Response rate (JAVELIN Merkel 200: Part B)

Outcomes by RECIST v1.1, per IRC Assessment JM200 Part B	Analysis May 2019
Number of patients, n	116
Median follow-up, mths	21.2 (range: 14.9-36.6)
Median duration of treatment, wks	24 (range: 2-154)
Confirmed ORR, % (95% CI)	39.7 (30.7, 49.2)
Confirmed BOR, n (%)	
Complete response	19 (16.4)
Partial response	27 (23.3)
Stable disease	12 (10.3)
Progressive disease	48 (41.4)
Non-CR/Non-PD	1 (0.9)
Non-evaluable	9 (7.8)*
Response durability	
Patients with durable response, n	35
Durable response rate (95% CI)	30.2% (22.0-39.4%)‡
Response duration	
Median DOR (95% CI), mths†	18.2 (11.3, NE)
Proportion of responses with duration ≥3 mths, % (95% CI)	89 (75-95)
Proportion of responses with duration ≥6 mths, % (95% CI)†	78 (63-87)
Proportion of responses with duration ≥12 mths, % (95% CI)†	66 (50-78)
Proportion of responses with duration ≥15 mths, % (95% CI)†	61 (44-74)

* No postbaseline assessments due to early death (n=4) or other reasons (n=2), no adequate baseline assessment (n=2), or all postbaseline assessments had overall response of NE (n=1)
† Based on Kaplan-Meier estimates
‡ Proportion of patients with a response lasting ≥6 months

Adverse effects of treatment and immune-related adverse events

Information concerning adverse effects (AEs) of treatment and immune-related adverse events may be found in Table 8 and Table 9 respectively.

All patients had an AE of any grade; 70 (60.3%) had a grade ≥3 AE. 94 patients (81.0%) had a treatment-related AE (TRAE) of any grade; 21 (18.1%) had a grade ≥3 TRAE. 17 patients (14.7%) had a serious TRAE, and 14 patients (12.1%) had a TRAE that led to treatment discontinuation. 15 patients (12.9%) had an AE that led to death, none of which were treatment related.

35 patients (30.2%) had an immune-related AE (irAE) of any grade; 7 patients (6.0%) had a grade ≥3 irAE. Infusion-related reactions (IRRs, identified via an expanded definition) occurred in 34 patients (29.3%). 1 patient (0.9%) had a grade 3 IRR; no grade 4 or 5 IRRs occurred.

Table 8: A8: TRAEs (any grade in ≥10% of pts or grade ≥3 in any pt) (JAVELIN Merkel 200: Part B)

N=116	Any grade		Grade ≥3	
	n	%	n	%
Any TRAE, n (%)*	94	81.0	21	18.1
Fatigue	24	20.7	1	0.9
Pruritus	15	12.9	1	0.9
Asthenia	16	13.8	0	-
Chills	12	10.3	0	-
Lipase increased	6	5.2	4	3.4
Decreased appetite	6	5.2	1	0.9
ALT increased	5	4.3	1	0.9
Amylase increased	3	2.6	3	2.6
AST increased	2	1.7	1	0.9
Autoimmune nephritis	1	0.9	1	0.9
Autoimmune neuropathy	1	0.9	1	0.9
Cholangitis	1	0.9	1	0.9
Colitis	1	0.9	1	0.9
Dehydration	1	0.9	1	0.9

Dermatitis psoriasiform	1	0.9	1	0.9
Gait disturbance	1	0.9	1	0.9
Liver function test increased	1	0.9	1	0.9
Paraneoplastic encephalomyelitis	1	0.9	1	0.9
Paraneoplastic syndrome	1	0.9	1	0.9
Polyneuropathy in malignant disease	1	0.9	1	0.9
Troponin increased	1	0.9	1	0.9
Tumor lysis syndrome	1	0.9	1	0.9
Any IRR†	34	29.3	1	0.9

* The incidence of treatment-related IRRs based on the single MedDRA Preferred Term is not listed.
† Includes AEs (irrespective of relatedness) categorized as IRR, drug hypersensitivity, or hypersensitivity reaction that occurred on the day of infusion or day after infusion, in addition to signs and symptoms of IRR that occurred on the same day of infusion and resolved in ≤2 days.
Key: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; IRR, infusion-related reaction; TRAE, treatment-related adverse event.

Table 9: A8: irAEs (any grade in ≥5% of pts or grade ≥3 in any pt) (JAVELIN Merkel 200: Part B)

N=116	Any grade		Grade ≥3	
	n	%	n	%
Any irAE, n (%)	35	30.2	7	6.0
Pruritus	9	7.8	1	0.9
Rash maculopapular	6	5.2	0	
ALT increased	3	2.6	1	0.9
Autoimmune nephritis	1	0.9	1	0.9
Autoimmune neuropathy	1	0.9	1	0.9
Dermatitis psoriasiform	1	0.9	1	0.9
Diabetes mellitus	1	0.9	1	0.9
Liver function test increased	1	0.9	1	0.9

* The incidence of treatment-related IRRs based on the single MedDRA Preferred Term is not listed.
Key: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; IRR, infusion-related reaction; TRAE, treatment-related adverse event.

Health-related quality of life

As the data cut used to inform this CDF review comes from a primary analysis of JAVELIN Merkel 200: Part B (data cut-off: 2 May 2019), a comprehensive summary

of the available HRQoL data is not yet available. However, a summary of the HRQoL data is provided below:

In JAVELIN Merkel 200: Part B, patients were administered two HRQoL questionnaires: the EQ-5D-5L and the FACT-M. THE EQ-5D is a generic preference-based measure of HRQoL that is commonly used in clinical trials. The FACT-M (Functional Assessment of Cancer Therapy – Melanoma) questionnaire is a melanoma-specific measure. FACT-M has been validated in patients with melanoma which shares many similarities with Merkel cell carcinoma.

Of the n=116 participants in the full analysis set (i.e. the intention-to-treat population), n=104 participants (89.7%) completed the EQ-5D and n=102 participants (87.9%) completed FACT-M questionnaires at the Screening visit. The proportion of participants completing each questionnaire generally stayed ≥ 60% throughout the treatment period and was similar for each questionnaire. The completion rate at the End-of-Treatment visit was 79.6% for both the EQ-5D and FACT-M.

No major changes in participants' health status overall were found with the EQ-5D Visual Analogue Score, the EQ-5D Index score or the FACT-M scores while on treatment, suggesting no deterioration in subjects' HRQoL over time. A summarised view of reported EQ-5D VAS scores is provided in Table 10 and Figure 7.

Table 10: A8: Results for EQ-5D VAS (JAVELIN Merkel 200: Part B)

	Value at time of assessment		Change from baseline
	N	Median (Std)	LS-Mean [95%-CI][a]
EQ-5D VAS			
Baseline	[redacted]	[redacted]	[redacted]
Week 7	[redacted]	[redacted]	[redacted]
Week 13	[redacted]	[redacted]	[redacted]
Week 19	[redacted]	[redacted]	[redacted]
Week 25	[redacted]	[redacted]	[redacted]
Week 31	[redacted]	[redacted]	[redacted]
Week 37	[redacted]	[redacted]	[redacted]
Week 43	[redacted]	[redacted]	[redacted]

	Value at time of assessment		Change from baseline
	N	Median (Std)	LS-Mean [95%-CI][a]
Week 49	[redacted]	[redacted]	[redacted]
Week 55	[redacted]	[redacted]	[redacted]
Week 61	[redacted]	[redacted]	[redacted]
Week 67	[redacted]	[redacted]	[redacted]
Week 73	[redacted]	[redacted]	[redacted]
Week 79	[redacted]	[redacted]	[redacted]
Week 85	[redacted]	[redacted]	[redacted]
Week 91	[redacted]	[redacted]	[redacted]
Week 97	[redacted]	[redacted]	[redacted]
Week 103	[redacted]	[redacted]	[redacted]
Week 109	[redacted]	[redacted]	[redacted]
Week 115	[redacted]	[redacted]	[redacted]
Week 121	[redacted]	[redacted]	[redacted]
Week 127	[redacted]	[redacted]	[redacted]
Week 133	[redacted]	[redacted]	[redacted]
Week 139	[redacted]	[redacted]	[redacted]
Week 145	[redacted]	[redacted]	[redacted]
Week 151	[redacted]	[redacted]	[redacted]
Overall[b]	[redacted]	[redacted]	[redacted]

a: Based on a linear mixed model, which contains the initial value, the visits (as a categorical variable) and the regression constant (intercept) as covariates. At each time of the survey, the available data of all patients in the patient-reported outcomes population (n=85) are included in the analysis. Positive values correspond to an improvement in the state of health.
b: LS-Mean Estimator of the average change compared to the baseline over the duration of the survey based on a linear mixed model, which contains the baseline and the regression constant (intercept) as covariates; p-value of the t-test to investigate whether the average change compared to the initial value differs significantly from zero. Positive values correspond to an improvement in the state of health.
Key: EQ-5D: EuroQol-5 dimensions questionnaire; CI: confidence interval; LS: Least Squares; NE: not estimable; Std: standard deviation; VAS: Visual analogue scale.

Figure 7: A8: Change (LS mean [95% CI]) of the EQ-5D VAS (JAVELIN Merkel 200: Part B)

Further information concerning the incorporation of the EQ-5D data into the updated economic model is provided in Appendix 4 of the company submission.

Subgroups

A9. Please provide subgroup results by sex (male/female) from Javelin Merkel 200: Part B separately for the outcomes PFS, OS, response rate and time on treatment.

Please find the requested subgroup results below:

- PFS: Figure 8
- OS: Figure 9
- Response rate: Table 11
- Time on treatment: Figure 10

Figure 8: A9: PFS from JM200: Part B, subgroup results by sex (male/female)

Figure 9: A9: OS from JM200: Part B, subgroup results by sex (male/female)

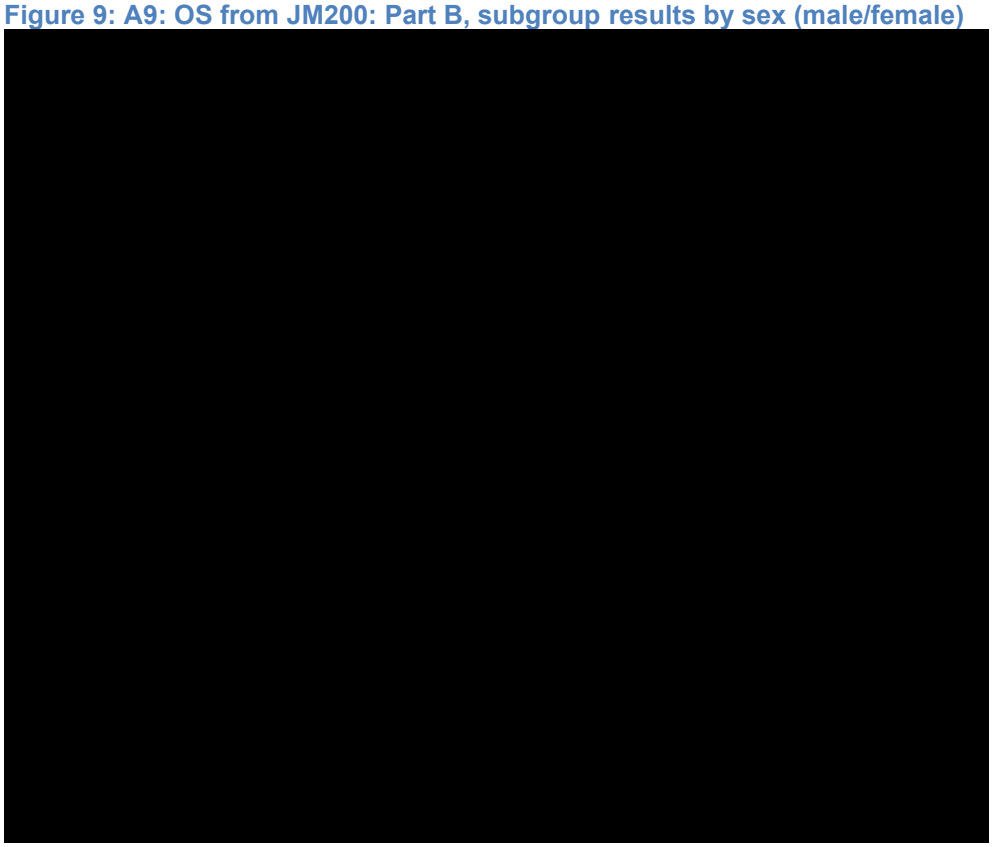


Table 11: A9: Response rate from JM200: Part B, subgroup results by sex (male/female)

Outcomes by RECIST v1.1, per IRC Assessment	Males	Females
Number of patients, n	81	35
Confirmed ORR, % (95% CI)	[redacted]	[redacted]
Confirmed BOR, n (%)		
CR	[redacted]	[redacted]
PR	[redacted]	[redacted]
SD	[redacted]	[redacted]
PD	[redacted]	[redacted]
Non-CR/Non-PD	[redacted]	[redacted]
NE	[redacted]	[redacted]

Key: BOR, best overall response; CR, complete response; IRC, independent review committee; NE, not evaluable; RECIST, response evaluation criteria in solid tumors; ORR, objective response rate; PD, progressive disease; PR, partial response; SD, stable disease.

Figure 10: A9: ToT from JM200: Part B, subgroup results by sex (male/female)

[redacted]

A10. Please provide the results for PFS and OS for each of the following subgroups in Javelin Merkel 200: Part B:

- Age (<80 years and ≥80 years);
- ECOG status at baseline;
- Merkel cell polyomavirus (MCPyV) status at baseline;
- Tumour PD-L1 expression status.

Please find the requested subgroup results below:

- Age (<80 years and ≥80 years)
  - PFS: Figure 11
  - OS: Figure 12
- ECOG status at baseline
  - PFS: Figure 13
  - OS: Figure 14
- Merkel cell polyomavirus (MCPyV) status at baseline
  - PFS: Figure 15
  - OS: Figure 16
- Tumour PD-L1 expression status
  - PFS: Figure 17
  - OS: Figure 18

Following the clarification teleconference call held on 19 March 2020, an additional analysis was suggested concerning age stratified by patients aged <75 years versus ≥75 years (to allow for comparison to Study 100070-Obs001). The results of this analysis are also provided below:

- Age (<75 years and ≥75 years)
  - PFS: Figure 19
  - OS: Figure 20

In addition, median outcomes are reported for each of these outcomes in Table 12.

Figure 11: A10: PFS from JM200: Part B, subgroup results by age (<80 and ≥80 years)

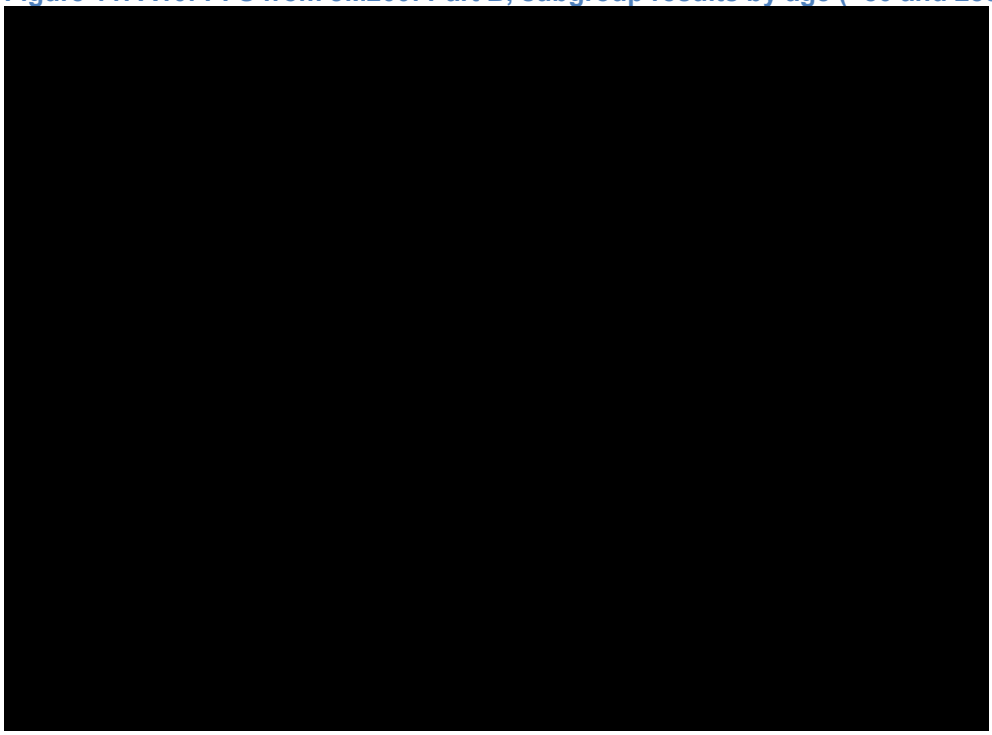


Figure 12: A10: OS from JM200: Part B, subgroup results by age (<80 and ≥80 years)

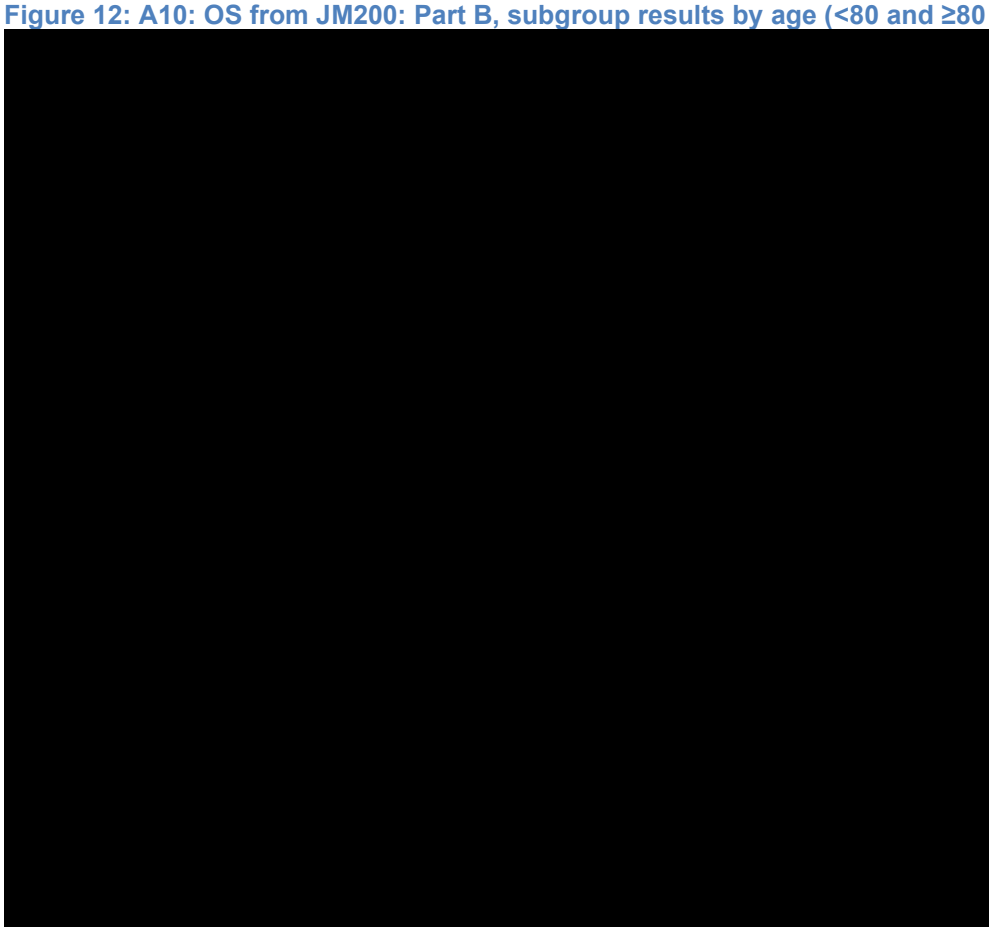


Figure 13: A10: PFS from JM200: Part B, subgroup results by ECOG status

Figure 14: A10: OS from JM200: Part B, subgroup results by ECOG status

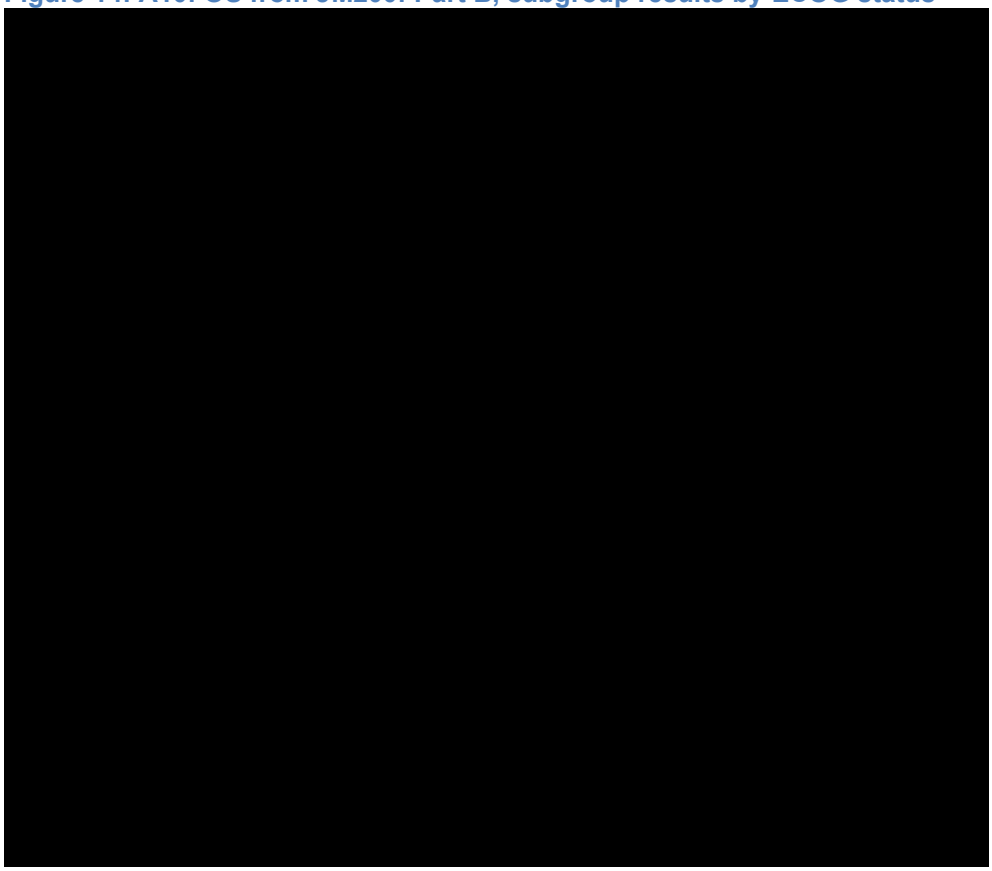


Figure 15: A10: PFS from JM200: Part B, subgroup results by MCPyV status

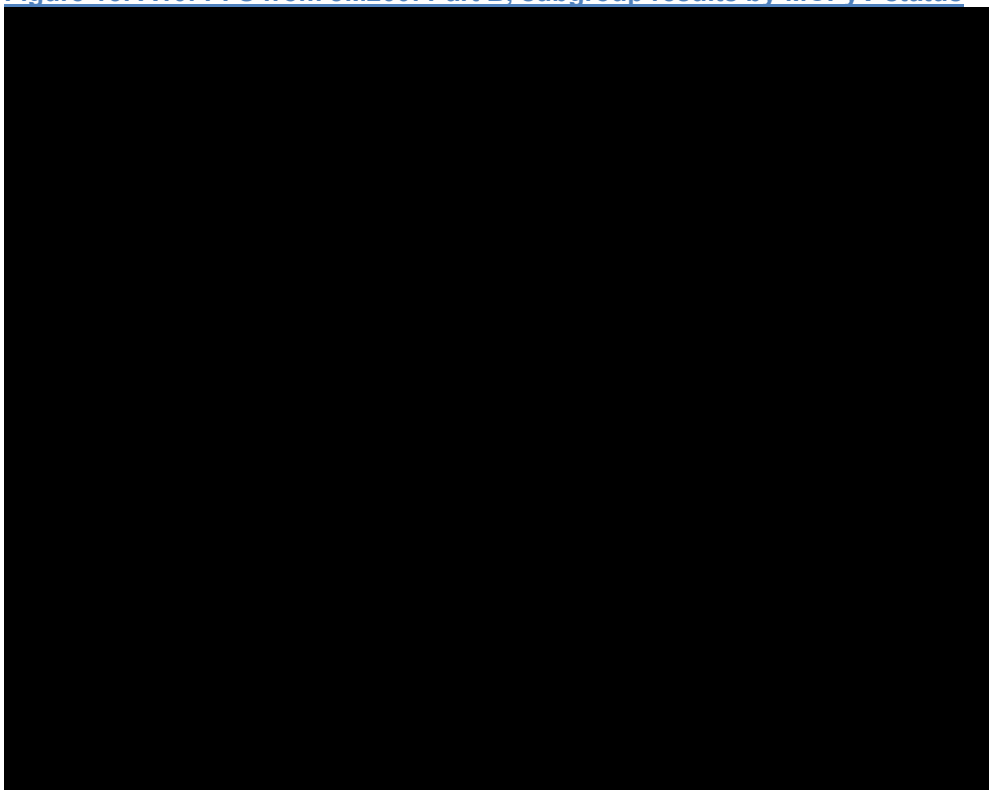


Figure 16: A10: OS from JM200: Part B, subgroup results by MCPyV status

Figure 17: A10: PFS from JM200: Part B, subgroup results by PD-L1 status

Figure 18: A10: OS from JM200: Part B, subgroup results by PD-L1 status

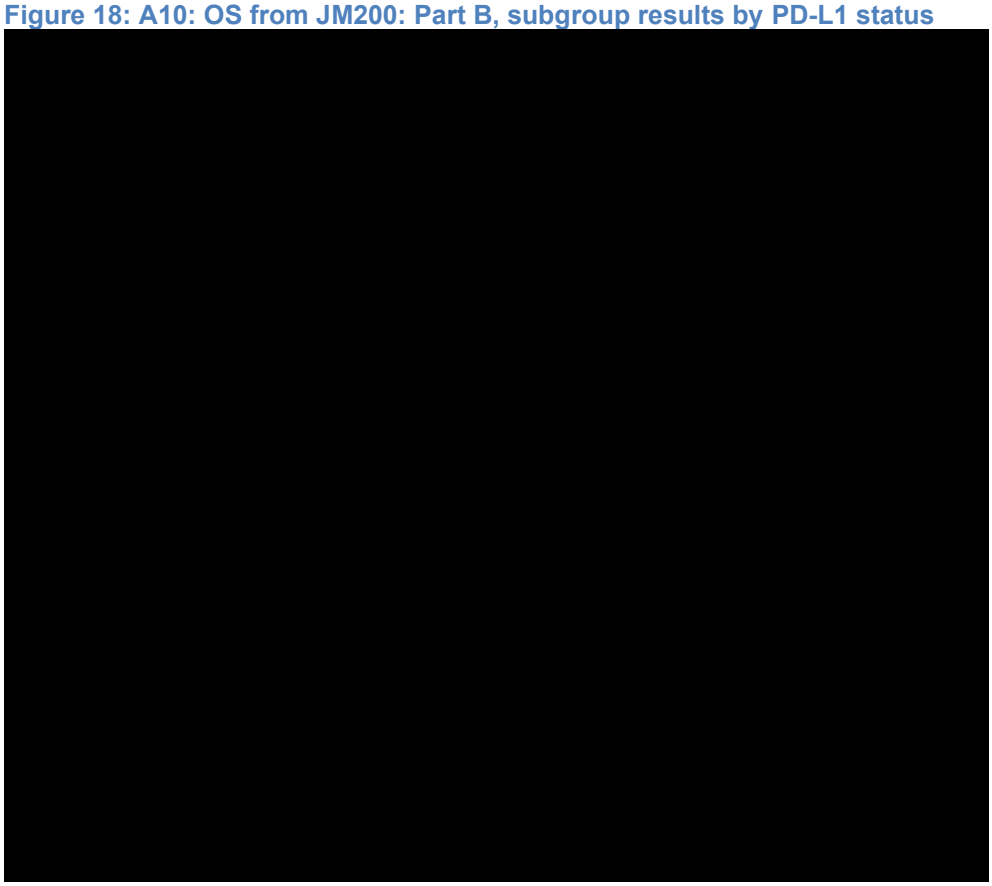


Figure 19: A10: PFS from JM200: Part B, subgroup results by age (<75 and ≥75 years)

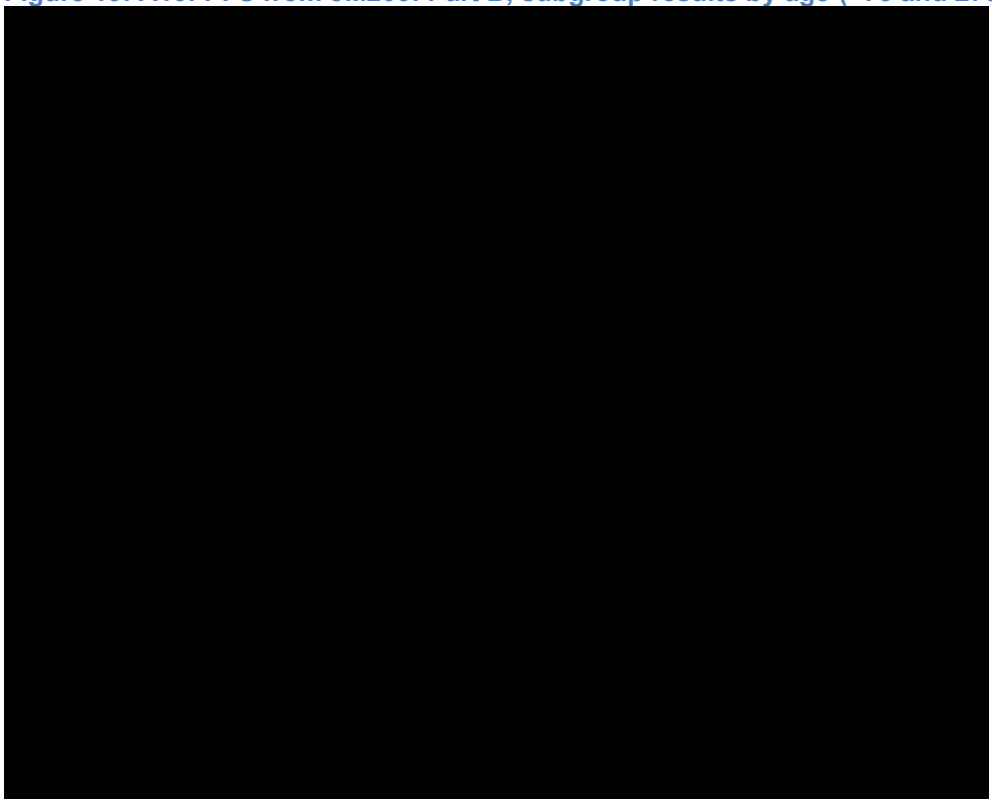


Figure 20: A10: OS from JM200: Part B, subgroup results by age (<75 and ≥75 years)

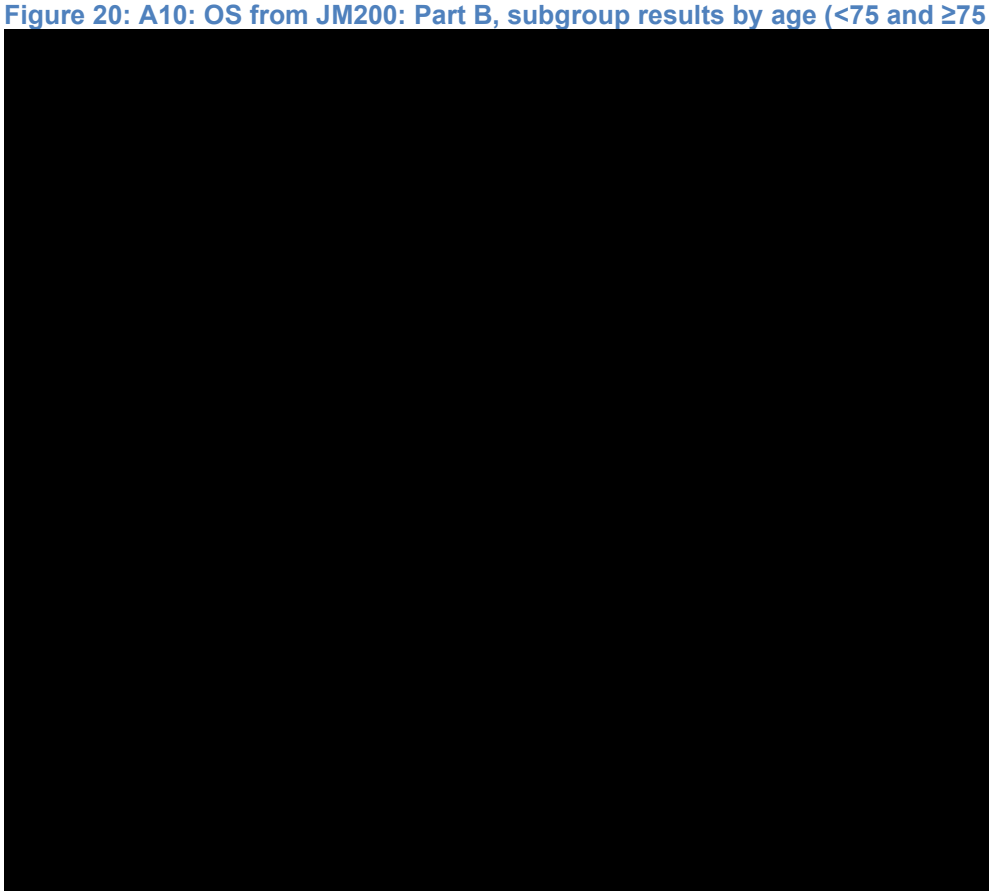


Table 12: A10: Median outcomes (subgroup analyses)

Subgroup	Result	Lower CI	Upper CI	Number
OS				
Aged <80 years	[redacted]	[redacted]	[redacted]	84
Aged >= 80 years	[redacted]	[redacted]	[redacted]	32
ECOG = 0	[redacted]	[redacted]	[redacted]	72
ECOG = 1	[redacted]	[redacted]	[redacted]	44
MCPyV = positive	[redacted]	[redacted]	[redacted]	70
MCPyV = negative	[redacted]	[redacted]	[redacted]	37
MCPyV = not estimable	[redacted]	[redacted]	[redacted]	9
PD-L1 = positive	[redacted]	[redacted]	[redacted]	21
PD-L1 = negative	[redacted]	[redacted]	[redacted]	87
PD-L1 = not estimable	[redacted]	[redacted]	[redacted]	8
Aged <75 years	[redacted]	[redacted]	[redacted]	59
Aged >= 75 years	[redacted]	[redacted]	[redacted]	57
PFS				
Aged <80 years	[redacted]	[redacted]	[redacted]	84
Aged >= 80 years	[redacted]	[redacted]	[redacted]	32
ECOG = 0	[redacted]	[redacted]	[redacted]	72
ECOG = 1	[redacted]	[redacted]	[redacted]	44
MCPyV = positive	[redacted]	[redacted]	[redacted]	70
MCPyV = negative	[redacted]	[redacted]	[redacted]	37
MCPyV = not estimable	[redacted]	[redacted]	[redacted]	9
PD-L1 = positive	[redacted]	[redacted]	[redacted]	21
PD-L1 = negative	[redacted]	[redacted]	[redacted]	87
PD-L1 = not estimable	[redacted]	[redacted]	[redacted]	8
Aged <75 years	[redacted]	[redacted]	[redacted]	59
Aged >= 75 years	[redacted]	[redacted]	[redacted]	57

Key: NE, not evaluable

Subsequent and concomitant therapies

A11. Priority question. Please provide details of subsequent treatments received by first-line patients post-progression in:

a) Javelin Merkel 200: Part B;

b) Study 100070-Obs001 overall population;

c) Study 100070-Obs001 immunocompetent subgroup.

Please see below a pdf detailing the output from the clinical study report for JAVELIN Merkel 200: Part B concerning subsequent therapies. Please note that reporting within JAVELIN Merkel 200: Part B is based on participants with at least one subsequent therapy, and the values reported in this table are not mutually-exclusive. Of the total n=116 population, n=40 patients (34.5%) had at least one subsequent therapy, most of whom received carboplatin and/or etoposide (n=35).

emr100070_003_Foll
ow-up treatments.p

For Study 100070-Obs001, the corresponding publication by Cowey *et al.* provides information concerning subsequent therapies that were recorded (provided as part of the submission reference pack, filename *007 Cowey (2017).pdf*).

Of the total n=67 patients, n=20 had a corresponding record of receiving treatment in the second-line and beyond setting. Of the n=20 patients who received at least one subsequent line of therapy, n=14 patients were immunocompetent (and thus comprised the 'primary analysis population' considered within the study publication). Most patients went on to receive either topotecan or a combination of vincristine + cyclophosphamide + doxorubicin (also known as 'CAV').

The breakdown of treatments received for each of these population is provided within Table 3 of the Cowey *et al.* publication.

A12. Please provide details of any concomitant therapies received by first-line patients in:

a) Javelin Merkel 200: Part B;

b) Study 100070-Obs001 overall population;

c) Study 100070-Obs001 immunocompetent subgroup.

Information concerning concomitant therapies in Study 100070-Obs001 are not available. However, please see below a pdf detailing the output from the clinical study report for JAVELIN Merkel 200: Part B concerning concomitant therapies.

emr100070-003-part B-PA-concomitant m

Avelumab versus chemotherapy

A13. Priority question. Please provide three separate Kaplan-Meier figures with both the results from the May 2019 data-cut from Javelin Merkel 200: Part B (n = 116) and the first-line immunocompetent subgroup of patients in Study 100070-Obs001 (n = 51) for the following outcomes:

a) PFS;

b) OS;

c) Time on treatment.

Please find the requested plots below:

- PFS: Figure 21
- OS: Figure 22
- ToT: Figure 23

Figure 21: A13: PFS, JM200: Part B versus Obs001 (immunocompetent)

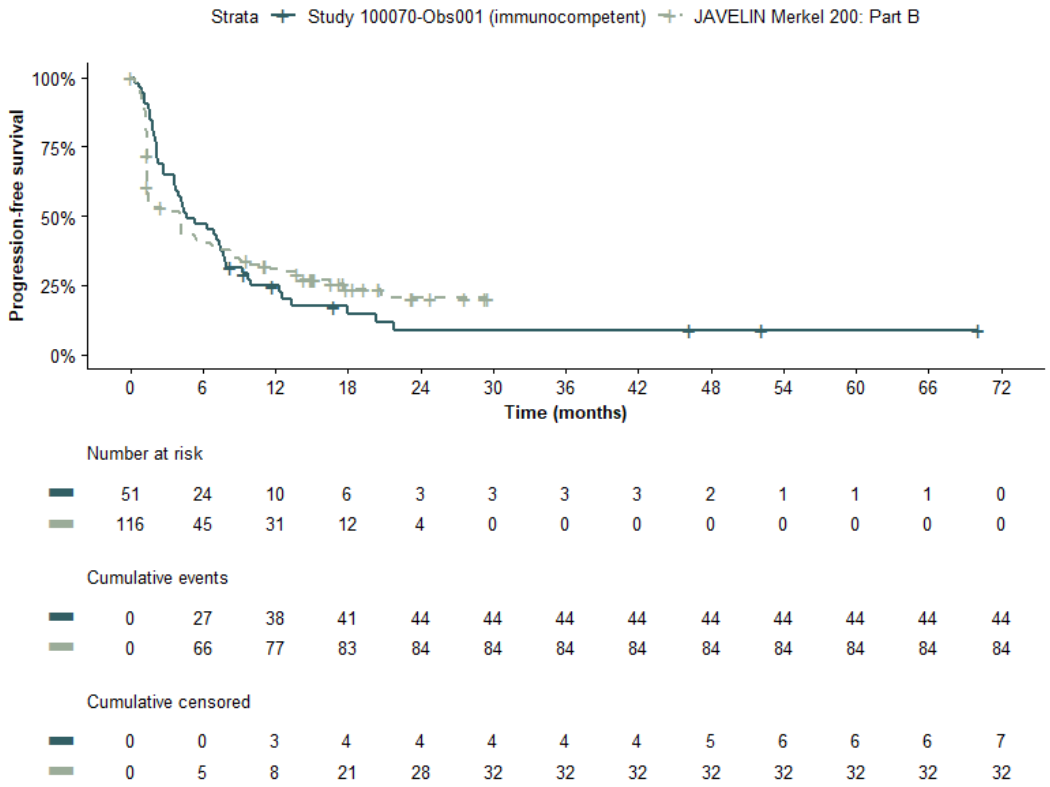


Figure 22: A13: OS, JM200: Part B versus Obs001 (immunocompetent)

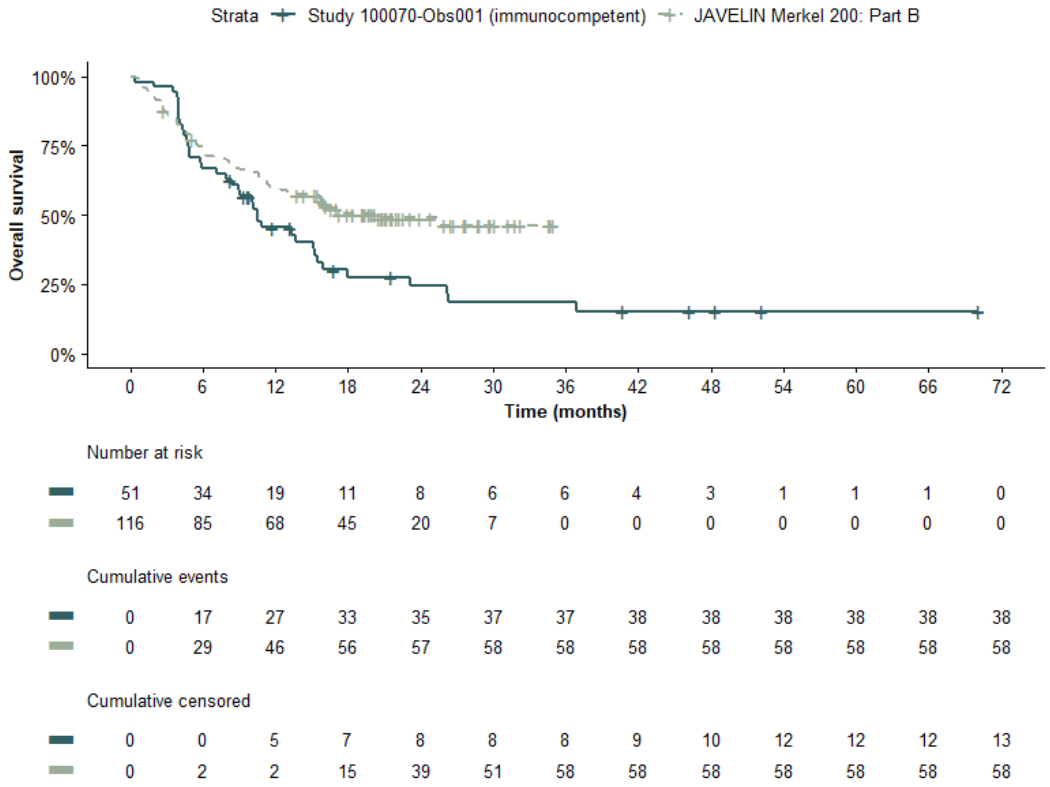


Figure 23: A13: ToT, JM200: Part B versus Obs001 (immunocompetent)

Study 100070-Obs001

A14. Priority question: Please provide Kaplan-Meier figures including the number of patients at risk, number of patients censored and the numbers of patients who have had an event for the first-line immunocompetent subgroup of patients in Study 100070-Obs001 (n = 51) for the following outcomes:

a) **PFS;**

b) **OS;**

c) **Time on treatment.**

Please see the corresponding Kaplan-Meier figures below for each outcome:

- PFS: Figure 24
- OS: Figure 25
- ToT: Figure 26

For completeness, two plots are presented for each outcome. The first plot contains estimates for the whole population (n=67), the immunocompetent subgroup (n=51), and the immunocompromised subgroup (n=16); whereas the second contains only the plot for the immunocompetent subgroup.

