NICE National Institute for Health and Care Excellence

Rapid Review

Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer (review of TA885) [ID6279]

Committee Papers

NICE National Institute for Health and Care Excellence

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Rapid Review

Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer (review of TA885) [ID6279]

Contents:

The following documents are made available to stakeholders:

Access the **final scope and final stakeholder list** on the NICE website.

1. **Company submission** from MSD
2. **Clarification questions and company responses**
3. **External Assessment Report** prepared by CRD and CHE Technology Assessment Group, University of York

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

Confidential

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Rapid review

Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer [TA885]

Company evidence submission

May 2023

File name	Version	Contains confidential information	Date
TA885 Rapid Review Pembrolizumab Cervical Cancer Redacted.docx	**1.0**	**Yes**	**31st May 2023**

Confidential

Executive Summary

This report contains updated clinical and cost-effectiveness results based on the KEYNOTE-826 Final Analysis, representing an additional 17 months of follow up on the Interim Analysis data previously considered by the NICE committee.

Briefly, the trends observed within the interim analysis have continued; an extensive plateau is observed in Progression Free Survival and is apparent in Overall Survival in the pembrolizumab combination arm. Median Overall Survival has now been reached in the pembrolizumab combination arm. As before, these trends are driven by lower event rates among Complete and Partial Responders. Time on Treatment has lengthened slightly as some patients who were previously censored now have more observations available. Adverse Event profiles remain similar to the Interim Analysis.

The Final Analysis data validate the company's spline-based model, considered by the committee at ACM2. This was the most optimistic of the survival analyses submitted by the company and a scenario in which pembrolizumab combination was comfortably cost-effective. All other models considered by the committee at ACM2 underestimated the Progression Free and Overall Survival that has now been observed.

[redacted]

The cost-effectiveness model has been updated with new survival analysis having been conducted. In all other respects, the cost-effectiveness model remains consistent with latest appraisal assumptions and committee preferences, as stated in 3.14 of the Final Appraisal Determination document (FAD). The additional maturity of the data has helped to narrow the range of appropriate and credible methods for extrapolation of PFS, TTP and PPS. Non-survival parameters such as utility values, Adverse Events and use of subsequent treatments have also been updated using the latest data from KEYNOTE-826.

Including the Commercial Arrangement for pembrolizumab, the model produces a base case incremental cost-effectiveness ratio (ICER) of [redacted] gained for pembrolizumab combination with scenario analyses in a range of [redacted] gained and a ~95% probability of being cost-effective versus a threshold of £50,000/QALY in

Confidential

probabilistic sensitivity analysis. These ICERs are significantly below the relevant NICE decision threshold indicating that minimal decision uncertainty remains following the Final Analysis from KEYNOTE-826.

Confidential

Structure of Document

Introduction (Section 1) – a brief summary of the TA885 Final Appraisal Determination (FAD).

Clinical Effectiveness (Section 2) – key clinical outcomes, including Progression Free Survival (PFS), Time to Progression (TTP), Overall Survival (OS), Response Rate (RR), subsequent treatments, Post Progression Survival (PPS) and Adverse Events (AEs) from the KEYNOTE-826 Final Analysis (FA).

Cost-effectiveness (Section 3) – comparison of the FA data with the Interim Analysis 1 (IA1) data and economic modelling predictions previously considered by the NICE committee. Updated survival analysis/survival curve selection for PFS, TTP and PPS. Updated Time on Treatment (ToT) and HRQoL analysis.

Results (Section 3.7) – cost-effectiveness results from the updated model including sensitivity and scenario analyses. Discussion and conclusions from the analysis (Section 4).

Confidential

1. Introduction

The appraisal 'Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer' (ID3798) resulted in the publication of TA885 on the 3rd of May 2023 (1). This was a recommendation for the indication to enter the Cancer Drugs Fund.

From the Final Appraisal Document issued on 29th March 2023: (2)

The committee concluded that the most plausible ICERs may be within the range usually considered a cost-effective use of resource when the end of life modifier was applied, but these were associated with high uncertainty. Collecting more evidence may reduce this uncertainty.

This rapid review provides the NICE committee with updated clinical effectiveness data from the Final Analysis from the KEYNOTE-826 trial, which has been implemented in the cost effectiveness model, in line with the committee's preferences outlined in the TA885 FAD.