Figure 24: A14: PFS from Obs001, subgroup results by immunosuppression

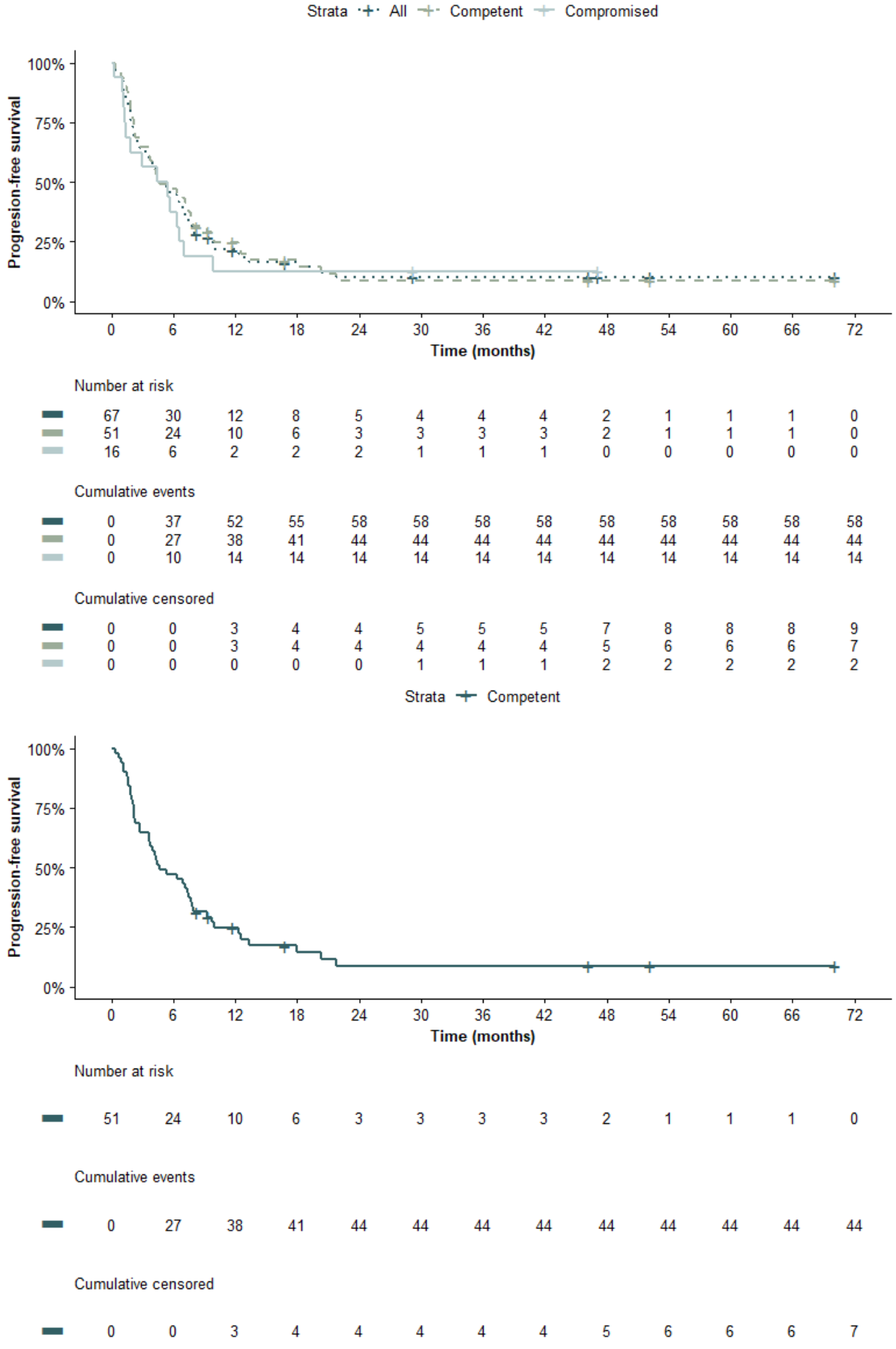


Figure 25: A14: OS from Obs001, subgroup results by immunosuppression

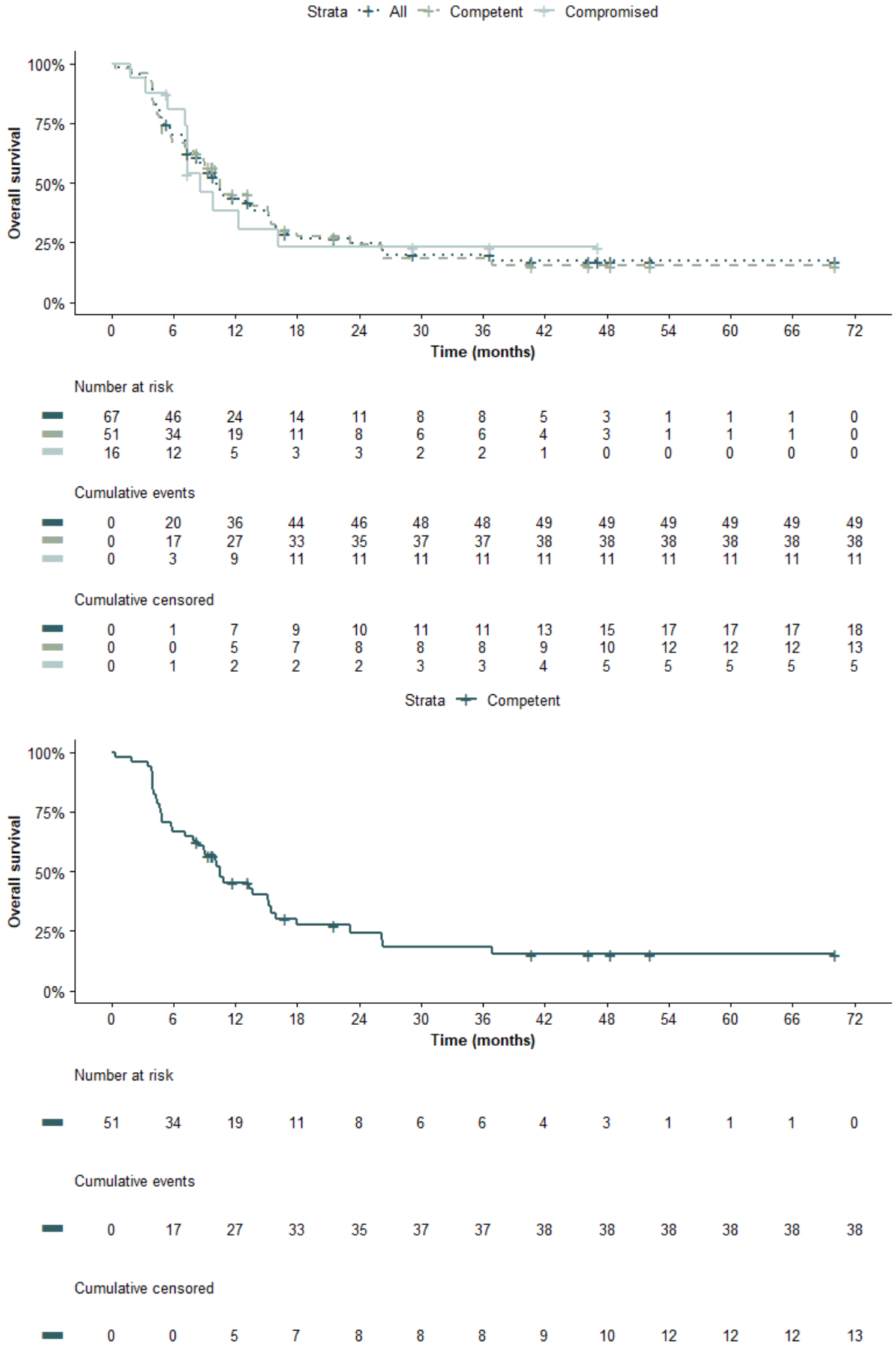
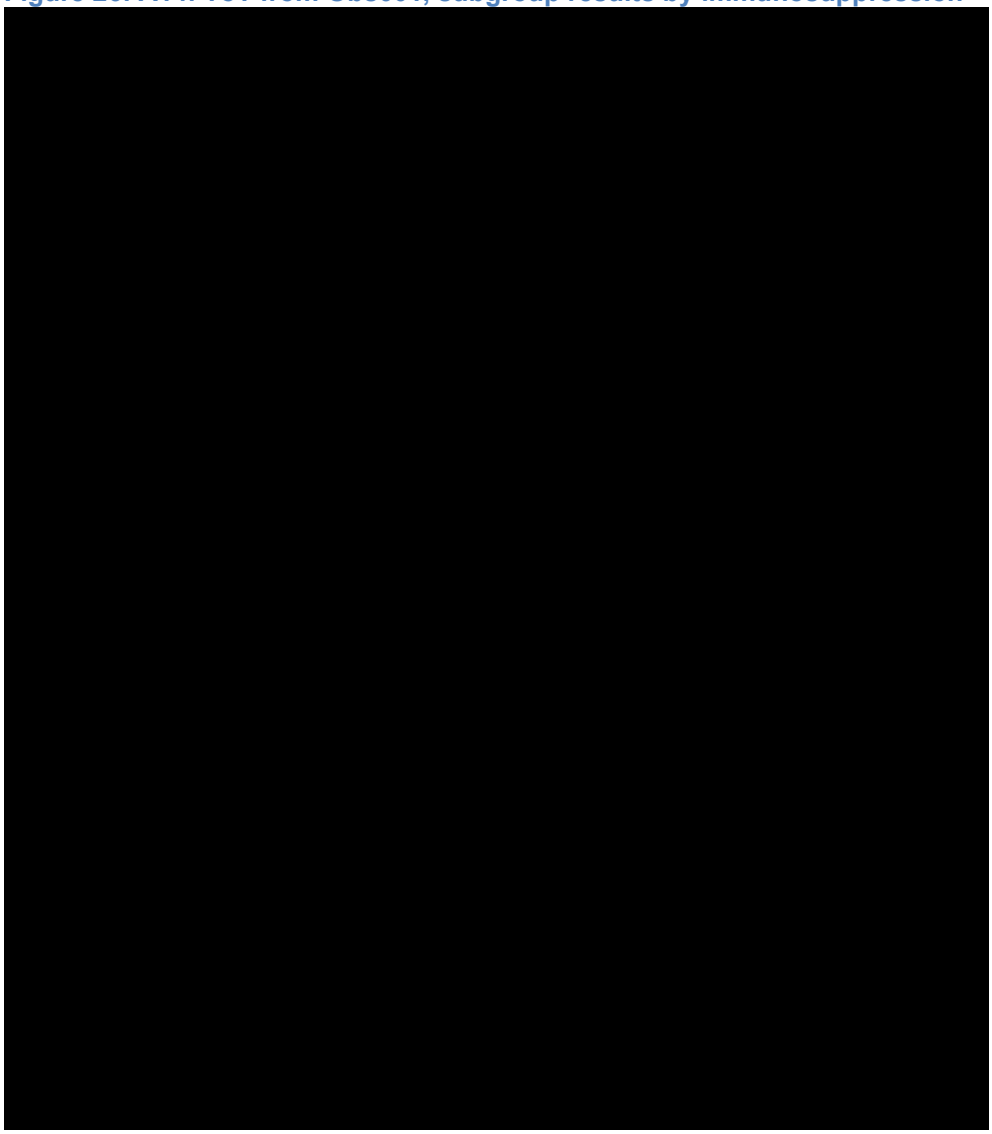


Figure 26: A14: ToT from Obs001, subgroup results by immunosuppression



A15. Priority question. Please provide the results as requested in Questions A4 to A7, for the first-line immunocompetent subgroup of patients in Study 100070-Obs001 (n = 51) and add additional rows or columns to the tables below to account for longer follow-up.

a) Time on treatment

Please see the requested information in Table 13.

Table 13: A15: ToT numbers (Study 100070-Obs001, immunocompetent)

Outcome	Result	95% CI	Number in analysis
Median ToT, months (95% CI)	[redacted]	[redacted]	51
6-month ToT rate, % (95% CI)	[redacted]	[redacted]	[redacted]

12-month ToT rate, % (95% CI)	[redacted]	[redacted]	[redacted]
15-month ToT rate, % (95% CI)	[redacted]	[redacted]	[redacted]

b) Number of patients at risk

Please see the requested information in Table 14.

Table 14: A15: OS and PFS numbers (Study 100070-Obs001, immunocompetent)

Outcome	Metric	Time intervals (months)											
		0	6	12	18	24	30	36	42	48	54	60	66	72
OS	At risk	[redacted]												
	Censored	[redacted]												
	Events	[redacted]												
PFS	At risk	[redacted]												
	Censored	[redacted]												
	Events	[redacted]												

c) Please provide details of the reasons for censoring for the PFS and OS analyses:

Reasons for per-patient censoring are not available for Study 100070-Obs001. However, please see below the relevant description of reasons for censoring adopted within the study:

Progression-free survival

"PFS was measured from the index date of treatment to the date of progression, or date of death due to any cause, or date of initiation of new regimen, censoring patients who were still alive and did not progress at the last office visit date without clinical or radiographic evidence of progression. PFS was estimated in months using the Kaplan-Meier method with 95% CIs, and summary tables of the number of events and censored patients at each month interval included."

Overall survival

"OS was defined as the interval between the index date of treatment and the date of death from any cause as documented in the [Social Security Death Index] *or*

[iKnowMed]*/chart review. Patients who did not die were censored on the study end date or the last visit date available, whichever occurred first. If a death date is known beyond the study end date, then the study end date was used as the censoring date. OS was estimated in months using the Kaplan-Meier method with 95% CIs, and summary tables of the number of events and censored patients at each month interval included."*

d) PFS and OS results

Please see the requested information in Table 15.

Table 15: A15: OS and PFS outcomes (Study 100070-Obs001, immunocompetent)

Outcome	Result	95% CI	Number in analysis
Median PFS, months (95% CI)	4.63	(2.79, 7.66)	51
6-month PFS rate, % (95% CI)	47.1%	(33.0%, 60.0%)	[redacted]
12-month PFS rate, % (95% CI)	24.8%	(13.8%, 37.4%)	[redacted]
15-month PFS rate, % (95% CI)	17.3%	(8.1%, 29.5%)	[redacted]
Median OS, months (95% CI)	10.51	(7.16, 15.24)	51
6-month OS rate, % (95% CI)	66.7%	(52.0%, 77.8%)	[redacted]
12-month OS rate, % (95% CI)	45.3%	(31.0%, 58.6%)	[redacted]
15-month OS rate, % (95% CI)	40.3%	(26.2%, 53.9%)	[redacted]

Key: NE, not evaluable

Clinical systematic literature review

A16. Please provide a full list of references for the 36 included studies identified in the updated clinical systematic literature review (SLR; Clinical effectiveness review: updated search, Figure 1: PRISMA)

Table 16 contains a summary of the included studies identified in the updated clinical SLR. The corresponding publication type (i.e. full text or abstract only) is provided alongside the reference number ("Ref") in the reference list within the clinical SLR document.

Table 16: A16: Included studies identified in the updated clinical SLR

#	Reference	Type	Ref

1	D'Angelo SP, Russell J, Lebbe C, Chmielowski B, Gambichler T, Grob JJ, et al. Efficacy and safety of first-line avelumab treatment in patients with stage IV metastatic merkel cell carcinoma a preplanned interim analysis of a clinical trial. JAMA Oncology. 2018;4:e180077.	Full text	5
2	Kaufman HL, Russell JS, Hamid O, Bhatia S, Terheyden P, D'Angelo SP, et al. Updated efficacy of avelumab in patients with previously treated metastatic Merkel cell carcinoma after >=1 year of follow-up: JAVELIN Merkel 200, a phase 2 clinical trial. Journal for ImmunoTherapy of Cancer. 2018;6:7.	Full text	6
3	Becker JC, Lorenz E, Ugurel S, Eigentler TK, Kiecker F, Pfohler C, et al. Evaluation of real-world treatment outcomes in patients with distant metastatic Merkel cell carcinoma following second-line chemotherapy in Europe. Oncotarget. 2017;8:79731-41.	Full text	7
4	Cowey C, Mahnke L, Espirito J, Helwig C, Oksen D, Bharmal M. Real-world treatment outcomes in patients with metastatic Merkel cell carcinoma treated with chemotherapy in the USA. Future Oncology. 2017;13:1699-710.	Full text	8
5	Rabinowits G, Lezcano C, Catalano PJ, McHugh P, Becker H, Reilly MM, et al. Cabozantinib in Patients with Advanced Merkel Cell Carcinoma. Oncologist. 2018;23:814-21.	Full text	9
6	Bharmal M, Fofana F, Barbosa CD, Williams P, Mahnke L, Marrel A, et al. Psychometric properties of the FACT-M questionnaire in patients with Merkel cell carcinoma. Health and quality of life outcomes. 2017;15:247.	Full text	10
7	Bharmal M, Hunger M, Schlichting M. PCN364 - EVALUATING PSYCHOMETRIC PROPERTIES OF UK EQ-5D-5L SCORING ALGORITHMS IN METASTATIC MERKEL CELL CARCINOMA. Value in Health. 2018;21:S76.	Abstract only	11
8	Bharmal M, Lambert J, Russell JS, Lebbe C, Chmielowski B, Hennessy M, et al. Patient (pt) experiences with avelumab in treatmentnaive metastatic Merkel cell carcinoma (mMCC): Qualitative interview findings from a registrational clinical trial. Journal of Clinical Oncology. 2019;37.	Abstract only	12
9	Bharmal M, Marrel A, Hennessy M, Fofana F, Lambert J, Arnould B. Comparative effectiveness of avelumab versus chemotherapy in Merkel cell carcinoma: innovative use of patient insights. Journal of comparative effectiveness research. 2018;7:881-90.	Full text	13
10	Bharmal M, Nolte S, Henry-Szatkowski M, Hennessy M, Schlichting M. PCN483 CONFIRMING THE PSYCHOMETRIC PERFORMANCE OF THE FUNCTIONAL ASSESSMENT OF CANCER THERAPY–MELANOMA (FACT-M) QUESTIONNAIRE IN PATIENTS WITH MERKEL CELL CARCINOMA (MCC). Value in Health. 2019;22:S531.	Abstract only	14
11	Bullement A, Amin A, Stapelkamp C, Willis A, Lilley C, Hatswell AJ, et al. MO2 - MODELLING OVERALL SURVIVAL IN IMMUNOTHERAPY USING PARAMETRIC TECHNIQUES: AVELUMAB IN PREVIOUSLY TREATED METASTATIC MERKEL CELL CARCINOMA. Value in Health. 2018;21:S11.	Abstract only	15
12	Bullement A, D'Angelo SP, Amin A, Stapelkamp C, Willis A, Lilley C, et al. Predicting overall survival in patients (pts) with treatment-naive metastatic Merkel cell carcinoma (mMCC) treated with avelumab. Journal of Clinical Oncology. 2018;36.	Abstract only	16
13	D'Angelo SP, Nolte S, Schlichting M, Henry-Szatkowski M, Hennessy M, Bharmal M. Health-related quality of life in patients with metastatic Merkel cell carcinoma receiving second-line or later avelumab treatment: 36-month follow-up data. Annals of Oncology. 2019;30:v538.	Abstract only	17
14	D'Angelo SP, Fofana F, Schlichting M, Henry-Szatkowski M, Hennessy M, Bharmal M. Responder analysis based on patient-reported outcomes (PROs) and clinical endpoints (CEPs) in patients (pts) with metastatic Merkel cell carcinoma (mMCC) treated with avelumab. Annals of Oncology. 2018;29:viii457-viii8.	Abstract only	18
15	D'Angelo SP, Hunger M, Brohl AS, Nghiem P, Bhatia S, Hamid O, et al. Early objective response to avelumab treatment is associated with improved overall survival in patients with metastatic Merkel cell carcinoma. Cancer immunology, immunotherapy : CII. 2019;68:609-18.	Full text	19
16	D'Angelo SP, Russell J, Hassel JC, Lebbe C, Chmielowski B, Rabinowits G, et al. First-line (1L) avelumab treatment in patients (pts) with metastatic Merkel cell carcinoma (mMCC): preliminary data from an ongoing study. Journal of clinical oncology Conference: 2017 annual meeting of the american society of clinical oncology, ASCO United states. 2017;35.	Abstract only	20
17	Kaufman H, Hunger M, Hennessy M, Schlichting M, Bharmal M. Minimal Impact on Patients’ Health Utilities Associated with Adverse Events in Metastatic Merkel Cell Carcinoma Patients on Treatment with Avelumab. Value in Health. 2017;20(9):A448.	Abstract only	21

18	Kaufman H, Mahnke L, von Heydebreck A, Bharmal M. Association Between Tumour Lesion Size and Health-Related Quality of Life Outcomes in Patients with Metastatic Merkel Cell Carcinoma Treated with Avelumab. Value in Health. 2017;20(9):A455.	Abstract only	22
19	Kaufman HL, Hunger M, Hennessy M, Schlichting M, Bharmal M. Nonprogression with avelumab treatment associated with gains in quality of life in metastatic Merkel cell carcinoma. Future Oncology. 2018;14:255-66.	Full text	23
20	Kaufman HL, Russell J, Hamid O, Bhatia S, Terheyden P, D'Angelo SP, et al. Avelumab in chemotherapy-refractory metastatic Merkel cell carcinoma: Subgroup analysis of efficacy. Journal of Clinical Oncology. 2017;35(7_suppl):80-.	Abstract only	24
21	Lanitis T, Proskorovsky I, Ambavane A, Hunger M, Zheng Y, Bharmal M, et al. Survival Analysis in Patients with Metastatic Merkel Cell Carcinoma Treated with Avelumab. Advances in Therapy. 2019;36:2327-41.	Full text	25
22	Nghiem P, Bhatia S, Brohl AS, Hamid O, Mehnert JM, Terheyden P, et al. Two-year efficacy and safety update from JAVELIN Merkel 200 part A: A registrational study of avelumab in metastatic Merkel cell carcinoma progressed on chemotherapy. Journal of Clinical Oncology. 2018;36(15_suppl):9507-.	Abstract only	26
23	Shapiro I, Grote HJ, D'Urso V, Von Heydebreck A, Mahnke L, Kaufman H, et al. Exploratory biomarker analysis in avelumab-treated patients with metastatic Merkel cell carcinoma progressed after chemotherapy. Journal of clinical oncology Conference: 2017 annual meeting of the american society of clinical oncology, ASCO United states. 2017;35.	Abstract only	27
24	Walker J, Kasturi V, Lebbe C, Sandhu SK, Grignani G, Hennessy MG, et al. Second-line avelumab treatment of patients (pts) with metastatic Merkel cell carcinoma (mMCC): Experience from a global expanded access program (EAP). Journal of Clinical Oncology. 2018;36(15_suppl):9537-.	Abstract only	28
25	Ascierto PA, Nathan P, Kasturi V, Dirix LY, Fenig E, Hennessy M, et al. Avelumab in European patients (pts) with metastatic Merkel cell carcinoma (mMCC): Experience from an ad hoc expanded access program (EAP). Annals of Oncology. 2018;29:x29.	Abstract only	29
26	Nathan P, Kasturi V, Dirix L, Fenig E, Ascierto PA, Hennessy M, et al. Avelumab in European patients (pts) with metastatic Merkel cell carcinoma (mMCC): Experience from an ad-hoc expanded access program (EAP). Annals of Oncology. 2018;29:viii460-viii1.	Abstract only	30
27	Levy S, Aarts MJB, Eskens FA, Keymeulen K, Been L, Grunhagen DJ, et al. Avelumab for advanced Merkel cell carcinoma in the Netherlands: A nationwide survey. Annals of Oncology. 2019;30:v337-v8.	Abstract only	31
28	Al Homsi MU, Mostafa M, Fahim K. Favorable Response to Treatment with Avelumab in an HIV-Positive Patient with Advanced Merkel Cell Carcinoma Previously Refractory to Chemotherapy. Case reports in oncology. 2018;11:467-75.	Abstract only	32
29	Klink AJ, Phatak H, Bharmal M, Kaufman J, Feinberg B. Merkel Cell Cancer: Poor Response To Chemotherapy Exposes Significant Unmet Need. Value in Health. 2017;20(9):A415.	Abstract only	33
30	Zheng Y, Kim R, Yu T, Dreyfus J, Gayle JA, Wassel CL, et al. PCN317 REAL-WORLD STUDY OF METASTATIC MERKEL CELL CARCINOMA PATIENTS RECEIVING CHECKPOINT INHIBITORS (CPIS) VS. CHEMOTHERAPY TREATMENTS. Value in Health. 2019;22:S117.	Abstract only	34
31	Chang JWC, Chang YY, Huang YL, Lo YF, Ho TY, Huang YT, et al. Merkel cell carcinoma in Taiwan: A series of 24 cases and literature review. Medicine. 2019;98:e17538.	Full text	35
32	Lopiccolo J, Schollenberger MD, Dakhil S, Rosner S, Ali O, Sharfman WH, et al. Rescue therapy for patients with anti-PD-1-refractory Merkel cell carcinoma: A multicenter, retrospective case series. Journal for ImmunoTherapy of Cancer. 2019;7:170.	Full text	36
33	Winkler JK, Dimitrakopoulou-Strauss A, Sachpekidis C, Enk A, Hassel JC. Ipilimumab has efficacy in metastatic Merkel cell carcinoma: a case series of five patients. Journal of the European Academy of Dermatology and Venereology : JEADV. 2017;31:e389-e91.	Full text	37
34	Ferrarotto R, Mata J, Mott F, Bhosale P, Rubin ML, Altan M, et al. Safety and interim results from a phase II, single-arm study of atezolizumab and bevacizumab in Merkel cell carcinoma (MCC). Journal of Clinical Oncology. 2019;37(15_suppl):e21006-e.	Abstract only	38
35	Roche L, Murphy M, Power DG. Treatment of merkel cell carcinoma with pembrolizumab in a patient with psoriasis and psoriatic arthritis. Journal of the european academy of dermatology	Abstract only	39

	and venereology Conference: 13th congress of the european association of dermato-oncology, EADO 2017 Greece. 2017;31:96.		
36	Thiem A, Gran F, Kneitz H, Schummer P, Herz S, Schrama D, et al., editors. COINCIDENT METASTATIC MELANOMA AND MERKEL CELL CARCINOMA WITH COMPLETE REMISSION ON TREATMENT WITH PEMBROLIZUMAB. 24th World Congress of Dermatology; 2019; Milan (Italy).	Abstract only	40

Section B: Clarification on cost-effectiveness data

B1. Priority question. Please provide an option in the economic model to allow the application of the clinical effectiveness results for OS and PFS from the analysis requested in questions A1. Please also present the results of a scenario analysis using this effectiveness data in the base case model.

Based on the information provided in response to question A1, cost-effectiveness results were produced using seven different approaches to address potential imbalances between the two studies. Four approaches used weighting-based methods, and three used matching-based methods. Of the total seven approaches, four were considered the most suitable to inform the economic model, which are described below:

- **B1-1:** PSW, using SW, based on all patients with available data for age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1+), excluding immunosuppression as a variable
- **B1-2:** PSW, using SW, based on all patients with available data for age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1), excluding immunosuppression as a variable and excluding patients with an ECOG PS of 2 or more
- **B1-3:** PSM on age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1+), excluding immunosuppression as a variable
- **B1-4:** PSM on age (aged ≥75 vs. <75 years), sex (female vs. male), and ECOG PS (0 vs. 1+), excluding immunosuppression as a variable and excluding patients with an ECOG PS of 2 or more

The other three approaches were not considered the most suitable owing to the lack of immunocompromised patients in JM200: Part B to match/weight to patients in Study 100070-Obs001, and the need to maintain the same overall sample size in matching approaches where possible (hence a preference for SW over IPTW).

Please see the response to clarification question A1 for further details concerning the analytical approach to addressing potential imbalances between the groups.

To inform the economic model, the parameterisations of OS and PFS for both the avelumab and chemotherapy arm were re-fitted using weights determined by each analysis. For the weighting analyses, SW were derived using variables available for all patients meant that each patient was assigned a weight less than, equal to, or greater than 1 such that the total sample size for each cohort was maintained. In the case of the matching analyses, this was equivalent to removing patients assigned a weight of 0, and including patients with a weight of 1.

The same functional forms were assumed for each of the parametric curves, which were as follows:

- **OS, avelumab:** 1-knot odds spline-based model
- **PFS, avelumab:** 2-knot odds spline-based model
- **OS, chemotherapy:** Log-logistic model
- **PFS, chemotherapy:** Log-logistic model

The corresponding model results are provided in Table 17.

Table 17: B1: Cost-effectiveness results accounting for matching/ weighting

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
CEA 3 (company submission base-case analysis)							
Chemotherapy	11,116	1.94	1.32				
Avelumab	96,758	5.88	3.60	85,642	3.95	2.27	37,670
B1-1: PSW, all patients (stable weights), using age, sex, ECOG (0 vs. 1+), excluding immunosuppression as a variable							
Chemotherapy	11,947	2.19	1.49				
Avelumab	96,675	6.08	3.69	84,728	3.89	2.20	38,522
B1-2: PSW, all patients (stable weights), using age, sex, ECOG (0 vs. 1), excluding immunosuppression as a variable and removing ECOG 2+ pts							
Chemotherapy	12,022	2.20	1.50				
Avelumab	96,687	6.04	3.67	84,665	3.83	2.17	39,014
B1-3: PSM using age, sex, ECOG (0 vs. 1+), excluding immunosuppression as a variable							
Chemotherapy	11,481	1.85	1.27				
Avelumab	97,688	6.20	3.81	86,207	4.35	2.54	33,905

B1-4: PSM using age, sex, ECOG (0 vs. 1), excluding immunosuppression as a variable and removing ECOG 2+ pts							
Chemotherapy	11,559	1.92	1.31				
Avelumab	97,978	6.73	4.12	86,419	4.80	2.81	30,754

Key: CEA, cost-effectiveness analysis; ECOG, Eastern Cooperative Oncology Group Performance Status; ICER, incremental cost-effectiveness ratio; IPTW, inverse probability of treatment weighting; LYG, life years gained; QALYs, quality-adjusted life years.

Both matching analyses led to a large reduction in avelumab patients, in order to match to each of the chemotherapy patients. Both of the matching analyses led to a lower ICER, driven primarily by the subgroup identified from JM200: Part B exhibiting improved OS compared to the entire n=116 ITT population (see response to clarification question A1). Given that any matching analysis leads to the removal of the majority of patients in JM200: Part B, weighting analyses were considered more appropriate.

The weighting analyses caused the average survival for both arms to increase slightly, but led to a small decrease in the incremental survival gain. This difference in survival led to an increase in the ICER of between £852 (B1-1) and £1,344 (B1-2).

B2. Priority question. Please provide a suitable range of survival models fitted to the first-line OS and PFS data from the latest data-cut of the immunocompetent subgroup of Study 100070-Obs001 (n=51) and add options to allow these to be applied in the economic model for chemotherapy.

A range of parametric models have been fitted to the immunocompetent subgroup of Study 100070-Obs001, and may be selected to inform the chemotherapy arm within the economic model.

As may be inferred from the responses to clarification questions A14 and A15, there is relatively little change in the corresponding results of this analysis, as the curves for OS and PFS are very similar across the whole cohort and the immunocompetent subgroup. However, due to the reduced number of patients at risk within the immunocompetent subgroup (versus the whole population), relatively fewer patients are present in the tail of the Kaplan-Meier curves, leading to slightly lower longer-term estimates of both OS and PFS (e.g. 5-year OS using a log-logistic model is 7.6% for the whole population versus 6.8% for the immunocompetent population).

A comparison of the base-case analysis results (equivalent to CEA 3, per the company submission) is provided in Table 18.

Table 18: B2: Cost-effectiveness results using immunocompetent data only

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
CEA 3 (company submission base-case analysis)							
Chemotherapy	11,116	1.94	1.32				
Avelumab	96,758	5.88	3.60	85,642	3.95	2.27	37,670
B2-1: Analysis using immunocompetent subgroup only*							
Chemotherapy	11,499	1.83	1.25				
Avelumab	96,766	5.88	3.60	85,268	4.06	2.35	36,330

Key: BSC, best supportive care; CDF, Cancer Drugs Fund; CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Notes: * Please note that the total costs for avelumab change marginally, as the average duration of survival for the chemotherapy arm is used within the estimated costings of radiotherapy for all treatment arms.

Based on the noted differences in the estimated OS for each population, restricting the comparator group (to only immunocompetent patients) leads to a lower overall estimate of life-years gained (1.83 versus 1.94), which causes the ICER to reduce from £37,670 to £36,330.

B3. Priority question. Please provide a suitable range of survival models fitted to the SACT data for both OS and ToT and add options to allow these to be applied in the economic model for avelumab. Please also add an option for PFS to be modelled using these ToT curves and provide the results of a scenario analysis with each option applied.

As notified by Merck during the clarification TC on 19 March 2020, this analysis has not been provided. The justification is presented below.

In the company submission, several limitations associated with the SACT dataset were described, including limited sample size, follow-up, and incomplete information – most notably, that data for only two outcomes are available from this cohort: OS and ToT. A key uncertainty in the early JM200: Part B data available at the time of the original TA517 appraisal was the small sample size (n= [redacted] patients) and short follow-up ([redacted]). Whilst the SACT cohort comprises a slightly larger sample (n=52), this is markedly smaller than the cohort available in the latest data cut from JM200: Part B (n=116). The minimum follow-up in SACT is also longer (5

months) than the original Part B data cut, although this is three-time shorter than the minimum follow-up now available from JM200: Part B (15 months).

Additionally, we highlighted a number of concerns regarding the selection of SACT patients into the cohort, owing to the availability of avelumab in the second line, and the inclusion of patients who did not meet the eligibility criteria for access via the CDF (10% of patients in the SACT dataset had an ECOG PS of 2 or 3; and a further 12% of patients had an unknown or missing ECOG PS. Therefore, over 20% of the SACT cohort did not explicitly meet the ECOG PS inclusion criterion). While it remains unclear how this may have influenced outcomes, patients with an ECOG PS of 2 or more are expected to have a poorer prognosis then those with an ECOG PS of 0 or 1.

Due to the reasons outlined above, we do not consider SACT data appropriate to inform the economic model and have therefore not provided the analyses requested. Given that the ERG has access to the same information concerning the SACT cohort that is available to the company (i.e. the report from Public Health England), the ERG may wish to explore outcomes for the SACT cohort in more detail.

B4. Priority question. Please add options in the economic model to apply subsequent treatment costs for both treatment groups, and align these with the treatments received by patients in the clinical effectiveness studies used in each of the economic analyses. Please also include relevant data that aligns with the additional scenarios requested in the questions above.

As discussed during the clarification teleconference held between NICE, the ERG, and the company on 19 March 2020, incorporation of subsequent treatment costs can be challenging for a number of reasons, including:

- Data regarding subsequent treatment costs are not collected in sufficient detail to enable an exhaustive micro-costing approach to incorporating these costs within the model (e.g. duration of subsequent treatments, dosing, etc.)
- Some treatments given to patients are not licensed and/or recommended for use within a metastatic MCC population, and so assumptions are required to infer proxy regimens for the purpose of imputing ‘missing’ data points

    - More specifically, a small proportion of patients in JM200: Part B went on to receive further anti-PD-1/PD-L1/CTLA-4 treatment. The majority of these treatments (with the exception of avelumab) are not reimbursed within the UK for this patient population, and retreatment with avelumab following progression is not expected to occur in NHS clinical practice
    - In addition, each of the three non-avelumab anti-PD-1/PD-L1/CTLA-4 treatments included are each associated with a commercially-sensitive patient access scheme (PAS) discount. Therefore, list prices have been assumed, though are not indicative of the 'true' cost to the NHS
- Transitions to the 'progressed disease' state are not explicitly modelled within a partitioned-survival analysis (PartSA) structure, and so estimated entry to the progressed state is required to assign the estimated costs at relevant points in time within the model

In spite of these limitations, an exploratory analysis has been included within the economic model to consider the impact of subsequent therapy costs on the cost-effectiveness results. A short overview of the approach taken to include these costs is provided below, with corresponding results for a number of key scenarios relating to subsequent treatments.

The information provided in response to clarification question A11 were used to establish the proportion of patients who received a subsequent anticancer therapy. For patients in JM200: Part B, the costs of antineoplastic agents were considered. For chemotherapy regimens, an average duration of 12 weeks was assumed. For the anti-PD-1/PD-L1/CTLA-4 regimens, an assumed treatment duration of 6 months was applied (approximately half the expected duration of treatment for an average first-line avelumab patient).

Since publication of NICE TA517, the cost of topotecan is now available from the drugs and pharmaceutical electronic market information tool (eMIT), and so an option was included to apply this updated cost (for alignment with the other generic chemotherapies). To explore the impact of subsequent anti-PD-1/PD-L1/CTLA-4 treatment costs specifically, an option was included to assume the same cost as an

average subsequent chemotherapy regimen for these patients in lieu of the calculated anti-PD-1/PD-L1/CTLA-4 costs.

For administration costs, the cost of an intravenous administration appointment (per CEA 3) was included every 3 weeks. AE costs were assumed to be as per the chemotherapy arm, and applied for the equivalent duration estimated for each arm.

Within the model, the difference in occupancy of the PFS curve between cycles was used to apportion the time points at which subsequent therapy costs were applied. The overall proportion of patients expected to receive at least one subsequent therapy was then multiplied by the proportion of patients who experience a PFS event each cycle. This value was then multiplied by the average cost per treated patient. For simplicity, all cost categories (acquisition, administration, and resolution of AEs) were combined into one singular cost.

A comparison of the base-case analysis results (equivalent to CEA 3, per the company submission) is provided in Table 19. Three analyses are provided:

- **B4-1: Include subsequent therapies (all costs, per study sources):** Costs for subsequent therapies are included based on the reported information available from each study source (as well as required assumptions where necessary relating to treatment duration, dosing etc.)
- **B4-2: Include subsequent therapies (all costs, per study sources + updated topotecan cost):** As per above, with a reduced cost applied for topotecan (based on latest available data from the NHS eMIT)
- **B4-3: Include subsequent therapies (no anti-PD-1/PD-L1/CTLA-4 + updated topotecan cost):** As per the above, with costs for patients recorded as receiving anti-PD-1/PD-L1/CTLA-4 assumed to be equal to the average cost incurred by the comparator arm (based on data from Study 100070-Obs001)

An additional option was incorporated within the model to allow the user to specify a custom proportion of patients and average cost per treated patient (results not provided within this document).

Table 19: B4: Cost-effectiveness results including subsequent therapy costs

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
CEA 3 (company submission base-case analysis)							
Chemotherapy	11,116	1.94	1.32				
Avelumab	96,758	5.88	3.60	85,642	3.95	2.27	37,670
B4-1: Include subsequent therapies (all costs, per study sources)							
Chemotherapy	11,374	1.94	1.32				
Avelumab	98,216	5.88	3.60	86,842	3.95	2.27	38,198
B4-2: Include subsequent therapies (all costs, per study sources + updated topotecan cost)							
Chemotherapy	11,249	1.94	1.32				
Avelumab	98,211	5.88	3.60	86,962	3.95	2.27	38,251
B4-3: Include subsequent therapies (no anti-PD-1/PD-L1/CTLA-4 + updated topotecan cost)							
Chemotherapy	11,249	1.94	1.32				
Avelumab	96,933	5.88	3.60	85,684	3.95	2.27	37,689

Key: BSC, best supportive care; CDF, Cancer Drugs Fund; CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Notes: * Please note that the total costs for avelumab change marginally, as the average duration of survival for the chemotherapy arm is used within the estimated costings of radiotherapy for all treatment arms.

Depending on the approach used to incorporate subsequent therapy costs, the ICER increases by between approximately £19 (B4-3) and £581 (B4-2).

Section C: Textual clarification and additional points

C1. Please explain the statement "....with n=50 patients still alive at 2 years" in the CS on page 13 when in Figure 1 there are only 20 patients at risk at 24 months.

This is a typographical error. The correct value is included within Figure 1 (n=20), and the statement should read "with n=20 patients still alive at 2 years"

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Single technology appraisal

Avelumab for treating metastatic Merkel cell carcinoma

(CDF review TA517) – [ID1617]

Additional clarification question responses

April 2020

File name	Version	Contains confidential information	Date
2020-04-02_ID1617 Avelumab MCC Additional ClarQ responses_v1-0	v1-0	Yes	02-April-2020

Tables

Figures

Follow-up questions concerning propensity score weighting (PSW) analyses

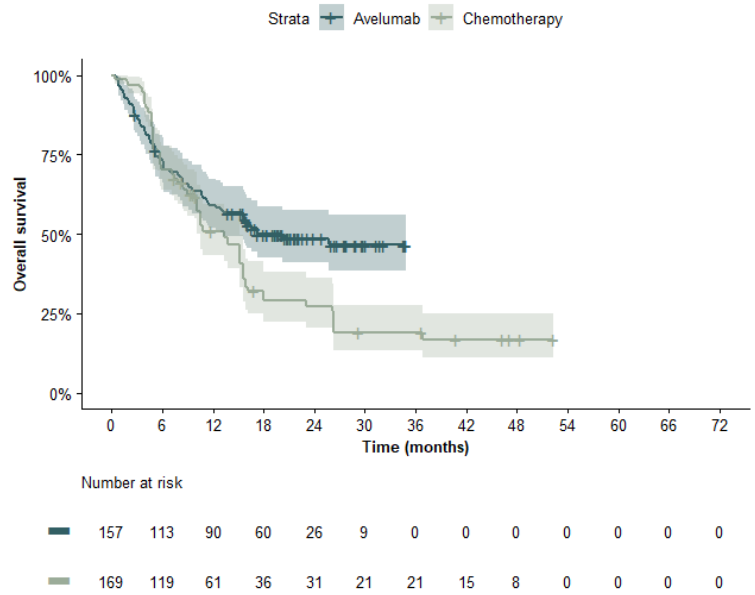
Kaplan-Meier curves

1) Kaplan-Meier plots for the outcome of OS with the matched populations for both JM200: Part B and Study 100070-001obs on the same figure for each of the PSW analyses (four figures in total).

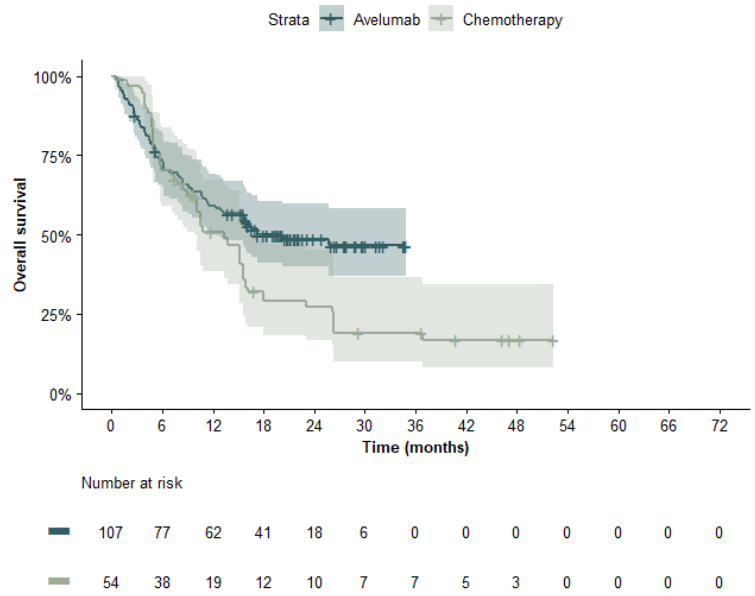
Please find the requested plots in Figure 1.

Figure 1: 1) Adjusted OS plots – JM200: Part B versus Study 100070-001

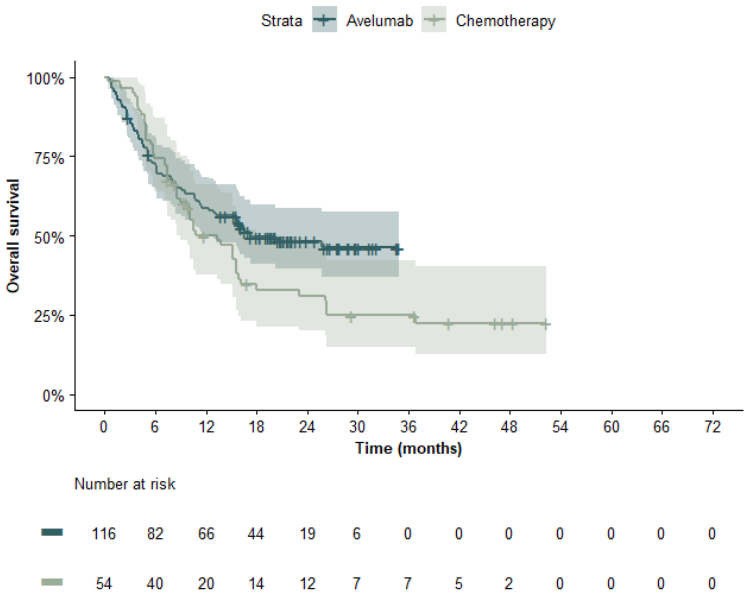
PSW 1



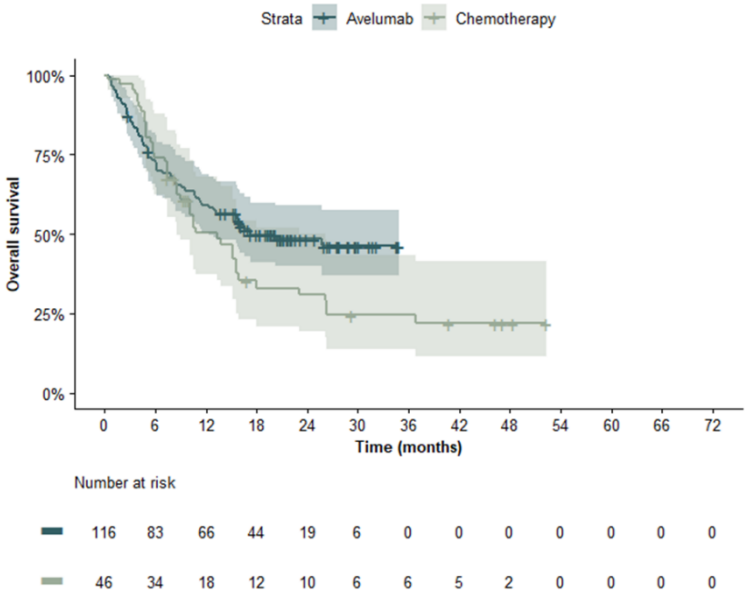
PSW 2



PSW 3


Strata Avelumab Chemotherapy
100%
75%
50%
25%
0%
Overall survival
0 6 12 18 24 30 36 42 48 54 60 66 72
Time (months)
Number at risk
116 82 66 44 19 6 0 0 0 0 0 0 0
54 40 20 14 12 7 7 5 2 0 0 0 0

PSW 4

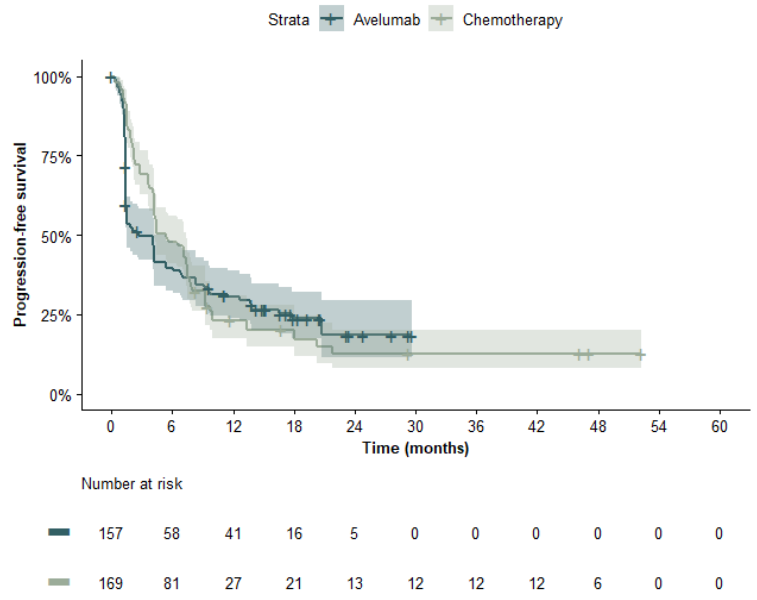

Strata Avelumab Chemotherapy
100%
75%
50%
25%
0%
Overall survival
0 6 12 18 24 30 36 42 48 54 60 66 72
Time (months)
Number at risk
116 83 66 44 19 6 0 0 0 0 0 0 0
46 34 18 12 10 6 6 5 2 0 0 0 0

2) Kaplan-Meier plots for the outcome of PFS with the matched populations for both JM200: Part B and Study 100070-001obs on the same figure for each of the PSW analyses (four figures in total).

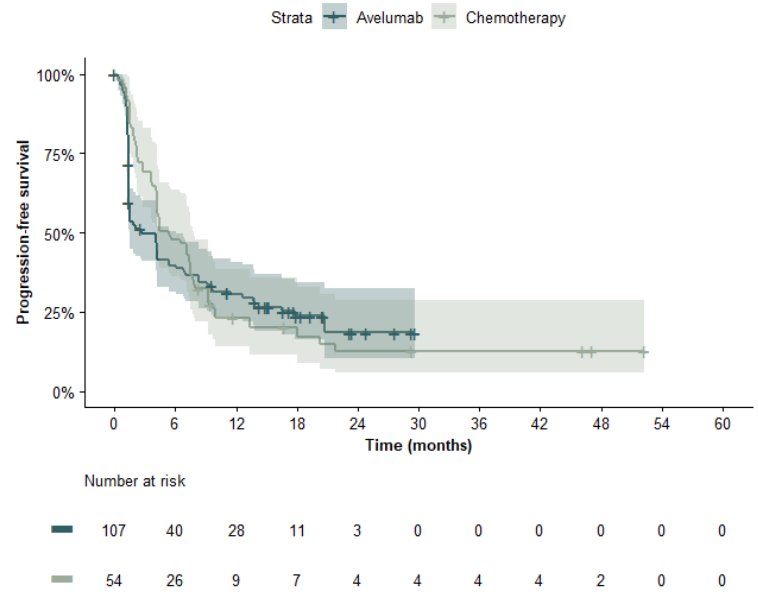
Please find the requested plots in Figure 2.

Figure 2: 2) Adjusted PFS plots – JM200: Part B versus Study 100070-001

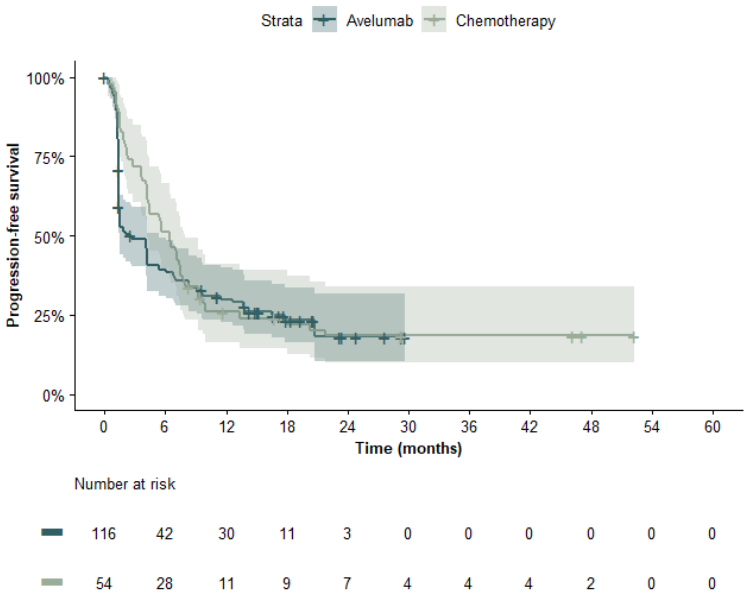
PSW 1



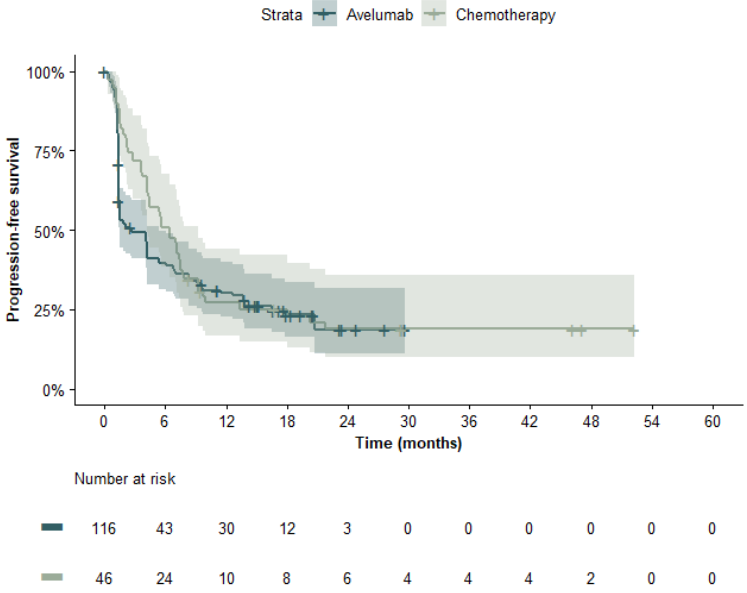
PSW 2



PSW 3


Strata
Avelumab
Chemotherapy
100%
75%
50%
25%
0%
Progression-free survival
0
6
12
18
24
30
36
42
48
54
60
Time (months)
Number at risk
116 42 30 11 3 0 0 0 0 0 0
54 28 11 9 7 4 4 4 2 0 0

PSW 4

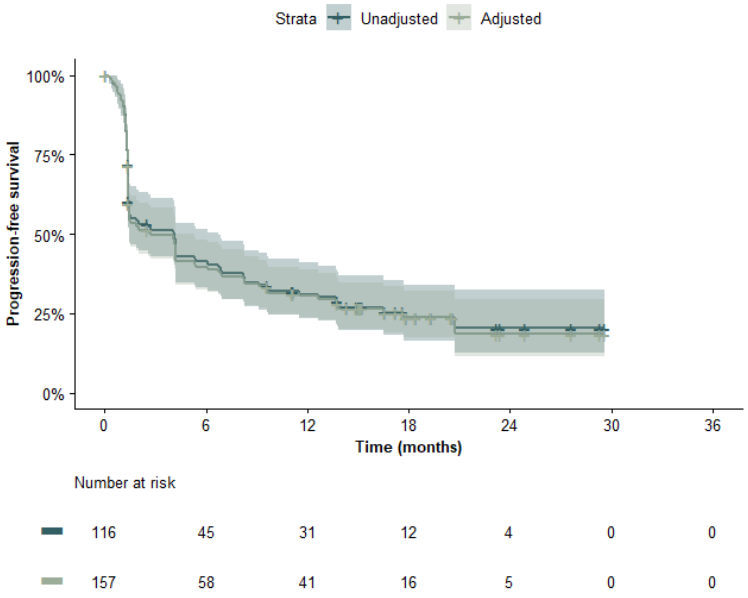

Strata
Avelumab
Chemotherapy
100%
75%
50%
25%
0%
Progression-free survival
0
6
12
18
24
30
36
42
48
54
60
Time (months)
Number at risk
116 43 30 12 3 0 0 0 0 0 0
46 24 10 8 6 4 4 4 2 0 0

3) Kaplan-Meier plots for the outcome of PFS with the matched and unmatched populations for each study (JM200: Part B and Study 100070-001obs) for each of the PSW analyses as already provided for OS (eight figures in total).

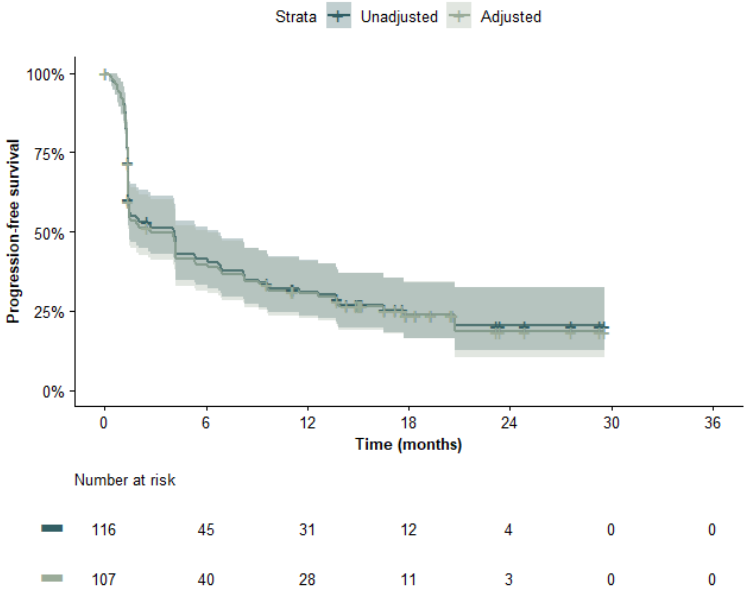
Please find the requested plots in Figure 3 (JM200: Part B) and Figure 4 (Study 100070-Obs001).

Figure 3: 3) Adjusted PFS plots – JM200: Part B

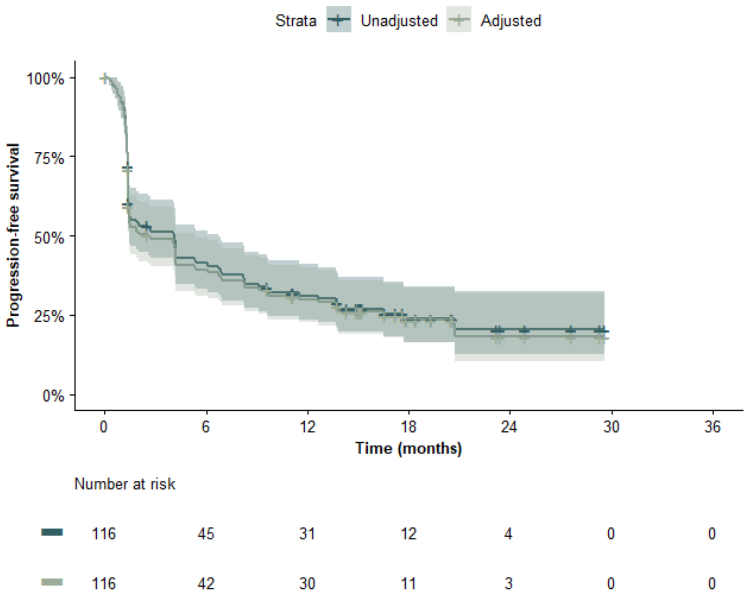
PSW 1



PSW 2



PSW 3


Strata Unadjusted Adjusted
100%
75%
50%
25%
0%
Progression-free survival
0
6
12
18
24
30
36
Time (months)
Number at risk
116 45 31 12 4 0 0
116 42 30 11 3 0 0

PSW 4

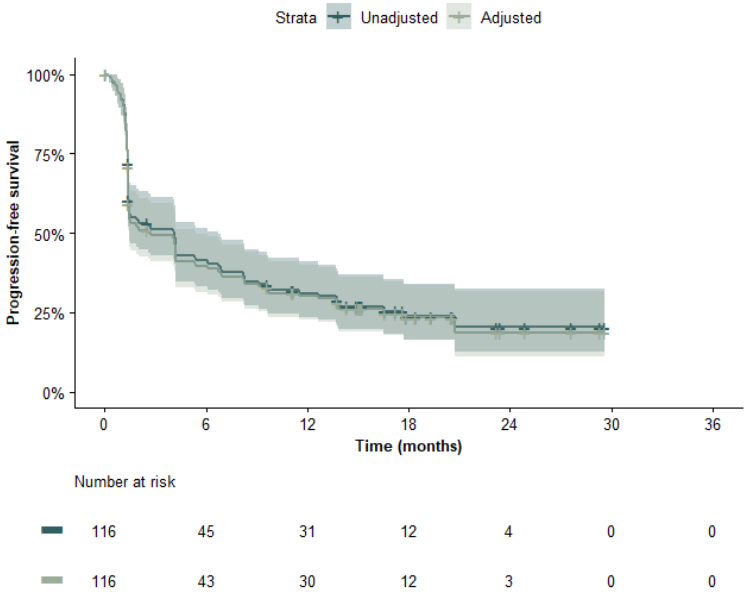
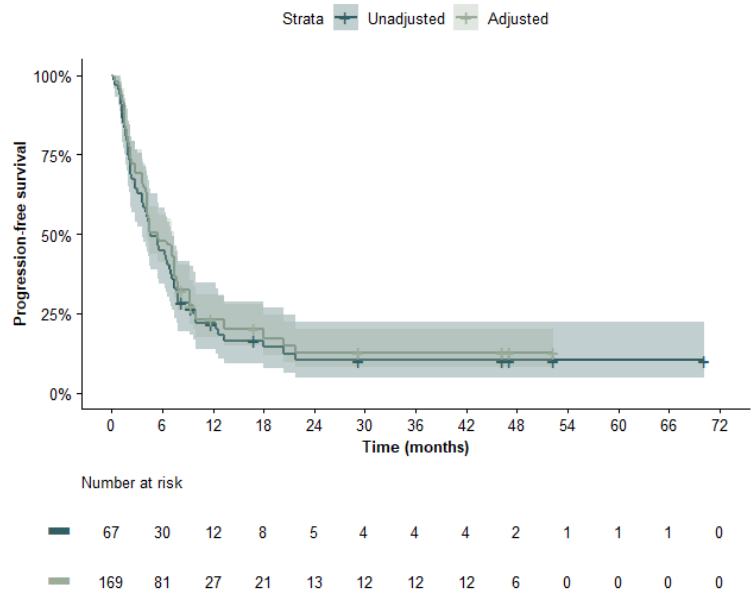
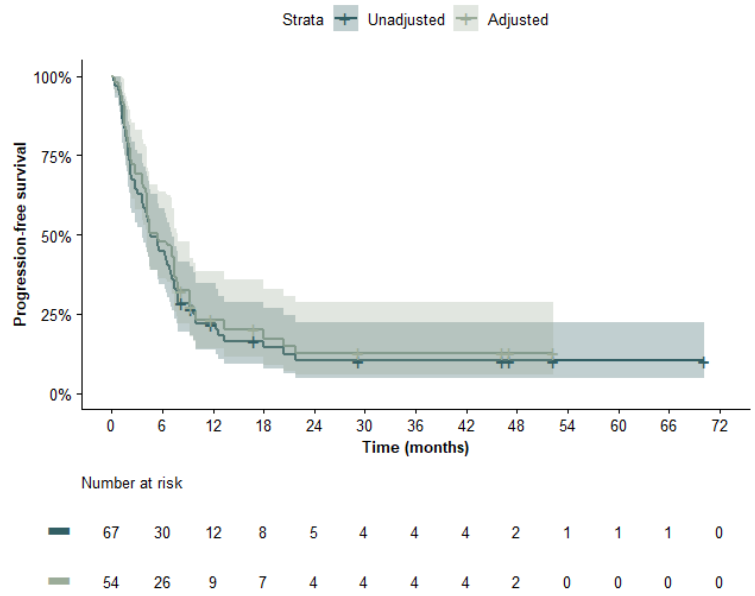

Strata Unadjusted Adjusted
100%
75%
50%
25%
0%
Progression-free survival
0
6
12
18
24
30
36
Time (months)
Number at risk
116 45 31 12 4 0 0
116 43 30 12 3 0 0

Figure 4: 3): Adjusted PFS plots – Study 100070-Obs001

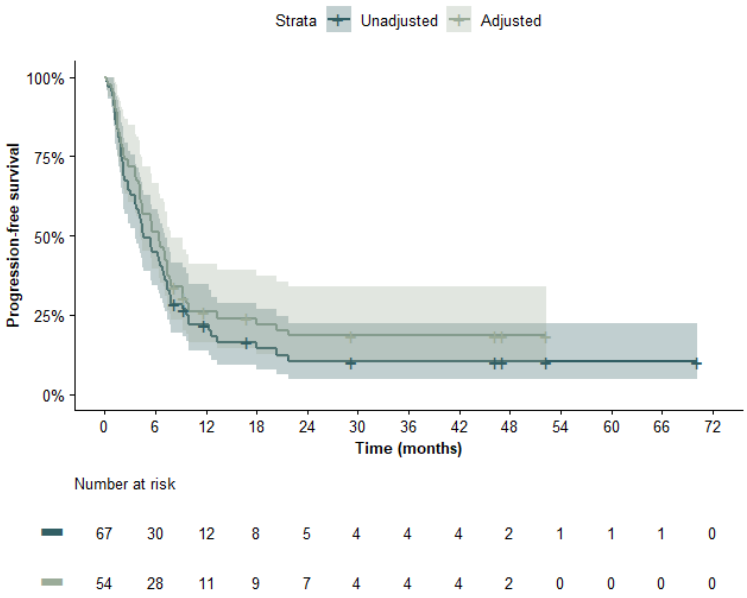
PSW 1



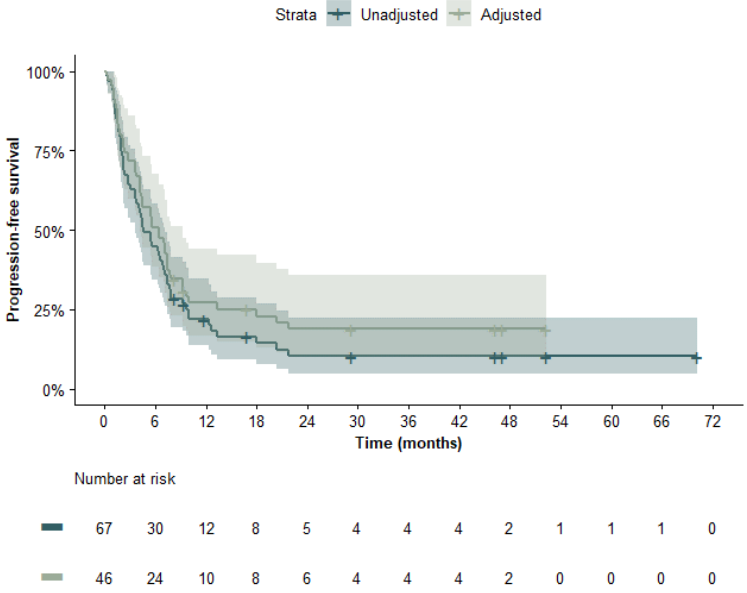
PSW 2



PSW 3


Strata Unadjusted Adjusted
100%
75%
50%
25%
0%
Progression-free survival
0 6 12 18 24 30 36 42 48 54 60 66 72
Time (months)
Number at risk
67 30 12 8 5 4 4 4 2 1 1 1 0
54 28 11 9 7 4 4 4 2 0 0 0 0

PSW 4


Strata Unadjusted Adjusted
100%
75%
50%
25%
0%
Progression-free survival
0 6 12 18 24 30 36 42 48 54 60 66 72
Time (months)
Number at risk
67 30 12 8 5 4 4 4 2 1 1 1 0
46 24 10 8 6 4 4 4 2 0 0 0 0

Baseline patient characteristics

4) Details of the patient characteristics for each study for each of the four PSW analyses.

Please see the patient characteristics included within the PSW analyses for each study arm in Table 1 (JM200: Part B) Table 2 (Study 100070-Obs001).

Table 1: 4) Baseline characteristics for JM200: Part B, re-weighted according to the PSW analyses conducted

Baseline characteristic		Unweighted		PSW 1		PSW 2		PSW 3		PSW 4	
		n	%	n	%	n	%	n	%	n	%
Age	<75	59.0	50.9	76.1	48.4	51.9	48.4	58.0	50.0	58.7	50.6
	>=75	57.0	49.1	81.0	51.6	55.3	51.6	58.0	50.0	57.3	49.4
Sex	Male	81.0	69.8	115.1	73.2	78.5	73.2	83.8	72.2	83.0	71.6
	Female	35.0	30.2	42.0	26.8	28.7	26.8	32.2	27.8	33.0	28.4
ECOG PS	0	72.0	62.1	83.0	52.8	56.6	52.8	58.7	50.6	61.6	53.1
	1	44.0	37.9	74.1	47.2	50.6	47.2	57.2	49.4	54.3	46.9
Immunocompetent	Yes	116.0	100.0	157.1	100.0	107.2	100.0	115.9	100.0	115.9	100.0

Table 2: 4) Baseline characteristics for Study 100070-Obs001, re-weighted according to the PSW analyses conducted

Baseline characteristic		Unweighted		PSW 1		PSW 2		PSW 3		PSW 4	
		n	%	n	%	n	%	n	%	n	%
Age	<75	32.0	47.8	73.7	43.6	23.4	43.6	25.5	47.0	21.8	47.1
	>=75	35.0	52.2	95.3	56.4	30.3	56.4	28.8	53.0	24.4	52.9
Sex	Male	53.0	79.1	121.5	71.9	38.6	71.9	38.7	71.4	32.9	71.2
	Female	14.0	20.9	47.5	28.1	15.1	28.1	15.5	28.6	13.3	28.8
ECOG PS	0	14.0	20.9	85.5	50.6	27.1	50.6	27.6	50.9	24.6	53.4
	1	32.0	47.8	70.0	41.4	22.2	41.4	21.5	39.6	21.5	46.6
	2+	8.0	11.9	13.5	8.0	4.3	8.0	5.2	9.5	0.0	0.0
	Missing	13.0	19.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Immunocompetent	Yes	51.0	76.1	156.0	92.3	49.5	92.3	40.9	75.5	35.8	77.6
	No	16.0	23.9	13.0	7.7	4.1	7.7	13.3	24.5	10.3	22.4

NICE National Institute for Health and Care Excellence

Patient organisation submission

Avelumab for treating metastatic Merkel cell carcinoma (CDF review of TA517) [ID1617]

Thank you for agreeing to give us your organisation's views on this technology and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

To help you give your views, please use this questionnaire with our guide for patient submissions.

You do not have to answer every question – they are prompts to guide you. The text boxes will expand as you type.

Information on completing this submission

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you intend to include journal articles in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 10 pages.

About you	
1.Your name	[redacted]

2. Name of organisation	Neuroendocrine Cancer UK (formerly NET Patient Foundation)
3. Job title or position	Patient support & information nurse specialist
4a. Brief description of the organisation (including who funds it). How many members does it have?	Neuroendocrine Cancer UK is a registered charity, established for the advocacy, education/ information and support of those affected by a Neuroendocrine Cancer diagnosis. We provide research and evidence based information resources for patients, families, and carers – as well as over 100 NHS Hospitals – including disease relevant medical and scientific organisations. We are donation dependent – details and further information can be found on the Charity Commission website – our registration number is 1092386. We have recently changed our name (from NET Patient Foundation) to more accurately reflect the cancer community we exist to support. We have 5 staff (= to 4WTE) and over 6000 members (patients, family members and healthcare professionals). The annual incidence of neuroendocrine cancer is rising – this is a global finding, not just within the UK, and may reflect better awareness and diagnostics – though cause and linked factors have yet to be identified for all cohorts within this heterogenous, multi site population. Merkel Cell Carcinoma has been linked to a specific virus as well as associated factors making it somewhat unique within Neuroendocrine malignancies (in that causal links have been identified) and a targeted immunotherapy Avelumab has been developed, which appears to offer some hope in treating this rare, highly aggressive disease.. The rarity of incidence (<6% of all Neuroendocrine Cancers) provides a challenge for both those diagnosed and those looking to provide effective treatments.
4b. Has the organisation received any funding from the manufacturer(s) of the technology and/or comparator	No

products in the last 12 months? [Relevant manufacturers are listed in the appraisal matrix.] If so, please state the name of manufacturer, amount, and purpose of funding.	
4c. Do you have any direct or indirect links with, or funding from, the tobacco industry?	No
5. How did you gather information about the experiences of patients and carers to include in your submission?	From face to face and online forums From QoL and trial publications/abstracts

Living with the condition	
6. What is it like to live with the condition? What do carers experience when caring for someone with the condition?	Merkel Cell Carcinoma is frightening – it is not just the impact of a cancer diagnosis but the visibility, potential disfigurement and observable rapid rate of change that can have both physical and psychological impact on those diagnosed. With standard treatment –hope is an emotion hard to sustain - chemotherapy is rarely effective and relapse rates are high – with little sustained response. Uncertainty abounds due to rarity of diagnosis, which not only limits clinical data and research – but also gives rise to the fear that lack of reaching target level of information will impact on decision-making about availability and accessibility of future treatment and options. Without Avelumab patients are faced with a decision to choose between chemotherapy or “doing nothing” This can lead to additional negative impact on those diagnosed as such a decision can lead to conflict within their support network – be it family +/- friends : despite poor impact of chemotherapy on mMCC – not proceeding with it, in the absence of alternatives, can be viewed as “giving up”. The flipside of this being the distress family members report at seeing their loved ones going through a “toxic therapy” for little benefit, because their loved one believes “anything is better than nothing” . . .despair is mentioned regularly. Family and friends describe hopelessness – as expressed in other cancer cohorts –but this is compounded by rarity – with limited accessible accurate and reliable information, expertise and support. . . “I’m already grieving, I try so hard not to as it is affecting our relationship . . .but I can see him . . can see it, taking him . . .”

Current treatment of the condition in the NHS	
7. What do patients or carers think of current treatments and care available on the NHS?	As per previous submission – treatments that were in place prior to Avelumab offered little hope – particularly in the metastatic group. Though numbers remain low on those treated with Avelumab it has been positively reported on by patients (and families) : My diagnosis was met with referral to the local hospice. However, I sought expert review and referral and started Immunotherapy (Avelumab). After 6 treatments my disease has shrunk to half the size and there has been no further progression of the cancer. . . the hospital has told me that fewer than 5 in this area are receiving this treatment, but all are showing similar responses . . .to think I could be in a hospice now.. .or . . . Concerns are raised about availability – many recognise this treatment may not offer complete cure – but it "offers hope" . Data published since last review shows that within the limited numbers available – responses are better maintained and have a positive impact on health related quality of life, when compared with available evidence on historic therapies such as chemotherapy or BSC (best supportive care).
8. Is there an unmet need for patients with this condition?	Yes – in the interim since last review there have been no comparator therapies developed that offer the outcomes seen in Avelumab. As data matures there is evidence of continued benefit from this therapy – supporting the step-change view of this therapy and its impact for those with mMCC
Advantages of the technology	
9. What do patients or carers think are the advantages of the technology?	Sustained response – visible effect and improvement in overall well-being.

Disadvantages of the technology	
10. What do patients or carers think are the disadvantages of the technology?	Uncertainty over sustained availability and positive impact of treatment. “Having regained hope its sometimes like waiting for the other shoe to drop . . .” Accessibility and luck – many view their access to Avelumab as a result of self-advocating for expert review or ‘luck” in that local service has an expert.
Patient population	
11. Are there any groups of patients who might benefit more or less from the technology than others? If so, please describe them and explain why.	Subsequent studies have not revealed priority groups : whilst we believe a trend has been commented on in subsequent abstracts with regards to PD-L1 status – the same documents do not report a significant impact of this on overall survival

Equality	
12. Are there any potential equality issues that should be taken into account when considering this condition and the technology?	In rare cancers there remains a risk that measures used to assess evidence can determine weight allocated to it – in that small number populations may not have the equivalent numbers and protocols as those of higher number. This factor needs to be taken into account to ensure patients are not discriminated against due to limits in incidence and therefore eligibility for trial inclusion and / or treatment. We need robust evidence – and alternatives to RCTs as a measure of value of evidence - need to be explored. Or HST adapted to fit needs of rare cancers. But this may be a policy/processes issue rather than equality.
Other issues	
13. Are there any other issues that you would like the committee to consider?	Assessment of QoL impact including tolerability and PFS is ongoing – this is to be encouraged to continue to provide meaningful data to inform decision-making. There have also been several papers published looking at cost-effectiveness of therapy and outcomes (patterns of survival) Bullement et al (2019), DAngelo et al (2019), Steuten et al (2019), Lantis et al (2019) amongst others
Key messages	
14. In up to 5 bullet points, please summarise the key messages of your submission: • Unmet need	

- Sustained response seen
- Safe, effective and durable
- Positive impact on QoL for both patients and carers (families)
-

Thank you for your time.

Please log in to your NICE Docs account to upload your completed submission.

……

Your privacy

The information that you provide on this form will be used to contact you about the topic above.

☐ Please tick this box if you would like to receive information about other NICE topics.

For more information about how we process your personal data please see our privacy notice.

……

Public Health England

Protecting and improving the nation's health

Avelumab for treating metastatic Merkel cell carcinoma

Data review

Commissioned by NHS England and NHS Improvement

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing and reduce health inequalities. We do this through world-leading science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health and Social Care, and a distinct delivery organisation with operational autonomy. We provide government, local government, the NHS, Parliament, industry and the public with evidence-based professional, scientific and delivery expertise and support.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Published March 2020
PHE publications
gateway number: GW-1050

PHE supports the UN
Sustainable Development Goals

Contents

Executive summary

Introduction

The National Institute for Health and Care Excellence (NICE) appraised the clinical and cost effectiveness of avelumab for the treatment of patients diagnosed with Merkel cell carcinoma. The appraisal committee highlighted clinical uncertainty around estimates of treatment duration and overall survival (OS) in the evidence submission. As a result, they recommended commissioning avelumab through the Cancer Drugs Fund (CDF), to allow a period of managed access, supported by additional data collection, to answer the clinical uncertainty.

NHS England and NHS Improvement commissioned Public Health England (PHE) to evaluate the real-world treatment effectiveness of avelumab, in the CDF population, during the managed access period. This report presents the results of the use of avelumab, in clinical practice, using the routinely collected Systemic Anti-Cancer Therapy (SACT) dataset.

This report, and the data presented, demonstrate the potential within the English health system to collect real-world data to inform decision-making about patient access to cancer treatments via the CDF. The opportunity to collect real-world data enables patients to get access to promising new treatments much earlier than might otherwise be the case, whilst further evidence is collected to address clinical uncertainty.

The NHS England and NHS Improvement and PHE partnership for collecting and following up real-world SACT data in the CDF in England has resulted in analysis of data for the full patient population, with 100% of patients and 100% of patient outcomes reported in the SACT dataset. PHE and NHS England and NHS Improvement are committed to providing world first high-quality real-world data on CDF cancer treatments to be appraised alongside the outcome data from the relevant clinical trials.

Methods

NHS England and NHS Improvements Blueteq® system was used to provide a reference list of all patients with an application for avelumab for Merkel cell carcinoma in the CDF. Patient NHS numbers were used to link Blueteq applications to PHE's routinely collected SACT data to provide SACT treatment history.

Between 1 March 2018 and 31 May 2019, 58 applications for avelumab were identified in the Blueteq system. Following appropriate exclusions (see Figures 1 and 2),

52 unique patients who received treatment were included in these analyses. All patients were traced to obtain their vital status using the personal demographics service (PDS)[1].

Results

All 52 (100%) unique patients with CDF applications were reported in the SACT dataset.

The median treatment duration for the analysis cohort was 6.0 months (182 days) [95% CI: 2.1, 9.8]. 46% [95% CI: 31%,60%] of patients were receiving treatment at 6 months and 31% [95% CI: 15%, 48%] of patients were receiving treatment at 12 months.

At data cut off, 88% (N=30) of patients were identified as no longer being on treatment. Of these, 100% (N=30) of patients had an outcome submitted by the treating trust to the SACT dataset that detailed the reason why a patient ended their treatment. 47% (N=14) of patients stopped treatment due to progression, 13% (N=4) of patients stopped treatment due to acute toxicity, 3% (N=1) of patients completed treatment as prescribed, 7% (N=2) of patients died on treatment and 30% (N=9) of patients died not on treatment.

The median overall survival was 11.8 months (358 days), confidence intervals could not be produced as insufficient events had occurred at the time of this report being produced. OS at 6 months was 58% [95% CI: 43%, 70%], survival at 12 months was 50% [95% CI: 33%, 64%].

Conclusion

This report analyses SACT real world data for patients treated with avelumab for the treatment of patients diagnosed with Merkel cell carcinoma in the CDF. It evaluates treatment duration, overall survival and treatment outcomes for all patients treated with avelumab for this indication.

Introduction

Avelumab is available through the Cancer Drugs Fund as a treatment option for metastatic Merkel cell carcinoma in adults who are chemotherapy naïve. Avelumab is available on the NHS, through routine commissioning, for people who have already received at least one round of chemotherapy.

Avelumab is recommended for use within the Cancer Drugs Fund, only if:

- they have not had chemotherapy for metastatic disease
- the conditions in the managed access agreement for avelumab are followed

Background to this report

The Public Health England and NHS England and NHS Improvement partnership on cancer data – using routinely collected data to support effective patient care

High quality and timely cancer data underpin NHS England NHS Improvement and Public Health England's (PHE's) ambitions of monitoring cancer care and outcomes across the patient pathway. The objective of the PHE and NHS England and NHS Improvement partnership on cancer data is to address mutually beneficial questions using Systemic Anti-Cancer Therapy (SACT) data collected by PHE. This includes NHS England and NHS Improvement commissioning PHE to produce routine outcome reports on patients receiving treatments funded through the Cancer Drugs Fund (CDF) during a period of managed access.

The CDF is a source of funding for cancer drugs in England[4]. From 29 July 2016 NHS England implemented a new approach to the appraisal of drugs funded by the CDF. The new CDF operates as a managed access scheme that provides patients with earlier access to new and promising treatments where there is uncertainty as to their clinical and cost effectiveness. During this period of managed access, ongoing data collection is used to answer the uncertainties raised by the NICE committee and inform drug reappraisal at the end of the CDF funding period[5].

PHE analyse data derived from patient-level information collected in the NHS, as part of the care and support of cancer patients. The data is collated, maintained, quality-assured and analysed by the National Cancer Registration and Analysis Service, which is part of PHE.

NICE Appraisal Committee appraisal of avelumab for metastatic Merkel cell carcinoma [TA517]

The NICE Appraisal Committee reviewed the evidence for the clinical and cost effectiveness of avelumab for treating metastatic Merkel cell carcinoma [TA517] and published guidance for this indication in April 2018[6]. Avelumab was recommended as an option for treating metastatic Merkel cell carcinoma in adults, only if they have had 1 or more lines of chemotherapy for metastatic disease.

Due to the clinical uncertainties identified by the committee and outlined below, the committee recommended commissioning of avelumab for those who have not had chemotherapy for metastatic disease through the CDF for a period of 19 months, March 2018 to October 2019.

During the CDF funding period, results from the ongoing clinical trials evaluating avelumab for those who have not had chemotherapy for metastatic disease are likely to answer the main

clinical uncertainties raised by the NICE committee. The ongoing trial that will support the evaluation of avelumab is JAVELIN 200[7]. Data collected from the JAVELIN 200 clinical trial will be the primary source of data collection.

Analysis of the SACT dataset will provide information on real-world treatment patterns and outcomes for avelumab for untreated metastatic Merkel cell carcinoma in England, during the CDF funding period. This will act as a secondary source of information alongside the results of the JAVELIN 200 clinical trial[7].

The key areas of uncertainty identified by the committee for reappraisal at the end of the CDF data collection are as follows:

- treatment duration for the use of avelumab for metastatic Merkel cell carcinoma
- overall survival from the start of a patient's first treatment with avelumab

Approach

Upon entry to the CDF, representatives from NHS England and NHS Improvement, NICE, PHE and the company (Merck Serono) formed a working group to agree the Data Collection Agreement (DCA)[6]. The DCA set out the real-world data to be collected and analysed to support the NICE reappraisal of avelumab. It also detailed the eligibility criteria for patient access to avelumab through the CDF and CDF entry and exit dates.

This report includes patients with approved CDF applications (via Blueteq®) for avelumab, followed up in the SACT dataset collected by PHE.

Methods

CDF applications - identification of the cohorts of interest

NHS England and NHS Improvement collects applications for CDF treatments through their online prior approval system (Blueteq®). The Blueteq application form captures essential baseline demographic and clinical characteristics of patients, needed for CDF evaluation purposes. Where appropriate, Blueteq data are included in this report.

Consultants must complete a Blueteq application form for every patient receiving CDF funded treatment. As part of the application form, consultants must confirm that a patient satisfies all clinical eligibility criteria to commence treatment PHE has access to the Blueteq database and key data items such as NHS numbers, primary diagnosis and drug information of all patients with an approved CDF application (which therefore met the treatment eligibility criteria).

The lawfulness of this processing is covered under Article 6(1)(e) of the EU General Data Protection Regulations (GDPR) (processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller). The processing of special categories of personal data is also covered under article 9(2)(h) of EU GDPR (processing is necessary for the purposes of preventive or occupational medicine). As NHS E & I do not have an exemption to the Common Law Duty of Confidentiality, NHS E & I cannot access the identifiable data directly. PHE, through the National Cancer Registration and Analysis Service, has permission to process confidential patient information though Regulation 2 of The Health Service (Control of Patient Information) Regulations 2002.

PHE collates data on all SACT prescribed drugs by NHS organisations in England, irrespective of the funding mechanism. The Blueteq extract is therefore essential to identify the cohort of patients whose treatment was funded by the CDF.

Avelumab clinical treatment criteria

The criteria for patient access to avelumab are:

- confirmed histological or cytological diagnosis of Merkel cell carcinoma
- patient has metastatic disease
- patient is treatment naïve to any systemic anti-cancer therapy for Merkel cell carcinoma and in particular to any immune checkpoint blockade therapies
- patient has an ECOG performance status of either 0 or 1. A patient with a performance status of 2 is not eligible for avelumab

- if a patient has brain metastases, then these have been treated and are stable
- avelumab is to be used as monotherapy only
- avelumab is to be continued until loss of clinical benefit or unacceptable toxicity or patient choice to stop treatment. Patients with radiological disease progression not associated with significant clinical deterioration (defined by a patient meeting all 3 of the following conditions: no new or worsening symptoms and no change in performance status for greater than 2 weeks and no need for salvage therapy) can continue treatment
- a formal medical review as to whether treatment with avelumab should continue or not will be scheduled to occur at least by the end of the first 8 weeks of treatment
- treatment breaks of up to 12 weeks beyond the expected cycle length of avelumab are allowed but solely to allow immune toxicities to settle
- avelumab is to be otherwise used as set out in its Summary of Product Characteristics

CDF applications - deduplication criteria

Before conducting any analysis on CDF treatments, the Blueteq data is examined to identify duplicate applications. The following deduplication rules are applied.

If 2 trusts apply for avelumab for the treatment of metastatic Merkel cell carcinoma for the same patient (identified using the patient's NHS number), and both applications have the same approval date, then the record where the CDF trust (the trust applying for CDF treatment) matches the SACT treating trust is selected.

If 2 trusts apply for avelumab for the treatment of metastatic Merkel cell carcinoma for the same patient, and the application dates are different, then the record where the approval date in the CDF is closest to the regimen start date in SACT is selected, even if the CDF trust did not match the SACT treating trust.

If 2 applications are submitted for avelumab for the treatment of metastatic Merkel cell carcinoma and the patient has no regimen start date in SACT capturing when the specific drug was delivered, then the earliest application in the CDF is selected.

Initial CDF cohorts

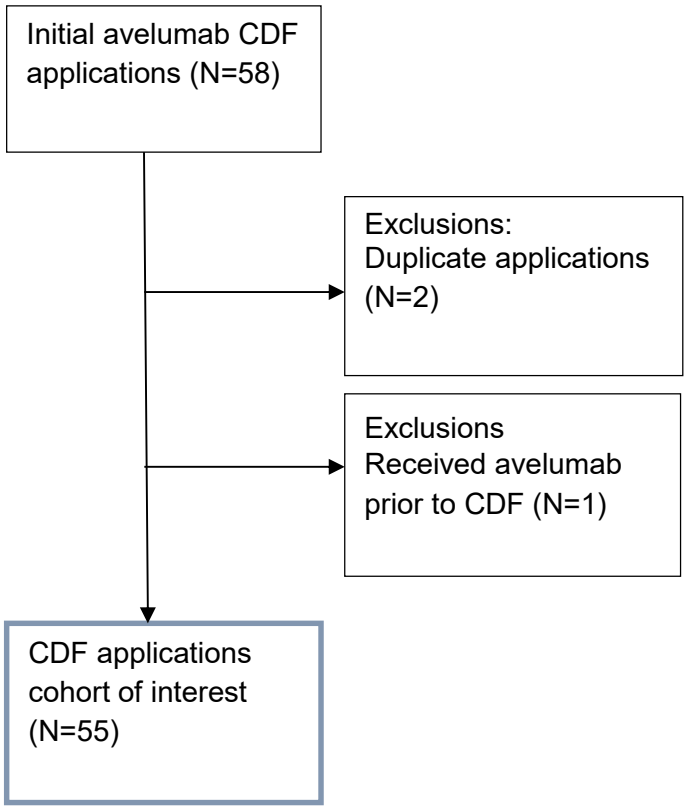
The analysis cohort is limited to the date avelumab entered the CDF for this indication onwards. Any treatments delivered before the CDF entry date are excluded as they are likely to be patients receiving treatment via an Early Access to Medicines Scheme (EAMS) or a compassionate access scheme run by the pharmaceutical company. These schemes may have different eligibility criteria compared to the clinical treatment criteria detailed in the CDF managed access agreement for this indication.

The CDF applications included in these analyses are from 1 March 2018 to 31 May 2019. A snapshot of SACT data was taken on 2 November 2019 and made available for analysis on the 11 November 2019. The snapshot includes SACT activity up to the 31 July 2019. Tracing the patients' vital status was carried out on 21 November 2019 using the personal demographics service (PDS)[1].

There were 58 applications for CDF funding for avelumab for metastatic Merkel cell carcinoma between 1 March 2018 and 31 May 2019 in the NHS England and NHS Improvement Blueteq database. Following deduplication this relates to 56 unique patients.

One patient was excluded from these analyses as they appeared to have received avelumab prior to the drug being available through the CDF.

Figure 1: Derivation of the cohort of interest from the initial CDF applications made for avelumab for metastatic Merkel cell carcinoma between 1 March 2018 and 31 May 2019



Linking CDF cohort to SACT

NHS numbers were used to link SACT records to CDF applications for avelumab in the Blueteq system. Information on treatments in SACT were examined to ensure the correct SACT treatment records were matched to the CDF application, this includes

information on treatment dates (regimen, cycle and administration dates) and primary diagnosis codes in SACT.

Addressing clinical uncertainties

Treatment duration

Treatment duration is calculated from the start of a patient's treatment to their last known treatment date in SACT.

Treatment start date is defined as the date the patient started their CDF treatment. This date is identified as the patient's earliest treatment date in the SACT dataset for the treatment of interest. Data items used to determine a patient's earliest treatment date are:

- start date of regimen – SACT data item #22
- start date of cycle – SACT data item #27
- administration date – SACT data item #34

The earliest of these dates is used as the treatment start date.

The same SACT data items (#22, #27, #34) [8] are used to identify a patient's final treatment date. The latest of these 3 dates is used as the patient's final treatment date.

Additional explanation of these dates is provided below:

Start date of regimen
A regimen defines the drugs used, their dosage and frequency of treatment. A regimen may contain many cycles. This date is generally only used if cycle or administration dates are missing.

Start date of cycle
A cycle is a period of time over which treatment is delivered. A cycle may contain several administrations of treatment, after each treatment administration, separated by an appropriate time delay. For example; a patient may be on a 3-week cycle with treatment being administered on the first and eighth day, but nothing on days 2 to 7 and days 9 to 20. The first day would be recorded as the "start day of cycle". The patient's next cycle would start on the 21st day.

Administration date
An administration is the date a patient is administered the treatment, which should coincide with when they receive treatment. Using the above example, the administrations for a single 3-week cycle would be on the first and eighth day. The next administration would be on the 21st day, which would be the start of their next cycle.

The interval between treatment start date and final treatment date is the patient's time on treatment.

All patients are then allocated a 'prescription length' which is a set number of days added to the final treatment date to allow for the fact that they are effectively still 'on treatment' between administrations. The prescription length should correspond to the typical interval between treatment administrations.

If a patient dies between administrations, then their censor date is their date of death and these patients are deemed to have died on treatment unless an outcome summary is submitted to the SACT database confirming that the patient ended treatment due to disease progression or toxicity before death.

Avelumab is administered intra-venously. As such, treatment is generally administered in a healthcare facility and healthcare professionals are able to confirm that treatment administration has taken place on a specified date. A duration of 13 days has been added to final treatment date for all patients, this represents the duration from a patient's last cycle to their next[8]. Avelumab is a 14-day cycle consisting of one administration.

Treatment duration is calculated for each patient as:

Treatment duration (days) = (final treatment date – treatment start date) + prescription length (days).

Once a patient's treatment duration has been calculated, the patient's treatment status is identified as one of the following:

No longer receiving treatment (event), if:

- the patient has died
- the outcome summary (SACT data item #41) detailing the reason for stopping treatment has been completed
- there is no further SACT records for the patient following a 3-month period

If none of the above apply, the patient is assumed to still be on treatment and is censored.

Overall survival (OS)

OS is calculated from the CDF treatment start date, not the date of a patient's cancer diagnosis. Survival from the treatment start date is calculated using the patient's earliest treatment date, as described above, and the patient's date of death or the date the patient was traced for their vital status.

All patients in the cohort of interest are submitted to the PDS to check their vital status (dead/alive). Patients are traced before any analysis takes place. The date of tracing is used as the date of follow-up (censoring) for patients who have not died.

OS is calculated for each patient as the interval between the earliest treatment date where a specific drug was given to the date of death or date of follow-up (censoring).

OS (days) = date of death (or follow-up) – treatment start date

The patient is flagged as either:

Dead (event): at the date of death recorded on the PDS.

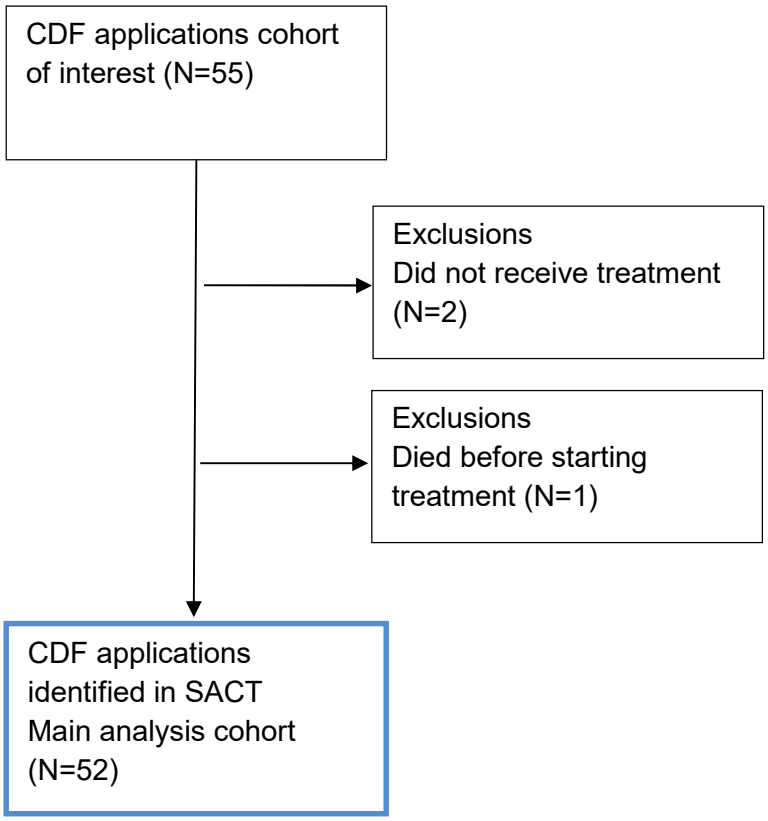
Alive (censored): at the date patients were traced for their vital status as patients are confirmed as alive on this date.

Results

Cohort of interest

Of the 55 new applications for CDF funding for avelumab for metastatic Merkel cell carcinoma, 2 patients did not receive treatment and one patient died before treatment[1] (see Figure 2).

Figure 2: Matched cohort - SACT data to CDF (Blueteq®) applications for avelumab for metastatic Merkel cell carcinoma between 1 March 2018 and 31 May 2019



A maximum of 52 avelumab records are expected in SACT for patients who were alive, eligible and confirmed to have commenced treatment (Figure 2). 100% (52/52) of these applicants for CDF funding have a treatment record in SACT.

[1] The 2 patients that did not receive treatment and one that died before treatment were confirmed with the relevant trusts by the PHE data liaison team.

[2] Figures may not sum to 100% due to rounding.

Completeness of SACT key variables

Table 1 presents the completeness of key data items required from SACT. Completeness is ≥88% for all key items and 100% for primary diagnosis, date of birth, gender, regimen and cycle dates.

Table 1: Completeness of key SACT data items for the avelumab cohort (N=52)

Variable	Completeness (%)
Primary diagnosis	100%
Date of birth (used to calculate age)	100%
Sex	100%
Start date of regimen	100%
Start date of cycle	100%
Administration date	100%
Performance status at start of regimen	88%

Table 2 presents the completeness of regimen outcome summary. A patient's outcome summary, detailing the reason why treatment was stopped, is only captured once a patient has completed their treatment. Therefore, the percentage completeness provided for outcome summary is for records where we assume treatment has stopped and an outcome is expected. Outcomes are expected if a patient has died, has an outcome in SACT stating why treatment has ended or has not received treatment with avelumab in at least 3 months. These criteria are designed to identify all cases where a patient is likely to have finished treatment. Based on these criteria, outcomes are expected for 30 patients. Of these, 30 have an outcome summary recorded in the SACT dataset 100% (30/30).

Table 2: Completeness of outcome summary for patients that have ended treatment (N=30)

Variable	Completeness (%)
Outcome summary of why treatment was stopped	100%

Patient characteristics

The median age of the 52 patients receiving avelumab for metastatic Merkel cell carcinoma was 75.5 years. The median age in males and females was 79.0 and 74.5 years respectively.

Table 3: Patient characteristics (N=52)

Patient characteristics[2]			
		Frequency (N)	**Percentage (%)**
Sex	Male	30	58%
	Female	22	42%
Age	<40	0	0%
	40-49	1	2%
	50-59	3	6%
	60-69	8	15%
	70-79	22	42%
	80+	18	35%
Performance status	0	7	13%
	1	34	65%
	2	4	8%
	3	1	2%
	4	0	0%
	Missing/unknown	6	12%

[2] Figures may not sum to 100% due to rounding.

Treatment duration

Of the 52 patients with CDF applications, 30 (58%) were identified as having completed treatment by 31 July 2019 (latest follow-up in SACT dataset). Patients are assumed to have completed treatment if they have died, have an outcome summary recorded in the SACT dataset or they have not received treatment with avelumab in at least 3 months (see Table 4). The median follow-up time in SACT was 2.7 months (82 days). The median follow-up time is the median observed time on treatment amongst all patients, including patients with ongoing and completed treatment.

Presently, 77% (N=108) of trusts submit their SACT return to the submission portal 2 months after the month's treatment activity has ended; this provides a maximum follow-up period of 16 months. 23% (N=32) of trusts submit their SACT return to the submission portal one month after the month's treatment activity has ended: this provides the maximum follow-up period of 17 months. SACT follow-up ends 31 July 2019.

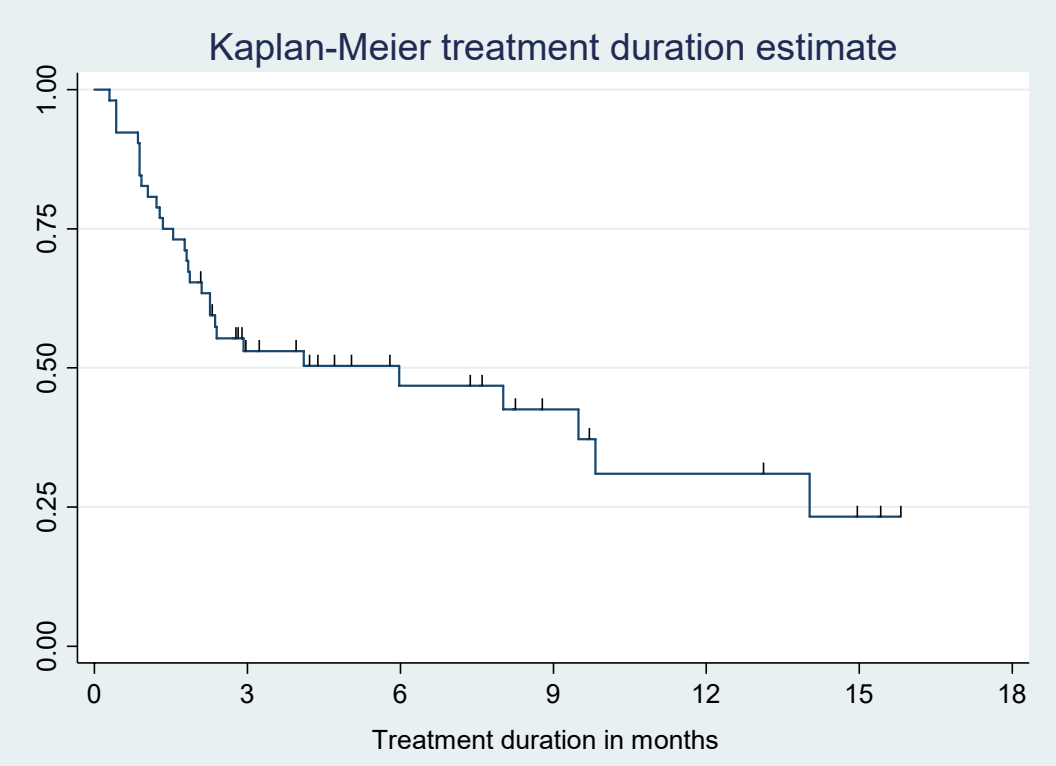
Table 4: Breakdown by patients' treatment status[3,4,5]

Patient status	Frequency (N)	Percentage (%)
Patient died – not on treatment	24	46%
Patient died – on treatment	2	4%
Treatment stopped	4	8%
Treatment ongoing	22	42%
Total	**52**	**100%**

[3] Figures may not sum to 100% due to rounding.
[4] Table 7 presents the outcome summary data reported by trusts. This includes patients from Table 4 that 'died on treatment', 'died not on treatment' and 'stopped treatment'.
[5] Deaths on treatment and deaths not on treatment are explained in the methodology paper available on the SACT website: www.chemodataset.nhs.uk/nhse_partnership

The Kaplan-Meier curve for ongoing treatment is shown in figure 3. The median treatment duration for all patients was 6.0 months (182 days) [95% CI: 2.1, 9.8] (N=52). 46% of patients were still receiving treatment at 6 months [95% CI: 31%, 60%], 31% of patients were still receiving treatment at 12 months [95% CI: 15%, 48%].

Figure 3: Kaplan-Meier treatment duration (N=52)



Tables 5 and 6 show the number of patients at risk, the number of patients that were censored and the number of patients that ended treatment (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for all patients for treatment duration was 16 months (486 days).

Table 5: Number of patients at risk, by quarterly breakpoints

Time intervals (months)	0-16	3-16	6-16	9-16	12-16	15-16
Number at risk	52	22	13	8	5	2

Table 6 shows that for all patients who received treatment, 22 were still on treatment (censored) at the date of follow-up and 30 had ended treatment (events).

Table 6: Number of patients at risk, by quarterly breakpoints split between patients that have ended treatment (events) and patients that are still on treatment (censored)

Time intervals (months)	0-16	3-16	6-16	9-16	12-16	15-16
Censored	22	16	9	5	4	2
Events	30	6	4	3	1	0

Table 7 gives a breakdown of a patient's treatment outcome recorded in SACT when a patient's treatment has come to an end. 58% (N=30) of patients had ended treatment at 31 July 2019.

Table 7: Treatment outcomes for patients that have ended treatment (N=30)[6,7]

Outcome	Frequency (N)	Percentage (%)
Stopped treatment – progression of disease	14	47%
Stopped treatment – acute chemotherapy toxicity	4	13%
Stopped treatment – treatment completed as prescribed	1	3%
Stopped treatment – died not on treatment[8]	9	30%
Stopped treatment – died on treatment	2	7%
Total	**30**	**100%**

[6] Figures may not sum to 100% due to rounding.