2. Clinical effectiveness - long-term results data from the KEYNOTE-826 study

The NICE appraisal resulting in TA885 was based upon clinical effectiveness evidence from the first interim analysis (data cut off 3rd May 2021) of the KEYNOTE-826 clinical trial. The final planned analysis (FA) of KEYNOTE-826 reports outcomes up to the 3rd of October 2022, an additional 17 months of follow-up time. The analysis was performed with a median duration of follow-up of 28.6 months (range 0.5-46.5 months) in the pembrolizumab arm and 16.5 months (range 0.3-46.2) in the control arm for participants with a CPS ≥1.

For brevity, pembrolizumab plus chemotherapy with or without bevacizumab will be referred to as the pembrolizumab combination (Pem+SoC) arm and the control arm as the placebo combination (SoC) arm.

2.1. Disposition of patients

Table 1: Disposition of participants (CPS ≥ 1 population, ITT, Final analysis)

	Pembrolizumab combination		Placebo combination	
	n	(%)	n	(%)
Participants in population	273		275	
Status for Trial				
Discontinued	[redacted]	[redacted]	[redacted]	[redacted]
Death	[redacted]	[redacted]	[redacted]	[redacted]
Lost To Follow-Up	[redacted]	[redacted]	[redacted]	[redacted]
Withdrawal By Subject	[redacted]	[redacted]	[redacted]	[redacted]
Participants Ongoing	[redacted]	[redacted]	[redacted]	[redacted]
Status for Study Medication in Trial				
Started	272		275	
Completed	[redacted]	[redacted]	[redacted]	[redacted]
Discontinued	[redacted]	[redacted]	[redacted]	[redacted]
Adverse Event	[redacted]	[redacted]	[redacted]	[redacted]
Clinical Progression	[redacted]	[redacted]	[redacted]	[redacted]
Complete Response	[redacted]	[redacted]	[redacted]	[redacted]
Excluded Medication	[redacted]	[redacted]	[redacted]	[redacted]
Physician Decision	[redacted]	[redacted]	[redacted]	[redacted]
Progressive Disease	[redacted]	[redacted]	[redacted]	[redacted]
Protocol Violation	[redacted]	[redacted]	[redacted]	[redacted]
Withdrawal By Subject	[redacted]	[redacted]	[redacted]	[redacted]
Participants Ongoing	[redacted]	[redacted]	[redacted]	[redacted]

	Pembrolizumab combination		Placebo combination	
	n	(%)	n	(%)
If the overall count of participants is calculated and displayed within a section in the first row, then it is used as the denominator for the percentage calculation. Otherwise, participants in population is used as the denominator for the percentage calculation. Database cut off 3rd October 2022				

2.2. Progression Free Survival

Progression Free Survival (PFS) at FA was consistent with that observed at IA1 for both arms of the trial. The trends observed in IA1 have continued in the Final Analysis. Particularly notable is the extension of the plateau in PFS in the pembrolizumab combination arm. The hazard ratio at the final analysis was 0.58 (95% CI: 0.47, 0.71) and is confirmatory of the HR observed at IA1 0.62 (95% CI: 0.50, 0.77) (3). We note that the point estimate has improved slightly and the confidence interval is now tighter. The median PFS reported at the FA was 10.5 months (95% CI: 9.7, 12.3) and 8.2 months (95% CI: 6.3, 8.5) for pembrolizumab and control arms, respectively (Table 3), which is understandably almost exactly the same as in IA1.

Mean and median PFS in the patients in the pembrolizumab and placebo arms are summarised in Table 2.

Figure 1: Kaplan-Meier estimates of Progression Free Survival (CPS ≥ 1 population, ITT)

[redacted]

Table 2: Analysis of PFS (CPS ≥ 1 participants, ITT)

	Pembrolizumab combination (n = 273)	Placebo combination (n = 275)
Number of events, n (%)	[redacted]	[redacted]
Median PFS, months (95% CI, months) [a]	10.5 (9.7, 12.3)	8.2 (6.3, 8.5)
PFS HR (95% CI) [b]	0.58 (0.47, 0.71)	
p-value [c]	<0.0001	
6-month PFS, % (95% CI)	81.5 (76.2, 85.7)	67.1 (61.0, 72.4)
12-month PFS, % (95% CI)	45.6 (39.3, 51.6)	33.7 (27.9, 39.5)
18-month PFS, % (95% CI)	[redacted]	[redacted]
24-month PFS, % (95% CI)	[redacted]	[redacted]
Key: CI, confidence interval; CPS, combined positive score; FIGO, International Federation of Gynaecology and Obstetrics; HR hazard ratio; NR, not reached; OS, Overall Survival; PD-L1, Programmed death-ligand 1.		