[7] Table 7 presents the outcome summary data reported by trusts. This includes patients from Table 4 that 'died on treatment', 'died not on treatment' and 'stopped treatment'.

[8] 'Deaths on treatment' and 'deaths not on treatment are explained in the methodology paper available on the SACT website: www.chemodataset.nhs.uk/nhse_partnership

Table 8: Treatment outcomes and treatment status for patients that have ended treatment (N=30)

Outcome[9]	Patient died [10] not on treatment	Treatment stopped	Patient died on treatment
Stopped treatment – progression of disease	13	1	
Stopped treatment – acute chemotherapy toxicity	2	2	
Stopped treatment – treatment completed as prescribed		1	
Stopped treatment – died not on treatment	9		
Stopped treatment – died on treatment			2
Total	**24**	**4**	**2**

[9] Relates to outcomes submitted by the trust in table 7.

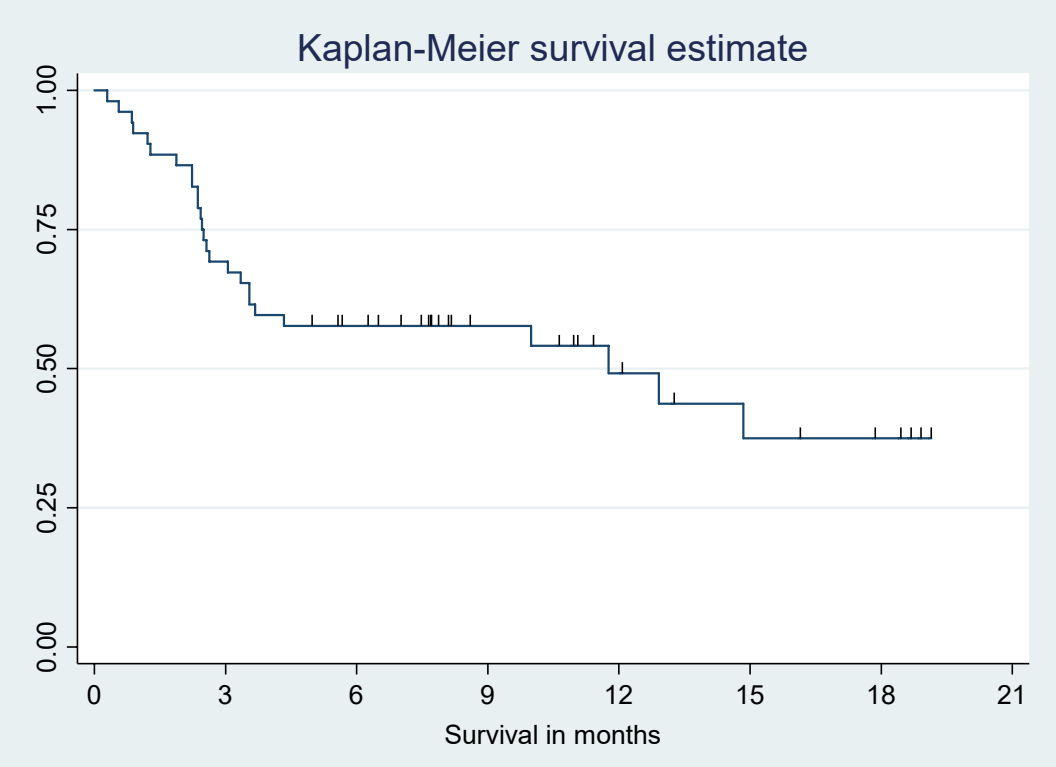
[10] Relates to treatment status in table 4 for those that have ended treatment.

Overall survival

Of the 52 patients with a treatment record in SACT, the minimum follow-up was 5 months (152 days) from the last CDF application to the date patients were traced for their vital status. Patients were traced for their vital status on 21 November 2019, this date was used as the follow-up date (censored date) if a patient is still alive. The median follow-up time was 6.3 months (191 days). The median follow-up is the patients' median observed time from the start of their treatment to death or censored date.

Figure 4 provides the Kaplan-Meier curve for overall survival, censored at 21 November 2019. The median survival for all patients was 11.8 months[11] (358 days). Survival at 6 months was 58% [95% CI: 43%, 70%], 12 months survival was 50% [95% CI: 33%, 64%].

Figure 4: Kaplan-Meier survival plot (N=52)



[11] Confidence intervals could not be produced as there was an insufficient number of events at the time this report was produced.

Table 9 and 10 show the number of patients at risk, the number of patients that were censored and the number of patients that died (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for survival was 20 months (608 days), all patients were traced on 21 November 2019.

Table 9: Includes the number of patients at risk, by quarterly breakpoints

Time intervals (months)	0-20	3-20	6-20	9-20	12-20	15-20	18-20
Number at risk	52	36	27	16	10	6	4

Table 10 shows that for all patients who received treatment, 26 were still alive (censored) at the date of follow-up and 26 had died (events).

Table 10: Number of patients at risk, those that have died (events) and those that are still alive (censored) by quarterly breakpoints

Time intervals (months)	0-21	3-21	6-21	9-21	12-21	15-21	18-21
Censored	26	26	23	12	8	6	4
Events	26	10	4	4	2	0	0

Sensitivity analyses

Cohort 1: 6-month SACT follow-up

Treatment duration

Sensitivity analyses was carried out on a cohort with at least 6 months follow-up in SACT. To identify the treatment duration cohort, CDF applications were limited from 1 March 2018 to 31 January 2019 and SACT activity was followed up to 31 July 2019. 30 patients (58%) were included in these analyses. The median follow-up time in SACT was 3.5 months (106 days).

The Kaplan-Meier curve for ongoing treatment is shown in figure 5. The median treatment duration for patients in this cohort was 2.9 months (88 days) [95% CI: 1.8, 9.8] (N=30). 43% of patients were still receiving treatment at 6 months [95% CI: 26%, 60%], 29% of patients were still receiving treatment at 12 months [95% CI: 13%, 47%].

Figure 5: Kaplan-Meier treatment duration (N=30)

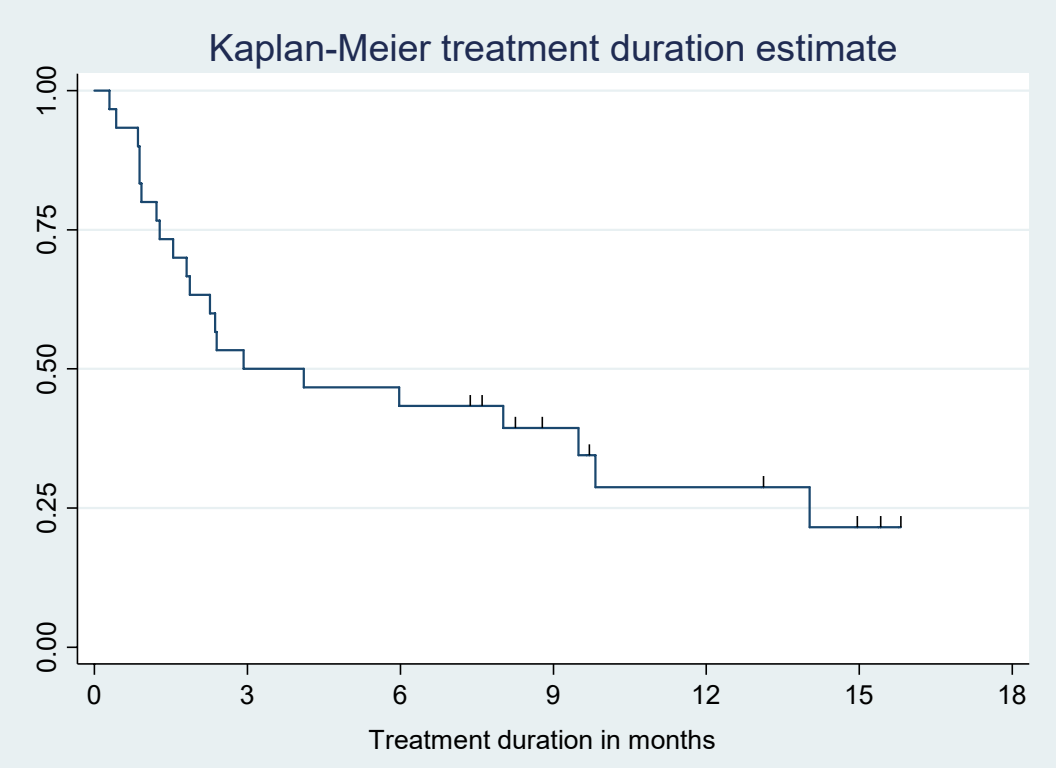


Table 11 and 12 show the number of patients at risk, the number of patients that were censored and the number of patients that ended treatment (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for treatment duration was 16 months (486 days). The minimum follow-up was 6 months.

Table 11: Number of patients at risk, by quarterly breakpoints

Time intervals (months)	0-16	3-16	6-16	9-16	12-16	15-16
Number at risk	30	15	13	8	5	2

Table 12 shows that for all patients who received treatment, 9 were still on treatment (censored) at the date of follow-up and 21 had ended treatment (events).

Table 12: Number of patients at risk, by quarterly breakpoints split between patients that have ended treatment (events) and patients that are still on treatment (censored)

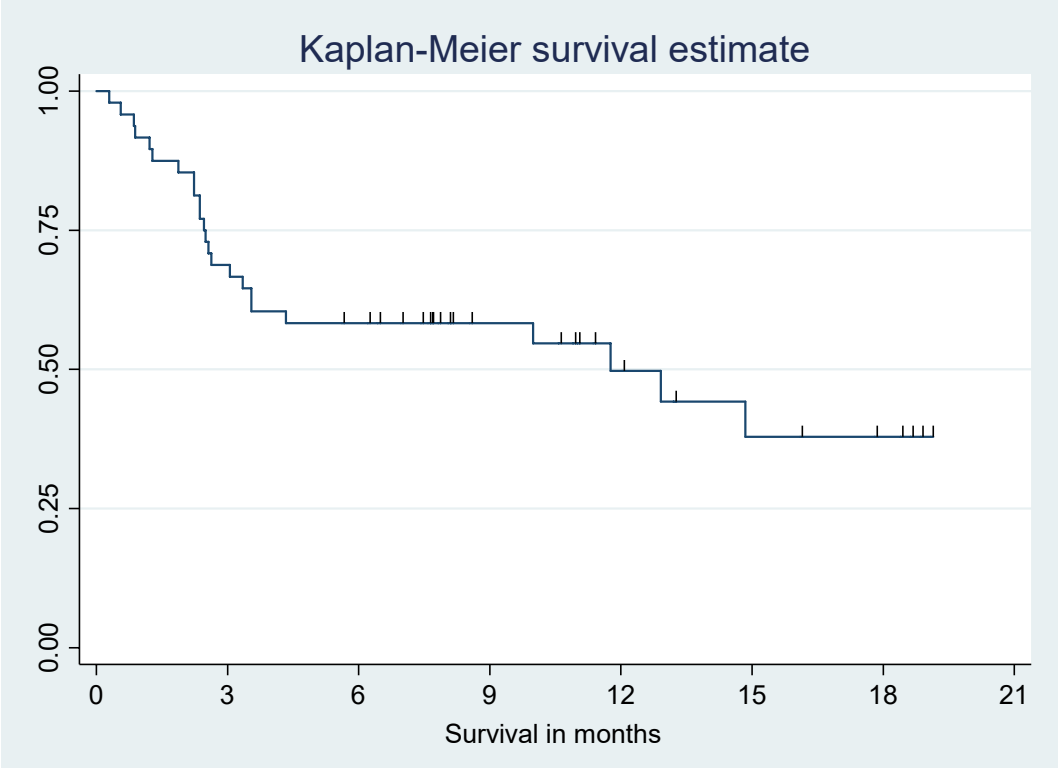
Time intervals (months)	0-16	3-16	6-16	9-16	12-16	15-16
Censored	9	9	9	5	4	2
Events	21	6	4	3	1	0

Overall survival

Sensitivity analyses was also carried out for OS on a cohort with at least 6 months follow-up in SACT. To identify the cohort, CDF applications were limited from 1 March 2018 to 21 May 2019. 48 patients (92%) were included in the survival analyses with all patients having a minimum follow-up of 6 months. Follow-up continued from treatment start date to date of tracing for vital status (21 November 2019). The median follow-up time was 7.2 months (219 days).

Figure 6 provides the Kaplan-Meier curve for overall survival, censored at 21 November 2019. The median survival for all patients was 11.8 months[12] (358 days). Survival at 6 months was 58% [95% CI: 43%, 71%], 12 months survival was 50% [95% CI: 33%, 65%].

Figure 6: Kaplan-Meier survival plot (N=48)



[12] Confidence intervals could not be produced as there was an insufficient number of events at the time this report was produced.

Table 13 and 14 show the number of patients at risk, the number of patients that were censored and the number of patients that died (events) from the time patients started treatment to the end of the follow-up period. The maximum follow-up period for survival was 20 months (608 days), all patients were traced on 21 November 2019.

Table 13: Includes the number of patients at risk, by quarterly breakpoints

Time intervals (months)	0-20	3-20	6-20	9-20	12-20	15-20	18-20
Number at risk	48	33	27	16	10	6	4

Table 14 shows that for all patients who received treatment, 24 were still alive (censored) at the date of follow-up and 24 had died (events).

Table 14: Number of patients at risk, those that have died (events) and those that are still alive (censored) by quarterly breakpoints

Time intervals (months)	0-20	3-20	6-20	9-20	12-20	15-20	18-20
Censored	24	24	23	12	8	6	4
Events	24	9	4	4	2	0	0

Table 15: Median treatment duration and overall survival, full cohort and sensitivity analysis

Metric	Standard analysis: Full cohort	Sensitivity analysis: 6 months follow-up cohort: treatment duration	Sensitivity analysis: 6 months follow-up cohort: OS
N	52	30	48
Median treatment duration	6.0 months (182 days) [95% CI: 2.1, 9.8]	2.9 months (88 days) [95% CI: 1.8, 9.8]	
Median OS	11.8 months[13] (358 days)		11.8 months (358 days)
Treatment at 6 months	46% [95% CI: 31%, 60%]	43% [95% CI: 26%, 60%]	
Treatment at 12 months	31% [95% CI: 15%, 48%]	29% [95% CI: 13%, 47%]	
OS at 6 months	58% [95% CI: 43%, 70%]		58% [95% CI: 43%, 71%]
OS at 12 months	50% [95% CI: 33%, 64%]		50% [95% CI: 33%, 65%]

[13] Confidence intervals could not be produced as there was an insufficient number of events at the time this report was produced.

Conclusions

52 patients received avelumab for the treatment of Merkel cell carcinoma [TA517] through the CDF in the reporting period (1 March 2018 and 31 May 2019). All patients were reported to the SACT dataset. An additional 3 patients with a CDF application, did not receive treatment or died before treatment. This was confirmed with the trusts responsible for those CDF applications by the team at PHE. All 52 patients receiving treatment in the approved indication were reported in the SACT dataset, giving a SACT ascertainment of 100%.

Patient characteristics from the SACT dataset show that proportionally more males received avelumab treatment compared to females (58% male, 42% female). Most of the cohort was aged 60+ (92%, N=48) and (79%, N=41) of patients had a performance status between 0 and 1 at the start of their regimen.

At the end of the data collection period, 88% (N=30) of patients were identified as no longer being on treatment. Of these, 100% (N=30) of patients had an outcome submitted by the treating trust to the SACT dataset which detailed the reason why a patient ended their treatment. 47% (N=14) of patients stopped treatment due to progression, 13% (N=4) of patients stopped treatment due to acute toxicity, 3% (N=1) of patients completed treatment as prescribed, 7% (N=2) of patients died on treatment and 30% (N=9) of patients died not on treatment.

The median treatment duration was 6.0 months (182 days) [95% CI: 2.1, 9.8]. The median follow-up was 83 days and the maximum follow-up was 16 months (486 days).

The median overall survival was 11.8 months (358 days) confidence intervals could not be produced as insufficient events had occurred at the time of this report being produced. The minimum follow-up was 5 months (152 days), the maximum follow-up was 20 months (608 days).

Sensitivity analyses were carried out to evaluate a cohort for which all patients had a minimum follow-up of 6 months. Results for this cohort showed a difference in treatment duration (full cohort = 6.0 months; sensitivity analysis cohort = 2.9 months). There was no difference in survival (full cohort = 11.8 months; sensitivity analysis cohort = 11.8 months) the difference in treatment duration and survival was not statistically significant.

References

1. The Personal Demographics Service (PDS) [Internet]. NHS Digital: 2019 [cited 2019 Dec]. Available from: https://digital.nhs.uk/demographics

2. Office for National Statistics. Cancer Registration Statistics, England: 2017. 2019 [cited 2019 Dec]. Available from: www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancerregistrationstatisticscancerregistrationstatisticsengland

3. National Institute for Health and Care Excellence: 2018 [cited 2019 Dec]. Available from: www.nice.org.uk/guidance/ta517/chapter/1-recommendations

4. Cancer Drugs Fund. [Internet]. NHS England and NHS Improvement: 2017 [cited 2019 Dec]. Available from: www.england.nhs.uk/cancer/cdf

5. Appraisal and funding of Cancer Drugs. NHS England and NHS Improvement: 2016 [cited 2019 Dec]. Available from: www.england.nhs.uk/wp-content/uploads/2013/04/cdf-sop.pdf

6. National Institute for Health and Care Excellence: 2018 [cited 2019 Dec]. Available from: www.nice.org.uk/guidance/ta517/resources

7. JAVELIN 200 clinical trial: 2019 [cited 2019 Dec]. Available from: https://clinicaltrials.gov/ct2/show/NCT02155647

8. CDF analytical methods. [Internet]. PHE: 2019 [cited 2019 Dec]. Available from: www.chemodataset.nhs.uk/nhse_partnership

9. Systemic Anti-Cancer Therapy [Internet]: SACT: 2019 [cited 2019 Dec]. Available from: www.chemodataset.nhs.uk/home/SACT

Avelumab for treating metastatic Merkel cell carcinoma (CDF review of TA517)

Cancer Drugs Fund Review

Source of funding

This report was commissioned by the National Institute for Health Research Systematic Reviews Programme as project number 16/134/09T.

Title:	Avelumab for treating metastatic Merkel cell carcinoma (CDF review of TA517)
Produced by:	BMJ Technology Assessment Group (BMJ TAG)
Authors:	
	Steve Edwards, Director of Health Technology Assessment, BMJ-TAG, London
	Victoria Wakefield, Principal Health Technology Assessment Analyst, BMJ-TAG, London
	Peter Cain, Senior Health Economist, BMJ-TAG, London
	Samantha Barton, Principal Health Technology Assessment Analyst, BMJ-TAG,
Correspondence to:	Steve Edwards, BMJ TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JR.
Date completed:	07/04/2020
Source of funding:	This report was commissioned by the National Institute for Health Research Systematic Reviews Programme as project number 16/134/09T.
Declared competing interests of the authors	No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.
Acknowledgments:	The ERG would like to thank Dr Ankit Rao, Consultant Medical Oncologist, Nottingham University Hospitals NHS Trust and Dr Kent Yip, Consultant in Clinical Oncology, Ipswich Hospital for providing clinical advice throughout the project, and for providing feedback on the clinical sections of the report.
Rider on responsibility for report:	The views expressed in this report are those of the authors and not necessarily those of the NIHR Systematic Reviews Programme. Any errors are the responsibility of the authors.
Report reference:	Edwards SJ, Wakefield V, Cain P and Barton S. Avelumab for treating metastatic Merkel cell carcinoma (CDF review of TA517): A Cancer Drugs Fund Review. BMJ Technology Assessment Group, 2020.

Contribution of authors:

Steve Edwards	Critical appraisal of the company's submission; provided feedback on all versions of the report. Guarantor of the report
Victoria Wakefield	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and drafted the summary, background and clinical results sections
Samantha Barton	Critical appraisal of the company's submission; critical appraisal of the clinical evidence; cross checking of company's search strategies; and assisted with drafting the clinical results sections
Peter Cain	Critical appraisal of the company's submission; critical appraisal of the economic model; cross checking of company's search strategies; critical appraisal of the economic evidence; carried out the economic analyses; and drafted the economic sections

All authors read and commented on draft versions of the ERG report.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AE	Adverse event
BOR	Best overall response
CDF	Cancer Drugs Fund
CQ	Clarification question
CI	Confidence Interval
CR	Complete response
CS	Company submission
DOR	Duration of response
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
ED-5D	EuroQol-5D
ERG	Evidence Review Group
FACT-M	Functional Assessment of Cancer Therapy – Melanoma
HRQoL	Health-related quality of life
IgG1	Immunoglobulin G1
IPTW	Inverse probability of treatment weighting
irAE	Immune-related adverse event
IRC	Independent Review Committee
IRR	Infusion-related reaction
JM	JAVELIN Merkel
MCC	Merkel cell carcinoma
mMCC	Metastatic Merkel cell carcinoma
NHS	National Health Service
NICE	National Institute of Health and Care Excellence
ORR	Overall response rate
OS	Overall Survival
PD-L1	Protein programmed-death ligand-1
PFS	Progression-free survival
PR	Partial response
PSM	Propensity score matching
PSW	Propensity score weighting
q2w	Once every 2 weeks
SACT	Systemic Anti-Cancer Therapy
SLR	Systematic literature review
SW	Stabilised weights
ToE	Terms of Engagement

ToT	Time on treatment
TRAE	Treatment-related adverse event
TSD	Technical Support Document

1 Executive summary

1.1 Critique of the adherence to committees preferred assumptions from the Terms of Engagement in the company's submission

In general, the Evidence Review Group (ERG) considers that the company has adhered to the committee's preferred assumptions from the Terms of Engagement (ToE). The clinical data presented by the company includes the ToE required later data cut from the company's single arm trial of avelumab in treatment naïve metastatic Merkel Cell Carcinoma (mMCC) patients, Javelin Merkel 200: Part B (JM200: Part B). In addition, the company presented a summary of the observational data that were also required to be collected by Public Health England during the period of managed access for avelumab, hereafter referred to as the Systemic Anti-Cancer Therapy (SACT) data set.

The data from JM200: Part B now comprises a minimum follow-up period of 15 months and includes more mature data for overall survival (OS), progression free survival (PFS) and time on treatment (ToT). The SACT data set comprises a minimum of 5 months follow-up and reports only on OS and ToT. The ERG is concerned that the population in Javelin Merkel 200: Part B may be slightly younger, comprise of more males and have more favourable performance status (PS), i.e. lower ECOG PS, than expected in clinical practice in England. The ERG's clinical experts reported that they considered the SACT data set more representative of patients in clinical practice.

The economic analyses provided by the company are generally in adherence with the ToE, with updated survival modelling and utilities based on the treatment-naïve population data from Javelin Merkel 200 trial. The model structure has remained the same and all corrections made in the model for TA517 have been carried forward into the CDF review analyses.

1.2 Summary of the key issues in the clinical effectiveness evidence

The ERG considers the key issues with the clinical effectiveness evidence to be:

- The population of JM200: Part B does not accurately reflect the patients likely to receive avelumab for first line treatment of mMCC in England; the SACT data set comprises a closer match to expected patient characteristics in clinical practice in England.
- The SACT data set does not provide information on PFS, HRQoL, response rate or adverse effects of treatment.
- The naïve comparison of avelumab in JM200: Part B with chemotherapy in Study 100070-Obs001 maybe confounded by the inclusion of patients with immunosuppression in Study

100070-Obs001 as all patients in JM200: Part B are immunocompetent. The ERG thus prefers the use of the subgroup of immunocompetent patients from Study 100070-Obs001 in the naïve comparison although the ERG still considers it to be potentially unreliable because of the imbalances in other patient characteristics between the two studies, the small number of patients in the studies, and the uncertainty caused by unmeasured variables that may be effect modifiers or prognostic indicators.

- The ERG's preferred propensity scored analysis for avelumab versus chemotherapy (propensity score weighting analysis 4 [PSW4]) omits immune status as a characteristic for matching and thus may overestimate the benefit of avelumab but nevertheless is the ERG's preferred source of data for use in the economic model for avelumab versus chemotherapy.
- The marketing authorisation approved licensed dose of avelumab was changed in November 2019 to a flat dose of 800mg for all patients rather than the weight-based dose used in the original CS for TA517 and in the JM200 clinical trial.

1.3 Summary of the key issues in the cost effectiveness evidence

The key uncertainties in the company's analyses lie in the estimation of treatment effectiveness as the company's base case relies on a naïve comparison to estimate OS and PFS for avelumab and chemotherapy. The company provided a range of different adjusted analyses to account for the imbalances in baseline characteristics across the studies for avelumab and chemotherapy, adjusting for age, sex, ECOG score, and immunocompetency status. However, no single analysis adjusted for all imbalances simultaneously.

To do this, the company provided propensity score matched (PSM) analyses as requested by the ERG, as well as propensity score weighted (PSW) analyses. The company considered the PSM to lose too much data in achieving a suitable match, and therefore, provided the weighted analyses to avoid this loss of data. Within these analyses, the company explored the exclusion of the immunocompetency variable from the estimation of propensity scores as well as the exclusion of patients with an ECOG score of 2 or more. This was because the Javelin Merkel 200 trial used to inform the avelumab group did not included immunosuppressed patients or patients with an ECOG score of 2 or more so no balance could be achieved with these variables.

The ERG considered the PSW analysis that excluded immunocompetency from the propensity score estimation and removed patients with an ECOG score of 2 or more provided the best balance across studies. However, the analysis is still likely to overestimate the benefits in favour of avelumab as the chemotherapy group still has the proportion of patients who are immunosuppressed who are likely

to have worse outcomes. Therefore, even the ERG's preferred base case ICER remains uncertain and may be underestimated.

Another issue that the ERG considered inappropriate was the company's adjustment to the avelumab time-on-treatment (ToT) curve based on the discontinuation rates for the treatment experienced population after 15 months. The company justified this because the minimum follow-up for the treatment-naïve population was 15 months; however, the ERG considers that the rates for the treatment-experienced population beyond 15 months cannot be expected to be reflective of treatment-naïve population. Therefore, the ERG removed this adjustment in the ERG's preferred base case analysis and used only the treatment-naïve based time-on-treatment curves.

A final area of uncertainty that the ERG noted is in the change of dosing to a flat 800mg dose compared with the previous weight-based dose (10mg/kg), which resulted in an average of 4.25x 200mg vials (849mg in total) required per administration including wastage. Given that the Javelin Merkel 200 trial treatment effects were based on weight-based dosing , the ERG considers the acquisition costs applied in the economic model should be weight-based to align with the effectiveness estimates. The ERG notes that this would, however, not reflect clinical practice, which will now be based on the 800mg flat dose.

1.4 Summary of ERG's preferred assumptions and resulting ICER

The ERG's preferred assumptions are as follows:

1. PSW analyses for OS and PFS with immunocompetency excluded from the estimation of propensity scores and patients with ECOG score 2 or more removed;
2. 1-knot hazard spline for OS;
3. 3-knot odds spline for PFS;
4. 3-knot hazard spline for ToT;
5. Removing the adjustment to the ToT curve using the treatment-experienced population data;
6. Weight-based acquisition costs for avelumab in line with effectiveness data.

The results of the ERG's preferred base case analysis are given in Table 1.

Table 1. ICER resulting from ERG's preferred assumptions

Intervention	Total costs	Total QALYs	Δ costs	Δ QALYs	ICER £/QALY
Chemotherapy	[redacted]	1.50	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year.

1.5 Summary of ERG's key scenarios and resulting ICER

The ERG also presents the results of two alternative scenario analyses that potentially resolve some issues relating to the immunocompetency imbalance and the differentiation from the expected population in clinical practice, but then bring other limitations in terms of imbalance of other characteristics. These analyses are the company's naïve comparison using the immunocompetent subgroup, and the ERG's naïve comparison using the SACT data for avelumab. The results are presented in Table 2 and Table 3, respectively.

Table 2. ICER resulting from immunocompetent naïve comparison

Intervention	Total costs	Total QALYs	Δ costs	Δ QALYs	ICER £/QALY
Chemotherapy	[redacted]	1.25	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year.

Table 3. ICER resulting from SACT naïve comparison

Intervention	Total costs	Total QALYs	Δ costs	Δ QALYs	ICER £/QALY
Chemotherapy	[redacted]	1.32	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; QALY, quality adjusted life year.

2 Introduction and Background

2.1 Introduction

Merkel cell carcinoma (MCC) is a rare neuroendocrine skin cancer that is more common in the elderly. Occurring more frequently on sites of the skin that receive greater exposure to the sun, MCC tends to metastasise at an early stage. The visible appearance of the cancer can cause considerable psychosocial distress to patients with symptoms dependent on the site of the primary tumour and metastases. There are currently very few treatment options for patients with metastatic MCC (mMCC) and it is generally associated with a poor prognosis.

Avelumab (Bavencio®) is a human, Immunoglobulin G1 (IgG1) lambda monoclonal antibody that inhibits the immune checkpoint protein, programmed-death ligand-1 (PD-L1) which is found in cancer cells. Avelumab was granted a conditional European Medicines Agency marketing authorisation on 18 September 2017.[1]

Avelumab is currently recommended for use within the Cancer Drugs Fund (CDF) for untreated mMCC in adults who have not had chemotherapy for metastatic disease (TA517)[2], and has been available in the CDF for untreated mMCC since April 2018. Avelumab is also a recommended treatment option in routine National Health Service (NHS) practice for previously treated mMCC in adults who have had one or more lines of chemotherapy for metastatic disease (TA517). Here, this report comprises a review of the latest clinical and cost-effectiveness evidence for avelumab in untreated mMCC.

2.2 Background

The clinical-effectiveness evidence for avelumab in the original company submission (CS) for TA517[2] was taken from the JAVELIN Merkel 200 (JM200) clinical trial, a single-arm trial of mMCC patients. JM200 comprised two cohorts of patients: Part A which was patients with treatment experienced mMCC (i.e. those who have previously received chemotherapy) and Part B which comprised patients with treatment naïve mMCC (i.e. those who have not previously received chemotherapy). This CDF review primarily focuses on updated data analyses from Part B of JM200.

The ERG notes that the company reported in the CS that the marketing authorisation approved licensed dose of avelumab was changed in November 2019 to a flat dose of 800mg for all patients rather than the weight-based dose used in the original CS for TA517 and in the JM200 clinical trial. The avelumab dose in JM200 was a target dose of 10mg/kg once every two weeks (q2w), through 1-hour intravenous infusions. The company reported that a dose of 800mg would directly correspond to the same dose with 10mg/kg weight-based dosing as it was based on the 80kg median body

weight for adults from three studies in patients with various tumour types (studies EMR100070-001, EMR100070-002 and EMR100070-003).

2.3 Data collection

The data collection required and specified in the Terms of Engagement(ToE) document[3] was as follows:

- The primary source for data collection required under the managed access agreement was the JM200: Part B trial. The expected data analysis is based on the trial protocol including the reporting of overall survival (OS) and treatment duration using the final analysis data-cut.
- Observational data were also required to be collected during the period of managed access via the Systemic Anti-Cancer Therapy (SACT) data set to support the data collected in the JM200: Part B clinical trial. SACT was required to collect data on OS and duration of therapy and Public Health England were required to provide a summary of the observational data collected.

The Evidence Review Group (ERG) notes that the specified data have been collected and are discussed in Section 3 of this report. The JM200: Part B data now comprise a minimum follow-up period of 15 months and the SACT data set had a minimum of 5 months follow-up.

2.4 Critique of company's adherence to committees preferred assumptions from the Terms of Engagement

In general, the ERG considers that the company has adhered to the committees preferred assumptions from the Terms of Engagement[3]. The ERG's critique of the company's adherence to the committees preferred assumptions from the Terms of Engagement is provided in Table 4.

Table 4. Preferred assumptions from Terms of Engagement[3]

Area	Assumption	Terms of Engagement	Addressed by company submission	ERG comment
Population	The recommendation in TA517 is an optimised CDF recommendation.	The CDF review will focus on the population that were recommended under the managed access agreement. That is, adult with mMCC who have not had chemotherapy for metastatic disease.	Yes	First-line therapy for mMCC population focussed on by the company.
Comparators	The most appropriate comparator for first-line treatment is chemotherapy.	CDF review will focus on the comparison with chemotherapy	Yes	The company submission focuses on the comparison of avelumab versus chemotherapy for first-line treatment of mMCC.
Generalisability of JAVELIN	The committee also noted that the marketing authorisation has been granted conditionally for the first-line group because of the immaturity of the data. The committee concluded that the JAVELIN results should be interpreted with caution as there were some unanswered questions about the generalisability of JAVELIN.	In light of new evidence, the CDF review will assess the generalizability of JAVELIN	Yes, although the ERG has some concerns about the generalisability of Javelin Merkel 200: Part B.	Updated analysis of Javelin Merkel 200: Part B is provided by the company and now includes N=116 patients with at least 15 months' follow-up. The ERG is concerned that the population in Javelin Merkel 200: Part B may be slightly younger, comprise of more males and have a lower ECOG PS than expected in clinical practice in England.
Comparative clinical data	The committee noted that there were no direct comparative data, and concluded that, although uncertain, the 2-part observational Study 100070-Obs001 provided the most appropriate comparator data.	If further comparative data is available, it should be explored in the CDF review.	N/A, no new comparative data identified by the updated systematic literature review.	The company updated their clinical effectiveness systematic literature review from the original company submission, and it did not identify any new comparative clinical data suitable for inclusion. However, data from a naïve pooled analysis of seven studies used in the original CS to inform PFS and OS in the economic model

				were used again which the ERG has concerns about due to clinical heterogeneity between the studies.
Indirect comparison	The results from the naive indirect comparison are highly uncertain.	In light of new evidence, the CDF review will consider the naive indirect comparison.	Yes, appropriate clinical data provided during clarification to stage to enable naïve indirect comparison between JM200: Part B and Study 100070-Obs001.	The ERG also requested a propensity score matching analysis during the clarification stage as the ERG is concerned about the lack of robustness of the naïve comparison.
Overall survival	The overall survival data may be confounded by the use of subsequent treatments, and no data on subsequent treatments were recorded as part of the trial. The committee noted that the overall survival data were still relatively immature.	The committee are expecting updated overall survival evidence in the CDF review	Yes.	The company now have a minimum of 15 months' follow-up for all patients in JM200: Part B and updated analyses for overall survival were provided by the company.
Model structure	The company's model structure is suitable for decision making.	It is anticipated that the model structure will not change for the CDF review	Yes	The company provided their updated analyses in a new version of the electronic model but the model structure was unchanged.
Extrapolation of survival	Due to the limitations of the data for first-line treatment, the company used estimates for survival derived from the second line and beyond model in its original first line model. The committee were concerned that the progression-free and overall survival estimates for first line treatment were based on clinical assumptions, not direct evidence. It was also aware that first-line treatment was most sensitive to the	It is anticipated that the method used to extrapolate survival will be explored in the CDF review	Yes	The company provide updated survival modelling for the avelumab group based on treatment-naïve patients from the Javelin Merkel 200 trial.

	hazard ratio chosen for overall survival. The committee concluded that the survival estimates for first-line treatment are highly uncertain.			
Time-on-treatment	The company assumed that two-thirds of patients would stop treatment after 2 years (and all remaining patients would stop treatment after 5 years). The clinical experts explained that they expect 95% of patients having avelumab to stop treatment by 2 years. The ERG considered the time-on-treatment extrapolation without truncation at 2 years. The committee consider both the company's and the ERG's assumptions in its decision-making.	It is anticipated further evidence on time-on-treatment will be available for the review and this assumption will be reconsidered in light of new evidence.	Yes	The company provided survival models based on the treatment-naïve population data from the Javelin Merkel 200 trial.
Utilities	The committee acknowledged that the utility values were implausibly high but it noted that, because the same utilities were applied regardless of treatment group, only the difference between health states mattered. The committee concluded that it could accept the company's utility values but acknowledged that these were very high.	It is anticipated that the same utility values will be used in the CDF review.	Yes	The ERG notes that the ToE document stated that if further evidence were available, an exploration of the most appropriate utilities should be performed and that the company has conducted additional sensitivity analysis using alternative utility values. The company provided updated analyses in the economic model using data from the treatment-naïve population of the Javelin Merkel 200 trial. The company also provided analyses using

				the treatment-experienced population as per TA517, as well as an analysis based on the combined data set.
Model corrections	The following model corrections were made: •adding the cost of premedication •added administration costs (approximately £43) •corrected an error in the calculation of background mortality.	It is anticipated that these corrections will be included in the CDF review	Yes	All corrections made in TA517 were carried forward into the analyses presented in the CDF review.
End of life	Avelumab meets the end-of-life criteria	Not reported	Yes	The company have stated that avelumab meets the criteria for end-of-life and their base case results are in line with this statement.
Abbreviations: CDF, Cancer Drugs Fund; ERG, Evidence Review Group; mMCC, metastatic Merkel Cell carcinoma; N/A, not applicable; ToE, terms of engagement.				

3 Clinical effectiveness

3.1 Critique of new clinical evidence

The primary evidence submitted by the company in the company submission (CS) is additional follow-up data from the single arm Javelin Merkel 200: Part B (JM200: Part B) clinical trial of avelumab in patients with metastatic Merkel cell carcinoma (mMCC) with no prior systemic therapy for metastatic disease (treatment-naïve mMCC). In total, there are now data for N=116 treatment naïve mMCC patients with a minimum of 15 months' follow-up in JM200: Part B (data cut: May 2019). In addition, the CS included an overview of the evidence on avelumab for treatment naïve mMCC collected via the Systemic Anti-Cancer Therapy (SACT) database[4] which comprised patients receiving avelumab in National Health Service (NHS) practice. The SACT data comprises only N=52 treatment-naïve mMCC patients, and the minimum follow-up for these patients is 5 months (data cut: November 2019 for OS and July 2019 for ToT).The Evidence Review Group (ERG) notes that only data for overall survival (OS) and time-on-treatment (ToT) are available from the SACT data set, with no data available on progression-free survival (PFS). The ERG also notes that the company reports that the SACT data cannot be utilised in the current economic model, although the ERG considers that it is possible to utilise the SACT data and provides this in scenario analyses (See Section 4.1.5). The data from JM200: Part B and the SACT database are discussed in more detail in the subsections below but the ERG considers it important to highlight that the ERG's clinical experts considered the SACT data set patient population to be more representative of patients likely to receive avelumab in clinical practice in England.

3.1.1 Javelin Merkel 200: Part B - study overview

The data for the first line (1L; treatment naïve) population in the original CS were derived from a pre-planned interim analysis of a data-cut from 24 March 2017, which comprised N=39 patients. These data were used for three analyses: efficacy for patients with ≥3-month follow-up (n=29), efficacy for patients with 6-month follow-up (n=14), and efficacy and safety endpoints for the full 39 patients. The planned sample size for the primary analysis of the 1L cohort was 112 patients and further analyses were planned for 2018. Patients in JM200: Part B were enrolled only in the USA and Europe. Avelumab was given intravenously at the 10mg/kg dose with the recommended paracetamol and antihistamine premedications. Treatment with avelumab was continued for between 6 and 12 months, or longer in agreement with the Sponsor if there was a complete response (CR). Treatment was discontinued on disease progression or study withdrawal for any reason including intolerable

toxicity. The primary efficacy endpoint was durable response which was defined as an objective response (CR or partial response [PR]) according to RECIST version 1.1, determined by the independent endpoint review committee, with a duration of at least 6 months.

The analysis of JM200: Part B in the new company submission comprises 116 patients with a minimum of 15-months' follow-up using a data-cut from May 2019. The ERG notes that the planned sample size of 112 has now been achieved. The baseline characteristics of the patients in the updated analysis of JM200: Part B (May 2019 data-cut) are presented in Table 5. The ERG's clinical experts reported that they would expect a gender split closer to 50:50 and the age of patients to be higher. In addition, clinical experts reported that there may be some patients in clinical practice with an Eastern Cooperative Oncology Group (ECOG) score of 2 or even higher that could benefit from avelumab although these patients were not included in JM200: Part B.

Table 5. Baseline characteristics from JAVELIN Merkel 200: Part B (May 2019 data-cut; adapted from Table 1 from the company's response to clarification dated 27 March 2020)

Baseline characteristic	Value(s) n (%)
Gender	
•Male	81 (70%)
•Female	35 (30%)
Age	
•<40	0 (0%)
•40–49	4 (3%)
•50–59	7 (6%)
•60–69	27 (23%)
•70–79	46 (40%)
•80+	32 (28%)
Median	74.0 years
ECOG	
•0	72 (62%)
•1	44 (38%)
•≥ 2	0 (0%)
Region (e.g. Europe), n (%)	
•North America	29 (25.0)
•Western Europe	75 (64.7)
•Australia	9 (7.8)
•Asia	3 (2.6)
Number of patients from UK sites, n (%)	0 (0%)

Site of primary tumour, n (%)	
•Skin	104 (89.7)
•Lymph node	1 (0.9)
•Not reported	11 (9.5)
Tumour size (cm)	n=57 (49.1%)
•Median	3.2
•Range	0.6 to 25.0
Time from initial diagnosis to study entry in months, median (range)	10.6 (0.7 to 120.9)
Presence of distal metastases, n (%)	Not available
Presence of visceral metastases, n (%)	
•Yes	79 (68.1)
•No	35 (30.2)
•Not evaluable	2 (1.7)
Presence of lymph node metastases, n (%)	
•Yes	25 (21.6)
•No	89 (76.7)
•Not evaluable	2 (1.7)
Tumour PD-L1 expression, n (%)	
•Positive	21 (18.1)
•Negative	87 (75.0)
•Not evaluable	8 (6.9)
Tumour MCPyV status, n (%)	
•Positive	70 (60.3)
•Negative	37 (31.9)
•Not evaluable	9 (7.8)
Abbreviations: ECOG, Eastern Cooperative Oncology Group.	

The ERG requested details at clarification from the company about the subsequent therapies received by patients in JM200: Part B. Of the total 116 patients, 40 patients (34.5%) had at least one subsequent therapy. The most frequent subsequent therapies were etoposide (15.5%), carboplatin (13.8%) and avelumab (8.6%). The ERG is unclear why avelumab is reported as a subsequent therapy when all patients received it as their primary therapy in JM200: Part B and the ERG's clinical experts reported that patients who respond to treatment would not routinely discontinue it although patients who had electively discontinued treatment after stable disease may be restarted on avelumab at signs of disease progression. The ERG also notes that 17 patients (14.7%) received at least one subsequent anti-cancer radiotherapy treatment. The ERG's clinical experts reported that the most frequently used chemotherapy regimens in mMCC are carboplatin or cisplatin with or

without etoposide with carboplatin preferred by some clinicians due to tolerance issues with cisplatin. The ERG therefore considers the subsequent therapy use of single agent etoposide not to be consistent with expected clinical practice in England.

3.1.2 Javelin Merkel 200 - Part B: updated results

The company reported that OS, PFS, ToT and health-related quality of life (HRQoL) data were available from the updated analysis of JM200-Part B, although HRQoL data were only provided directly during the clarification response.

3.1.2.1 Time on treatment

ToT from the May 2019 data-cut data is presented in Figure 1 and includes [redacted] patients who were censored in the analysis, albeit beyond 15 months (Appendix 9.1, Table 31). The median ToT was [redacted] and [redacted] of patients were still receiving treatment at 12 months, [redacted] of patients remaining on avelumab (Appendix 9.1, Table 31 and Table 32).

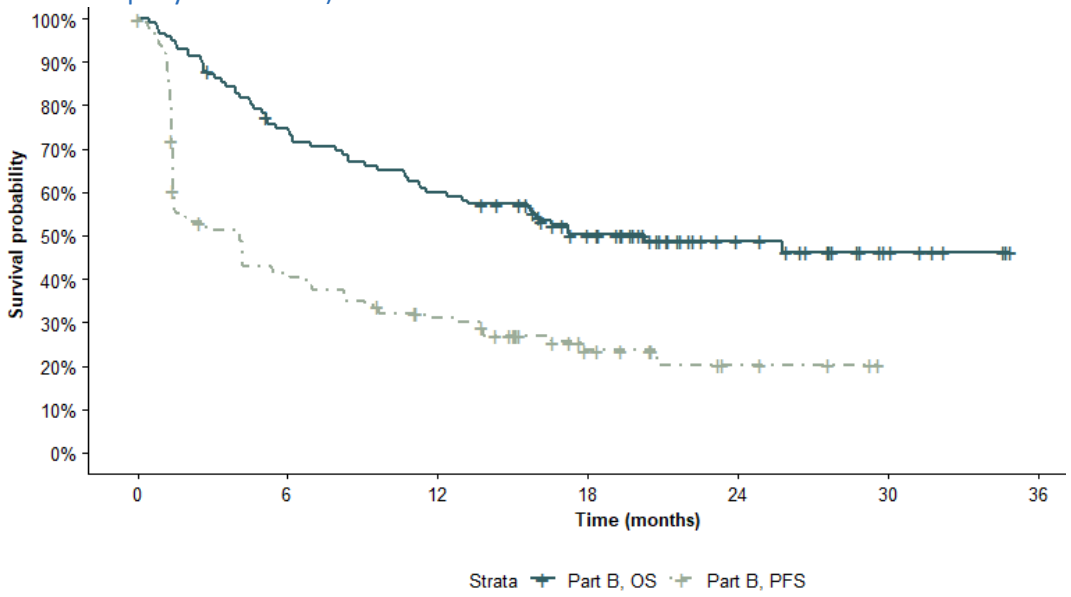
Figure 1. JAVELIN Merkel 200 Part B Kaplan-Meier time on treatment (ToT; reproduced from Figure 3 from the company submission)

[redacted]

3.1.2.2 OS and PFS

The May 2019 data-cut for OS and PFS in JM200: Part B comprised 116 patients with a minimum of 15 months' follow-up (Figure 2). The ERG considers it important to highlight that in the May 2019 data-cut the plateaus for both OS and PFS are lower compared to the data-cut used in the original CS for TA517, therefore long-term OS and PFS rates with the May 2019 data-cut are lower. The ERG notes that in the May 2019 data-cut only [redacted] patients have been censored in the PFS analysis, although [redacted] patients are censored in the OS analysis with [redacted] censored by 24 months (Appendix 9.1 Table 33). The company provided a breakdown of the reasons why patients were censored in the OS and PFS analyses and the ERG notes that the most common reason [redacted] (Appendix 9.1, Table 34).

Figure 2. JM200: Part B, overall survival (OS) and progression-free survival (PFS; reproduced from Figure 1 from the company submission)



At risk

Part B, OS	116	85	68	45	20	7	0
Part B, PFS	116	45	31	12	4	0	0

Median PFS using the May 2019 data-cut is 4.1 months and PFS rates at 6 and 12 months are 41.3% and 31.0%, respectively (Table 6). The median OS is 20.3 months and OS rates at 6 and 12 months are 74.8% and 59.9% (Table 6). The ERG notes that in the March 2017 data-cut used in the original

CS, median OS had not been reached and median PFS was [redacted] months although these data were immature and based only on [redacted] patients.

Table 6. OS and PFS outcomes from JAVELIN Merkel 200: Part B (reproduced from Table 6 from the company's response to clarification dated 27 March 2020)

Outcome	Result	95% CI	Number in analysis
Median PFS, months (95% CI)	4.11	(1.48 to 6.74)	116
•6-month PFS rate, % (95% CI)	41.3%	(33.1% to 51.5%)	[redacted]
•12-month PFS rate, % (95% CI)	31.0%	(23.5% to 41.0%)	[redacted]
•15-month PFS rate, % (95% CI)	27.0%	(19.8% to 36.8%)	[redacted]
Median OS, months (95% CI)	20.3	(13.0 to NE)	116
•6-month OS rate, % (95% CI)	74.8%	(67.3% to 83.2%)	[redacted]
•12-month OS rate, % (95% CI)	59.9%	(51.5% to 69.6%)	[redacted]
•15-month OS rate, % (95% CI)	57.2%	(48.8% to 67.1%)	[redacted]

Notes: 95% confidence interval limits were estimated using the default settings in the statistical software R. For the 'Number in analysis' column, we have populated this assuming the total sample size for the median value, and the number at risk for the values at a specific time point.
Abbreviations: NE, not evaluable; OS, overall survival; PFS, progression-free survival.[redacted]

3.1.2.3 Response rate

The May 2019 results from JM200: Part B show that 39.7% of patients had a confirmed objective response (CR or PR) to avelumab as assessed by the independent review committee using RECIST v1.1 criteria (Table 7). In addition, 10.3% of patients had stable disease as their best overall response. The median duration of response was 18.2 months and for 61% of patients the duration of response was ≥15 months (note these figures are both based on Kaplan-Meier estimates).

Table 7. Response rate in JAVELIN Merkel 200: Part B (reproduced from Table 7 from the company's response to clarification dated 27 March 2020)

Outcomes by RECIST v1.1, per IRC Assessment JM200 Part B	Analysis May 2019
Number of patients, n	116
Median follow-up, mths	21.2 (range: 14.9 to 36.6)
Median duration of treatment, wks	24 (range: 2 to 154)

Confirmed ORR, % (95% CI)	39.7 (30.7 to 49.2)
Confirmed BOR, n (%)	
•Complete response	19 (16.4)
•Partial response	27 (23.3)
•Stable disease	12 (10.3)
•Progressive disease	48 (41.4)
•Non-CR/Non-PD	1 (0.9)
•Non-evaluable	9 (7.8)[a]
Response durability	
•Patients with durable response, n	35
•Durable response rate (95% CI)	30.2% (22.0 to 39.4%)[b]
Response duration	
•Median DOR (95% CI), mths†	18.2 (11.3 to NE)
•Proportion of responses with duration ≥3 mths, % (95% CI)	89 (75 to 95)
•Proportion of responses with duration ≥6 mths, % (95% CI)[c]	78 (63 to 87)
•Proportion of responses with duration ≥12 mths, % (95% CI)[c]	66 (50 to 78)
•Proportion of responses with duration ≥15 mths, % (95% CI)[c]	61 (44 to 74)

[a] No postbaseline assessments due to early death (n=4) or other reasons (n=2), no adequate baseline assessment (n=2), or all postbaseline assessments had overall response of NE (n=1).
[b] Based on Kaplan–Meier estimates.
[c] Proportion of patients with a response lasting ≥6 months.
Abbreviations: BOR, best overall response; CI, confidence interval; CR, complete response; DOR, duration of response; IRC, Independent Review Committee; ORR, overall response rate; PD, progressive disease.