Confidential

Notes: [a] From product-limit (Kaplan–Meier) method for censored data. [b] Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10). [C] One-sided p-value based on log-rank test stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10).
Source: KEYNOTE-826 Clinical study report. Database cut off 3rd October 2022

2.3. Time to Progression

Time to Progression (TTP) is defined as the time from randomization to the first documented disease progression. It differs from PFS in that deaths are considered censors rather than events.

An additional [redacted] and [redacted] patients had progressed in the 17 months between IA1 and FA in the pembrolizumab and control arms, respectively. Consistent with the IA1 data, the median TTP was [redacted] in the pembrolizumab arm and [redacted] in the control arm. The Kaplan Meier curves in Figure 2 shows the apparent slowing of TTP in the pembrolizumab arm from ~1 year.

Figure 2: Kaplan Meier Curves of TTP Based on Investigator Assessment (CPS ≥1)

[redacted]

2.4. Overall Survival

Overall Survival (OS) results at FA were consistent with that observed at IA1 for pembrolizumab and control arms, with a plateau continuing to emerge in the pembrolizumab combination arm. The hazard ratio at the final analysis was 0.60 (95% CI: 0.49, 0.74) and is confirmatory of the HR observed at IA1, 0.64 (0.50, 0.81) (3). The median OS, which had not been reached at IA1, reported at the FA was 28.6 months (95% CI: 22.1, 38.0) and 16.5 months (95% CI: 14.5, 20.0) for pembrolizumab and control arms, respectively (Table 3).

Figure 3: Kaplan-Meier estimates of Overall Survival (CPS ≥ 1 population, ITT)

[redacted]

Confidential

Table 3: Analysis of OS (CPS ≥ 1 participants, ITT)

	Pembrolizumab combination (n = 273)	Placebo combination (n = 275)
Number of events, n (%)	[redacted]	[redacted]
Median OS, months (95% CI, months) [a]	28.6 (22.1, 38.0)	16.5 (14.5, 20.0)
OS HR (95% CI) [b]	0.60 (0.49, 0.74)	
p-value [c]	<0.0001	
6-month OS, % (95% CI)	91.9 (88.0, 94.6)	85.5 (80.7, 89.1)
12-month OS, % (95% CI)	[redacted]	[redacted]
18-month OS, % (95% CI)	[redacted]	[redacted]
24-month OS, % (95% CI)	53.5 (47.4, 59.2)	39.4 (33.6, 45.2)

Key: CI, confidence interval; CPS, combined positive score; FIGO, International Federation of Gynaecology and Obstetrics; HR hazard ratio; NR, not reached; OS, Overall Survival; PD-L1, Programmed death-ligand 1.
Notes: [a] From product-limit (Kaplan–Meier) method for censored data. [b] Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10). [C] One-sided p-value based on log-rank test stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10).
Source: KEYNOTE-826 Clinical study report. Database cut off 3rd October 2022

2.5. Response Rate

Table 4 presents the best Overall Response Rate (ORR) per (per RECIST 1.1) by investigator assessment.

The proportion of CPS ≥ 1 participants who achieved complete response (CR) or partial response (PR) was significantly greater in the pembrolizumab group than those treated with placebo ([redacted]% versus [redacted]%, respectively).

Compared with the IA1 results the number of participants who achieved a complete response increased in both arms, by eight and four in the pembrolizumab and placebo combination arms, respectively. The majority of these were participants who had been recorded in the partial response group at the IA1 data cut and moved to complete response.