3.1.2.4 Health-related quality of life

The company provided a summary of the HRQoL data from the May 2019 data-cut off JM200: Part B in their response to clarification questions and reported that a comprehensive summary of the available HRQoL data is not yet available. However, detailed data for the EQ-5D Visual Analogue Score (VAS) were provided in the clarification response to Question A8 (dated 27 March 2020).

In JAVELIN Merkel 200: Part B, patients were administered two HRQoL questionnaires: the EQ-5D-5L, a generic tool and the FACT-M (Functional Assessment of Cancer Therapy – Melanoma) questionnaire, a melanoma-specific measure of HRQoL. The company reported that while the FACT-M is designed and validated for use in patients with melanoma it is also useful in Merkel cell carcinoma (MCC) due to similarities in the two types of cancer.

The company reported that [redacted] of patients ([redacted]) completed the EQ-5D at screening and [redacted] of patients ([redacted]) completed the FACT-M questionnaires at the Screening visit. The completion rate at the End-of-Treatment visit was reported by the company as [redacted] for both the EQ-5D and FACT-M and they stated that completion rates during the treatment period were similar for both questionnaires (mostly [redacted]). The company reported that the EQ-5D VAS, the EQ-5D Index score or the FACT-M scores found no major changes in participants' health status overall while on treatment with avelumab, suggesting no change in HRQoL over time. A results table and Kaplan-Meier plot summarising the EQ-5D VAS scores was provided by the company (CQ response A8) and the ERG notes that the EQ-5D data from JM200: Part B were used to inform HRQoL in the economic model (see Section 4.1.7 for more detail).

3.1.2.5 Adverse events

The company provided results for adverse events (AEs) from the updated JM200: Part B data in their clarification response and reported that all 116 patients had an AE of any grade with 60.3% (70 patients) experiencing a grade ≥3 AE. Treatment-related AEs (TRAEs) occurred in 94 patients (81.0%) and 21 patients (18.1%) had a grade ≥3 TRAE. The most frequent TRAE was fatigue (20.7% patients) and the only grade ≥3 TRAEs that occurred in more than 1 patient were raised lipase (4 patients) and raised amylase (3 patients, Table 8). Infusion-related reactions (IRRs) occurred in 34 patients (29.3%) with 1 patient (0.9%) who had a grade 3 IRR and none with grade 4 or 5 IRRs.

The company reported that 14.7% of patients had a serious TRAE, and 12.1% had a TRAE that led to treatment discontinuation. There were 15 patients (12.9%) who had an AE that led to death, although none of these were considered treatment related.

Table 8. TRAEs (any grade in ≥10% of pts) occurring in JAVELIN Merkel 200: Part B (adapted from Table 8 from the company's response to clarification dated 27 March 2020)

Adverse effect	Any grade		Grade ≥3	
	N	%	N	%
Any TRAE, n (%)[a]	94	81.0	21	18.1

Fatigue	24	20.7	1	0.9
Pruritus	15	12.9	1	0.9
Asthenia	16	13.8	0	-
Chills	12	10.3	0	-
Lipase increased	6	5.2	4	3.4
Decreased appetite	6	5.2	1	0.9
ALT increased	5	4.3	1	0.9
Amylase increased	3	2.6	3	2.6
AST increased	2	1.7	1	0.9
Any IRR[b]	34	29.3	1	0.9

[a] The incidence of treatment-related IRRs based on the single MedDRA Preferred Term is not listed.
[b] Includes AEs (irrespective of relatedness) categorized as IRR, drug hypersensitivity, or hypersensitivity reaction that occurred on the day of infusion or day after infusion, in addition to signs and symptoms of IRR that occurred on the same day of infusion and resolved in ≤2 days.
Abbreviations: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; IRR, infusion-related reaction; TRAE, treatment-related adverse event.

Table 9 provides a summary of the 35 patients (30.2%) had an immune-related AE (irAE) of any grade. The ERG notes that 20% of the irAEs were grade ≥3 (7 patients [6.0%]).

Table 9: IrAEs (any grade in ≥5% of patients or grade ≥3 in any patient) in JAVELIN Merkel 200: Part B (reproduced from Table 9 from the company's response to clarification dated 27 March 2020)

Adverse effect	Any grade		Grade ≥3	
	N	%	N	%
Any irAE, n (%)	35	30.2	7	6.0
Pruritus	9	7.8	1	0.9
Rash maculopapular	6	5.2	0	
ALT increased	3	2.6	1	0.9
Autoimmune nephritis	1	0.9	1	0.9
Autoimmune neuropathy	1	0.9	1	0.9
Dermatitis psoriasiform	1	0.9	1	0.9
Diabetes mellitus	1	0.9	1	0.9
Liver function test increased	1	0.9	1	0.9

Abbreviations: AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; IRR, infusion-related reaction; TRAE, treatment-related adverse event.

3.1.2.6 Subgroups

In response to clarification questions the company provided results by subgroups for key subgroups identified by the ERG's clinical experts or using the original CS (Clarification response A9 and A10). The ERG notes that the subgroup analyses were not sufficiently powered to detect significant differences between subgroups in treatment effect but the ERG reports the direction of any trends.

In terms of sex, the male subgroup [redacted] compared with the female subgroup (clarification response A9). The ERG also notes that the male subgroup was more than double the size of the female subgroup (81 males and female 35 females) and the ERG's clinical experts reported that this split was not reflective of clinical practice (where closer to a 50:50 split would be expected).

The Kaplan-Meier plots for [redacted]

3.1.3 *Systemic Anti-Cancer Therapy (SACT) data*

As discussed above, the results for OS and ToT were presented in the CS from the SACT data[4] which comprised 52 treatment naïve mMCC patients treated in the NHS (November 2019 data cut). The ERG notes that there is more than double the number of patients in the latest analysis of JM200:Part B (N=116) compared to the SACT data-set (N=52) and minimum follow-up is substantially longer in JM200:Part B (minimum 15 months compared to 5 months for SACT), with outcome data for PFS as well as OS and ToT. However, the ERG's clinical experts considered the baseline characteristics of patients in the SACT data set to be more reflective of patients in England likely to receive avelumab than the patients in JM200:Part B. This is because male to female ratio is closer to the anticipated 50:50, a higher proportion of ≥ 80 years patients and a higher proportion of ECOG PS 1 and above (Table 10).

Table 10. Baseline demographics, JM200: Part B versus SACT (reproduced from Table 5 from the company submission)

Characteristic		JM200: Part B	SACT
Sample size		116	52
Follow-up		Minimum: 15 months Median: 16 months	Minimum: 5 months Median: 6 months
Sex	Male	81 (70%)	30 (58%)
	Female	35 (30%)	22 (42%)

Age	<40	0 (0%)	0 (0%)
	40-49	4 (3%)	1 (2%)
	50-59	7 (6%)	3 (6%)
	60-69	27 (23%)	8 (15%)
	70-79	46 (40%)	22 (42%)
	80+	32 (28%)	18 (35%)
	Median	74.0 years	75.5 years
ECOG PS	0	72 (62%)	7 (13%)
	1	44 (38%)	34 (65%)
	2	0 (0%)	4 (8%)
	3	0 (0%)	1 (2%)
	4	0 (0%)	0 (0%)
	Missing/ unknown	0 (0%)	6 (12%)

Abbreviations: ECOG PS, Eastern Cooperative Oncology Group Performance Status; JM200, JAVELIN Merkel 200; SACT, Systemic Anti-Cancer Therapy.

The company argues that the SACT data set is not entirely representative of the first-line eligible patients in clinical practice as some patients may have previously been deemed ineligible for treatment with first-line chemotherapy (e.g. due to its toxicities), and therefore managed with best supportive care prior to avelumab. This means some patients in the SACT data set may have a poorer prognosis compared to newly-diagnosed mMCC patients although the ERG is unsure whether this limitation also affects the JM200:Part B data set.

The company's reasons for using the latest JM200: Part B data for their preferred data set for both the clinical and cost-effectiveness analysis in this CDF review include:

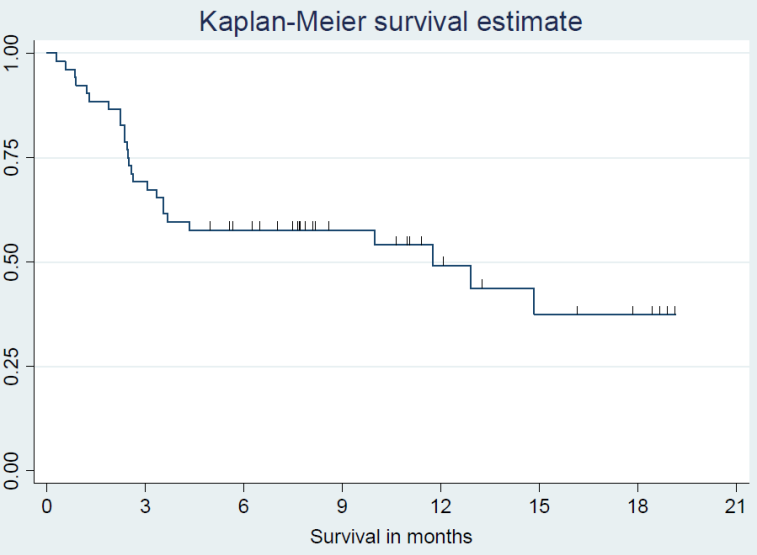
- JM200: Part B reports on all outcomes specified in the final NICE scope for TA517, whereas the SACT data set only contains data for OS and ToT.
- The patient population in JM200: Part B comprised patients with ECOG PS of 0 or 1, and a life expectancy of at least 12 weeks, which the company consider to be sufficient life expectancy to benefit from cancer immunotherapy such as avelumab. In contrast, 10% of the SACT data set had an ECOG PS of 2 or 3 and 12% had unknown ECOG PS. In addition, the company considers the OS curve for the SACT cohort is suggestive that some patients may not have sufficient life expectancy to benefit from immunotherapy.
- JM200: Part B comprises a larger sample of patients compared with the SACT data set (N=116 versus N=52) and JM200: Part B also comprises longer follow-up (a minimum follow-up in JM200: Part B of 15 months versus 5 months in SACT).

3.1.3.1 SACT data OS

Median OS in the SACT data set was 11.8 months, with estimated OS rates at 6 and 12 months of 58% and 50%, respectively (Figure 3). In contrast, median OS in JM200: Part B was longer (20.3 months) and estimated OS rates at 6 and 12 months were higher (74.8% and 59.9%, respectively). The ERG notes that the median follow-up time for OS was shorter at 6.3 months (191 days) for the SACT data set, compared with 15.9 months (483 days) in JM200: Part B. In addition, the SACT data set are based on a smaller sample of patients compared to JM200: Part B (n=52, versus n=116, respectively). In JM200: Part B, patients with an ECOG PS greater than 1 were excluded which the company reported they would also have expected to be the case in the SACT data set, yet 10% of patients in the SACT data set had an ECOG PS of 2 or 3. The ERG notes that patients with an ECOG PS of 2 or more would be expected to have a poorer prognosis then those with an ECOG PS of 0 or 1 and thus OS would be expected to be lower in the SACT data set. However, the ERG's clinical experts also reported that they would expect the ECOG PS of patients in clinical practice to be more similar to that of patients in the SACT data set compared to JM200: Part B. Therefore, the ERG considers the main limitation with the SACT data set for OS is the shorter follow-up – this leads to immature results for OS that is further limited by the smaller sample size of the SACT data compared to the JM200: Part B data set.

The company also argues that the SACT data set may have included patients with a life expectancy that is unlikely to be sufficient to derive benefit from avelumab due to its mechanism of action as an immunotherapy – and consider this to be one of the reasons for lower OS rates compared to in JM200: Part B. The company considers the evidence in support of this argument to be that 33% of patients died within the first three months, whereas only 13% of patients in JM200: Part B cohort died in the first three months of avelumab. The ERG notes that JM200: Part B excluded patients with a life expectancy of less than 12 weeks. The ERG's clinical experts considered patients treated in the SACT data set would reflect patients expected to receive avelumab in clinical practice and life expectancy would have been taken into consideration when selecting treatment. The experts also considered it to be hard to accurately predict a patient's life expectancy in routine clinical practice.

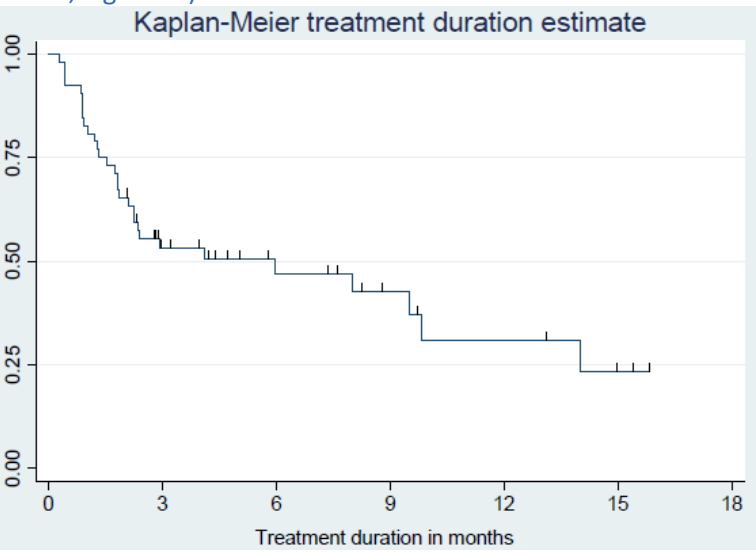
Figure 3. SACT, OS (reproduced from Avelumab for treating metastatic Merkel cell carcinoma – data review, Figure 4[4])



3.1.3.2 SACT data ToT

The ERG notes that the maximum follow-up time in the SACT data set for the outcome of ToT is approximately 16 months. Median ToT in the SACT data set was 6.0 months (Figure 4) which is [redacted] than the [redacted] reported in the latest JM200:Part B data. However, the ERG considers the proportion of patients who were still receiving avelumab at 6 and 12 months [redacted] the SACT data and JM200: Part B (6 months: 46% and [redacted] respectively; 12_months: 31% and_[redacted] respectively).

Figure 4. SACT database ToT (reproduced from Avelumab for treating metastatic Merkel cell carcinoma – data review, Figure 3[4])



3.2 Critique of comparison with chemotherapy

In the ToE document it is reported that, "the committee noted that there were no direct comparative data, and concluded that, although uncertain, the 2-part observational Study 100070-Obs001[5] provided the most appropriate comparator data". In addition, it was reported in the ToE that any new comparative data identified in the CDF review should be explored. The ERG notes that the company conducted an update to the clinical effectiveness systematic literature review (SLR) provided in the original CS with revised searches conducted in January 2020 and that the company reported that the updated SLR did not identify any new comparator studies. The details of the SLR were provided in an embedded document in Appendix 2 of the CS. The company provided limited details of the methods of the revised SLR but, having reviewed the 36 included studies, the ERG agrees with the company that Study 100070-Obs001[5] remains the best source of comparator data for chemotherapy of the studies identified.

The company conducted a naïve comparison of avelumab with chemotherapy in the original CS using JM200: Part B and Study 100070-Obs001. In addition, supporting data from a further study, Iyer 2016[6], was provided in the original CS. The ERG acknowledges that the ToE only requires the company to use Study 100070-Obs001 for the comparator in data in this CDF review in the absence of new comparative clinical data. The ERG requested the company conduct a propensity score matching analysis during the clarification stage as the ERG remains concerned about the methodological robustness of using a naïve unadjusted comparison.

The ERG notes that in the economic model in the original CS the company used a naïve pooled analysis of seven studies identified from the SLRs and their own studies to provide PFS and OS data on chemotherapy in 1L mMCC patients (Study 100070-Obs001;[5]Iyer 2016;[7] Voog 1999;[8] Satpute 2014;[9] Santamaria-Barria 2013;[10] Fields 2011;[11] Allen 2005[12]). The company considered that all seven chemotherapy studies had similar outcomes thus justified the naïve pooling. The company fitted parametric curves to inform the base-case analysis in 1L patients and reported that the naïve pooling resulted in, "increased patient numbers for analysis, and likely the most generalisable results". The ERG however, considered that the approach is likely to introduce unnecessary heterogeneity into the analysis although it is not possible to predict the likely direction of the resulting bias. The ERG also notes that in the ToE it was specified that Study 100070-Obs001 should

be used to provide the comparator data for chemotherapy in the absence of new data and that this naïve pooled analysis was not referred to in the ToE.

The ERG also considers it important to highlight that it considers the immunocompetent subgroup from Study 100070-Obs001 (n=51) to be more appropriate for any comparison with JM200: Part B than the full study population of Study 100070-Obs001 (n=67). This is because the ERG's clinical experts suggested there may be differences in responses between immunocompetent and immunocompromised patients. The ERG notes that the population of JM200: Part B comprises only immunocompetent patients. The ERG also notes from the subgroup results of Study 100070-Obs001, that [redacted] (CQ response A14, Figure 15). In addition, the ERG considers the subgroup results to suggest immunocompromised patients have a [redacted] compared to immunocompetent patients, although there is [redacted]. The ERG acknowledges that Study 100070-Obs001 was not powered to detect subgroup differences in outcomes but nevertheless the ERG considers the immunocompetent subgroup of Study 100070-Obs001 is the most appropriate source of comparator data on chemotherapy compared to JM200: Part B.

3.2.1 Naïve comparison of avelumab versus chemotherapy

There were only 51 patients in the immunocompetent subgroup of Study 100070-Obs001 compared to 116 patients in JM200: Part B (all immunocompetent). The most frequently used chemotherapy regimens in the immunocompetent subgroup of Study 100070-Obs001 were carboplatin + etoposide (62.7% patients) and cisplatin + etoposide (17.6% patients), which the ERG's clinical experts reported was inline with clinical practice for mMCC in England. The median ToT in the immunocompetent subgroup of Study 100070-Obs001 was [redacted] months which is [redacted] than the median ToT of [redacted] in JM200: Part B. In addition, the ERG notes that [redacted] in the immunocompetent subgroup of Study 100070-Obs001 were still on treatment at 6 months compared to [redacted] of patients in JM200: Part B.

In terms of subsequent therapies, the ERG notes that 27% of patients (n=14) in the immunocompetent subgroup of Study 100070-Obs001 received at least one subsequent line of therapy after discontinuing from their original chemotherapy compared to 34.5% (n=40) in JM200: Part B. The most common second line treatments in the immunocompetent subgroup of Study

100070-Obs001 were topotecan (42.9%) and a combination of vincristine + cyclophosphamide + doxorubicin (28.6%), whereas in JM200: Part B, etoposide (15.5%) and carboplatin (13.8%) were more common. The ERG's clinical experts reported that these subsequent treatments were broadly in line with clinical practice in England although as already discussed, single agent etoposide is not commonly used.

Kaplan-Meier plots for the naïve comparison of avelumab with chemotherapy using JM200: Part B and the immunocompetent subgroup of Study 100070-Obs001 for the outcomes of ToT, PFS and OS are presented below (Figure 5, Figure 6 and Figure 7; respectively). The ERG notes that [redacted]

Figure 5. Time on treatment, JM200: Part B versus Study 100070-Obs001 immunocompetent subgroup (reproduced from Figure 23 from the company's response to clarification dated 27 March 2020)

[redacted]

Figure 6. Progression-free survival, JM200: Part B versus Study 100070-Obs001 immunocompetent subgroup (reproduced from Figure 21 from the company's response to clarification dated 27 March 2020)

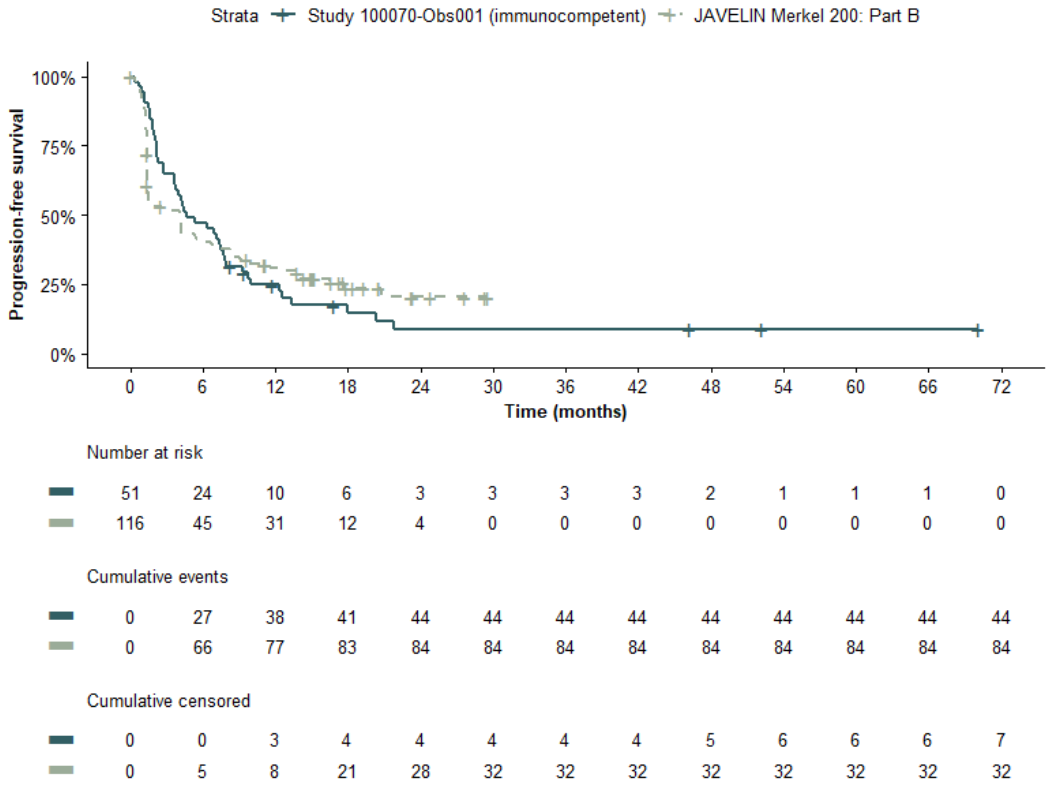
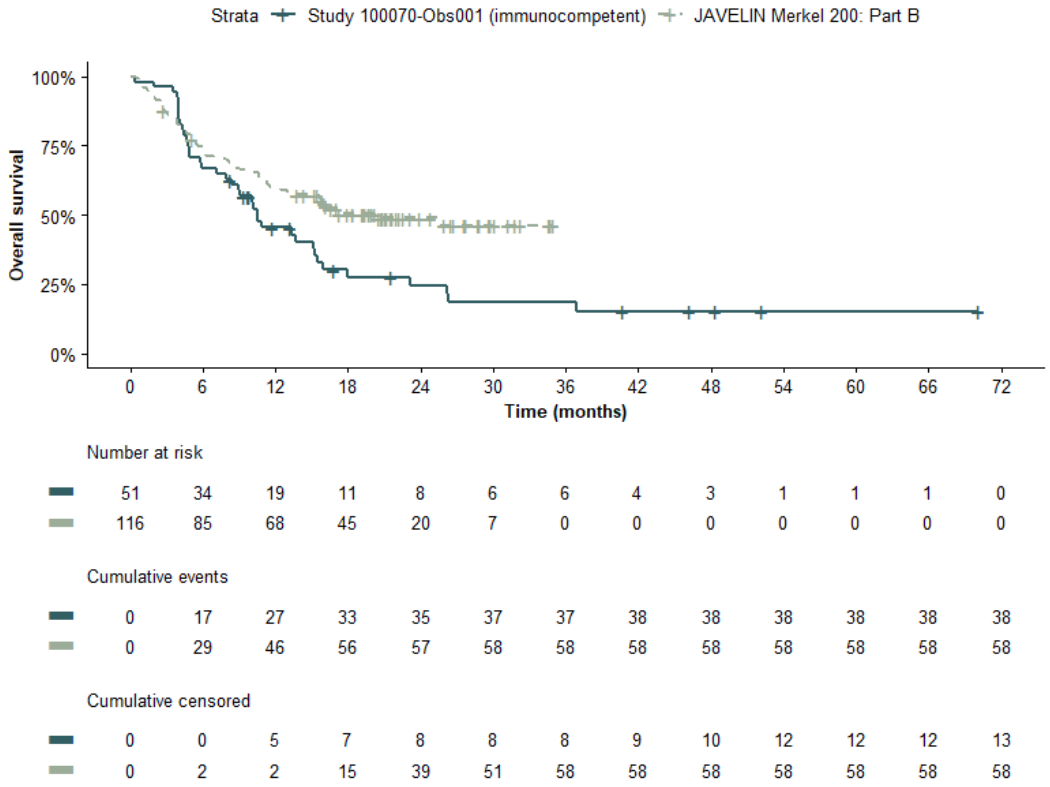


Figure 7. Overall survival, JM200: Part B versus Study 100070-Obs001 immunocompetent subgroup (reproduced from Figure 22 from the company's response to clarification dated 27 March 2020)



The median PFS for chemotherapy in the immunocompetent subgroup of Study 100070-Obs001 was slightly longer than the median PFS for avelumab in JM200: Part B (4.63 months versus 4.11 months) and the 6-month PFS rate was slightly higher in the immunocompetent subgroup of Study 100070-Obs001, although from 12-months onwards the PFS rate was higher in JM200: Part B (Table 11). However, median OS was longer in JM200: Part B than in the immunocompetent subgroup of Study 100070-Obs001 (20.3 months versus 10.5 months, respectively) and OS rates were consistently higher at 6, 12 and 15 months in JM200: Part B (Table 11).

Table 11. Progression-free survival (PFS) and overall survival (OS) for JM200: Part B and Study 100070-Obs001, immunocompetent subgroup Study 100070-Obs001 (adapted from Table 6 from the company's response to clarification dated 27 March 2020)

Outcome	JM200: Part B			Study 100070-Obs001, immunocompetent		
	Result	95% CI	Number in analysis	Result	95% CI	Number in analysis
Median PFS, months (95% CI)	4.11	(1.48 to 6.74)	116	4.63	(2.79 to 7.66)	51
•6-month PFS rate, % (95% CI)	41.3%	(33.1% to 51.5%)	[redacted]	47.1%	(33.0% to 60.0%)	[redacted]
•12-month PFS rate, % (95% CI)	31.0%	(23.5% to 41.0%)	[redacted]	24.8%	(13.8% to 37.4%)	[redacted]
•15-month PFS rate, % (95% CI)	27.0%	(19.8% to 36.8%)	[redacted]	17.3%	(8.1% to 29.5%)	[redacted]
Median OS, months (95% CI)	20.3	(13.0 to NE)	116	10.51	(7.16 to 15.24)	51
•6-month OS rate, % (95% CI)	74.8%	(67.3% to 83.2%)	[redacted]	66.7%	(52.0% to 77.8%)	[redacted]
•12-month OS rate, % (95% CI)	59.9%	(51.5% to 69.6%)	[redacted]	45.3%	(31.0% to 58.6%)	[redacted]
•15-month OS rate, % (95% CI)	57.2%	(48.8% to 67.1%)	[redacted]	40.3%	(26.2% to 53.9%)	[redacted]

Notes: 95% confidence interval limits were estimated using the default settings in the statistical software R. For the 'Number in analysis' column, we have populated this assuming the total sample size for the median value, and the number at risk for the values at a specific time point.
Abbreviations: CI, confidence interval; NE, not evaluable; OS, overall survival; PFS, progression-free survival.

3.3 Propensity score matching analysis of avelumab versus chemotherapy

In response to a clarification question from the ERG, the company conducted a series of propensity score analyses to compare avelumab and chemotherapy efficacy using JM200: Part B and Study 100070-Obs001. The company presented results for two different methods: propensity score matching (PSM), and propensity score weighting (PSW) and conducted a total of seven different sensitivity analyses (three with PSM and four with PSW); the different analyses are summarised below. The sensitivity analyses in the PSW method included using alternative methods, with inverse probability of treatment weighting (IPTW) used for one analysis and stabilised weights (SW) used for the other three analyses. The patient characteristics considered in the analyses were related to the

ones requested by the ERG (based on clinical expert opinion) that had available data from both studies (age, sex, ECOG PS and immune status) and therefore the company reported matching based on tumour PD-L1 expression status, MCPyV status and tumour burden could not be conducted. The company also reported that 13 patients from Study 100070-Obs001 had to be excluded from the propensity score matching analyses where ECOG status was included due to not having baseline ECOG PS data available.

- PSM analyses conducted by the company:
    - PSM 1. Age (aged ≥75 vs <75 years), sex (female vs male), ECOG PS (0 vs 1+), and immune status (immunocompetent versus immunocompromised);
    - PSM 2. Age (aged ≥75 vs <75 years), sex (female vs male), and ECOG PS (0 vs 1+), excluding immunosuppression as a variable;
    - PSM 3. Age (aged ≥75 vs <75 years), sex (female vs male), and ECOG PS (0 vs 1+), excluding immunosuppression as a variable and excluding patients with an ECOG PS of 2 or more.
- PSW analyses conducted by the company:
    - PSW 1. IPTW, based on all patients with available data for age (aged ≥75 vs <75 years), sex (female vs male), ECOG PS (0 vs 1+), and immune status;
    - PSW 2. SW, based on all patients with available data for age (aged ≥75 vs <75 years), sex (female vs male), ECOG PS (0 vs 1+), and immune status;
    - PSW 3. SW, based on all patients with available data for age (aged ≥75 vs <75 years), sex (female vs male), and ECOG PS (0 vs 1+), excluding immunosuppression as a variable;
    - PSW 4. SW, based on all patients with available data for age (aged ≥75 vs <75 years), sex (female vs male), and ECOG PS (0 vs 1), excluding immunosuppression as a variable and excluding patients with an ECOG PS of 2 or more.

The ERG considers the methods used by the company to conduct the propensity score analyses to be broadly consistent with those recommended in NICE DSU TSD17; further details on the company's propensity score analyses are provided in the company's clarification response to Question A1 (dated 27 March 2020).

The ERG notes that a limitation of the PSM analyses is the omission of patients from the analyses and that given the already small study size of Study 100070-Obs001 it has the greatest impact on the

availability of data for chemotherapy. The ERG therefore prefers the PSW analyses as they include data from a greater number of patients despite the omission of the 13 patients with no ECOG PS baseline data.

The ERG's preferred analysis is PSW4 as it maintains all patients in the analysis and has the best balance in baseline characteristics after matching for all characteristics matched other than immune status (i.e. age, ECOG PS and sex [Table 12]). The ERG thus presents the results of only OS and PFS for PSW4 below although the remaining PSW characteristics and Kaplan-Meier plots are presented in Appendix 9.2.1 to 9.2.3. The ERG acknowledges that its decision to focus on PSW4 is in stark contrast to the ERG's preferred naïve comparison where the immunocompetent subgroup of Study 100070-Obs001 was selected. The ERG still stands by its preference for the immunocompetent subgroup of Study 100070-Obs001 in the naïve comparison but considers PSW4 has the best balance in baseline characteristics after matching and enables matching of other potentially important characteristics that may impact on treatment efficacy despite the omission of matching for immune status. The ERG also considers PSW4 to be more reliable than the company's naïve comparison of JM200: Part B and Study 100070-Obs001 as the naïve comparison does not account for as many imbalances in the patient characteristics between the two studies.

The ERG also considers the potential direction of any bias resulting from differences in immune status to be predictable in PSW4 through the use of the study level subgroup data from Study 100070-Obs001. In particular, the ERG considers OS for chemotherapy is likely to be underestimated given that in the immunocompetent subgroup of Study 100070-Obs001 median OS is 10.5 months (N=51), whereas for the full study population OS is 10.1 months (N=67).[13] Median PFS is 4.6 months for both the immunocompetent subgroup and full study population of Study 100070-Obs001. Subgroup data for median OS and PFS are not reported for the immunosuppressed subgroup and therefore the exact impact of immune status on the results of PSW4 are unknown.

Table 12. Baseline characteristics for JM200: Part B and Study 100070-Obs001, re-weighted in PSW4 analysis (adapted from Tables 1 and 2 from the company's response to clarification dated 02 April 2020)

Baseline characteristic		JM200: Part B				Study 100070-Obs001			
		Unweighted		PSW 4		Unweighted		PSW 4	
		n	%	n	%	n	%	N	%
Age	<75	59.0	50.9	[redacted]	[redacted]	32.0	47.8	[redacted]	[redacted]
	>=75	57.0	49.1	[redacted]	[redacted]	35.0	52.2	[redacted]	[redacted]
Sex	Male	81.0	69.8	[redacted]	[redacted]	53.0	79.1	[redacted]	[redacted]
	Female	35.0	30.2	[redacted]	[redacted]	14.0	20.9	[redacted]	[redacted]

ECOG PS	0	72.0	62.1	[redacted]	[redacted]	14.0	20.9	[redacted]	[redacted]
	1	44.0	37.9	[redacted]	[redacted]	32.0	47.8	[redacted]	[redacted]
	2+	0	0	[redacted]	[redacted]	8.0	11.9	[redacted]	[redacted]
	Missing	0	0	[redacted]	[redacted]	13.0	19.4	[redacted]	[redacted]
Immunocompetent	Yes	116.0	100.0	[redacted]	[redacted]	51.0	76.1	[redacted]	[redacted]
	No	0	0	[redacted]	[redacted]	16.0	23.9	[redacted]	[redacted]

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PSW, propensity score weighting.

The resulting Kaplan-Meier plots after matching for PFS and OS in PSW 4 are presented as Figure 8 and Figure 9. The ERG notes that in both the Kaplan-Meier plots for PFS and OS [redacted] The ERG considers it important to highlight again that the data for chemotherapy from Study 100070- Obs001 includes weighting for [redacted]% of patients who are immunocompromised and that JM200: Part B has no immunocompromised patients. The ERG therefore considers the resulting efficacy estimates for chemotherapy may be skewed by the inclusion of immunocompromised patients and that the likely direction of bias in the analyses would favour avelumab. The ERG also considers it important to highlight that the adjusted PSW4 analysis resulted in [redacted] for chemotherapy compared to the use of the unadjusted full study population of Study 100070- Obs001 (Appendix 0, Figure 31 and Figure 33). [redacted] for avelumab compared to the unadjusted JM200: Part B trial level estimates (Appendix 0, Figure 30 and Figure 32).

Figure 8. Adjusted progression-free survival plot for PSW4 - JMM200: Part B versus Study 100070-Obs001 (reproduced from Figure 2 from the company's response to clarification dated 02 April 2020)

Figure 9. Adjusted overall survival plot for PSW4 – JM200: Part B versus Study 100070- Obs001 (reproduced from Figure 1 from the company's response to clarification dated 02 April 2020)

3.4 Conclusions of the clinical effectiveness section

In general, the ERG considers that the company has adhered to the committee's preferred assumptions from the ToE. The clinical data presented by the company includes the ToE required later data cut from JM 200: Part B and the observational SACT data that were also required to be collected by Public Health England during the period of managed access for avelumab. The ERG agrees that the company has focussed on the required first line mMCC population and the key comparator of chemotherapy.

The ERG notes that the marketing authorisation approved licensed dose of avelumab was changed in November 2019 to a flat dose of 800mg for all patients rather than the weight-based dose used in the original CS for TA517 and in the JM200 clinical trial. The company reported that the change in dose does not impact on the resulting efficacy of safety estimates derived from the weight-based dose as the new dose is considered to be equivalent.

JM200: Part B now comprises a minimum follow-up period of 15 months and includes more mature data for OS, PFS and ToT. The SACT data set comprises a minimum of 5 months follow-up and reports only on OS and ToT. The ERG considers there is still uncertainty in the clinical data despite the now more mature data from JM200: Part B. The ERG is concerned that the population in Javelin Merkel 200: Part B may be slightly younger, comprise of more males and have more favourable ECOG PS than expected in clinical practice in England. The ERG's clinical experts reported that they considered the SACT data set to be more representative of patients in clinical practice in England who would receive avelumab. However, the SACT data set does not provide information on PFS, HRQoL, response rate or adverse effects of treatment and the data for OS are immature with median OS not yet reached. In addition, the sample size from the SACT is smaller (N=52) compared to the latest analysis of JM200:Part B (N=116).

The ERG notes that in the ToE it is requested that the company conduct a naïve comparison of avelumab with chemotherapy using Study 100070-Obs001 to inform chemotherapy. However, the ERG is concerned that the use of the full trial population of Study 100070-Obs001 maybe confounded by the inclusion of patients with immunosuppression, especially given that immunosuppression was an exclusion factor in JM200: Part B, the study informing avelumab in the company's economic model. The ERG considers the subgroup analyses by immune status in Study 100070-Obs001 suggest

[redacted]

[redacted] As such, the ERG prefers the use of the subgroup of immunocompetent patients from Study 100070-Obs001 in the naïve comparison of avelumab versus chemotherapy. However, the ERG considers the company's naïve comparison of the JM200: Part B and Study 100070-Obs001, to be unreliable because of the imbalances in the patient characteristics between the two studies, the small number of patients in the studies, and the uncertainty caused by unmeasured variables that may be effect modifiers or prognostic indicators.

In response to clarification questions the company conducted propensity score adjusted analyses for the comparison of avelumab versus chemotherapy. The ERG's preferred analysis of the options presented was PSW4 which included adjustments for age (aged ≥75 vs <75 years), sex (female vs male), and ECOG PS (0 vs 1; patients with an ECOG PS of 2 or more were excluded). However, PSW4 included patients irrespective of immune status which the ERG would have preferred to be adjusted for in the analyses and the ERG thus considers PSW4 may overestimate the benefit of avelumab in comparison to chemotherapy.

[redacted] The company did not present any analyses of avelumab versus chemotherapy using the SACT data set although the ERG's clinical experts consider the SACT data set to be a better match to patients in clinical practice in England.

In summary, the ERG considers the results of any comparison of avelumab with chemotherapy to be uncertain due to the single-arm nature of the trials. In addition, the ERG is concerned about the potential mismatch in patient characteristics between JM200: Part B and the patients expected in clinical practice in England. The ERG nevertheless considers PSW4 the most robust comparison of the options presented by the company.

4 Cost effectiveness

4.1 Summary and critique of the company's submitted economic evaluation by the ERG

The company's submission (CS) for the CDF review of avelumab was largely performed using the same approach as per their original submission with the key areas updated as follows:

- Progression-free survival (PFS) and overall survival (OS) modelling for avelumab is now based on the treatment-naïve population data from the Javelin Merkel 200 trial rather than the treatment-experienced population data;
- Utility values are now based on the treatment-naïve population data;
- Avelumab acquisition costs now based on a flat-dose of 800mg rather than a weight-based dose of 10mg/kg, to align with the newly approved dose.

The following subsections will outline the company's latest submission in more detail, focusing mostly on the key areas that have changed since the original submission or that differ from the company's methods for the treatment-experienced population, for which avelumab was approved by NICE for routine commissioning.

4.1.1 Population

The company's updated analyses for the CDF review focus on the treatment-naïve population from their original submission whom were eligible for treatment funded through the CDF. That is, patients with metastatic Merkel cell carcinoma (mMCC) who have not received chemotherapy for metastatic disease.

The modelled population in the company's base case analysis is based on a combination of the Javelin Merkel 200 trial for avelumab and the Study 100070-Obs001 for the chemotherapy group, which were compared naïvely to inform the economic model. As discussed in Section 3.2, there were imbalances in the baseline characteristics between the two studies, so the ERG requested some additional analyses to adjust for key imbalances that were considered to be prognostic of disease progression and death. These analyses are discussed further in Section 4.1.5.

Another point to note regarding the population, is that the real-world Systemic Anti-Cancer Therapy (SACT) data population – those who received avelumab via the CDF – also differed from that of the Javelin Merkel 200 trial population. The ERG's clinical experts considered the SACT population to be

more reflective of the population expected to receive avelumab in clinical practice; however, the company considered the use of the SACT data to be unreliable due to the immaturity in the data and mismatch of the population treated with the Javelin Merkel 200 trial population. The ERG performed a scenario analysis using these data to test the impact on the results. This is discussed further in Section 4.1.5.

4.1.2 Interventions and comparators

Avelumab dosing has changed since the original submission and is now administered as an 800mg flat dose, using four 200mg vials. Previously a weight-based dose was given at a dose of 10mg/kg, which resulted in an average of 4.25x 200mg vials (849mg in total) required per administration, including wastage. This estimate was based on the method of moments, which estimates the distribution of the actual number of vials received by patients in order to calculate an accurate estimate of the mean dose for the population in the Javelin Merkel 200 trial.

The implications of the change to a new flat-dose of avelumab are discussed in terms of treatment effects and costs in Section 4.1.5 and Section 4.1.8, respectively.

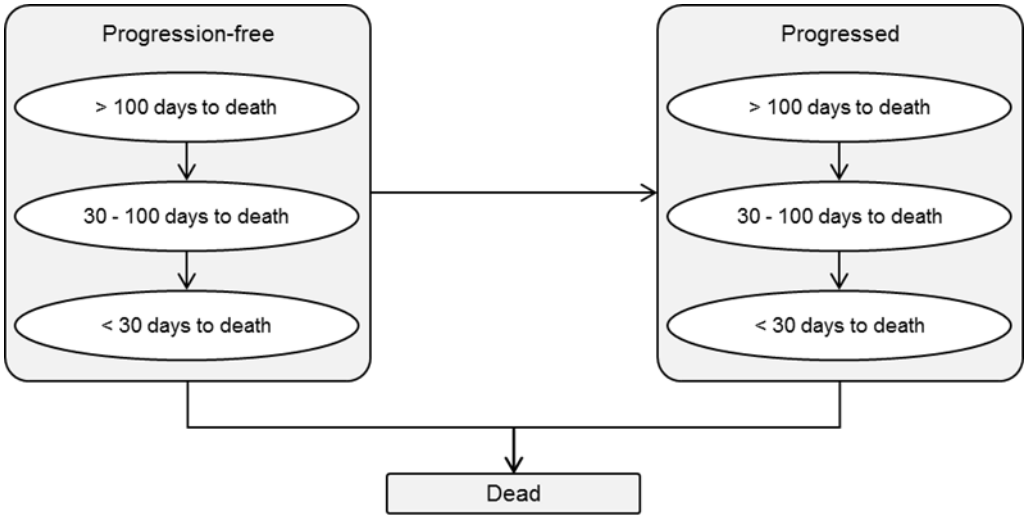
Chemotherapy is the comparator to avelumab for the treatment-naïve population and this has remained unchanged from the original submission.

4.1.3 Modelling approach and model structure

The model structure used for this CDF review is unchanged from that used in the original submission, which was accepted by the committee for the treatment-experienced population. The model is based on a partitioned survival structure with states for progression-free disease, progressed disease and death, and the proportions of patients in each state are estimated in weekly cycles. The progression-free and progressed disease states are split into three sub-states, which determine how close patients are to death. These sub-states, in the original submission were defined as "> 100 days to death", "30-100 days to death", and "<30 days to death", as shown in Figure 10. This approach was chosen to allow the application of utility decrements as patients approach death. The company's updated analysis using different time points for the groupings of time-to-death based on an updated utility analysis, which resulted in states defined as ">266 days to death", "35-266 days to death", and "<35 days to death". This is described further in Section 4.1.7.

The ERG considers the company's model structure to be suitable for decision making.

Figure 10. Model structure from the original submission



4.1.4 *Perspective, time horizon and discounting*

The perspective of the economic analysis is the same as in the original submission; that is, from the perspective of the NHS and personal social services.

The time horizon of the model is 40 years, which is considered to cover a lifetime. This was accepted by committee in the original submission for the treatment-experienced population and the ERG considers it to be reasonable.

Discounting was applied at an annual rate of 3.5% for both costs and QALYs as per the NICE reference case.

4.1.5 *Treatment effectiveness*

The methods for estimating treatment effectiveness for the treatment-naïve population have been amended since the company's original submission, which focused on estimating a relative effect between the treatment-naïve and treatment-experienced population outcomes of the Javelin Merkel 200 trial. These effects, measured as hazard ratios (HRs), were then applied to more robust survival models fitted to the treatment-experienced data to estimate the long-term outcomes for OS and PFS for treatment-naïve patients in the avelumab group. The ERG considered the company's original approach to be no more reliable than fitting survival models to the sparse data from the treatment-naïve population. A naïve comparison using Study 100070-Obs001, an observational study conducted by the company, was used to inform the chemotherapy group in the model.

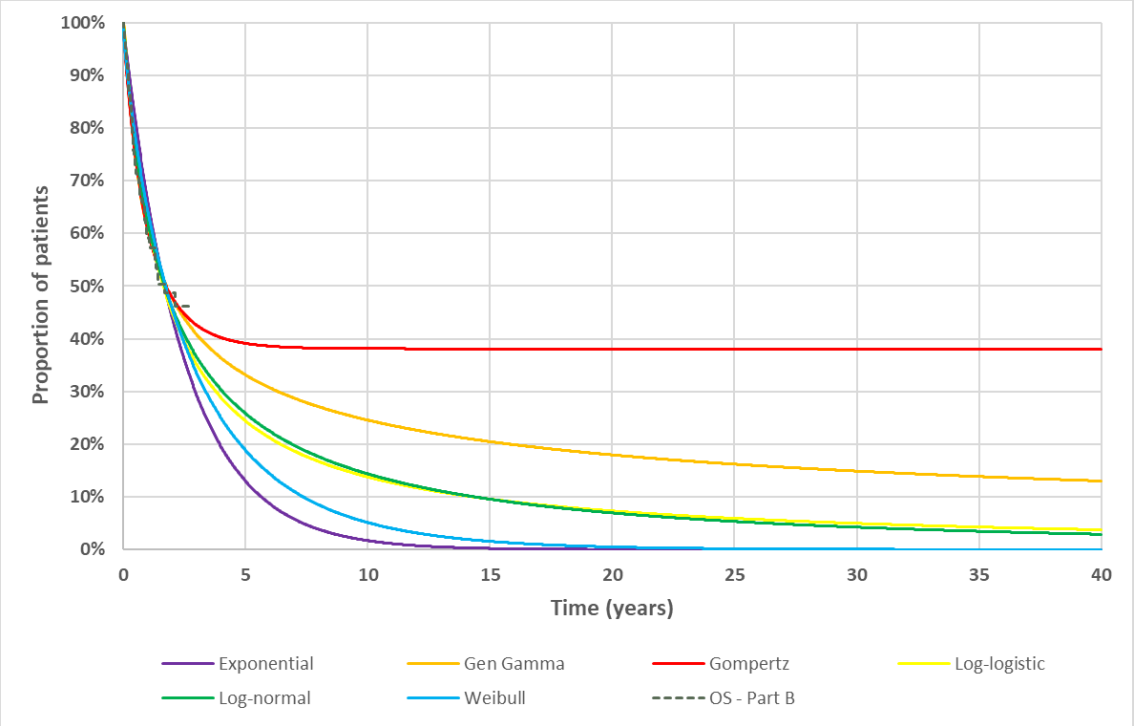
For the updated analyses provided for the CDF review, the company still relied on a naïve comparison, which used the same study to inform the chemotherapy group, but this time the company fitted survival models directly to the updated data from the treatment-naïve group of the Javelin Merkel 200 trial to inform the outcomes for the avelumab group. The methods applied by the company specifically to estimate long-term OS and PFS for the CDF review are described and critiqued in the following subsections.

4.1.5.1 Overall survival

To estimate the expected OS for avelumab, the company fitted a range of standard and spline-based parametric survival models to the updated treatment-naïve population data from the Javelin Merkel 200 trial. The company found that most of the standard models such as log-normal and log-logistic provided implausible extrapolations that overlapped with the latest Kaplan-Meier (KM) data from the treatment-experienced population of the Javelin Merkel 200 trial. An exception to this was the more flexible three-parameter generalised gamma model. The company found that the spline-based models provided a good fit but importantly also produced more plausible extrapolations that remained above the treatment-experienced KM data as would be expected. The standard parametric curves are shown in Figure 11 and the spline-base curves are shown in Figure 12.

To determine the best fitting curve, the company considered the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) as well as the visual fit and the plausibility of the extrapolations. The AIC and BIC statistics are given in Table 13, showing that the 1-knot normal spline had the lowest AIC and the log-normal had the lowest BIC. The company opted for the 1-knot odds spline given that it had essentially as good a fit as the 1-knot normal-based spline but the company considered the extrapolation, which was more favourable to avelumab, to be more plausible.

Figure 11. Standard overall survival models for avelumab



Abbreviations: OS, overall survival.