Confidential

Table 4: Summary of Best Objective Response (Confirmed) based on investigator assessment per RECIST 1.1 (CPS ≥1 participants, ITT)

	Pembrolizumab combination (n=273)	**Placebo combination (n=275)**
Complete response (CR) n (%)	[redacted]	[redacted]
Partial response (PR) n (%)	[redacted]	[redacted]
Objective response (CR+PR) n (%)	[redacted]	[redacted]
Stable disease n (%)	[redacted]	[redacted]
Progressive disease n (%)	9 (3.3)	29 (10.5)
Not evaluable n (%)	1 (0.4)	2 (0.7)
No assessment n (%)	19 (7.0)	18 (6.5)
Notes: Investigator assessed responses per RECIST 1.1 (confirmed) are included in this table. Database Cutoff Date: 3rd October 2022		

Figure 4 and Figure 5 show the OS and PFS Kaplan Meir curves by responder category. The two-year OS and PFS among complete responders in the pembrolizumab combination arm was [redacted]% and [redacted]%, respectively. It can be seen that there have been very few additional PFS and OS events among patients who responded well to treatment with pembrolizumab combination, particularly complete responses between IA1 and the FA.

Confidential

Figure 4: Kaplan Meir estimates of Overall Survival by response category (CPS ≥1 population)

Figure 5: Kaplan Meir estimates of Progression Free Survival by response category (CPS ≥1 population)

Confidential

2.6. Subsequent treatments

Table 5: Utilization of Subsequent Oncologic Therapies by Line and Type Participants who completed or discontinued from Study Treatment (CPS ≥1, APaT)

	Pembrolizumab combination (n = 256)	Placebo combination (n = 263)
Participants who did not receive subsequent systemic oncologic therapies, n (%)	[redacted]	[redacted]
Participants who died	[redacted]	[redacted]
Participants who are alive	[redacted]	[redacted]
Participants who received subsequent systemic oncologic therapies	[redacted]	[redacted]
1st subsequent line	[redacted]	[redacted]
2nd subsequent line	[redacted]	[redacted]
3rd subsequent line	[redacted]	[redacted]
4th subsequent line	[redacted]	[redacted]
5th subsequent line	[redacted]	[redacted]
6th subsequent line	[redacted]	[redacted]
Every participant is counted once for each applicable row and column. Percentages are based on the number of participants who completed or discontinued study treatment. Database cut off 3rd October 2022		

Overall, the proportion of progressed patients receiving subsequent oncologic therapies is somewhat higher in both arms than was estimated at the MSD UK advisory board (see Section 3.6.2; 65% across both arms).

Although they were not commonly available at the trial sites, we examined the data to see how many patients had received subsequent treatment with immune checkpoint inhibitors.

[redacted] in the control arm and [redacted] in the pembrolizumab arm had immunotherapies as subsequent treatment. In the UK there are no immunotherapies are currently available for advanced cervical cancer. It is therefore likely that the trial overestimates outcomes (particularly PPS) in the control arm vs. what would be seen in the UK setting.

2.7. Post Progression Survival

The median Post Progression Survival was [redacted] months (95% CI: [redacted]) and [redacted] months (95% CI: [redacted]) for the pembrolizumab combination and placebo combination, respectively.

Figure 6: Kaplan Meier curves of Post Progression Survival based on investigator assessment (CPS ≥1)

2.8. Adverse Events

The percentage of participants with a reported adverse event (AE), drug related AE or death due to AE was the same as IA1 in both arms. In the remaining categories the percentage increased by one or two percentage points (see table 13 in the Company Submission, document B).

Table 6: Summary of adverse events (APaT)

	Pembrolizumab combination (n = 307)	Placebo combination (n = 309)
Participants with…		
≥1 AE	305 (99.3)	307 (99.4)
≥ 1 drug related AE	298 (97.1)	300 (97.1)
≥ 1 grade 3-5 AE	253 (82.4)	233 (75.4)
≥ 1 grade 3-5 drug related AE	212 (69.1)	201 (65.0)
≥ 1 SAE [a]	157 (51.1)	132 (42.7)
≥ 1 Drug related SAE [a]	94 (30.6)	73 (23.6)
Deaths due to drug related AE [b]	2 (0.7)	4 (1.3)
Discontinued any drug due to an AE	125 (40.7)	91 (29.4)

AE: Adverse event. SAE: Serious adverse event.
[a] SAEs/Drug related SAEs up to 90 days after last dose as determined by the investigator
[b] AEs resulting in death up to 90 days after last dose
Database cut off 3rd October 2022

3. Cost-effectiveness

The cost-effectiveness model uses all the assumptions agreed by the committee in the FAD e.g. treatment effect waning from 3-5 years, independent curves for the PPS health state and progression-based utilities. The only difference is that the model's

Confidential

parameters have been updated to include the Final Analysis data from KEYNOTE-826.

3.1. Comparison of survival data between IA1 and FA

Below are figures with the IA1 and FA Kaplan-Meier curves for PFS, TTP, OS and TTP overlaid on top of one another. It can be seen that for PFS, TTP and OS the FA data are confirmatory of those seen at IA1. PPS is somewhat different, and this is explained in detail in Section 3.3.2.

Figure 7: Time to Progression comparison – IA1 and FA

Figure 8: Progression Free Survival comparison – IA1 and FA

Figure 9: Post Progression Survival comparison – IA1 and FA

Figure 10: Overall Survival Comparison – IA1 and FA

Confidential

3.2. IA1 economic model vs. FA KM data

In section 3.7 of the FAD, the committee identify the method of extrapolation for PFS and TTP as a principle uncertainty in the appraisal.

Figure 11 (TTP) and Figure 12 (TTP) compare the different extrapolations presented at ACM2 (two-piece, one-piece, Response Based Model (RBM) and 2-knot spline) with the Final Analysis KM data from KEYNTOE-826.

It can be seen that all models considered at ACM2 except the 2-knot spline underpredict the PFS and TTP observed in the Final Analysis, particularly in the pembrolizumab arm. It is notable that the economic model using the 2-knot spline produced an incremental cost-effectiveness ratio (ICER) significantly below £50,000/QALY gained.

Figure 11: Progression-free Survival comparison – IA1 modelled (shared at ACM2) and FA KM

Figure 12: Time To Progression comparison – IA1 modelled (shared at ACM2) and FA KM

3.3. Survival analysis

3.3.1. Progression-free Survival and Time To Progression

As discussed in section 3.7 of the FAD, the company submitted several approaches for extrapolating PFS and TTP; one-piece, two-piece, spline-based models and an exploratory response-based model. The reason such an array of approaches was explored was to help identify an approach that appropriately fitted the complex shape of the hazard function for these outcomes; the heterogeneous hazard rates experienced by patients in different response categories (see Figure 5) meant that none of the standard curves that calculate survival as a simple function of time (once-piece models) were able to fit the KM data well. The committee noted that a range of approaches capable of capturing complex hazard functions was helpful for decision making. The exploratory Response Based Model aside, we have repeated the same approaches already considered by the committee.

Confidential

3.3.1.1. One-piece models

It can be seen in the graph below (Figure 13 and Figure 14) that no one-piece models are able to capture the complex hazard function for PFS and TTP for the pembrolizumab combination arm. The visual fit is now much poorer than the (already poorly-fitting) curves considered during the original appraisal.

There are some one-piece curves that provide a reasonable visual fit to the standard of care arm but the company feel very strongly that this should not be used as a reason to model the pembrolizumab combination arm using one-piece curves.

The statistical fit for one-piece models versus the more flexible spline-based approaches is compared via the Akaike Information Criterion (AIC) later on in the document and finds that all spline models provide lower AIC than one-piece models for both arms (Table 11).

The evidence on visual and statistical fit strongly suggests that one-piece curves are unsuitable for modelling the pembrolizumab combination arm. We have therefore not reported any results using them but have implemented them as an option in the economic model due to their methodological orthodoxy.

Figure 13: PFS KM FA vs. One-piece model

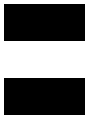
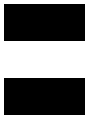


Figure 14: TTP KM FA vs. One-piece model



3.3.1.2. Two-piece models

The company's original base-case used KM data for PFS and TTP up to 37-weeks and extrapolated from that point using one-piece curves. The selection of the cut-point was based on a variety of factors that are detailed in the original submission (statistical and visual fit to KM, assessment of smooth spline hazard trajectory, clinical plausibility of longer term estimates), for example both arms saw a peak in the smooth spline hazard rate at about 37 weeks followed by steady decline. A 46-week cut-point was

Confidential

examined in sensitivity analysis but received little discussion in the original appraisal so we do not spend time discussing it here. The 46-week data have, however, been implemented in the economic model and are used in a scenario analysis.