Figure 12. Spline-based overall survival models for avelumab

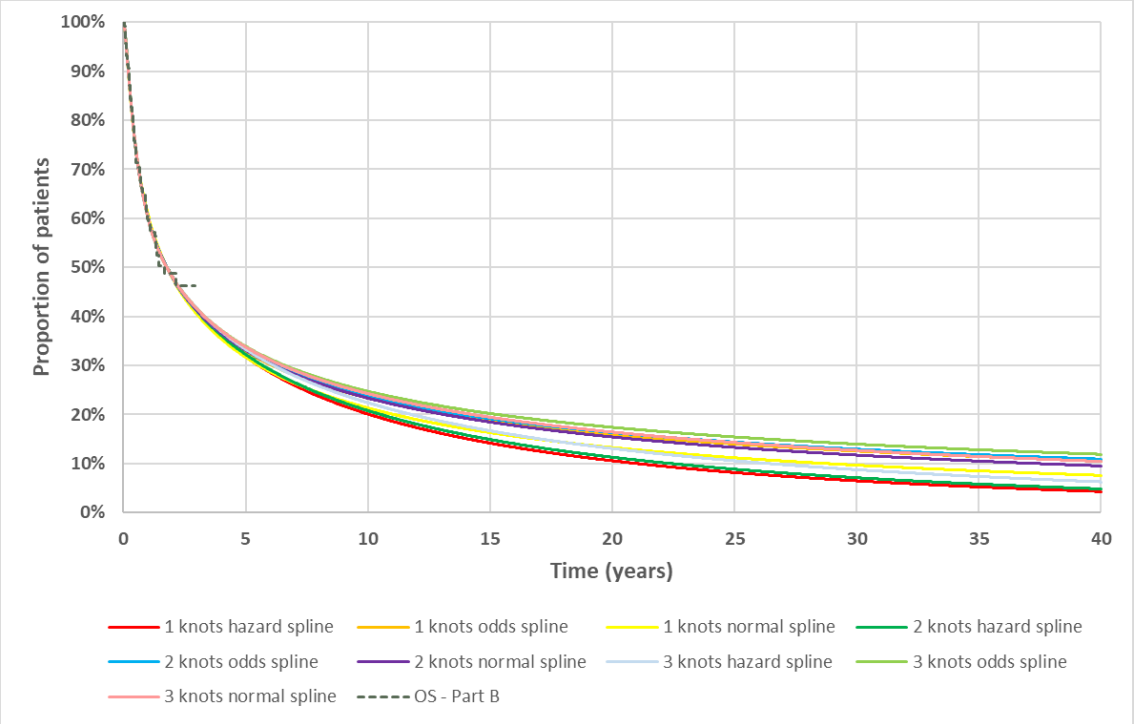


Table 13. Statistical goodness-of-fit scores for OS in Javelin Merkel 200

Model	AIC	BIC
1-knot, odds	501.76	510.02
1-knot, normal	**501.46**	509.72
1-knot, hazard	501.90	510.16
2-knots, odds	503.77	514.78
2-knots, normal	503.30	514.32
2-knots, hazard	503.90	514.92
3-knots, odds	505.68	519.45
3-knots, normal	505.25	519.01
3-knots, hazard	505.77	519.54
Exponential	510.52	513.27
Weibull	509.90	515.41
Gompertz	503.08	508.59
Log-logistic	505.37	510.87
Lognormal	502.04	**507.55**
Generalised gamma	501.05	509.31

Abbreviations: AIC, Akaike's information criterion; BIC, Bayesian information criterion; OS, overall survival.

The ERG considers it reasonable to discount the log-normal and log-logistic curves based on the implausible crossing of the treatment-experienced KM data, although notes that the tail of the KM curve is somewhat uncertain due to low numbers at risk. The ERG considers it reasonable, therefore, to focus the choice of OS curve on the 1-knot splines but notes important differences in the extrapolations produced by the hazard-, normal- and odds-based splines. The company's base case ICER changes from [redacted] per QALY to [redacted] per QALY and [redacted] per QALY for the normal-based and hazard-based splines, respectively. Given that there is uncertainty in the company's naïve comparison of treatment effects between avelumab and chemotherapy, the ERG considers it may be more appropriate to opt for a more conservative approach using the hazard-based 1-knot spline.

As a result of the uncertainty in the company's naïve treatment comparison, the ERG requested the company to provide an adjusted analysis based on propensity score matching, which may avoid the need to consider a conservative approach in the company's naïve analysis. This is discussed after a brief description and critique of the company's approach to estimating OS for the chemotherapy group.

The company conducted their own observational study (Study 100070-Obs001) to provide OS data for the chemotherapy group, but for the base case analysis they pooled the data with some additional chemotherapy studies. To extrapolate these data beyond the follow-up period the

company applied the same approach as described above for the avelumab group, although as the standard parametric models were considered sufficient, the spline-based models were not fitted. The curves fitted for OS are given in Figure 13, and the AIC and BIC statistics are given in Table 14.

Figure 13. Parametric curves fitted to OS data from Study 100070-Obs001

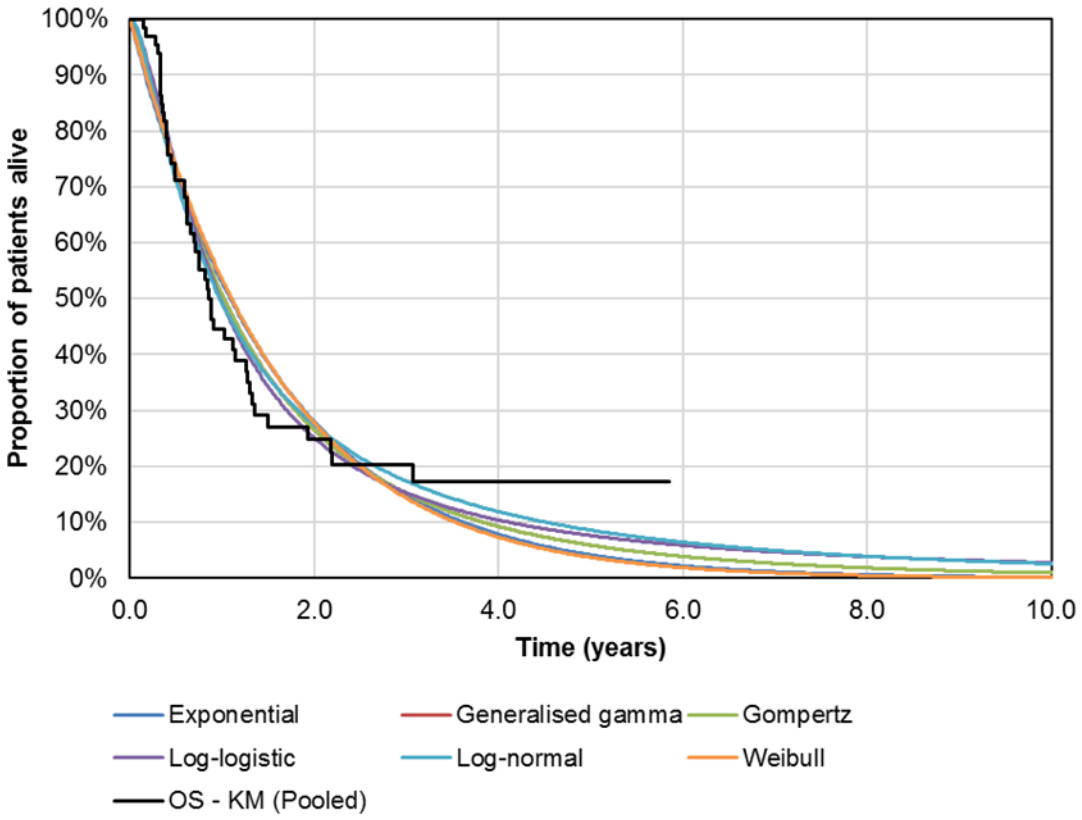


Table 14. Statistical goodness-of-fit scores for OS in Study 100070-Obs001

Model	AIC	BIC
Exponential	2,712	2,716
Weibull	2,714	2,721
Log-logistic	2,689	2,696
Log-normal	2,700	2,707
Generalised Gamma	2,699	2,709

Abbreviations: AIC, Akaike's information criterion; BIC, Bayesian information criterion; PFS, progression-free survival.

The ERG considers that the curves do not track the KM data particularly well in two parts; namely, in the initial few months and then for the duration of the second year. All extrapolations appear to be well below the KM plot towards the end of the follow-up period, although the data for the tail become more uncertain. Despite this, the ERG considers that it may have been useful to explore the

fit of more flexible spline-based models. As a minimum, the ERG considers it may be more suitable to choose the log-normal curve which appears to track the tail of the KM plot more closely. This change increases the company's base case ICER from [redacted] per QALY to [redacted] per QALY.

Before considering the curves for the ERG's preferred base case, we now consider the adjusted analyses that the company provided in response to clarification questions, as mentioned above.

The ERG requested the company to provide an adjusted comparison between Javelin Merkel 200 and Study 100070-Obs001 based on propensity score matching (PSM), with adjustment for age, sex, ECOG performance score, and immunocompetency. In response to this request, the company provided a range of analyses based on PSM as well as propensity score weighting (PSW). Within each of these two approaches provided by the company, additional analyses were also provided that considered matching/weighting without immunocompetency included as a variable, and also with patients who had an ECOG score of 2 or more excluded from the chemotherapy group, as there were no patients in Javelin Merkel 200 with an ECOG score greater than 1. A full discussion of these analyses is provided in Section 3.3.

The company's PSM analyses resulted in a large amount of data being excluded, particularly from the Javelin Merkel 200 trial in order to provide a good match. The PSW analyses avoid the need to remove patients to provide a suitable match, therefore, the ERG considers the company's PSW analyses to be preferable. For these analyses, the ERG submitted a further request for the company to provide "adjusted" baseline characteristics to assess the balance for the key variables across the two trials after adjustment. These are discussed further in Section 3.3.

All of the PSW analyses appear to result in a greater proportion of younger patients in the avelumab group, which may, therefore, overestimate the benefits in favour of avelumab. On the whole, in terms of age and sex, the analyses that do not include immunocompetency as a variable in the propensity score estimation provide the most balanced values across the two trials. This is also the case for ECOG score, particular for the analysis that also excludes those with a score of 2 or more in the chemotherapy group. The only potentially less preferable aspect of the latter analysis is that the chemotherapy group has a large proportion of patients who are immunosuppressed. However, this makes the analysis overly optimistic but with a clearer direction of effect as the chemotherapy outcomes are likely to be underestimated relative to avelumab with the inclusion of immunosuppressed patients. This analysis may, therefore, produce a useful lower bound for the ICER. There is also still an imbalance in favour of avelumab for age, which adds further evidence that the results of this analysis may still overestimate the relative effectiveness of avelumab compared to chemotherapy. Even this more balanced analysis may, therefore, still underestimate the ICER.

As an alternative approach, the ERG requested the company to provide a naïve comparison using the SACT OS data for the avelumab group. The ERG's clinical experts considered the SACT population to be more reflective of the population expected to receive avelumab in clinical practice, so the ERG considered this a useful scenario analysis. However, as the data are immature the company did not consider this analysis to be reliable and, therefore, did not provide this analysis.

The ERG notes the company's concerns with the data but considers it worthwhile to assess the impact of using the potentially more reflective population data within the economic analysis. Therefore, the ERG conducted the analysis by digitising the SACT OS KM data and fitting survival curves using R software,[14] following the same general approach as the company. Standard parametric curves and spline-based parametric curves are shown in Figure 14 and Figure 15, respectively. The AIC and BIC statistics are given in Table 15.

Figure 14. Standard parametric curves fitted to SACT OS avelumab data

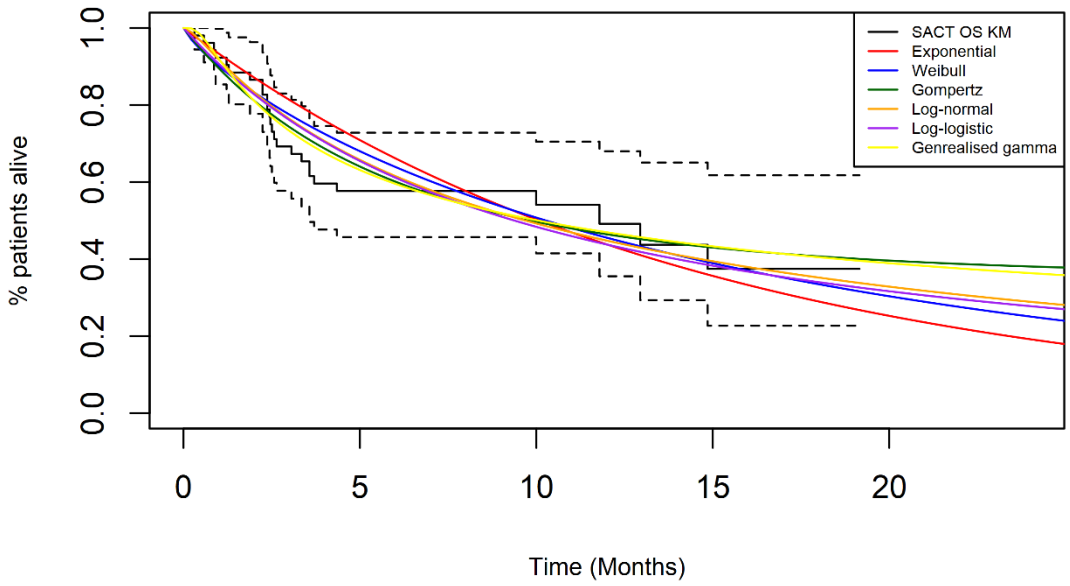


Figure 15. Spline-base curves fitted to SACT OS avelumab data

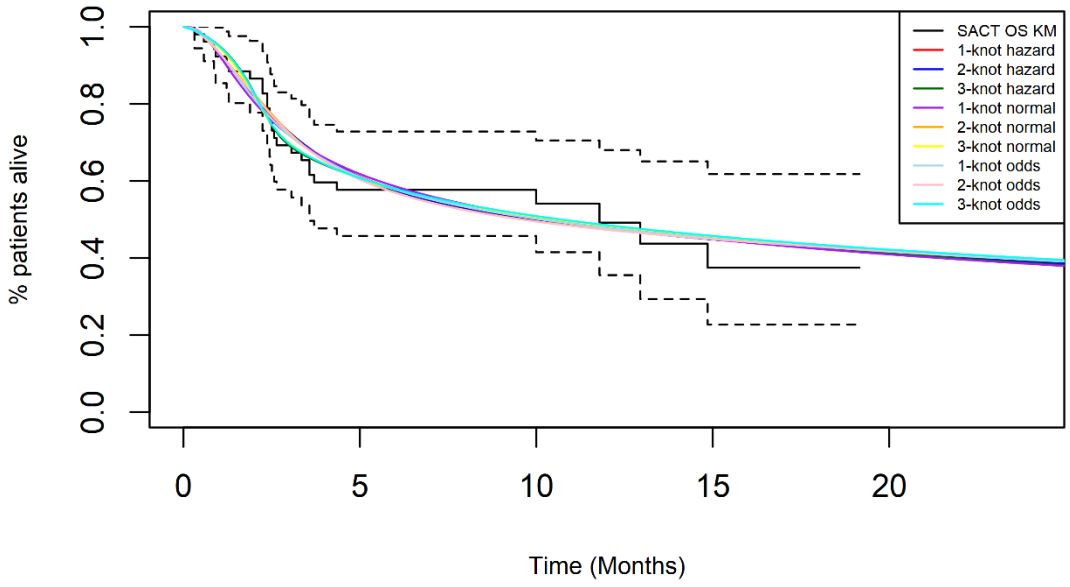


Table 15. Statistical goodness-of-fit scores for the SACT OS curves

Model	AIC	BIC
1-knot, odds	186.62	192.47
1-knot, normal	187.12	192.98
1-knot, hazard	**186.61**	**192.46**

2-knots, odds	188.26	196.06
2-knots, normal	187.81	195.61
2-knots, hazard	188.56	196.37
3-knots, odds	188.18	197.94
3-knots, normal	188.34	198.09
3-knots, hazard	187.72	197.48
Exponential	193.19	195.14
Weibull	193.60	197.50
Gompertz	190.30	194.20
Log-logistic	190.98	194.89
Lognormal	188.93	192.83
Generalised gamma	188.17	194.02

Abbreviations: AIC, Akaike's information criterion; BIC, Bayesian information criterion; OS, overall survival.

The ERG considers the spline-based models to provide a superior fit to the data and considers the 1-knot hazard spline, which has the lowest AIC and BIC, to be suitable to estimate OS in a scenario analysis. Scenario analyses using this curve are given in Section 6.

The next subsection will discuss the company's approach to estimating and extrapolating PFS outcomes.

4.1.5.2 Progression-free survival

The company applied the same general approach to estimate and extrapolate PFS outcomes as they did for OS. The standard parametric curves fitted to the PFS data from the Javelin Merkel 200 trial are given in Figure 16 and the spline-based curves are given in Figure 17. AIC and BIC statistics are given in Table 16, showing that the best fitting curve was the 3-knot odds-based spline, as determined by both AIC and BIC.

Figure 16. Standard parametric curves fitted to PFS data from Javelin Merkel 200 trial

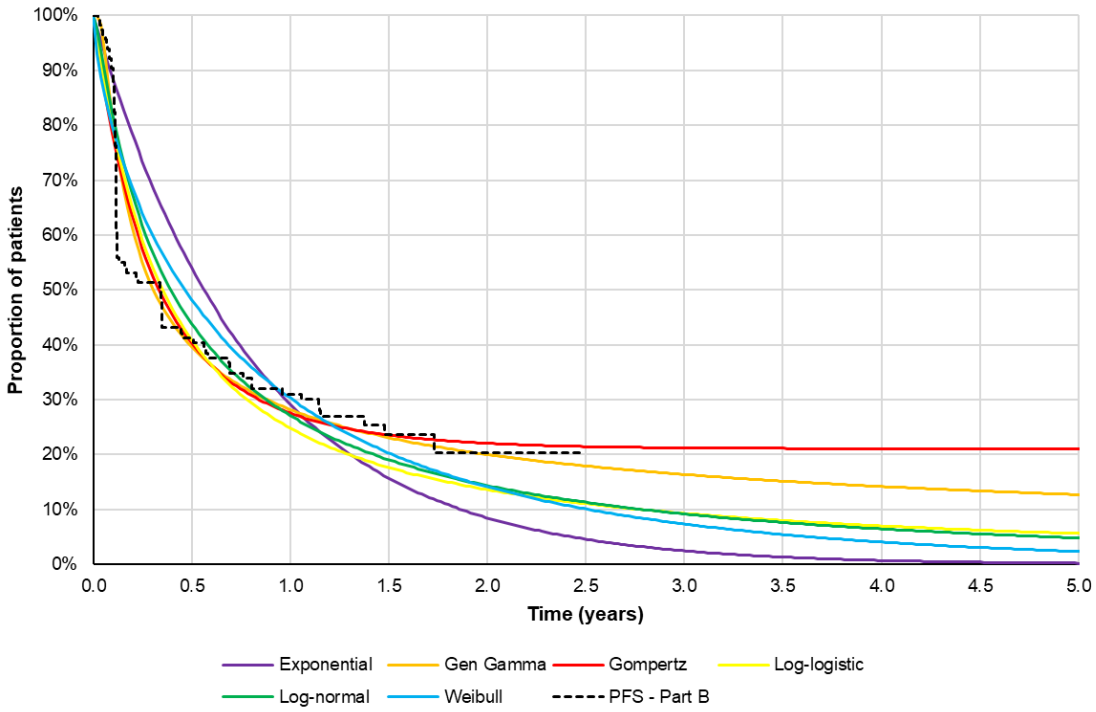


Figure 17. Spline-base curves fitted to PFS data from Javelin Merkel 200 trial

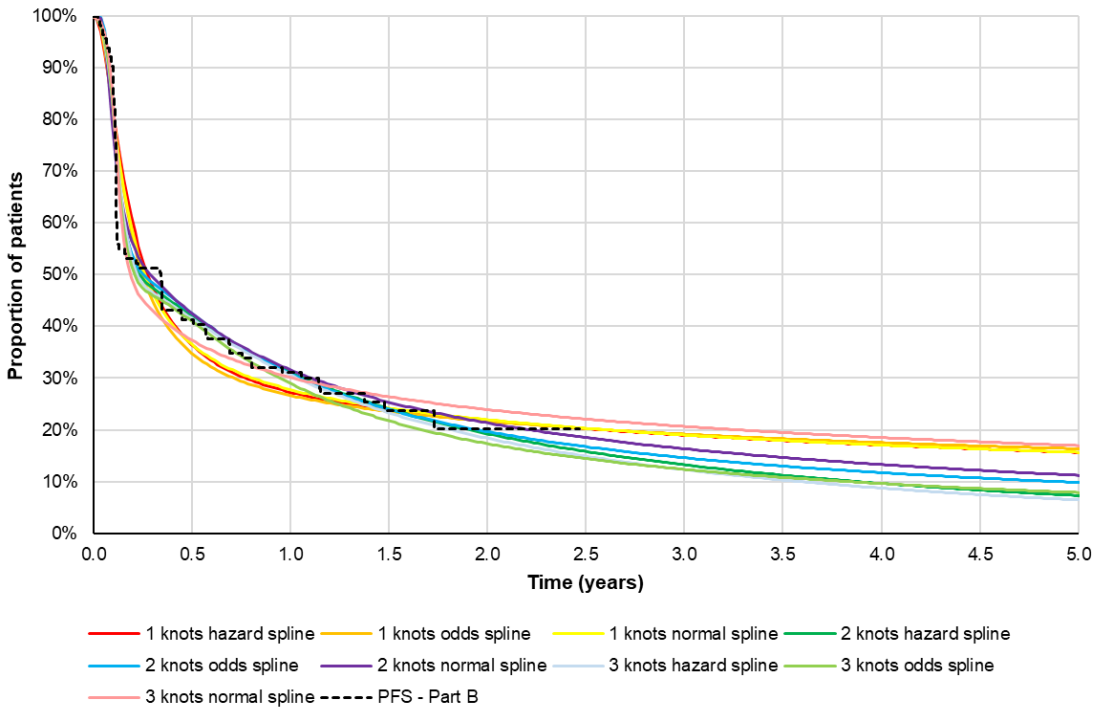


Table 16. Statistical goodness-of-fit scores for PFS

Model	AIC	BIC
1-knot, odds	483.25	491.51
1-knot, normal	481.43	489.69
1-knot, hazard	488.24	496.50
2-knots, odds	462.71	473.72
2-knots, normal	473.65	484.66
2-knots, hazard	463.94	474.95
3-knots, odds	**455.04**	**468.81**
3-knots, normal	-	-
3-knots, hazard	461.70	475.46
Exponential	552.28	555.03
Weibull	536.83	542.34
Gompertz	514.33	519.83
Log-logistic	517.22	522.73
Lognormal	512.03	517.54
Generalised gamma	486.91	495.17

Abbreviations: AIC, Akaike's information criterion; BIC, Bayesian information criterion; PFS, progression-free survival.

Despite the optimal goodness-of-fit statistics for the 3-knot splines, the company chose to use the 2-knot odds spline for their base case, as they considered the visual fit of the curve to be preferable. The ERG disagrees with this and considers the 2-knot spline to underestimate the KM data between 0.5 years and 1 year, and then more importantly appears to overestimate the KM data for the tail. The ERG considers, at least based on the naïve comparison used in the company's base case, that the 3-knot odds spline provides a better extrapolation as well as being the best fit to the data.

As per the OS modelling, in response to clarification questions the company also provided adjusted analyses due to the imbalances across the trials used for the naïve comparison. Before discussing these adjusted analyses, the ERG will first discuss the company's methods to estimate PFS outcomes for chemotherapy in the naïve comparison used for the company's base case.

The company applied the same approach to estimate PFS outcomes for chemotherapy as they did for OS, using data from Study 100070-Obs001 pooled with additional studies (See Section 3.2) and fitting parametric survival curves to extrapolate. As per the OS modelling, the company considered the standard parametric models to be sufficient and so did not consider the more flexible spline-based models.

The PFS parametric curves for chemotherapy are given in Figure 18 and the AIC and BIC statistics are given in Table 17, showing the log-logistic as the best fitting curve based on both AIC and BIC. The company chose the log-logistic curve for their base case and the ERG considers this reasonable for the company's naïve comparison. However, as per the OS analysis, the ERG considers an adjusted analysis to be preferable.

Figure 18. Parametric curves fitted to PFS data from Study 100070-Obs001

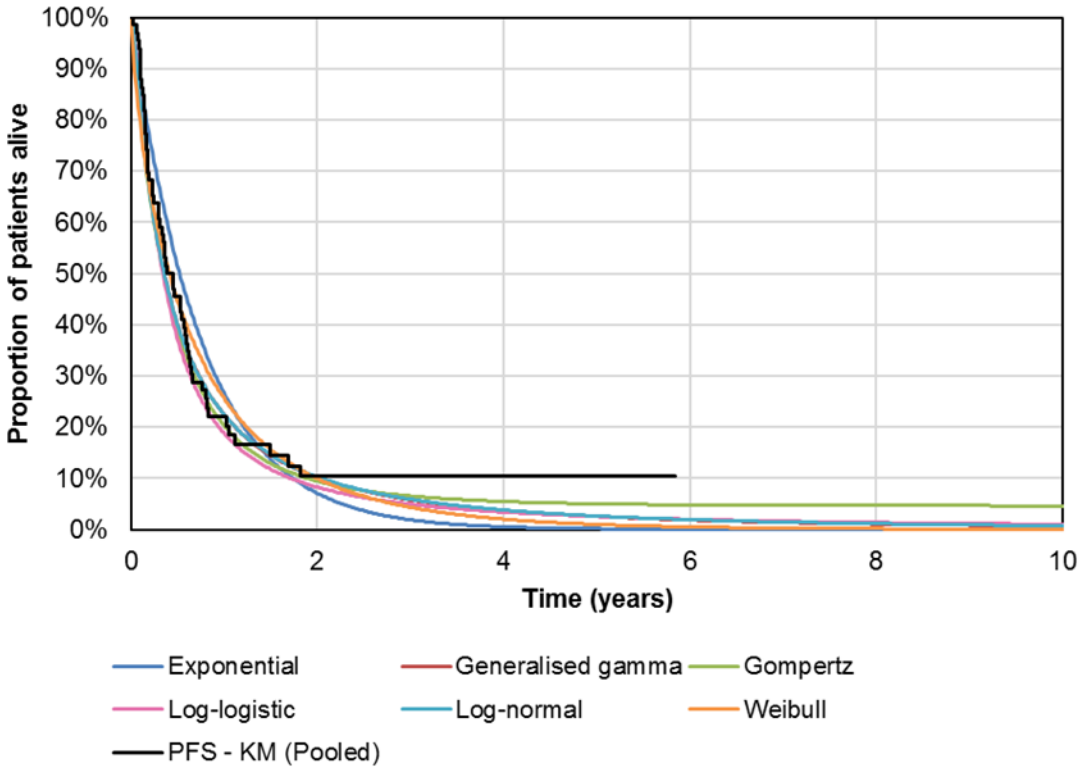


Table 17. Statistical goodness-of-fit scores for PFS in Study 100070-Obs001

Model	AIC	BIC
Exponential	1604	1607
Weibull	1586	1592
Gompertz	1558	1564
Log-logistic	1549	1554
Log-normal	1560	1565
Generalised Gamma	1562	1570

Abbreviations: AIC, Akaike's information criterion; BIC, Bayesian information criterion; PFS, progression-free survival.

The adjustments applied to the OS analysis apply equally to the PFS analysis and for the reasons outlined previously for the OS analysis, the ERG considers the PSW4 weighted analysis is preferable

as it provided the best balance in characteristics across the two trials. As with the OS analysis, it may also be considered to overestimate the relative effectiveness in favour of avelumab based on the inclusion of immunosuppressed patients and a lower proportion of younger patients in the chemotherapy study compared to the avelumab trial.

As described previously for OS, the ERG considered the possibility of informing PFS with data from SACT. Although PFS data were no collected in SACT, the ERG considered using TTD collected from SACT as a proxy for PFS. Therefore, the ERG conducted scenario analyses using the curves fitted to the TTD data, which is discussed further in Section 4.1.5.3. The scenario analyses are given in Section 6.

4.1.5.3 Time on treatment

The company's approach to estimating time-on-treatment for avelumab used parametric curves fitted to the treatment-naïve data for the first 15 months and then switched to the extrapolation of curves fitted to the treatment-experienced population data from Javelin Merkel 200. This was justified by the company because the minimum follow-up for the treatment-naïve population was 15 months, whereas for the treatment-experienced population it was 36 months. The company's adjusted curve is compared against the unadjusted curve fitted to the treatment-naïve population data from Javelin Merkel 200. Both company's chosen curve for the treatment-naïve population (as well as the treatment-experienced population) was the Weibull, although this did not have the best AIC and BIC statistics, as shown in Table 18.

Figure 19. Time on treatment extrapolation (CS, Appendix 3, page 39, Figure 15).

Table 18. Statistical goodness-of-fit scores for time on treatment

Model	AIC	BIC
1-knot, odds	1,217.40	1,225.66
1-knot, normal	1,217.00	1,225.26
1-knot, hazard	1,216.10	1,224.36
2-knots, odds	1,219.27	1,230.28
2-knots, normal	1,218.05	1,229.06
2-knots, hazard	1,218.18	1,229.20
3-knots, odds	1,210.65	1,224.41
3-knots, normal	1,210.65	1,224.42
3-knots, hazard	**1,209.55**	1,223.31
Exponential	1,242.81	1,245.56
Weibull	1,216.39	1,221.90
Gompertz	1,225.60	1,231.11
Log-logistic	1,215.42	**1,220.93**
Lognormal	1,217.38	1,222.89
Generalised gamma	1,216.77	1,225.04

Abbreviations: AIC, Akaike's information criterion; BIC, Bayesian information criterion.

Although the ERG considers the choice of curve to be reasonable, the ERG considers that the curves fitted directly to the treatment-naïve population data should be used for the economic analysis and

should not be adjusted by the treatment-experienced population data on the basis of longer minimum follow-up. The ERG considers it reasonable to stop treatment at 5 years as this is likely to happen in clinical practice based on the ERG's clinical expert input.

As discussed previously for OS and PFS, the ERG considered the use of SACT data to provide scenario analyses in the economic model using the potentially more reflective population from SACT. The ERG fitted survival curves to the TTD data collected in SACT using R software, following the same general approach as the company did for their analyses. Standard parametric and spline-based parametric curves fitted to the SACT TTD data are shown in Figure 20 and Figure 21, respectively.

Figure 20. Standard parametric curves fitted to SACT TTD avelumab data

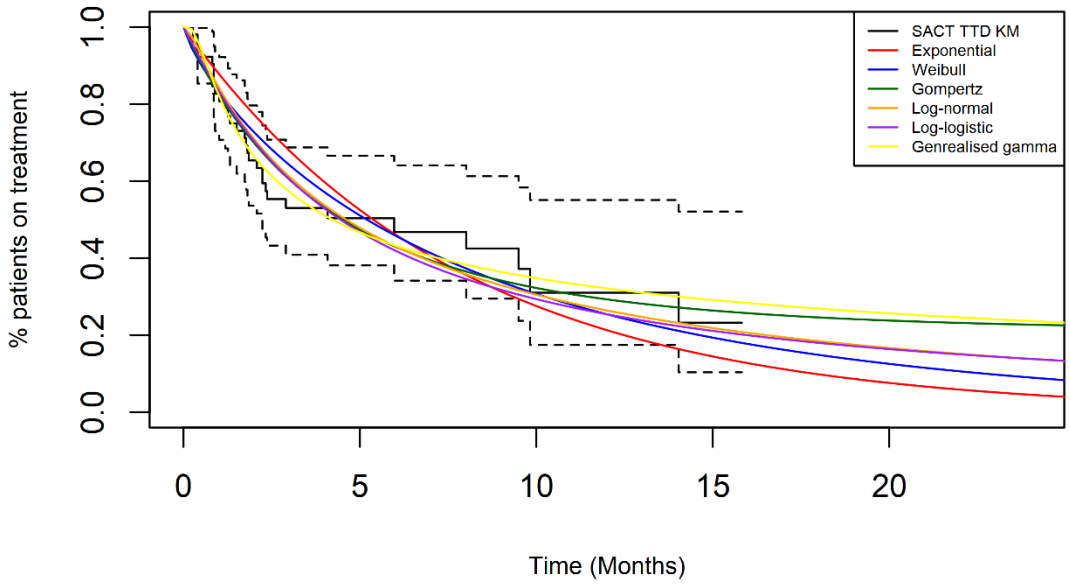


Figure 21. Spline-based curves fitted to SACT TTD avelumab data

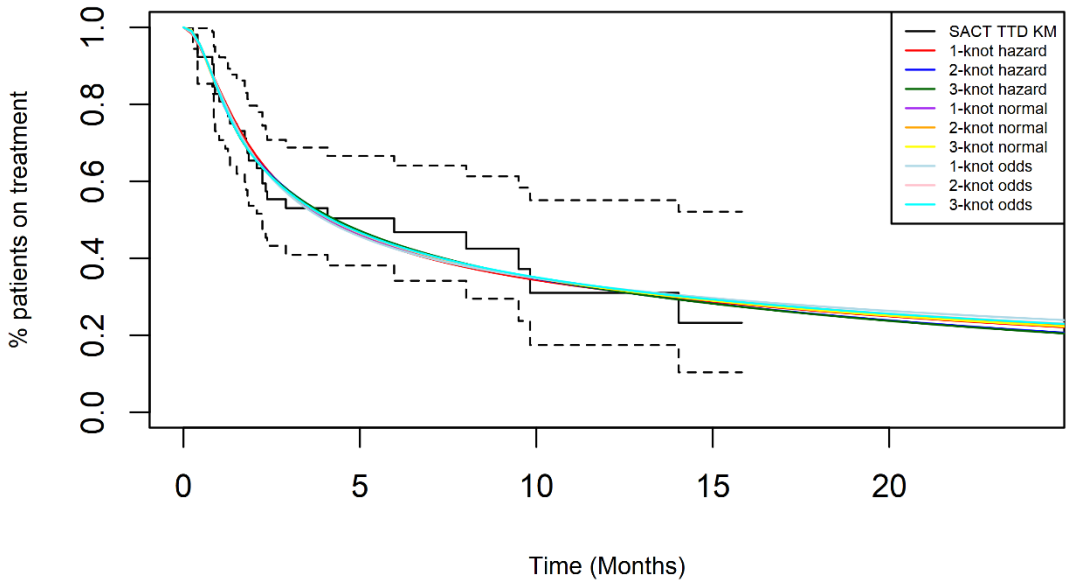


Table 19. Statistical goodness-of-fit scores for SACT TTD curves

Model	AIC	BIC
1-knot, odds	178.16	184.01
1-knot, normal	177.38	183.23
1-knot, hazard	178.40	184.26
2-knots, odds	179.83	187.63
2-knots, normal	179.35	187.16
2-knots, hazard	179.68	187.49
3-knots, odds	181.82	191.57
3-knots, normal	181.34	191.09
3-knots, hazard	181.63	191.39
Exponential	184.95	186.90
Weibull	184.87	188.77
Gompertz	181.87	185.77
Log-logistic	181.20	185.11
Lognormal	178.96	**182.86**
Generalised gamma	**177.36**	183.21

Abbreviations: AIC, Akaike's information criterion; BIC, Bayesian information criterion.

The ERG chose the 1-knot normal spline as the most plausible fit as the AIC and BIC statistics were very close to the best curves for each statistic, and the curve resulted in a plausible extrapolation. Scenario analyses using this curve to inform TTD in the economic model are given in Section 6.

4.1.6 Adverse events

Adverse events (AEs) were included in the model based on grade 3 or 4 AEs that were experienced by at least 5% of patients in either the avelumab or chemotherapy trials. This approach was used in the original submission and was accepted by the committee for the treatment-experienced population. The ERG considers the company's approach to be reasonable and also notes that AEs are not a key driver of the cost effectiveness results.

4.1.7 Health-related quality of life

The Javelin Merkel 200 trial collected EQ-5D-5L data for both the treatment-experienced and treatment-naïve populations. In the company's original submission, as data were limited for the treatment-naïve population, the company used data from the treatment-experienced population to inform the health state utility values (HSUVs) for the treatment-naïve population. Using these data, the company firstly used a crosswalk algorithm to estimate EQ-5D-3L scores as per the NICE preferred approach and then performed a regression analysis to estimate effects for health state and time to death on these utility scores to fully inform the economic model.

This approach was accepted by the committee for the treatment-experienced population in the original appraisal and the company used the same approach for the treatment-naïve population. Given the updated data for the treatment-naïve population, the company were able to provide three alternative scenario analyses with regard to utility values, using the following populations from the Javelin Merkel 200 trial:

- Treatment-experienced population (as per the original submission);
- Treatment-naïve population;
- Combined population with a prior treatment variable included.

The company's approach for this CDF review was largely the same; however, the methods used to estimate the best fitting groupings for the time-to-death approaches were more sophisticated and included more constraints to determine the most suitable time points. The constraints applied to estimate the time-to-death health states within the regression analysis were as follows:

- Health states must be defined in multiples of 7 days to align with the model cycles;
- Health states must be at least 14 days long;

- Health states that are further away from death must have a duration at least as long as those that are closer to death;
- Health states were required to be informed by at least 10 utility values to ensure a reliable estimation;
- Changes in utility between distinct health states must be at least 0.08, based on the minimum important difference in utility for cancer.[15]

The company used mean absolute error (MAE), root mean squared error (RMSE) and quasi-information criterion (QIC) to determine the goodness-of-fit of the resulting models. The best fitting model was identified with cut points of <35 days, 35-266 days and >266 days. The company also provided a sensitivity analysis based loosely around the cut-points from the original submission but more aligned with the model cycles by constraining them to multiples of 7 days. For the this, the company used <28 days instead of the original <30 days cut-off, and 84 days instead of the upper 100 days cut-off. The ERG is unsure why the company did not use a value of 98 days for the upper limit for this sensitivity analysis. The resulting utility values from each of these analyses is given in Table 20. A comparison of the resulting HSUVs from each analysis is given in Table 21.

Table 20. Results of the company's utility analyses

Label	Health state	Utility value/coefficient		
		Treatment-naïve	Treatment-experienced	Combined population
Original TA517 values	> 100 days to death		0.7744	
	30-100 days to death		0.7540	
	< 30 days to death		0.7082	
Optimal cut-points	> 266 days to death	0.8128	0.7561	0.8019
	35-266 days to death	0.6893	0.6943	0.7096
	< 35 days to death	0.4206	0.4174	0.4411
	Coefficient: Treatment-experienced			-0.0348
Sensitivity analysis cut-points	> 84 days to death	0.7837	0.7494	0.7839
	28-84 days to death	0.6487	0.6208	0.6525
	< 28 days to death	0.3951	0.2804	0.3513
	Coefficient: Treatment-experienced			-0.0349

Table 21. HSUVs applied in economic model

Description	Average utility value
Original TA517 values	0.7744
Optimal cut-points (35, 266), Combined population	0.7753
Optimal cut-points (35, 266), Treatment-naïve	0.7453
Optimal cut-points (35, 266), Treatment-experienced	0.7401
Sensitivity analysis cut-points (28, 84), Combined population	0.7881

Sensitivity analysis cut-points (28, 84), Treatment-naïve	0.7750
Sensitivity analysis cut-points (28, 84), Treatment-experienced	0.7958

The company also provided an alternative health-state based approach to estimating utilities in their original submission, and this analysis appears to have been updated using the updated utility data from the Javelin Merkel 200 trial, although the company have not described this fully in their submission. The company provided options to apply progression-free and progressed disease utilities based on either the treatment-experienced population (as per the original submission), the treatment-naïve population or the combined data from Javelin Merkel 200. The values for each approach are given in Table 22.

Table 22. Progression-based HSUVs

Health state	Treatment-experienced	Treatment-naïve	Combined
Progression-free	0.7350	0.7746	0.7587
Progressed disease	0.6843	0.7265	0.7090

The ERG considers the company's time-to-death approach to be reasonable as it captures the deterioration in utility over time as patients get closer to death. The ERG also considers the updates to the approach taken to estimate the groupings applied to this approach are reasonable and considers the estimates used in the model to be reliable.

4.1.8 Resource use and costs

The company's approach to estimating resource use and costs was largely the same as the approach used in the original submission, which was accepted by committee for the treatment-experienced population. Two key aspects which have now either changed since the original submission or differed from the accepted approach taken for the treatment-experienced analysis, are the dosing of avelumab and the application of subsequent treatments.

The company's updated analyses have been updated to include the recently approved new dosing for avelumab, which is now a flat dose of 800mg. This approach has led to a reduction in acquisition costs as the original weight-based dosing resulted in an average of 4.25x 200mg vials (849mg in total, including wastage) required per administration. Given that this latter approach was based on the doses that patients actually received in the Javelin Merkel 200 trial, the ERG considers that this approach should have been applied in the economic model so that the costs are in line with the effects measure in the trial. Applying this approach in the model increases the company's base case

ICER from [redacted] per QALY to [redacted] per QALY. The ERG has included this change in the ERG's preferred base case analysis presented in Section 6.4.

The company's original submission did not include subsequent treatments in the economic analysis, so the ERG requested the company to include subsequent treatments that aligned with the studies used to inform the treatment effectiveness. The company provided this in response to clarification questions but noted some limitations given that duration of treatment had to be assumed and some treatments had confidential patient access scheme discounts that were not available.

The company also provided two alternative approaches where the costs of topotecan were assumed to be equivalent to other chemotherapy costs and another where no monoclonal antibody treatments were included as well as the change to topotecan costs. These analyses increased the ICER by £19 to £581 depending on the analysis chosen.

5 Cost effectiveness results

5.1 Company's cost effectiveness results

The company's base case results are presented in Table 23, showing an ICER of [redacted] per QALY gained for avelumab compared to chemotherapy.

Table 23. Company's deterministic cost effectiveness results

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Chemotherapy	11,116	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life-year.

5.1.1 Company's sensitivity analyses

The company provided a probabilistic sensitivity analysis (PSA) based on 1,000 samples, to assess the impact of parameter when all parameters are varied simultaneously in the economic model. The results of the PSA are given in Table 24, showing a slightly decreased ICER of [redacted] per QALY compared to the deterministic base case ICER. The results of all 1,000 sampled results are presented on the cost effectiveness plane in Figure 22.

Table 24. Company's probabilistic cost effectiveness results

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Chemotherapy	[redacted]	1.95	1.33	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER, incremental cost effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life-year.

Figure 22. [redacted]

The company also conducted a range of one-way sensitivity analyses to assess the impact of varying each parameter individually. The results of these are shown in the tornado plot in Figure 23.

Figure 23. [redacted]

5.1.2 Company's scenario analyses

The company presented a range of scenario analyses in the CS using alternative approaches for OS, ToT and utility values. These scenarios are described alongside the results in Table 25.

Table 25. Company's key scenario analyses

Scenario and cross reference	Scenario detail	Brief rationale	Impact on base-case ICER (£) (+/- from base case)
Base case			[redacted]
OS for avelumab	Use generalised gamma model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (generalised gamma) which projects higher OS versus the base-case analysis	[redacted]
	Use 1-knot normal spline-based model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (1-knot normal spline) which projects lower OS versus the base-case analysis	[redacted]
Clinical expectation of ToT	Assume all patients discontinued by 10 years, with no interim capping at 2 years	Approach intended to serve as an upper bound of potential long-term treatment with avelumab. In practice, discontinuation with avelumab is expected to occur before 5 years	[redacted]
	Assume one-third of patients continue treatment after 2 years, and all discontinue by 5 years	Approach aligned with clinical expert opinion at the time of the original TA517 CS, and is still expected to be broadly representative of clinical practice	[redacted]
Utility values	Use original utility values from TA517	Allows for assessment of impact on cost-effectiveness results through updating utility values	[redacted]
	Use only data from JM200: Part B to inform utility values	Allows exploration of using utility values derived only from a treatment-naïve metastatic MCC population	[redacted]

Abbreviations: BSC, best supportive care; CS, company submission; ICER, incremental cost-effectiveness ratio; JM200, JAVELIN Merkel 200; OS, overall survival.

In response to clarification questions, the company also provided the following range of additional analyses:

- Range of PSM and PSW analyses for PFS and OS (Table 26);
- Immunocompetent subgroup analysis (Table 27); and,
- A range of scenarios to include subsequent treatments (Table 28).

Table 26. Scenario analyses with alternative adjusted OS and PFS analyses

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company's base-case analysis							
Chemotherapy	[redacted]	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PSW, all patients (stable weights), using age, sex, ECOG (0 vs. 1+), excluding immunosuppression as a variable							
Chemotherapy	[redacted]	2.19	1.49	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PSW, all patients (stable weights), using age, sex, ECOG (0 vs 1), excluding immunosuppression as a variable and removing ECOG 2+ pts							
Chemotherapy	[redacted]	2.20	1.50	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PSM using age, sex, ECOG (0 vs 1+), excluding immunosuppression as a variable							
Chemotherapy	[redacted]	1.85	1.27	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PSM using age, sex, ECOG (0 vs 1), excluding immunosuppression as a variable and removing ECOG 2+ pts							
Chemotherapy	[redacted]	1.92	1.31	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abbreviations: ECOG, Eastern Cooperative Oncology Group Performance Status; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PSM, propensity score matched; PSW, propensity score weighted; QALYs, quality-adjusted life years.							

Table 27. Scenario analyses with alternative subgroup OS and PFS analysis

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company's base-case analysis							
Chemotherapy	[redacted]	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Analysis using only the immunocompetent subgroup							
Chemotherapy	[redacted]	1.83	1.25	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abbreviations: CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.							

Table 28. Scenario analyses with alternative subsequent treatment approaches applied

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company's base-case analysis							
Chemotherapy	[redacted]	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Include all subsequent therapies as per study sources							
Chemotherapy	[redacted]	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Update topotecan costs to standard chemotherapy cost.							

Chemotherapy	[redacted]	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Remove monoclonal antibodies and update topotecan costs							
Chemotherapy	[redacted]	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.							

5.1.3 Model validation and face validity check

The company provided their updated analyses in a new version of the economic model that the company had improved in terms of performance and usability. The ERG considered the model to be sound and suitable for decision making after errors were corrected in the original submission (TA517) and these corrections have been carried through to the CDF review.

The company provided options in the updated economic model to reproduce the results of the original analyses from TA517, which further validates the reliability of the company's updated model.

6 Additional economic analysis undertaken by the ERG

6.1 Model corrections

The ERG had no further corrections to make in the company's revised analysis and considers the company's model to be sound and suitable for decision making.

6.2 Exploratory and sensitivity analyses undertaken by the ERG

The ERG conducted a range of scenario analyses to assess the impact of different assumptions applied for OS, PFS and ToT in terms of the best fitting curves as well as the assumption of weight-based dosing. The ERG also explored the impact of using the SACT data to inform OS and TTD, as well as using the TTD data as a proxy to inform PFS. The ERG applied the curves described in Section 4.1.5 for OS, PFS, and TTD, individually, and then applied all changes in a combined scenario analysis. The results are given in 6.3.

6.3 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

Table 29. Results of the ERG's scenario analyses

	Results per patient	Avelumab	Chemotherapy	Incremental value
0	**Company base case**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
1	**Weight-based dosing for avelumab acquisition costs**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
2	**OS: 1-knot hazard spline**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
3	**PFS: 3-knot odds spline**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
4	**ToT: 3-knot hazard spline**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
5	**ToT: No change to treatment-experienced data**			
	Total costs (£)	[redacted]	[redacted]	[redacted]

	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
6	**OS: ERG's SACT OS curves applied**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
7	**PFS: ERG's SACT TTD curves used as a proxy for PFS**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
8	**ToT: ERG's SACT TTD curves applied**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]
9	**ERG's SACT curves applied for OS, PFS and TTD (Scenarios 6+7+8)**			
	Total costs (£)	[redacted]	[redacted]	[redacted]
	QALYs	[redacted]	[redacted]	[redacted]
	ICER			[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

6.4 ERG preferred assumptions

The ERG's preferred base case includes the following changes compared to the company's base case analysis:

1. PSW analyses for OS and PFS with immunocompetency excluded from the estimation of propensity scores and patients with ECOG score 2 or more removed;
2. 1-knot hazard spline for OS;
3. 3-knot odds spline for PFS;
4. 3-knot hazard spline for ToT;
5. Removing the adjustment to the ToT curve using the treatment-experienced population data;
6. Weight-based acquisition costs for avelumab in line with effectiveness data.

These changes resulted in an ICER of [redacted] per QALY for the ERG's preferred base case. The cumulative results as each change is applied are given in Table 30.

Table 30. ERG's preferred model assumptions

Preferred assumption	Section in ERG report	Cumulative ICER £/QALY
Company base case	-	[redacted]

PSW analyses for OS and PFS with immunocompetency excluded from the estimation of propensity scores and patients with ECOG score 2 or more removed	Section 4.1.5	[redacted]
1-knot hazard spline for OS	Section 4.1.5	[redacted]
3-knot odds spline for PFS	Section 4.1.5	[redacted]
3-knot hazard spline for ToT	Section 4.1.5	[redacted]
Removing the adjustment to the ToT curve using the treatment-experienced population data	Section 4.1.5	[redacted]
Weight-based acquisition costs for avelumab	Section 4.1.8	[redacted]
ERG's preferred base case ICER		[redacted]

Abbreviations: ERG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

6.5 Conclusions of the cost effectiveness section

The company's analyses initially had a large degree of uncertainty as they were based on a naïve comparison of treatment effects. In response to clarification questions the company reduced this uncertainty by providing a range of adjusted analyses using propensity scores to match or weight the populations in the different studies in order to achieve balance in the patient characteristics.

The PSW analysis chosen by the ERG in their preferred base case achieved a better balance in the patient characteristics but notably still had a proportion of patients in the chemotherapy group who were immunosuppressed. This variable could not be adjusted properly as the avelumab study included no immunosuppressed patients. This imbalance is likely to underestimate the effectiveness of chemotherapy relative to avelumab, and therefore, the ERG's preferred base case ICER may still be an underestimate of the true ICER. Another slight imbalance that may underestimate the ICER is that there were more younger people in the avelumab study, therefore, potentially further overestimating the relative benefit of avelumab compared to chemotherapy.

Another key point to consider is how reflective the population in the ERG's preferred weighted analysis is compared to the population expected to receive avelumab in clinical practice. The ERG's clinical experts considered the SACT data to be more reflective of clinical practice and that showed that avelumab patients performed less well in terms of OS compared with those in the Javelin Merkel 200 trial. The impact of these population differences on the chemotherapy group is unclear so this remains an outstanding area of uncertainty.

The remainder of the company's analysis was generally sound with one additional area of uncertainty being in the dosing of avelumab. If administration with a flat dose of 800mg can provide the same effectiveness as the weight-based dosing, then the ERG's preferred base case ICER may be slightly overestimated. The ERG's preferred base case with a flat dose applied is [redacted] per QALY compared to [redacted] per QALY with weight-based dosing. This is, however, still an area of uncertainty.

7 End of Life

In terms of meeting the NICE end of life criteria, both the company's and the ERG's preferred analyses clearly meet the gain in survival of 3 months required. The company's base case analysis generates an expected gain of [redacted], while the ERG's preferred base case generates a gain in survival of [redacted]. Although the ERG considers the ERG's preferred approach may still overestimate the relative benefit in favour of avelumab, the gain will still almost certainly be above the threshold of 3 months.

In terms of the baseline survival for the chemotherapy group, the results are much closer to the threshold of 2 years. The company's base case analysis results in mean life-years (LYs) for the chemotherapy group of 1.94 years, whereas the ERG's preferred base case results in mean LYs of 2.20 years.

Although the ERG considers the relative effectiveness estimates in the ERG's preferred base case are more reliable, the ERG notes that the baseline characteristics of the company's studies used in the PSW analyses are different from those expected in clinical practice so the baseline survival estimates may not be fully reflective of clinical practice.

Given that the SACT data – a population considered by the ERG's clinical experts to be more reflective of the population expected in clinical practice – demonstrated worse outcomes for avelumab than the Javelin Merkel 200 trial, then it is likely that chemotherapy outcomes in a population similar to the SACT data may have worse outcomes than those estimated in the PSW analyses, which were matched to the Javelin Merkel 200 population. It is possible, therefore, that the expected survival for patients on chemotherapy is less than 2 years. However, there is a large degree of uncertainty with no reliable estimate of the true expected survival.

8 References

1. European MEdicines Agency (EMA). Avelumab Summary of Product Characteristics, 2019. Available from: https://www.ema.europa.eu/en/documents/product-information/bavencio-epar-product-information_en.pdf. Date accessed: April 2020.
2. National Institue for Health and Care Excellence. Avelumab for treating metastatic Merkel cell carcinoma [TA517], 2018. Available from: https://www.nice.org.uk/guidance/ta517/chapter/1-Recommendations. Date accessed: 06 Apr 2020.
3. National Institue for Health and Care Excellence. Terms of engagement for CDF review: Avelumab for treating metastatic Merkel cell carcinoma (TA517). 2018.
4. Public Health England. Avelumab for treating metastatic Merkel cell carcinoma - data review. 2019.
5. Cowey L BJ, Bharmal M. Retrospective Observational Study to Evaluate Treatment Outcomes in Patients with Metastatic Merkel Cell Carcinoma Following Chemotherapy Observational Study Protocol/Analysis Plan. 2016.
6. Iyer JG, Blom A, Doumani R, Lewis C, Tarabadkar ES, Anderson A, et al. Response rates and durability of chemotherapy among 62 patients with metastatic Merkel cell carcinoma. *Cancer Med* 2016; **5**: 2294-301.
7. Iyer JG, Blom A, Doumani R, Lewis C, Tarabadkar ES, Anderson A, et al. Response rates and durability of chemotherapy among 62 patients with metastatic Merkel cell carcinoma. *Cancer Med* 2016; **5**: 2294-301.
8. Voog E, Biron P, Martin J-P, Blay J-Y. Chemotherapy for patients with locally advanced or metastatic Merkel cell carcinoma. *Cancer* 1999; **85**: 2589-95.
9. Satpute S AN, Einhorn LH. Role of platinum-based chemotherapy for Merkel cell tumor in adjuvant and metastatic settings. *Journal of Clinical Oncology* 2014; **32**: A9049.
10. Santamaria-Barria JA, Boland GM, Yeap BY, Nardi V, Dias-Santagata D, Cusack JC. Merkel Cell Carcinoma: 30-Year Experience from a Single Institution. *Annals of Surgical Oncology* 2012; **20**: 1365-73.
11. Fields RC, Busam KJ, Chou JF, Panageas KS, Pulitzer MP, Allen PJ, et al. Five Hundred Patients With Merkel Cell Carcinoma Evaluated at a Single Institution. *Annals of Surgery* 2011; **254**: 465-75.
12. Allen PJ, Bowne WB, Jaques DP, Brennan MF, Busam K, Coit DG. Merkel Cell Carcinoma: Prognosis and Treatment of Patients From a Single Institution. *Journal of Clinical Oncology* 2005; **23**: 2300-9.
13. Cowey CL, Mahnke L, Espirito J, Helwig C, Oksen D, Bharmal M. Real-world treatment outcomes in patients with metastatic Merkel cell carcinoma treated with chemotherapy in the USA. *Future Oncol* 2017; **13**: 1699-710.
14. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
15. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes* 2007; **5**: 70.

9 Appendices

9.1 Additional clinical results tables

Table 31. ToT numbers for JAVELIN Merkel 200: Part B (reproduced from Table 2 from the company's response to clarification dated 27 March 2020)

Time intervals (months)	0	6	12	18	24	30	36
ToT number at risk	[redacted]						
ToT censored							
ToT events (ended treatment)							

Abbreviation: ToT, time on treatment.

Table 32. ToT results for JAVELIN Merkel 200: Part B (reproduced from Table 3 from the company's response to clarification dated 27 March 2020)

Outcome	Result	95 % CI	Number in analysis
Median ToT, months	[redacted]		
•6-month ToT rate, %			
•12-month ToT rate, %			
•15-month ToT rate, %			

Notes: 95% confidence interval limits were estimated using the default settings in the statistical software *R*. For the 'Number in analysis' column, we have populated this assuming the total sample size for the median value, and the number at risk for the values at a specific time point.
Abbreviations: CI, confidence interval; ToT, time on treatment.

Table 33. OS and PFS numbers for JAVELIN Merkel 200: Part B (reproduced from Table 4 from the company's response to clarification dated 27 March 2020)

Time intervals (months)	0	6	12	18	24	30	36
OS number at risk	116	85	68	45	20	7	0
OS censored	[redacted]						
OS events							
PFS number at risk	116	45	31	12	4	0	0
PFS censored	[redacted]						
PFS events							

Abbreviations: OS, overall survival; PFS, progression-free survival.

Table 34. OS and PFS reasons for censoring in JAVELIN Merkel 200: Part B (reproduced from Table 5 from the company's response to clarification dated 27 March 2020)

Outcome	Reasons for censoring
OS	[redacted]
PFS	[redacted]

Table 35. Median outcomes for subgroup analyses in JM200: Part B (reproduced from Table 12 from the company's response to clarification dated 27 March 2020)

Subgroup	Result	Lower CI	Upper CI	Number
OS				
Aged <80 years	[redacted]	[redacted]	[redacted]	84
Aged >= 80 years	[redacted]	[redacted]	[redacted]	32
ECOG = 0	[redacted]	[redacted]	[redacted]	72
ECOG = 1	[redacted]	[redacted]	[redacted]	44
MCPyV = positive	[redacted]	[redacted]	[redacted]	70
MCPyV = negative	[redacted]	[redacted]	[redacted]	37
MCPyV = not estimable	[redacted]	[redacted]	[redacted]	9
PD-L1 = positive	[redacted]	[redacted]	[redacted]	21
PD-L1 = negative	[redacted]	[redacted]	[redacted]	87
PD-L1 = not estimable	[redacted]	[redacted]	[redacted]	8
Aged <75 years	[redacted]	[redacted]	[redacted]	59
Aged >= 75 years	[redacted]	[redacted]	[redacted]	57
PFS				
Aged <80 years	[redacted]	[redacted]	[redacted]	84
Aged >= 80 years	[redacted]	[redacted]	[redacted]	32
ECOG = 0	[redacted]	[redacted]	[redacted]	72
ECOG = 1	[redacted]	[redacted]	[redacted]	44

MCPyV = positive	[redacted]	[redacted]	[redacted]	70
MCPyV = negative	[redacted]	[redacted]	[redacted]	37
MCPyV = not estimable	[redacted]	[redacted]	[redacted]	9
PD-L1 = positive	[redacted]	[redacted]	[redacted]	21
PD-L1 = negative	[redacted]	[redacted]	[redacted]	87
PD-L1 = not estimable	[redacted]	[redacted]	[redacted]	8
Aged <75 years	[redacted]	[redacted]	[redacted]	59
Aged >= 75 years	[redacted]	[redacted]	[redacted]	57

Abbreviations: CI, confidence interval; ECOG, Eastern Cooperative Oncology Group; NE, not evaluable; OS, overall survival; PD-L1, Protein programmed-death ligand-1; PFS, progression-free survival.

9.2 Propensity score matching - additional tables and figures

9.2.1 Baseline characteristics

Table 36. Baseline characteristics for JM200: Part B and Study 100070-Obs001 re-weighted as per PSW1 (adapted from Tables 1 and 2 from the company's response to clarification dated 02 April 2020)

Characteristic		JM200: Part B		Study 100070-Obs001	
		N	%	n	%
Age	<75	[redacted]	[redacted]	[redacted]	[redacted]
	>=75	[redacted]	[redacted]	[redacted]	[redacted]
Sex	Male	[redacted]	[redacted]	[redacted]	[redacted]
	Female	[redacted]	[redacted]	[redacted]	[redacted]
ECOG PS	0	[redacted]	[redacted]	[redacted]	[redacted]
	1	[redacted]	[redacted]	[redacted]	[redacted]
	2+	[redacted]	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]	[redacted]
Immunocompetent	Yes	[redacted]	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PS, performance status; PSW, propensity score weighting.

Table 37. Baseline characteristics for JM200: Part B and Study 100070-Obs001 re-weighted as per PSW2 (adapted from Tables 1 and 2 from the company's response to clarification dated 02 April 2020)

Characteristic		JM200: Part B		Study 100070-Obs001	
		n	%	n	%
Age	<75	[redacted]	[redacted]	[redacted]	[redacted]
	>=75	[redacted]	[redacted]	[redacted]	[redacted]
Sex	Male	[redacted]	[redacted]	[redacted]	[redacted]

	Female	[redacted]	[redacted]	[redacted]	[redacted]
ECOG PS	0	[redacted]	[redacted]	[redacted]	[redacted]
	1	[redacted]	[redacted]	[redacted]	[redacted]
	2+	[redacted]	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]	[redacted]
Immunocompetent	Yes	[redacted]	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PS, performance status; PSW, propensity score weighting.

Table 38. Baseline characteristics for JM200: Part B and Study 100070-Obs001 re-weighted as per PSW3 (adapted from Tables 1 and 2 from the company's response to clarification dated 02 April 2020)

Characteristic		JM200: Part B		Study 100070-Obs001	
		n	%	n	%
Age	<75	[redacted]	[redacted]	[redacted]	[redacted]
	>=75	[redacted]	[redacted]	[redacted]	[redacted]
Sex	Male	[redacted]	[redacted]	[redacted]	[redacted]
	Female	[redacted]	[redacted]	[redacted]	[redacted]
ECOG PS	0	[redacted]	[redacted]	[redacted]	[redacted]
	1	[redacted]	[redacted]	[redacted]	[redacted]
	2+	[redacted]	[redacted]	[redacted]	[redacted]
	Missing	[redacted]	[redacted]	[redacted]	[redacted]
Immunocompetent	Yes	[redacted]	[redacted]	[redacted]	[redacted]
	No	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ECOG, Eastern Cooperative Oncology Group; PS, performance status; PSW, propensity score weighting.

9.2.2 OS

Figure 24. Adjusted OS plot for PSW1 – JM200: Part B versus Study 100070-001 (reproduced from Figure 1 from the company's response to clarification dated 02 April 2020)

Figure 25. Adjusted OS plot for PSW2 – JM200: Part B versus Study 100070-001 (reproduced from Figure 1 from the company's response to clarification dated 02 April 2020)

Figure 26. Adjusted OS plot for PSW3 – JM200: Part B versus Study 100070-001 (reproduced from Figure 1 from the company's response to clarification dated 02 April 2020)

9.2.3 PFS

Figure 27. Adjusted OS plot for PSW1 – JM200: Part B versus Study 100070-001 (reproduced from Figure 2 from the company's response to clarification dated 02 April 2020)

Figure 28. Adjusted OS plot for PSW2 – JM200: Part B versus Study 100070-001 (reproduced from Figure 2 from the company's response to clarification dated 02 April 2020)

Figure 29. Adjusted OS plot for PSW4 – JM200: Part B versus Study 100070-001 (reproduced from Figure 2 from the company's response to clarification dated 02 April 2020)

9.2.4 Unadjusted versus adjusted study level results for PSW4

Figure 30. Adjusted PFS plot for PSW4 – JM200: Part B (reproduced from Figure 3 from the company's response to clarification dated 02 April 2020)

Figure 31. Adjusted PFS plot for PSW4 – Study 100070-001 (reproduced from Figure 4 from the company's response to clarification dated 02 April 2020)

Figure 32. Adjusted OS plot for PSW4 – JM200: Part B (reproduced from Figure 5 from the company's response to clarification dated 27 March 2020)

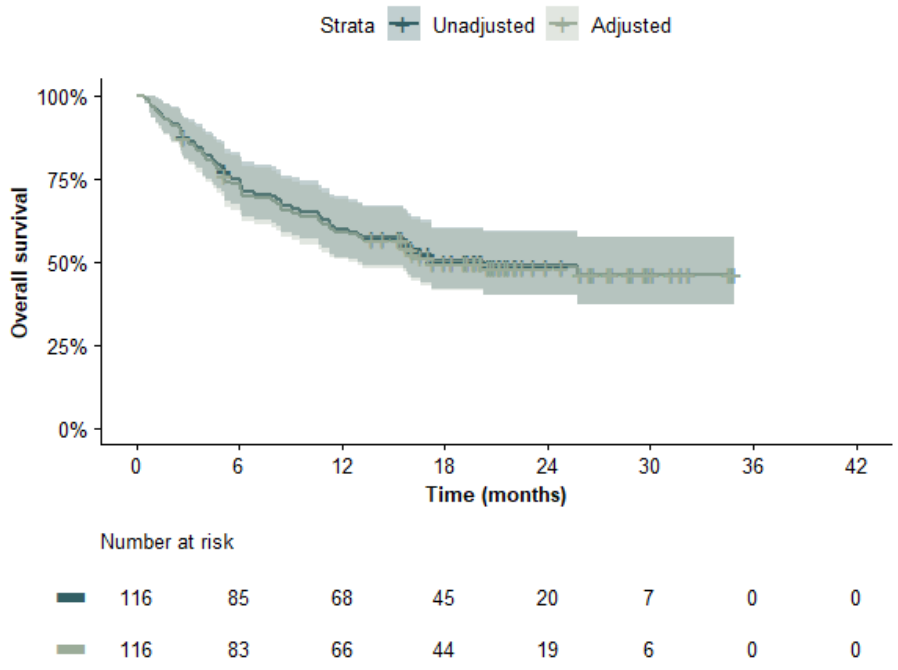
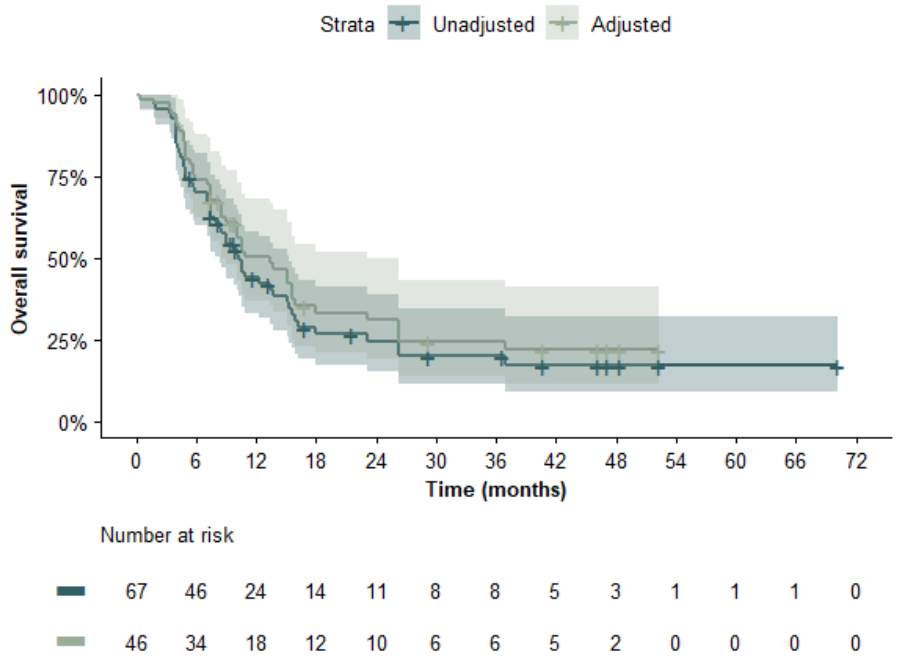


Figure 33. Adjusted OS plot for PSW4 - Study 100070-Obs001 (reproduced from Figure 6 from the company's response to clarification dated 27 March 2020)



Avelumab for treating metastatic Merkel cell carcinoma (CDF review of TA517)

Cancer Drugs Fund Review Addendum

December 2020

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 16/134/09T.

1 Introduction

The Evidence Review Group (ERG) has provided an addendum to the ERG report for the Cancer Drugs Fund (CDF) review of avelumab for treating metastatic Merkel cell carcinoma (TA517) as the company has provided an addendum to their submission, which includes revised base case results with a patient access scheme (PAS) discount of [redacted] on the list price of avelumab applied.

Section 2 of the addendum provides the company's base case results, sensitivity and scenario analyses with the PAS discount applied and Section 3 presents the ERG's scenarios and base-case, also inclusive of the PAS discount. All the results in the addendum supersede results presented in the ERG report, which are based on the list price of avelumab.

2 Company base case results with PAS

The company's base case results, with the [redacted] patient access scheme (PAS) discount applied, are presented in Table 1, showing an incremental cost-effectiveness ratio (ICER) of £17,947 per quality-adjusted life-year (QALY) gained for avelumab compared to chemotherapy.

Table 1. Company's deterministic cost effectiveness results (Table 1 of the company's Results addendum with PAS)

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Chemotherapy	11,116	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947

Abbreviations: ICER, incremental cost effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life-year.

2.1 Company sensitivity analysis

The company provided a probabilistic sensitivity analysis (PSA) based on 1,000 samples, to assess the impact of parameter when all parameters are varied simultaneously in the economic model. The results of the PSA are given in Table 2, showing a slightly decreased ICER of £17,939 per QALY compared to the deterministic base case ICER. The results of all 1,000 sampled results are presented on the cost effectiveness plane in Figure 1.

Table 2. Company's probabilistic cost effectiveness results (Table 2 of the company's Results addendum with PAS)

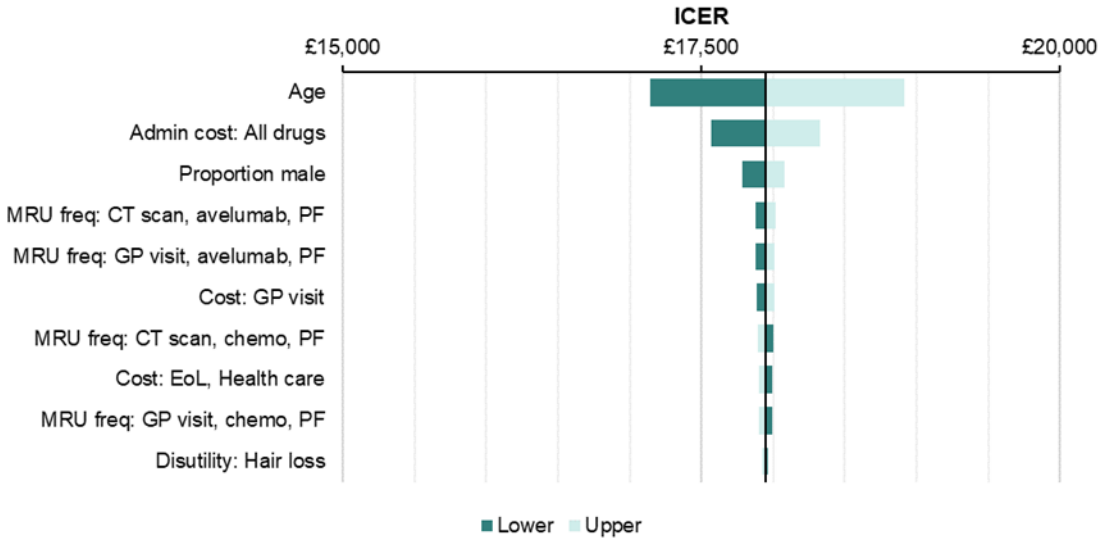
Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Chemotherapy	[redacted]	1.95	1.33	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,939

Abbreviations: ICER, incremental cost effectiveness ratio; LYG, life-years gained; QALY, quality-adjusted life-year.

Figure 1. Scatterplot of probabilistic results, with PAS (Figure 1 of the company's Results addendum with PAS)

The company also conducted a range of one-way sensitivity analyses to assess the impact of varying each parameter individually. The results of these are shown in the tornado plot in Figure 2.

Figure 2. One-way sensitivity analyses, with PAS (Figure of the company's Results addendum with PAS)



2.1.1 *Company's scenario analyses*

The company presented a range of scenario analyses in the company submission using alternative approaches for overall survival (OS), time on treatment (ToT) and utility values. These scenarios are described alongside the results in Table 3.

Table 3. Company's key scenario analyses (Table 3 of the company's Results addendum with PAS)

Scenario and cross reference	Scenario detail	Brief rationale	Impact on base-case ICER (£) (+/- from base case)
Base case			**17,947**
OS for avelumab	Use generalised gamma model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (generalised gamma) which projects higher OS versus the base-case analysis	17,363 (-584)
	Use 1-knot normal spline-based model to estimate OS for patients treated with avelumab	This scenario explores the use of a model (1-knot normal spline) which projects lower OS versus the base-case analysis	19,276 (+1,329)
Clinical expectation of ToT	Assume all patients discontinued by 10 years, with no interim capping at 2 years	Approach intended to serve as an upper bound of potential long-term treatment with avelumab. In practice, discontinuation with avelumab is expected to occur before 5 years	18,895 (+948)
	Assume one-third of patients continue treatment after 2 years, and all discontinue by 5 years	Approach aligned with clinical expert opinion at the time of the original TA517 CS, and is still expected to be broadly representative of clinical practice	15,278 (-2,671)
Utility values	Use original utility values from TA517	Allows for assessment of impact on cost-effectiveness results through updating utility values	18,655 (+708)
	Use only data from JM200: Part B to inform utility values (28 & 84 cut off points)	Allows exploration of using utility values derived only from a treatment-naïve metastatic MCC population	18,395 (+448)

Abbreviations: BSC, best supportive care; CS, company submission; ICER, incremental cost-effectiveness ratio; JM200, JAVELIN Merkel 200; OS, overall survival.

In response to the ERG clarification questions, the company also provided the following range of additional analyses:

- Range of propensity score matching (PSM) and propensity score weighting (PSW) analyses for PFS and OS (Table 4);
- Immunocompetent subgroup analysis (Table 5); and,
- A range of scenarios to include subsequent treatments (Table 6).

Table 4. Scenario analyses with alternative adjusted OS and PFS analyses

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company's base-case analysis							
Chemotherapy	11,116	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947
PSW, all patients (stable weights), using age, sex, ECOG (0 vs 1+), excluding immunosuppression as a variable							
Chemotherapy	11,947	2.19	1.49	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,135
PSW, all patients (stable weights), using age, sex, ECOG (0 vs 1), excluding immunosuppression as a variable and removing ECOG 2+ pts							
Chemotherapy	12,022	2.20	1.50	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,352
PSM using age, sex, ECOG (0 vs 1+), excluding immunosuppression as a variable							
Chemotherapy	11,481	1.85	1.27	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	16,269
PSM using age, sex, ECOG (0 vs 1), excluding immunosuppression as a variable and removing ECOG 2+ pts							
Chemotherapy	11,559	1.92	1.31	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	14,797

Abbreviations: ECOG, Eastern Cooperative Oncology Group Performance Status; ICER, incremental cost-effectiveness ratio; LYG, life years gained; PSM, propensity score matched; PSW, propensity score weighted; QALYs, quality-adjusted life years.

Table 5. Scenario analyses with alternative subgroup OS and PFS analysis

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company's base-case analysis							
Chemotherapy	11,116	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947
Analysis using only the immunocompetent subgroup							
Chemotherapy	11,499	1.83	1.25	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,225

Abbreviations: CEA, cost-effectiveness analysis; ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Table 6. Scenario analyses with alternative subsequent treatment approaches applied

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company's base-case analysis							
Chemotherapy	11,116	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947
Include all subsequent therapies as per study sources							
Chemotherapy	11,374	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,474
Update topotecan costs to standard chemotherapy cost							
Chemotherapy	11,249	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	18,527
Remove monoclonal antibodies and update topotecan costs							
Chemotherapy	11,249	1.94	1.32	-	-	-	-
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,966

Abbreviations: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

3 ERG preferred analysis

3.1 Exploratory and sensitivity analyses undertaken by the ERG

The Evidence Review Group (ERG) conducted a range of scenario analyses to assess the impact of different assumptions applied for overall survival (OS), progression-free survival (PFS) and time on treatment (ToT) in terms of the best fitting curves as well as the assumption of weight-based dosing. The ERG also explored the impact of using the SACT data to inform OS and TTD, as well as using the TTD data as a proxy to inform PFS. The ERG applied the curves described in Section **Error! Reference source not found.** of the ERG report for OS, PFS, and TTD, individually, and then applied all changes in a combined scenario analysis. The results are presented in Table 7.

Table 7. Results of the ERG's scenario analyses

	Results per patient	Avelumab	Chemotherapy	Incremental value
0	**Company base case**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**17,947**
1	**Weight-based dosing for avelumab acquisition costs**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**18,938**
2	**OS: 1-knot hazard spline**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**20,097**
3	**PFS: 3-knot odds spline**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**17,852**

4	**ToT: 3-knot hazard spline**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**18,290**
5	**ToT: No change to treatment-experienced data**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**19,332**
6	**OS: ERG's SACT OS curves applied**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**24,957**
7	**PFS: ERG's SACT TTD curves used as a proxy for PFS**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**18,243**
8	**ToT: ERG's SACT TTD curves applied**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**16,852**
9	**ERG's SACT curves applied for OS, PFS and TTD (Scenarios 6+7+8)**			
	Total costs (£)	[redacted]	11,116	[redacted]
	QALYs	[redacted]	1.32	[redacted]
	ICER			**23,485**

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year; ToT, time on treatment.

3.2 ERG preferred assumptions

The ERG's preferred base case includes the following changes compared to the company's base case analysis:

Scenario 2a. PSW analyses for OS and PFS with immunocompetency excluded from the estimation of propensity scores and patients with ECOG score 2 or more removed;

Scenario 3. 1-knot hazard spline for OS;

Scenario 4. 3-knot odds spline for PFS;

Scenario 5. 3-knot hazard spline for ToT;

Scenario 6. Removing the adjustment to the ToT curve using the treatment-experienced population data;

Scenario 7. Weight-based acquisition costs for avelumab in line with effectiveness data.

These changes resulted in an ICER of £21,958 per QALY for the ERG's preferred base case. The cumulative results as each change is applied are given in Table 8. NICE requested additional scenarios to be performed and these are also presented in Table 8.

Table 8. ERG's and NICE preferred model assumptions

	Scenario	Technical engagement Issue number	ICER (£/QALY)	Change from company base case ICER
0	**Company base case**	**-**	**17,947**	**-**
1	SACT dataset: ERG's curves for OS, PFS and ToT applied to SACT data (n=52) instead of the updated JAVELIN 1L data	1	23,485	+5,538
2a	Propensity score weighting (PSW): PSW analyses for OS and PFS using updated JAVELIN 1L data for avelumab (n=116) and company's part A 1L study for chemotherapy (n=67) instead of a naïve comparison. PSW4 (n=162) applied: with adjustments for age (aged ≥75 vs <75 years), sex	2	18,352	+405

	(female vs male), and ECOG PS (0 vs 1).			
2b	Using only immunocompetent patients in the company's part A 1L study (n=51) for OS and PFS instead of the pooled naïve estimate for chemotherapy.	2	17,225	-722
3	Using 1-knot hazard spline for OS instead of 1-knot odds spline for JAVELIN 1L data	3	20,097	+2,150
4	Using 3-knot odds spline for PFS instead of 2-knot odds spline for JAVELIN 1L data	4	17,852	-95
5	Using 3-knot hazard spline for ToT instead of Weibull for JAVELIN 1L data	5	18,290	+343
6	Removing the adjustment to the ToT curve using the JAVELIN 2L+ data	5	19,332	+1,385
7	Weight-based avelumab dose instead of flat dose of 800 mg	6	18,938	+991
8	**Cumulative changes with PSW4: 2a + 3 to 6 (flat dose)**	**-**	**20,780**	**+2,833**
9	**Cumulative changes with PSW4: 2a + 3 to 7 (weight-base dose) - ERG's base case ICER**	**-**	**21,958**	**+4,011**
10	**Cumulative changes with immunocompetent group: 2b + 3 to 6 (flat dose)**	**-**	**19,832**	**+1,885**
11	**Cumulative changes with immunocompetent group: 2b + 3 to 7 (weight-base dose)**	**-**	**20,914**	**+2,967**
12	**Cumulative changes with SACT dataset: 1 + 2b (flat dose)**	**-**	**22,252**	**+4,305**
13	**Cumulative changes with SACT dataset: 1 + 2b + 7 (weight-base dose)**	**-**	**23,486**	**+5,539**

Abbreviations: 1L, first-line; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; mg, milligram; OS, overall survival; PFS, progression-free survival; PSW, propensity score weighting; QALY, quality adjusted life year; ToT, time on treatment.

National Institute for Health and Care Excellence
Centre for Health Technology Evaluation

ERG report – factual accuracy check

Avelumab for treating metastatic Merkel cell carcinoma [ID1617]

You are asked to check the ERG report to ensure there are no factual inaccuracies contained within it.

If you do identify any factual inaccuracies, you must inform NICE by **5pm on Friday 24 April 2020** using the below comments table. All factual errors will be highlighted in a report and presented to the Appraisal Committee and will subsequently be published on the NICE website with the committee papers.

The factual accuracy check form should act as a method of detailing any inaccuracies found and how and why they should be corrected.

Issue 1 Discussion of patient population

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Section 1.2, page 15: *"The population of JM200: Part B does not accurately reflect the patients likely to receive avelumab for first line treatment of mMCC in England; the SACT data set comprises a closer match to expected patient characteristics in clinical practice in England."* Merck Serono notes that for consistency with the remainder of the ERG's report, this text should be revised to clarify that the SACT cohort is considered by the ERG to be *more representative* of patients likely to receive avelumab for first line treatment of mMCC in England, versus the JM200: Part B cohort. Merck Serono considers it important to clarify that while both cohorts include patients that are likely to receive avelumab in practice, it is the ERG's opinion (based on clinical advice received) that the SACT cohort is more generalisable to the population expected to be treated in practice.	Merck Serono suggests the text be amended to the following: *"The SACT data set comprises a closer match to expected patient characteristics in clinical practice in England compared to the population of JM200: Part B."*	This minor modification is intended to clarify that the ERG considers the SACT cohort to be more generalisable to population expected to be treated in practice. The original wording has the potential to be misleading, as it may suggest the JM200 trial reflects a population that would not be treated in practice. This suggested revision is intended to avoid this potential misunderstanding while preserving the ERG's statement concerning generalisability.	Not a factual error.

Issue 2 Justification of ToT extrapolation adjustment

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Section 1.3, page 17: *"The company justified this because the minimum follow-up for the treatment-naïve population was 15 months; however, the*	Merck Serono suggests the text be amended to the following:	This edit to the ERG's report is intended to clarify the justification behind this model assumption made	Not a factual error

ERG considers that the rates for the treatment-experienced population beyond 15 months cannot be expected to be reflective of treatment-naïve population." Similar text may be found in Section 4.1.5.3, page 64: *"This was justified by the company because the minimum follow-up for the treatment-naïve population was 15 months, whereas for the treatment-experienced population it was 36 months."* Merck Serono wishes to clarify that the justification for using the treatment-experienced data was due to the unadjusted extrapolation of the treatment-naïve data yielding estimates of treatment duration that were higher than expected based on clinical input (i.e. resulting in patients projected to be on treatment for more than 5 years, and thus estimates leading up to 5 years were also higher than anticipated). While related to follow-up, the justification behind adjusting the extrapolation was driven primarily by the intention to adjust the curve so that it provided a more realistic extrapolation. Separately, the model assumes patients do not continue treatment beyond 5 years. However, the combination of the cap at 5 years and this adjustment leads to a less pronounced drop at the 5-year time point, which Merck Serono considered to be a more realistic extrapolation.	*"The company justified this as without adjustment, all models projected some patients to continue treatment longer than 5 years; however, the ERG considers that the rates for the treatment-experienced population beyond 15 months (the minimum follow-up in JM200: Part B) cannot be expected to be reflective of treatment-naïve population."* … and: *"This was justified by the company because all models projected some patients to continue treatment longer than 5 years, and that the minimum follow-up for the treatment-naïve population was 15 months, whereas for the treatment-experienced population it was 36 months."*	in the submitted base-case analysis (which was not driven solely due to the available follow-up data from JM200: Part B for this outcome). The revised text is intended to clarify why the adjustments were made, but without editing the ERG's critique.	

Issue 3 Discussion of weight-based versus fixed dose of avelumab

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Section 1.3, page 17:	Merck Serono suggests the text be amended to the following:	The fixed dose of 800mg addresses the wastage and administrative burden associated with the original	Thank you for highlighting this. The ERG report has been amended.

"A final area of uncertainty that the ERG noted is in the change of dosing to a flat 800mg dose compared the previous weight-based dose, which resulted in an expected dose of 849mg. Given that the Javelin Merkel 200 trial treatment effects were based on weight-based dosing with an average dose of 849mg, the ERG considers the acquisition costs applied in the economic model should be weight-based to align with the effectiveness estimates. The ERG notes that this would, however, not reflect clinical practice, which will now be based on the 800mg flat dose." There are also several other instances wherein the value of 849mg is stated: • Section 4.1, page 50: *"Avelumab acquisition costs now based on a flat-dose of 800mg rather than a weight-based dose of 849mg, to align with the newly approved dose."* • Section 4.1.2, page 51: *"Previously a weight-based dose was given at a dose of 10mg/kg, which resulted in an expected dose of 849mg"* • Section 4.1.8, page 70: *"This approach has led to a reduction in acquisition costs as the original weight-based dosing resulted in an expected dose, estimated from the distribution of full vials used, of 849mg."* The value of 849mg comes from the economic model wherein the methods of moments approach was used to reflect the weight-based	*"A final area of uncertainty that the ERG noted is in the change of dosing to a flat 800mg dose compared the previous weight-based dose (10mg/kg), which resulted in an average of 4.25x 200mg vials required per administration (versus 4x 200mg vials for a fixed dose of 800mg)."* Outside of this specific revision, Merck Serono requests all references to the value of 849mg dose be changed to consider the number of vials (i.e. 4.25) instead.	weight-based dose of avelumab without compromising safety or efficacy. The value of 849mg cited does not represent the mean expected dose for patients, though the equivalent number of vials (4.25) reflects the estimated number of vials required per administration based on the original weight-based dosing approach while accounting for wastage. This adjustment to the presentation of this difference in dosing is intended to more accurately describe the difference between the approaches – that is, the difference in the product used per administration (rather than the dose received). The mean weight of the trial population was 78.5kg, which is equivalent to a mean dose of 785mg (i.e. less than 800mg). The method of moments approach was implemented to account for wastage which would be omitted in a cost-per-mg approach, and resulted in an average administration of 4.25 vials. However, the mean dose received by patients was not equivalent to the number of vials used (as wastage is incorporated). The value of 849mg is presented within the economic model, but is otherwise not described within the	

dosing (per the original TA517 submission). Merck Serono wishes to highlight that the value of 849mg is <u>not</u> the "expected dose" – rather, it is the mean volume of product that was estimated to be used per administration based on the weight-based dose. This value includes product wastage, and so does not accurately reflect the dose administered to patients (though Merck Serono appreciates this was not described fully within its submission). Merck Serono highlights the importance of this distinction, as the mean dose received by patients is not equal to the value of 849mg as may be inadvertently inferred from the description provided in the ERG's report.		company submission (CS). For both transparency and accuracy of reporting, Merck Serono considers it more appropriate to discuss the difference in approaches based on the number of vials required per administration (avoiding issues relating to the difference between dose required versus dose costed).	

Issue 4 Approach taken to use SACT data in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Section 3.1, page 25: *"The ERG also notes that the company reports that the SACT data cannot be utilised in the current economic model, although the ERG considers that it is possible to utilise the SACT data and provides this in scenario analyses (See Section 4.1.5)."* For clarity, Merck Serono considers it important to note that while it is possible to generate a scenario using the SACT data, this is only possible if ToT is assumed to be equal to PFS, and that other model parameters derived using data from	Merck Serono suggests the text be amended to the following: *"The ERG also notes that the company reports that the SACT data cannot be utilised in the current economic model, although the ERG considers that it is possible (with some assumptions) to utilise the SACT data and provides this in scenario analyses (See Section 4.1.5)."*	This revision is intended to clarify that while it is possible to generate a scenario using these data, additional assumptions are required in order to do so – for example, assuming ToT is equal to PFS, and that AE rates etc. are equivalent to the trial cohort.	Not a factual error.

the JM200 trial (e.g. adverse event rates, utility values, weight etc.) are also assumed equal.			

Issue 5 Choice of comparator data source

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Section 3.2, pages 38-39: *"The ERG also notes that in the ToE it was specified that Study 100070-Obs001 should be used to provide the comparator data for chemotherapy in the absence of new data and that this naïve pooled analysis was not referred to in the ToE."* While this is true, Merck Serono wishes to highlight that it was not possible to meet the specified base case in the ToE document, as the base-case ICER corresponds to the pooled comparator, whereas the comparator dataset in the ToE document suggests using only Study 100070-Obs001 data. The consequences of this apparent discrepancy were not immediately clear to Merck Serono, and therefore this was not commented on within the CS.	Merck Serono suggests the text be amended to the following: *"The ERG also notes that in the ToE it was specified that Study 100070-Obs001 should be used to provide the comparator data for chemotherapy in the absence of new data and that this naïve pooled analysis was not referred to in the ToE. However, use of the naïve pooled analysis corresponded to the committee's preferred ICER results that were presented within the ToE document."*	This edit has been suggested to explain why this apparent discrepancy was introduced within the CS.	Not a factual error.

Issue 6 Terminology concerning the indirect comparison

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Section 3.4, pages 48-49: *"However, the ERG considers the company's naïve comparison of the JM200: Part B and Study 100070-Obs001,*	Merck Serono suggests the text be amended to the following:	The ERG's comment has been reworded to use specific terminology relating to the critique raised (i.e.	Not a factual error.

to be unreliable because of the imbalances in the patient characteristics between the two studies, the small number of patients in the studies, and the uncertainty caused by unmeasured variables that may be effect modifiers or prognostic indicators." Merck Serono understands the uncertainty associated with the comparison undertaken, yet use of the term "unreliable" has specific statistical connotations that do not apply here – for example, if a study is deemed "unreliable", the study conduct (e.g. its design) may be questioned.	*"However, the ERG considers the company's naïve comparison of the JM200: Part B and Study 100070-Obs001, to be subject to a number of limitations (and is therefore uncertain) because of the imbalances in the patient characteristics between the two studies, the small number of patients in the studies, and the uncertainty caused by unmeasured variables that may be effect modifiers or prognostic indicators."*	that the comparison is subject to several limitations and its outcome is therefore uncertain). While a relatively small point to comment upon, Merck Serono is concerned that the original phrasing may lead to misinterpretations concerning the study conduct relating to both JM200 and Study 100070-Obs001.	

Issue 7 Rationale provided concerning the use of SACT data in the model

Description of problem	Description of proposed amendment	Justification for amendment	ERG response
Section 4.1.5, page 58: *"However, as the data are immature the company did not consider this analysis to be reliable and, therefore, did not provide this analysis."* Merck Serono highlights that this was one reason provided as to why the SACT data were not considered appropriate for inclusion within the economic model. However, several other reasons were also provided (which Merck Serono appreciates the ERG may disagree with), which can be found on page 11 of the CS: • Avelumab is currently recommended by NICE and routinely commissioned in the second-line setting. In this	Merck Serono suggests the text be amended to the following: *"However, the company did not consider this analysis to be reliable (for several reasons, including data maturity and other context-specific issues described in CS Section A.6.1) and, therefore, did not provide this analysis."*	This edit is intended to reflect the CS content in terms of the rationale provided as to why the SACT data were deemed inappropriate for use within the model. The suggested revised text is intended to acknowledge that reasons other than data maturity were considered when deciding on the relevance of the	Not a factual error.

context, some patients who would have been candidates for treatment with first-line avelumab may have instead been treated with chemotherapy by clinicians who preferred to reserve avelumab for use after chemotherapy • Some patients included within the SACT cohort may have previously been deemed ineligible for treatment with first-line chemotherapy owing to its associated toxicities, and consequently managed with best supportive care prior to avelumab. This means that due to the lack of availability of other options, some patients may have had poorer prognosis compared to newly-diagnosed patients • Patients need to have sufficient life expectancy to benefit from immunotherapies (such as avelumab), due to the mechanism of action of these treatments. In JM200, patients with a life expectancy of at least 12 weeks were eligible for inclusion. The initial drop apparent in the OS curve for the SACT cohort indicates the inclusion of some patients who may not have sufficient life expectancy to benefit from immunotherapy For transparency, Merck Serono considers it necessary that the ERG's report acknowledge several reasons were presented concerning the potential use of SACT data within the model.		SACT data for inclusion within the model.	

NICE National Institute for Health and Care Excellence

Technical engagement response form

Avelumab for treating metastatic Merkel cell carcinoma (CDF Review of TA517) [ID1617]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments: **Thursday 7 January 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.
- If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.

- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Merck Serono and Pfizer Ltd. (Merck/Pfizer alliance)**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

<table>
<tr><th>Key issue</th><th>Does this response contain new evidence, data or analyses?</th><th>Response</th></tr>
<tr><td>Key issue 1: The ERG considers the updated JAVELIN data to be appropriate for use in the model. However, the NICE technical team would like to know how generalisable the JAVELIN trial is to clinical practice and how the Systemic Anti-Cancer Therapy database (SACT) data could be used for decision making.</td><td>No</td><td>We agree with the ERG that the most appropriate data source to inform the model concerning the safety and efficacy of avelumab comes from the pivotal JAVELIN Merkel 200 (JM200) study (Part B).

The JM200 is the largest clinical trial to date demonstrating the effectiveness and safety of a treatment in a Merkel cell carcinoma (MCC) patient cohort. Upon receiving the technical report, the trial data was further validated with a clinical expert who confirmed that the demographics of the JM200 study is representative of the UK mMCC patient population. Furthermore, the efficacy outcomes including OS and ORR are similar to outcomes expected in the clinical setting and the safety data reported in the study is equivalent to adverse events encountered. The SACT database collected real world data in a small patient cohort with a short follow-up period, outside of the clinical trial setting. The SACT data is useful in providing an additional data source to demonstrate the benefits of avelumab in 1L mMCC.

To assist the NICE appraisal committee with understanding the key differences between the JM200 and SACT populations, and how this may affect interpretations related to generalisability, please see the table below. The table summarises the key features of these two cohorts of patients, while also highlighting (in bold) which characteristics in the respective populations most accurately represents UK clinical practice.

<table>
<tr><th>JM200: Part B</th><th>SACT</th></tr>
<tr><td>The key efficacy endpoint of the study was durable response rate (DRR)* and included</td><td>The only data points collected in the SACT were treatment duration and OS.</td></tr>
</table>
</td></tr>
</table>

		<table><tr><td>**other efficacy endpoints such as BOR, PFS, OS and duration of response. These endpoints provide clinical importance to assess the effectiveness of avelumab.**</td><td></td></tr><tr><td>Median age 74.0 years</td><td>**Median age 75.5 years**</td></tr><tr><td>The clinical trial included 64.7% of patients recruited from Western Europe.</td><td>**UK-based population in England**</td></tr><tr><td>**Large sample size (n=116)**</td><td>Small sample size (n=52)</td></tr><tr><td>**Long follow-up (minimum 15 months)**</td><td>Short follow-up (minimum 5 months)</td></tr><tr><td>**All patients had ECOG 0-1**</td><td>Some patients with ECOG 2+ or unknown</td></tr><tr><td colspan="2">* Durable response rate (DRR)*=best overall response rate of ≥6 months</td></tr></table> Merck/Pfizer recognises that both sources have features that are generalisable to the NHS population, and others that are less generalisable. Both data sets also demonstrate the substantial benefit of avelumab in the 1L mMCC setting. However, on balance, the JM200 study data are considered to provide the most reasonable basis to inform decision making to demonstrate the benefit of avelumab in 1L mMCC. The SACT data provides an additional data source which can be used in conjunction with JM200 to support clinical decision-making.
Key issue 2: The ERG considers the propensity score weighting analysis 4 (PSW4 analysis), with adjustments for age, sex and Eastern Cooperative Oncology Group Performance Status (ECOG PS), appropriate to inform the avelumab versus chemotherapy	Yes	As noted in the ERG's report, the choice of propensity score weighting analysis has a relatively limited impact on the cost-effectiveness results (see ERG addendum Table 4). However, Merck/Pfizer acknowledges the importance of exploring the impact of adjusting for immunocompetency, as well as other potentially important characteristics. In Study 100070-Obs001, a total of n=13 patients were immunocompromised, versus n=51 patients who were immunocompetent. Immunocompetency is not expected to have a large impact on the outcome of treatment with chemotherapy, within the context of an mMCC population, as these patients were considered fit enough to receive chemotherapy. However, immunocompetency was an important inclusion criterion within JM200, owing to the mechanistic properties of avelumab.

comparison in the model. However, because immunocompetency was not included in the PSW4 analysis, and because of the ERG's concerns about JAVELIN generalisability (Issue 1), the NICE team considers that the following two scenarios are also relevant:

- a naïve comparison using updated JAVELIN data (n=116) and the immunocompetent subgroup of the company's observational chemotherapy study (n=51), and
- a naïve comparison using SACT data (n=52) and the

At clarification stage, a comparison between JM200: Part B (n=116) and the immunocompetent subgroup from Study 100070-Obs001 (n=51) was requested by the ERG and provided by Merck/Pfizer. However, subsequent to the provision of this analysis, a confidential simple patient access scheme (PAS) discount of [redacted] off the list price of avelumab was approved by NHS England. This was not reflected in the results presented by Merck/Pfizer. For completeness, the results of this analysis (including the PAS discount) are provided in the table below, compared to the company's preferred base-case analysis:

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company submission base-case analysis							
Chemotherapy	11,116	1.94	1.32				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947
Analysis using immunocompetent subgroup only*							
Chemotherapy	11,499	1.83	1.25				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,225**

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Notes: * Please note that the total costs for avelumab change marginally, as the average duration of survival for the chemotherapy arm is used within the estimated costings of radiotherapy for all treatment arms. **The ICER of £17,225 was also presented in the ERG's addendum in Table 5.

As described in response to clarification question B2, due to the reduced number of patients at risk within the immunocompetent subgroup (versus the whole population), relatively fewer patients are present in the tail of the Kaplan-Meier curves, leading to slightly lower longer-term estimates of both OS and PFS (e.g. 5-year OS using a log-logistic model is 7.6% for the whole population versus 6.8% for the immunocompetent population). The ICER is similar, but slightly lower than the base-case analysis.

immunocompetent subgroup (n=51)

The second analysis highlighted in this Key Issue was not requested by the ERG at the clarification stage but is possible to generate within the model edited by the ERG to inform its report. The corresponding results of the second analysis are provided in the table below for completeness. However, Merck/Pfizer urges caution when interpreting these results in line with the limitations of the SACT dataset described in response to Key Issue 1:

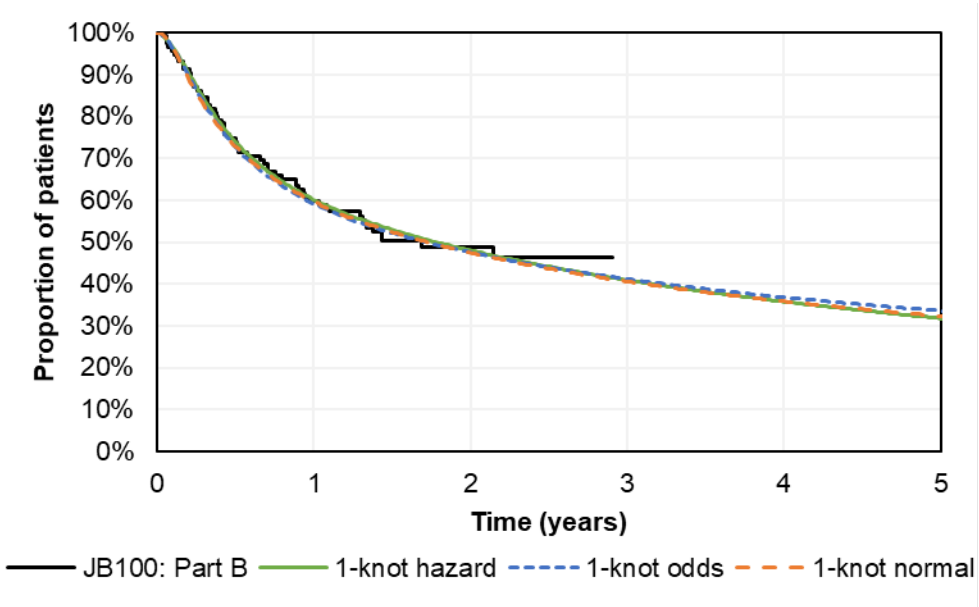
Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company submission base-case analysis							
Chemotherapy	11,116	1.94	1.32				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947
Analysis using immunocompetent subgroup only versus SACT data set*							
Chemotherapy	11,499	1.83	1.25				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	22,252

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; SACT, Systemic Anti-Cancer Therapy.

Notes: * Please note that the total costs for avelumab change marginally, as the average duration of survival for the chemotherapy arm is used within the estimated costings of radiotherapy for all treatment arms.

The total costs, QALYs, and LYs are identical to the results presented for the previous analysis, so are not discussed further here. When switching to using the SACT data set, the total costs for avelumab increase slightly, whereas the total QALYs and LYs decrease slightly – the combined effect of these changes leads to an increase in the ICER from £17,947 (company base-case analysis) to £22,252.

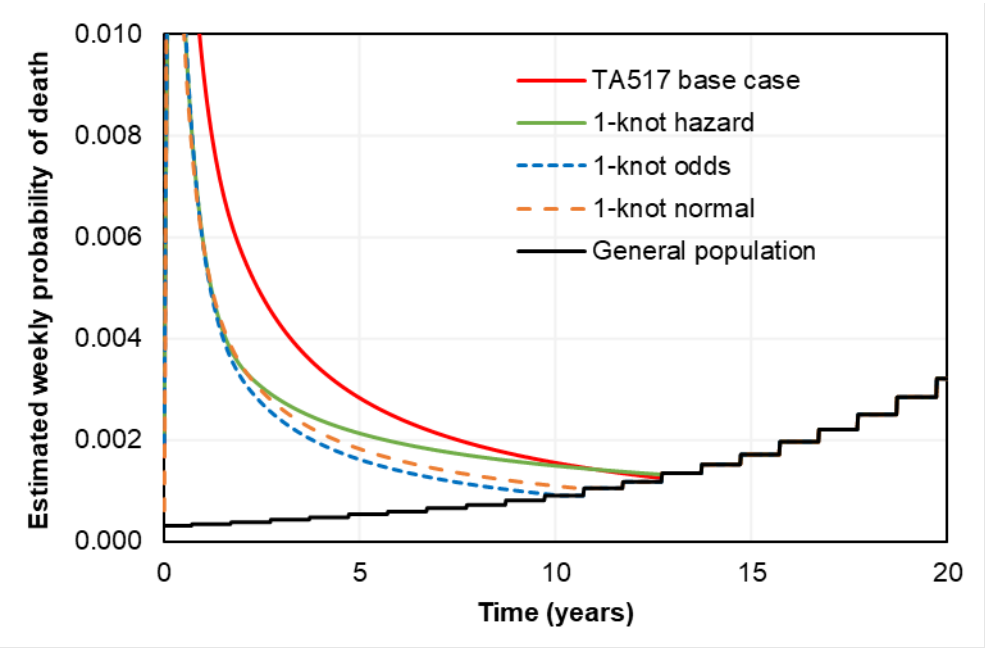
In summary, Merck/Pfizer do not expect immunocompetency to impact the outcome of treatment in the chemotherapy arm and therefore considers a comparison to the full

<table>
<tr><td></td><td></td><td>observational chemotherapy study as most appropriate. Limitations to comparing to the SACT data set have been documented above under Key Issue 1.</td></tr>
<tr><td>Key issue 3: The ERG agrees with the company to focus the choice of avelumab OS curve on the 1-knot spline models but notes important differences in the extrapolations produced by the hazard-, normal- and odds-based splines. Because there is uncertainty in the naïve comparison of the treatment effects between avelumab and chemotherapy, the ERG considers it may be more appropriate to choose a curve with more conservative extrapolation. The ERG chose the hazard-based 1-knot spline for OS.</td><td>No</td><td>The choice of the most appropriate extrapolation of overall survival is challenging within the context of maturing trial data. In choosing between the three 1-knot models, the difference in long-term survival may be difficult to validate, as all three models produce estimates that are broadly in keeping with advice provided to Merck/Pfizer, and produce near-identical fits to the Kaplan-Meier curve (shown in the diagram below):
 

It may instead be helpful to consider a comparison of the different survival models at specific time points in the longer term, as shown in the table below:
<table>
<tr><th>Time (years)</th><th>1-knot hazard</th><th>1-knot odds</th><th>1-knot normal</th></tr>
<tr><td>5</td><td>31.9%</td><td>33.7%</td><td>32.4%</td></tr>
</table></td></tr>
</table>

10	20.1%	24.6%	22.5%
15	13.8%	17.8%	16.2%
20	7.5%	9.7%	8.8%
25	2.4%	3.1%	2.9%
30	0.3%	0.4%	0.4%
35	0.0%	0.0%	0.0%
40	0.0%	0.0%	0.0%

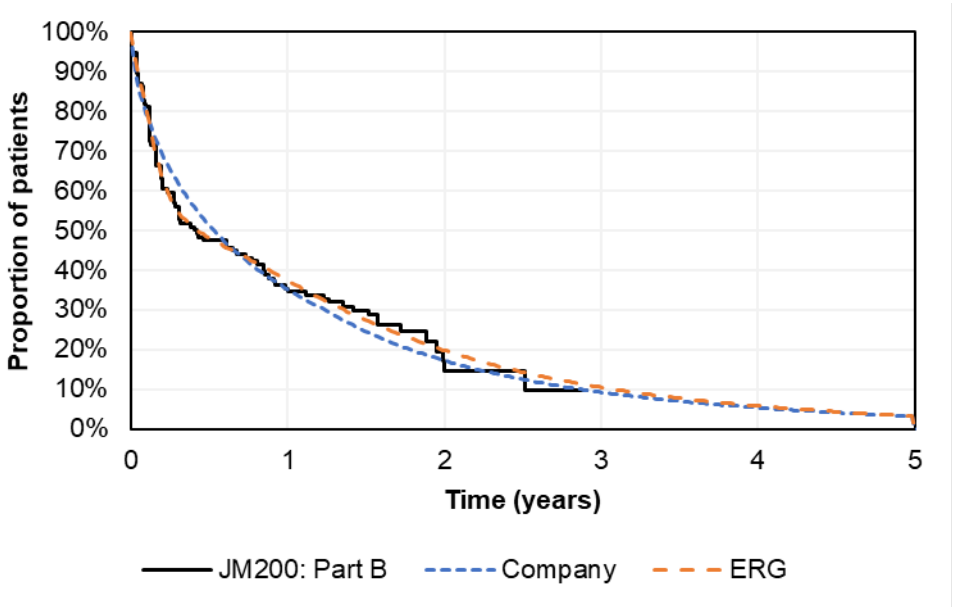
It should also be noted that a 1-knot odds-based spline was selected to inform the base-case analysis adopted in TA517 (for the 2L+ population). When comparing the estimated hazard of death inherent within the base-case analysis in TA517 for the 2L+ population versus each of the 1-knot models fitted to the JB100: Part B data, the following plot is produced:



The plot above demonstrates that the projected hazards produced by the 1-knot hazard spline result in an extrapolation which eventually produces an estimated hazard of death which exceeds that of the base-case analysis presented in TA517. While this does not occur until approximately 11 years, this is misaligned with clinical opinion that outcomes for patients treated in the 1L setting are expected to be better than those for a 2L+ population. Moreover, the shape of the hazard function implies a slightly different shape for the 1L versus 2L+ populations if a 1-knot hazard-based spline is selected.

Merck/Pfizer therefore considers its preferred base-case analysis (using a 1-knot odds-based spline) to provide a more appropriate estimation of OS versus the ERG's preferred 1-knot hazard model. However, each of the three 1-knot spline-based models may be helpful in decision making.

Key issue 4: The ERG considers the 2-knot spline to underestimate the KM data between 0.5 years and 1 year. It considers the 3-knot odds spline to provide a better extrapolation as well as being the best fit to the data. The ERG chose 3-knot odds spline curve for PFS. However, the ERG considers the PSW adjusted analysis to be preferable (Issue 2)	No	A small impact on the ICER is noted when the PFS extrapolation is changed from Merck/Pfizer preferred base-case analysis using a 2-knot spline model to the ERG's preferred 3-knot spline model. Merck/Pfizer considers there to be relatively little evidence to definitively reject either one of these models in favour of the other, and so considers both approaches to be suitable to inform decision making. The PSW-adjusted analyses affect the estimation of PFS for the chemotherapy arm (as all patients enrolled in JM200 were immunocompetent). Merck/Pfizer has responded in relation to the PSW analyses within the context of Key Issue 2.
Key issue 5: The ERG agrees with the company, that it is reasonable to stop treatment at 5 years as this is likely to happen in clinical practice and used the same assumption in its preferred base-case. However, the ERG considers that the curves fitted to 1L data should not be adjusted by 2L+ data because this is not	No	Merck/Pfizer acknowledges the view of the ERG regarding the potential issues with using non-1L data to inform the extrapolation of time-on-treatment (ToT). In Merck/Pfizer's base-case analysis, JM200: Part A (2L+) data were considered for use within the model to inform the rate of treatment discontinuation beyond the minimum follow-up period of JM200: Part B (1L). This approach was undertaken to supplement the limited data available for longer-term treatment discontinuation available from JM200: Part B (1L), with more mature data from JM200: Part A (2L+) while also maintaining a model based solely on JM200: Part B (1L) data for the earlier portion of the curve. The base-case approach was also proposed in keeping with recommendations from the ERG as part of TA517, wherein a Weibull model was selected to inform ToT for the ERG's preferred base-case analysis. The 3-knot hazard-based spline could be considered an extension of a Weibull model, but has a greater reliance on fit to the JB200: Part B data than Merck/Pfizer's preferred approach.

reflective of treatment-naïve population (1L). The ERG chose 3-knot hazard spline for the 1L data (2L+ data are not used in the ERG approach)		The figure below presents a comparison of Merck/Pfizer's and the ERG's preferred approach to modelling ToT. From this figure, it can be seen that the ERG's preferred model provides a relatively better visual fit to the Kaplan-Meier curve, which is unsurprising given its increased flexibility versus the Weibull-based approach chosen by Merck/Pfizer.  It could potentially be argued that the 3-knot spline-based approach overfits the Kaplan-Meier curve versus a more simplified model based on interpretation of BIC scores which apply a higher penalty for more complex models versus AIC scores. For example, the Weibull model (without adjustment) has a BIC score of 1,221.90 versus 1,223.31 for the 3-knot hazard-based spline model. However, the 3-knot spline-based model has a better AIC versus the Weibull model, and the AIC/BIC scores do not account for the incorporation of JM200: Part A data within Merck/Pfizer's preferred base-case analysis.

		Nevertheless, both models result in very similar mean ToT estimates (Merck/Pfizer: 12.59 months, versus ERG: 13.07 months), and so the impact on the ICER is small. Merck/Pfizer views its approach as being based on a relatively simple model but making use of external data (i.e. data from JM200: Part A), whereas the ERG's approach makes use of a more flexible modelling approach, without relying on external data. Therefore, both approaches may be helpful to inform decision making.

Additional issues

Please use the table below to respond to additional issues in the ERG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (e.g. at the clarification stage).

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response
Additional issue 1: Weight-based versus flat dosing	ERG report Section 1.3.	No	The ERG's report provides a scenario analysis using the weight-based dose of avelumab (per the JM 200 study), though the licensed dose of avelumab is now based on a flat dosing regimen. Merck/Pfizer highlights that since the time of its original submission (and production of the ERG's report), the SmPC has now been fully updated to reflect the flat dosing regimen, and NICE has published its guidance concerning the flat dose of avelumab as part of TA645 (avelumab with axitinib for untreated advanced renal cell carcinoma). With respect to the dosing regimen, the FAD for TA645 states: *"The committee highlighted that a weight-based dose for avelumab was used in JAVELIN Renal 101, whereas the licence specifies a fixed dose. The companies explained that they derived the fixed dose using pharmacokinetic and pharmacodynamic data, and taking into account similar approaches used historically. The Cancer Drugs Fund clinical lead advised that this approach was taken with other drugs for this disease area. The committee was aware that it could appraise drugs only within their marketing authorisation. It accepted that the licensed fixed dose would have similar effectiveness to the weight-based dose, and concluded that it would use the licensed dose in making decisions."* The publication of the FAD for TA645 is highlighted here for completeness. In conclusion, Merck/Pfizer disagree with the ERG's suggestion that weight-based acquisition costs (instead of the 800mg flat dose) should be applied in the economic model as this would not reflect the licensed dose of avelumab, how it will be used in clinical practice and is inconsistent with NICE guidance for other avelumab indications.

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

No change made to company's preferred base-case cost-effectiveness results.

NICE National Institute for Health and Care Excellence

Patient expert statement and technical engagement response form

Avelumab for treating metastatic Merkel cell carcinoma [ID1617]

Thank you for agreeing to give us your views on this treatment and its possible use in the NHS.

You can provide a unique perspective on conditions and their treatment that is not typically available from other sources.

About this Form

In **part 1** we are asking you to complete questions about living with or caring for a patient with the condition.

In **part 2** we are asking you to give your views on issues raised in the Evidence Review Group (ERG) report that are likely to be discussed by the committee.

The ERG report provides a critique of the company submission. In part 2 of this form we have included any of the issues raised by the ERG where we think having a patient perspective could help either:

- resolve any uncertainty that has been identified
  or
- provide missing or additional information that could help committee reach a collaborative decision in the face of uncertainty that cannot be resolved.
-

In **part 3** we are asking you to provide a few summary sentences on the main points contained in this document.

If you have any questions or need help with completing this form please email the public involvement team via pip@nice.org.uk (please include the ID number of your appraisal in any correspondence to the PIP team).

Please return this form by **5pm** on **Thursday 28 January 2021**

NICE National Institute for Health and Care Excellence

Please use this questionnaire with our hints and tips for patient experts. You can also refer to the Patient Organisation submission guide. **You do not have to answer every question** – they are prompts to guide you. There is also an opportunity to raise issues that are important to patients that you think have been missed and want to bring to the attention of the committee. The text boxes will expand as you type.

Important information on completing this expert statement

- Please do not embed documents (such as a PDF) in a submission because this may lead to the information being mislaid or make the submission unreadable
- We are committed to meeting the requirements of copyright legislation. If you want to include **journal articles** in your submission you must have copyright clearance for these articles. We can accept journal articles in NICE Docs.
- Your response should not be longer than 15 pages.

PART 1	
About you	
1.Your name	[redacted]
2. Are you (please tick all that apply):	☐ a patient with metastatic Merkel cell carcinoma? ☐ a patient with experience of the treatment being evaluated? ☐ a carer of a patient with metastatic Merkel cell carcinoma? X a patient organisation employee or volunteer? ☐ other (please specify):
3. Name of your nominating organisation.	Neuroendocrine Cancer UK
4. Has your nominating organisation provided a submission? Please tick all options that apply.	☐ No, (please review all the questions below and provide answers where possible) ☐ Yes, my nominating organisation has provided a submission ☐ I agree with it and **do not wish to** complete a patient expert statement X Yes, I authored / was a contributor to my nominating organisations submission X I agree with it and **do not wish to** complete this statement ☐ I agree with it and **will be** completing
5. How did you gather the information included in your	☐ I am drawing from personal experience.

statement? (please tick all that apply)	☐ I have other relevant knowledge/experience (e.g. I am drawing on others' experiences). Please specify what other experience: ☐ I have completed part 2 of the statement **after attending** the expert engagement teleconference ☐ I have completed part 2 of the statement **but was not able to attend** the expert engagement teleconference ☐ I have not completed part 2 of the statement

NICE National Institute for Health and Care Excellence

Technical engagement response form

Avelumab for treating metastatic Merkel cell carcinoma (CDF Review of TA517) [ID1617]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments: **Thursday 7 January 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.
- If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.

- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Neuroendocrine Cancer UK**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**Nothing to declare**

NICE National Institute for Health and Care Excellence

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	Does this response contain new evidence, data or analyses?	Response
Key issue 1: The ERG considers the updated JAVELIN data to be appropriate for use in the model. However, the NICE technical team would like to know how generalisable the JAVELIN trial is to clinical practice and how the Systemic Anti-Cancer Therapy database (SACT) data could be used for decision making.	YES	Generalisability of RCT data has been identified as problematic due to the strict criteria regarding recruitment that may reduce external validity – however consideration should also take into account the relative numbers included in both trial and SACT data as a proportion of total MCC population for whom this therapy would be suitable. Trial and SACT numbers may be small, relatively speaking, but then so too is the population for whom this therapy would apply – proportional representation. Combination of trial and real world data available (depending on assessment tool used) should provide a better understanding for decision-making, reducing some uncertainty: although criteria for considering treatment was less rigid (eg regarding PS) and data capture not as complete – Information from the SACT database on real world experience may help inform decision-making provided variables are taken into account. Also Walker et al (2020) who conclude : That the avelumab expanded access program for patients with mMCC demonstrated efficacy and safety in a real-world setting, consistent with the results from JAVELIN Merkel 200, and provided a treatment for patients with limited options. And more recent incidence data : Genus et al (2019) MCC UK incidence 0.62 per 100,000

		The annual incidence of MCC worldwide varies between 0.13 and 1.6 per 100,000 and appears to be increasing
Key issue 2: The ERG considers the propensity score weighting analysis 4 (PSW4 analysis), with adjustments for age, sex and Eastern Cooperative Oncology Group Performance Status (ECOG PS), appropriate to inform the avelumab versus chemotherapy comparison in the model. However, because immunocompetency was not included in the PSW4 analysis, and because of the ERG's concerns about JAVELIN generalisability (Issue 1), the NICE team considers that the following two scenarios are also relevant: • a naïve comparison using updated JAVELIN data (n=116) and the immunocompetent subgroup of the company's observational chemotherapy study (n=51), and • a naïve comparison using SACT data (n=52) and the	**YES/NO**	Clarification regarding whether naïve comparison would be direct or indirect

immunocompetent subgroup (n=51)		
Key issue 3: The ERG agrees with the company to focus the choice of avelumab OS curve on the 1-knot spline models but notes important differences in the extrapolations produced by the hazard-, normal- and odds-based splines. Because there is uncertainty in the naïve comparison of the treatment effects between avelumab and chemotherapy, the ERG considers it may be more appropriate to choose a curve with more conservative extrapolation. The ERG chose the hazard-based 1-knot spline for OS.	**YES/NO**	
Key issue 4: The ERG considers the 2-knot spline to underestimate the KM data between 0.5 years and 1 year. It considers the 3-knot odds spline to provide a better extrapolation as well as being the best fit to the data. The ERG chose 3-knot odds spline curve for PFS. However, the ERG considers the PSW adjusted analysis to be preferable (Issue 2)	**YES/NO**	

Key issue 5: The ERG agrees with the company, that it is reasonable to stop treatment at 5 years as this is likely to happen in clinical practice and used the same assumption in its preferred base-case. However, the ERG considers that the curves fitted to 1L data should not be adjusted by 2L+ data because this is not reflective of treatment-naïve population (1L). The ERG chose 3-knot hazard spline for the 1L data (2L+ data are not used in the ERG approach)	**YES/NO**	

NICE National Institute for Health and Care Excellence

Technical engagement response form

Avelumab for treating metastatic Merkel cell carcinoma (CDF Review of TA517) [ID1617]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments: **Thursday 7 January 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.
- If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.

- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**NCRI-ACP-RCP-RCR** The NCRI-ACP-RCP-RCR is grateful for the opportunity to respond to the above consultation. We have liaised with our experts and would like to make the following comments.
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

NICE National Institute for Health and Care Excellence

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	**Does this response contain new evidence, data or analyses?**	Response
Key issue 1: The ERG considers the updated JAVELIN data to be appropriate for use in the model. However, the NICE technical team would like to know how generalisable the JAVELIN trial is to clinical practice and how the Systemic Anti-Cancer Therapy database (SACT) data could be used for decision making.	No	Broadly speaking, the JAVELIN trial patient population does resemble those patients treated in real-world clinical practice, with a median age of 75 years, although our experts note that there was an excess of male patients (76%) and in view of the older age of this patient population, it seems unusual that 80% of patients had a WHO performance status of 0. About two-thirds of patients had visceral metastases which resembles clinical practice.
Key issue 2: The ERG considers the propensity score weighting analysis 4 (PSW4 analysis), with adjustments for age, sex and Eastern Cooperative Oncology Group Performance Status (ECOG PS), appropriate to inform the avelumab versus chemotherapy comparison in the model. However, because immunocompetency was not included in the PSW4 analysis, and because of the ERG's concerns about	No	These 2 comparisons do seem appropriate. Is it worth defining what immunocompetent refers to more precisely (i.e no CLL/haematological malignancy, no HIV, no bone marrow/organ transplant).

JAVELIN generalisability (Issue 1), the NICE team considers that the following two scenarios are also relevant: • a naïve comparison using updated JAVELIN data (n=116) and the immunocompetent subgroup of the company's observational chemotherapy study (n=51), and • a naïve comparison using SACT data (n=52) and the immunocompetent subgroup (n=51)		
Key issue 3: The ERG agrees with the company to focus the choice of avelumab OS curve on the 1-knot spline models but notes important differences in the extrapolations produced by the hazard-, normal- and odds-based splines. Because there is uncertainty in the naïve comparison of the treatment effects between avelumab and chemotherapy, the ERG considers it may be more appropriate to choose a curve with more conservative extrapolation. The	No	IA more detailed discussion needs to be had to consider the pros and cons of the models. It is not necessarily most appropriate to choose the most conservative model. Indeed, with immunotherapy where, in other cancers, we have seen long remissions and possible cures in a proportion of patients a less conservative and more optimistic model is appropriate. Therefore, it would in fact be better to wait longer before re-evaluating, so that there is longer follow up to allow a more robust and reliable modelling. Our experts question whether there a way in which the statisticians can model how much follow up would be appropriate in this sort of setting (where a proportion of patients on immunotherapy have long remissions- or a plateau on the survival graph).

ERG chose the hazard-based 1-knot spline for OS.		
Key issue 4: The ERG considers the 2-knot spline to underestimate the KM data between 0.5 years and 1 year. It considers the 3-knot odds spline to provide a better extrapolation as well as being the best fit to the data. The ERG chose 3-knot odds spline curve for PFS. However, the ERG considers the PSW adjusted analysis to be preferable (Issue 2)	No	In this setting, continuing the current guidance and revisiting the modelling when it is more robust (ie with longer follow up) would be appropriate.
Key issue 5: The ERG agrees with the company, that it is reasonable to stop treatment at 5 years as this is likely to happen in clinical practice and used the same assumption in its preferred base-case. However, the ERG considers that the curves fitted to 1L data should not be adjusted by 2L+ data because this is not reflective of treatment-naïve population (1L). The ERG chose 3-knot hazard spline for the 1L data (2L+ data are not used in the ERG approach)	No	5 years of continuous immunotherapy does seem like a prolonged period of treatment.

NICE National Institute for Health and Care Excellence

Technical engagement response form

Avelumab for treating metastatic Merkel cell carcinoma (CDF Review of TA517) [ID1617]

As a stakeholder you have been invited to comment on the ERG report for this appraisal. The ERG report and stakeholders' responses are used by the appraisal committee to help it make decisions at the appraisal committee meeting. Usually, only unresolved or uncertain key issues will be discussed at the meeting.

We need your comments and feedback on the key issues below. You do not have to provide a response to every issue. The text boxes will expand as you type. Please read the notes about completing this form. We cannot accept forms that are not filled in correctly. Your comments will be included in the committee papers in full and may also be summarised and presented in slides at the appraisal committee meeting.

Deadline for comments: **Thursday 7 January 2021**

Thank you for your time.

Please log in to your NICE Docs account to upload your completed form, as a Word document (not a PDF).

Notes on completing this form

- Please see the ERG report which summarises the background and submitted evidence, and presents the ERG's critique of the evidence and exploratory analyses. This will provide context and describe the questions below in greater detail.
- If you would like to comment on issues in the ERG report that have not been identified as key issues, you can do so in the 'Additional issues' section.
- If you are the company involved in this appraisal, please complete the 'Summary of changes to the company's cost-effectiveness estimates(s)' section if your response includes changes to your cost-effectiveness evidence.
- Please do not embed documents (such as PDFs or tables) because this may lead to the information being mislaid or make the response unreadable. Please type information directly into the form.
- Do not include medical information about yourself or another person that could identify you or the other person.
- Do not use abbreviations.
- Do not include attachments such as journal articles, letters or leaflets. For copyright reasons, we will have to return forms that have attachments without reading them. You can resubmit your form without attachments, but it must be sent by the deadline.
- If you provide journal articles to support your comments, you must have copyright clearance for these articles.

- Combine all comments from your organisation (if applicable) into 1 response. We cannot accept more than 1 set of comments from each organisation.
- Please underline all confidential information, and separately highlight information that is submitted under 'commercial in confidence' in turquoise, all information submitted under 'academic in confidence' in yellow, and all information submitted under 'depersonalised data' in pink. If confidential information is submitted, please also send a second version of your comments with that information replaced with the following text: 'academic/commercial in confidence information removed'. See the Guide to the processes of technology appraisal (sections 3.1.23 to 3.1.29) for more information.

We reserve the right to summarise and edit comments received during engagement, or not to publish them at all, if we consider the comments are too long, or publication would be unlawful or otherwise inappropriate.

Comments received during engagement are published in the interests of openness and transparency, and to promote understanding of how recommendations are developed. The comments are published as a record of the comments we received, and are not endorsed by NICE, its officers or advisory committees.

About you

Your name	[redacted]
Organisation name – stakeholder or respondent (if you are responding as an individual rather than a registered stakeholder please leave blank)	**Merck Serono and Pfizer Ltd. (Merck/Pfizer alliance)**
Disclosure Please disclose any past or current, direct or indirect links to, or funding from, the tobacco industry.	**None**

Key issues for engagement

Please use the table below to respond to questions raised in the ERG report on key issues. You may also provide additional comments on the key issue that you would like to raise but which do not address the specific questions.

Key issue	Does this response contain new evidence, data or analyses?	Response	ERG response
Key issue 1: The ERG considers the updated JAVELIN data to be appropriate for use in the model. However, the NICE technical team would like to know how generalisable the JAVELIN trial is to clinical practice and how the Systemic Anti-Cancer Therapy database (SACT) data could be used for decision making.	No	We agree with the ERG that the most appropriate data source to inform the model concerning the safety and efficacy of avelumab comes from the pivotal JAVELIN Merkel 200 (JM200) study (Part B). The JM200 is the largest clinical trial to date demonstrating the effectiveness and safety of a treatment in a Merkel cell carcinoma (MCC) patient cohort. Upon receiving the technical report, the trial data was further validated with a clinical expert who confirmed that the demographics of the JM200 study is representative of the UK mMCC patient population. Furthermore, the efficacy outcomes including OS and ORR are similar to outcomes expected in the clinical setting and the safety data reported in the study is equivalent to adverse events encountered. The SACT database collected real world data in a small patient cohort with a short follow-up period, outside of the clinical trial setting. The SACT data is useful in providing an additional data source to demonstrate the benefits of avelumab in 1L mMCC. To assist the NICE appraisal committee with understanding the key differences between the JM200 and SACT populations, and how this may affect interpretations related to generalisability, please see the table below. The table summarises the key features of these two cohorts of patients,	As discussed in the ERG report, the ERG and its clinical experts considers the SACT data set to comprise a closer match to expected patient characteristics in clinical practice in England. However, the SACT data set does not provide information on PFS, HRQoL, response rate or adverse effects of treatment.

<table>
<tr>
<td></td>
<td></td>
<td>while also highlighting (in bold) which characteristics in the respective populations most accurately represents UK clinical practice.

<table>
<tr><th>JM200: Part B</th><th>SACT</th></tr>
<tr><td>The key efficacy endpoint of the study was durable response rate (DRR)* and included other efficacy endpoints such as BOR, PFS, OS and duration of response. These endpoints provide clinical importance to assess the effectiveness of avelumab.</td><td>The only data points collected in the SACT were treatment duration and OS.</td></tr>
<tr><td>Median age 74.0 years</td><td>Median age 75.5 years</td></tr>
<tr><td>The clinical trial included 64.7% of patients recruited from Western Europe.</td><td>UK-based population in England</td></tr>
<tr><td>Large sample size (n=116)</td><td>Small sample size (n=52)</td></tr>
<tr><td>Long follow-up (minimum 15 months)</td><td>Short follow-up (minimum 5 months)</td></tr>
<tr><td>All patients had ECOG 0-1</td><td>Some patients with ECOG 2+ or unknown</td></tr>
<tr><td colspan="2">* Durable response rate (DRR)*=best overall response rate of ≥6 months</td></tr>
</table>

Merck/Pfizer recognises that both sources have features that are generalisable to the NHS population, and others that are less generalisable. Both data sets also demonstrate the substantial benefit of avelumab in the 1L mMCC setting. However, on balance, the JM200 study data are considered to provide the most reasonable basis to inform decision making to demonstrate the benefit of avelumab in 1L mMCC. The SACT data provides an additional data source which can be used in conjunction with JM200 to support clinical decision-making.</td>
<td>The ERG disagrees with the company’s view regarding ECOG status as the ERG’s clinical experts reported that they would expect the ECOG performance status of patients in clinical practice to be more similar to that of patients in the SACT data set compared to JM200: Part B.</td>
</tr>
<tr>
<td>Key issue 2: The ERG considers the propensity</td>
<td>Yes</td>
<td>As noted in the ERG’s report, the choice of propensity score weighting analysis has a relatively limited impact on the cost-effectiveness results</td>
<td>As mentioned in the ERG report,</td>
</tr>
</table>

score weighting analysis 4 (PSW4 analysis), with adjustments for age, sex and Eastern Cooperative Oncology Group Performance Status (ECOG PS), appropriate to inform the avelumab versus chemotherapy comparison in the model. However, because immunocompetency was not included in the PSW4 analysis, and because of the ERG's concerns about JAVELIN generalisability (Issue 1), the NICE team considers that the following two scenarios are also relevant:

- a naïve comparison using updated JAVELIN data (n=116) and the immunocompetent subgroup of the company's

(see ERG addendum Table 4). However, Merck/Pfizer acknowledges the importance of exploring the impact of adjusting for immunocompetency, as well as other potentially important characteristics.

In Study 100070-Obs001, a total of n=13 patients were immunocompromised, versus n=51 patients who were immunocompetent. Immunocompetency is not expected to have a large impact on the outcome of treatment with chemotherapy, within the context of an mMCC population, as these patients were considered fit enough to receive chemotherapy. However, immunocompetency was an important inclusion criterion within JM200, owing to the mechanistic properties of avelumab.

At clarification stage, a comparison between JM200: Part B (n=116) and the immunocompetent subgroup from Study 100070-Obs001 (n=51) was requested by the ERG and provided by Merck/Pfizer. However, subsequent to the provision of this analysis, a confidential simple patient access scheme (PAS) discount of [redacted] off the list price of avelumab was approved by NHS England. This was not reflected in the results presented by Merck/Pfizer. For completeness, the results of this analysis (including the PAS discount) are provided in the table below, compared to the company's preferred base-case analysis:

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company submission base-case analysis							
Chemotherapy	11,116	1.94	1.32				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947
Analysis using immunocompetent subgroup only*							
Chemotherapy	11,499	1.83	1.25				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,225**

the ERG considers the use of the subgroup of immunocompetent patients from Study 100070-Obs001 to be preferable in the naïve comparison. However, it is potentially unreliable because of the small number of patients in the studies, the imbalances in other patient characteristics between the two studies, and the uncertainty caused by unmeasured prognostic indicators or treatment effect modifiers. The ERG prefers the use of propensity score weighting rather than the naïve analyses and the ERG's

observational chemotherapy study (n=51), and

- a naïve comparison using SACT data (n=52) and the immunocompetent subgroup (n=51)

Key: ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years.

Notes: * Please note that the total costs for avelumab change marginally, as the average duration of survival for the chemotherapy arm is used within the estimated costings of radiotherapy for all treatment arms. **The ICER of £17,225 was also presented in the ERG's addendum in Table 5.

As described in response to clarification question B2, due to the reduced number of patients at risk within the immunocompetent subgroup (versus the whole population), relatively fewer patients are present in the tail of the Kaplan-Meier curves, leading to slightly lower longer-term estimates of both OS and PFS (e.g. 5-year OS using a log-logistic model is 7.6% for the whole population versus 6.8% for the immunocompetent population). The ICER is similar, but slightly lower than the base-case analysis.

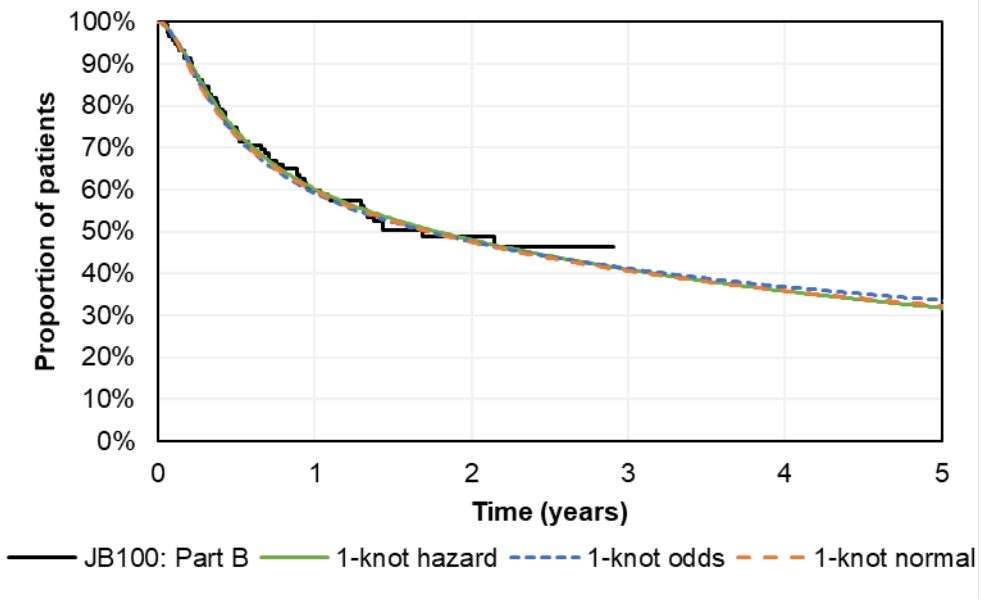
The second analysis highlighted in this Key Issue was not requested by the ERG at the clarification stage but is possible to generate within the model edited by the ERG to inform its report. The corresponding results of the second analysis are provided in the table below for completeness. However, Merck/Pfizer urges caution when interpreting these results in line with the limitations of the SACT dataset described in response to Key Issue 1:

Technologies	Total			Incremental			ICER (£/QALY)
	Costs (£)	LYG	QALYs	Costs (£)	LYG	QALYs	
Company submission base-case analysis							
Chemotherapy	11,116	1.94	1.32				
Avelumab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	17,947
Analysis using immunocompetent subgroup only versus SACT data set*							
Chemotherapy	11,499	1.83	1.25				

preferred analysis is PSW4 as it maintains all patients in the analysis and has the best balance in baseline characteristics after matching for all characteristics matched other than immune status. The ERG considers that from a cost-effectiveness perspective, using the immunocompetent subgroup from Study 100070-Obs001 does not have a substantial impact on the ICER, but nonetheless considers it was a key scenario to explore for decision making.

		<table><tr><td>Avelumab</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>[redacted]</td><td>22,252</td></tr></table> **Key:** ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years; SACT, Systemic Anti-Cancer Therapy. **Notes:** * Please note that the total costs for avelumab change marginally, as the average duration of survival for the chemotherapy arm is used within the estimated costings of radiotherapy for all treatment arms. The total costs, QALYs, and LYs are identical to the results presented for the previous analysis, so are not discussed further here. When switching to using the SACT data set, the total costs for avelumab increase slightly, whereas the total QALYs and LYs decrease slightly – the combined effect of these changes leads to an increase in the ICER from £17,947 (company base-case analysis) to £22,252. In summary, Merck/Pfizer do not expect immunocompetency to impact the outcome of treatment in the chemotherapy arm and therefore considers a comparison to the full observational chemotherapy study as most appropriate. Limitations to comparing to the SACT data set have been documented above under Key Issue 1.	
Key issue 3: The ERG agrees with the company to focus the choice of avelumab OS curve on the 1-knot spline models but notes important differences in the extrapolations produced by the hazard-, normal- and odds-based splines. Because there is uncertainty in the naïve	No	The choice of the most appropriate extrapolation of overall survival is challenging within the context of maturing trial data. In choosing between the three 1-knot models, the difference in long-term survival may be difficult to validate, as all three models produce estimates that are broadly in keeping with advice provided to Merck/Pfizer, and produce near-identical fits to the Kaplan-Meier curve (shown in the diagram below):	The ERG thanks the company for providing the hazards of death plot over time for the three 1-knot spline models compared with TA517 base case. As mentioned in the ERG report, there is uncertainty in the naïve comparison

comparison of the treatment effects between avelumab and chemotherapy, the ERG considers it may be more appropriate to choose a curve with more conservative extrapolation. The ERG chose the hazard-based 1-knot spline for OS.

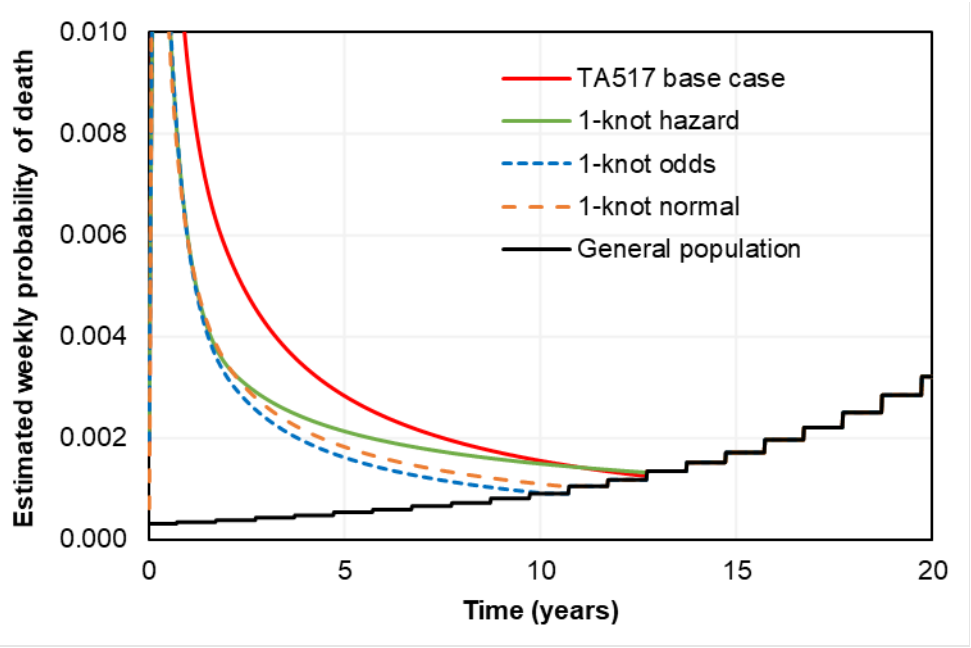


It may instead be helpful to consider a comparison of the different survival models at specific time points in the longer term, as shown in the table below:

Time (years)	1-knot hazard	1-knot odds	1-knot normal
5	31.9%	33.7%	32.4%
10	20.1%	24.6%	22.5%
15	13.8%	17.8%	16.2%
20	7.5%	9.7%	8.8%
25	2.4%	3.1%	2.9%
30	0.3%	0.4%	0.4%
35	0.0%	0.0%	0.0%
40	0.0%	0.0%	0.0%

of the treatment effects between avelumab and chemotherapy, as such the ERG considers it may be more appropriate to choose a curve with more conservative extrapolation. However, the ERG acknowledges that the three 1-knot models presented by the company are very similar and, on its own, the selection of model is unlikely make a difference to decision making.

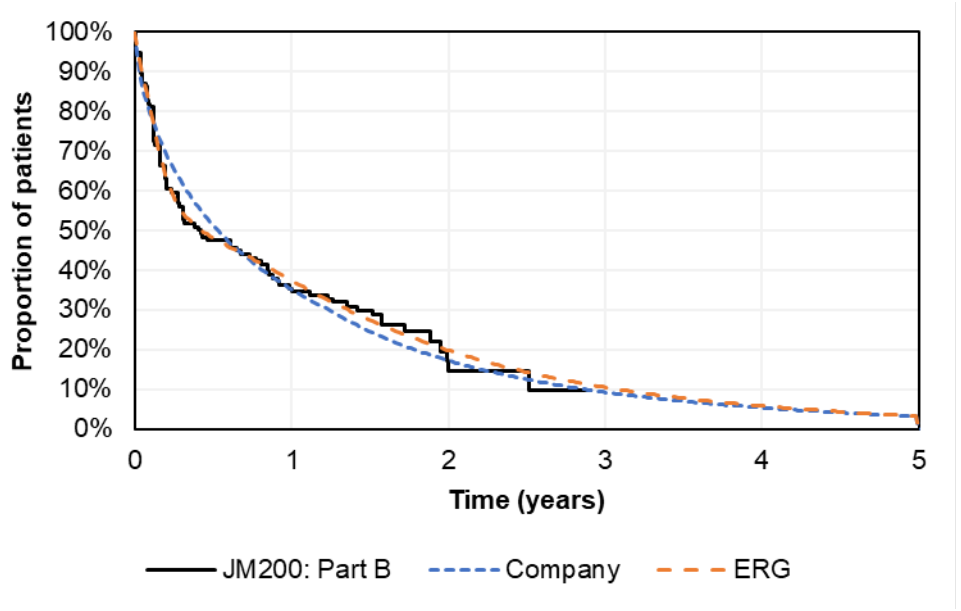
It should also be noted that a 1-knot odds-based spline was selected to inform the base-case analysis adopted in TA517 (for the 2L+ population). When comparing the estimated hazard of death inherent within the base-case analysis in TA517 for the 2L+ population versus each of the 1-knot models fitted to the JB100: Part B data, the following plot is produced:



The plot above demonstrates that the projected hazards produced by the 1-knot hazard spline result in an extrapolation which eventually produces an estimated hazard of death which exceeds that of the base-case analysis presented in TA517. While this does not occur until approximately 11 years, this is misaligned with clinical opinion that outcomes for patients treated in the 1L setting are expected to be better than those for a 2L+ population. Moreover, the shape of the hazard function implies a slightly different shape

		for the 1L versus 2L+ populations if a 1-knot hazard-based spline is selected. Merck/Pfizer therefore considers its preferred base-case analysis (using a 1-knot odds-based spline) to provide a more appropriate estimation of OS versus the ERG's preferred 1-knot hazard model. However, each of the three 1-knot spline-based models may be helpful in decision making.	
Key issue 4: The ERG considers the 2-knot spline to underestimate the KM data between 0.5 years and 1 year. It considers the 3-knot odds spline to provide a better extrapolation as well as being the best fit to the data. The ERG chose 3-knot odds spline curve for PFS. However, the ERG considers the PSW adjusted analysis to be preferable (Issue 2)	No	A small impact on the ICER is noted when the PFS extrapolation is changed from Merck/Pfizer preferred base-case analysis using a 2-knot spline model to the ERG's preferred 3-knot spline model. Merck/Pfizer considers there to be relatively little evidence to definitively reject either one of these models in favour of the other, and so considers both approaches to be suitable to inform decision making. The PSW-adjusted analyses affect the estimation of PFS for the chemotherapy arm (as all patients enrolled in JM200 were immunocompetent). Merck/Pfizer has responded in relation to the PSW analyses within the context of Key Issue 2.	The ERG's position is unchanged by the company's response to this issue. In terms of evidence to reject one model over the other, the ERG considers that the model fit statistics and visual inspection of the curves supports the use of the 3-knot odds spline model and is also the more conservative approach.
Key issue 5: The ERG agrees with the company, that it is reasonable to stop treatment at 5 years as this is likely to happen	No	Merck/Pfizer acknowledges the view of the ERG regarding the potential issues with using non-1L data to inform the extrapolation of time-on-treatment (ToT). In Merck/Pfizer's base-case analysis, JM200: Part A (2L+) data were considered for use within the model to inform the rate of treatment	The ERG's position is unchanged by the company's response to this issue.

in clinical practice and used the same assumption in its preferred base-case. However, the ERG considers that the curves fitted to 1L data should not be adjusted by 2L+ data because this is not reflective of treatment-naïve population (1L). The ERG chose 3-knot hazard spline for the 1L data (2L+ data are not used in the ERG approach)		discontinuation beyond the minimum follow-up period of JM200: Part B (1L). This approach was undertaken to supplement the limited data available for longer-term treatment discontinuation available from JM200: Part B (1L), with more mature data from JM200: Part A (2L+) while also maintaining a model based solely on JM200: Part B (1L) data for the earlier portion of the curve. The base-case approach was also proposed in keeping with recommendations from the ERG as part of TA517, wherein a Weibull model was selected to inform ToT for the ERG's preferred base-case analysis. The 3-knot hazard-based spline could be considered an extension of a Weibull model, but has a greater reliance on fit to the JB200: Part B data than Merck/Pfizer's preferred approach. The figure below presents a comparison of Merck/Pfizer's and the ERG's preferred approach to modelling ToT. From this figure, it can be seen that the ERG's preferred model provides a relatively better visual fit to the Kaplan-Meier curve, which is unsurprising given its increased flexibility versus the Weibull-based approach chosen by Merck/Pfizer.	



It could potentially be argued that the 3-knot spline-based approach overfits the Kaplan-Meier curve versus a more simplified model based on interpretation of BIC scores which apply a higher penalty for more complex models versus AIC scores. For example, the Weibull model (without adjustment) has a BIC score of 1,221.90 versus 1,223.31 for the 3-knot hazard-based spline model. However, the 3-knot spline-based model has a better AIC versus the Weibull model, and the AIC/BIC scores do not account for the incorporation of JM200: Part A data within Merck/Pfizer's preferred base-case analysis.

Nevertheless, both models result in very similar mean ToT estimates (Merck/Pfizer: 12.59 months, versus ERG: 13.07 months), and so the impact on the ICER is small. Merck/Pfizer views its approach as being based on a relatively simple model but making use of external data (i.e. data from JM200: Part A), whereas the ERG's approach makes use of a

		more flexible modelling approach, without relying on external data. Therefore, both approaches may be helpful to inform decision making.	

Additional issues

Please use the table below to respond to additional issues in the ERG report that have not been identified as key issues. Please do **not** use this table to repeat issues or comments that have been raised at an earlier point in this appraisal (e.g. at the clarification stage).

Issue from the ERG report	Relevant section(s) and/or page(s)	Does this response contain new evidence, data or analyses?	Response	ERG response

Additional issue 1: Weight-based versus flat dosing	ERG report Section 1.3.	No	The ERG's report provides a scenario analysis using the weight-based dose of avelumab (per the JM 200 study), though the licensed dose of avelumab is now based on a flat dosing regimen. Merck/Pfizer highlights that since the time of its original submission (and production of the ERG's report), the SmPC has now been fully updated to reflect the flat dosing regimen, and NICE has published its guidance concerning the flat dose of avelumab as part of TA645 (avelumab with axitinib for untreated advanced renal cell carcinoma). With respect to the dosing regimen, the FAD for TA645 states: *"The committee highlighted that a weight-based dose for avelumab was used in JAVELIN Renal 101, whereas the licence specifies a fixed dose. The companies explained that they derived the fixed dose using pharmacokinetic and pharmacodynamic data, and taking into account similar approaches used historically. The Cancer Drugs Fund clinical lead advised that this approach was taken with other drugs for this disease area. The committee was aware that it could appraise drugs only within their marketing authorisation. It accepted that the licensed fixed dose would have similar effectiveness to the weight-based dose, and concluded that it would use the licensed dose in making decisions."*	The ERG report addendum presents the ERG preferred assumptions using the avelumab flat dosing regimen in Table 11, scenario 8 for the committee's consideration and can be considered as the ERG's base case in light of the recommendations of dosing in the avelumab SmPC.

			The publication of the FAD for TA645 is highlighted here for completeness. In conclusion, Merck/Pfizer disagree with the ERG's suggestion that weight-based acquisition costs (instead of the 800mg flat dose) should be applied in the economic model as this would not reflect the licensed dose of avelumab, how it will be used in clinical practice and is inconsistent with NICE guidance for other avelumab indications.	

Summary of changes to the company's cost-effectiveness estimate(s)

Company: If you have made changes to the company's preferred cost-effectiveness estimate(s) in response to technical engagement, please complete the table below to summarise these changes.

No change made to company's preferred base-case cost-effectiveness results.