We have repeated the 37-week Two-piece analysis and it can be seen that several options provide good visual fit to the data in both arms. The criteria we used to assess which curve to implement beyond the cut point were:-

1. Statistical fit assessed by AIC/BIC
2. Visual Fit to the KM curve
3. Clinical plausibility of extrapolations

Table 7 shows a summary of the selection criteria for the two-piece models. The detail of each criteria is discussed below. As in the original submission we selected the same type of model for TTP as PFS, due to them being comprised of almost the same data.

Table 7: Selection criteria for 37-week two-piece PFS curves

-	Exponential	Weibull	Log-normal	Log-logistic	Gompertz	Generalised gamma
Visual fit for Pembro+SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Visual fit for SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AIC/BIC for Pembro+SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AIC/BIC for SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical Plausibility Pembro+SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Clinical Plausibility SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

For the SoC arm, there is no obvious difference in AIC between the options whereas for the pembrolizumab combination arm the generalised gamma and Gompertz models have the lowest AIC.

Table 8: AIC for 37-week two-piece models for PFS

	Pembrolizumab + SoC		SoC	
Model	**AIC**	**AIC Rank**	**AIC**	**AIC Rank**
Exponential	[redacted]	[redacted]	[redacted]	[redacted]

Confidential

Weibull	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]

We implemented the curves in the economic model, with settings consistent with the committee's preferences outlined in section 3.10 of the FAD (e.g. treatment effect waning 3-5 years), examined the 5-year, 10-year, 15-year and 20-year OS projections and validated them using the clinical opinion that had been provided at the MSD UK advisory board (comprising seven UK clinicians currently treating advanced cervical cancer in the NHS) and by experts at NICE ACM1 and ACM2 to determine longer term plausibility. The original estimates that were deemed plausible were 7% OS at 20 years for the pembrolizumab+SoC arm and 1.3% OS in the SoC arm, although the clinical advisors noted that this was slightly optimistic and 0.5%-1% might be more plausible. The only model pairs with plausible 20-year survival were log-normal, log-logistic and generalised gamma.

Table 9: Landmark survival estimates from economic model - 37-week two-piece models

37-week two-piece models		**Overall Survival estimate from economic model**			
Model type	**Arm**	**5-years**	**10-years**	**15-years**	**20-years**
Exponential	Pem+SoC	[redacted]	[redacted]	[redacted]	[redacted]
	SoC	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	Pem+SoC	[redacted]	[redacted]	[redacted]	[redacted]
	SoC	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	Pem+SoC	[redacted]	[redacted]	[redacted]	[redacted]
	SoC	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	Pem+SoC	[redacted]	[redacted]	[redacted]	[redacted]
	SoC	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	Pem+SoC	[redacted]	[redacted]	[redacted]	[redacted]
	SoC	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma	Pem+SoC	[redacted]	[redacted]	[redacted]	[redacted]
	SoC	[redacted]	[redacted]	[redacted]	[redacted]

Confidential

Based on the above criteria, we selected the log-normal (close to lowest AIC for SoC and central to a pack of plausible curves in the Pem+SoC arm) as the base case and generalised gamma (lowest AIC for Pem+SoC) as a sensitivity analysis.

Figure 15: PFS FA two-piece models

[redacted]

[redacted]

3.3.1.3. Spline-based models

We followed the same standard methodology for fitting spline models to the data as in the ACM2 submission; 1, 2 and 3 knot spline models on the normal, hazard and odds scales were fitted, providing nine options for each KM curve. The locations of the knots were predetermined by the R package (flexsurvspline) at the relevant standard event quantiles as at ACM2 (4).

To assess which curves to implement in the economic model we examined:-

1. Visual fit to the KM and smooth hazard curves
2. Statistical fit assessed via AIC
3. Clinical plausibility of long-term survival vs. estimates confirmed as plausible at the UK advisory board and NICE ACM meetings

The spline models provide significantly improved visual fit to the data over one-piece models, particularly in the pembrolizumab arm (Figure 16).

Figure 16: PFS visual fits of spline and one-piece models

[redacted]

[redacted]

We also examined the visual fit to the smooth spline hazards observed in the trial. These plots did not provide a strong rationale to pick between models, with all models appearing to capture the trend in hazards in the latter part of the trial well. We note

Confidential

that the 1-knot models fit the quadratic shape of the hazard function with the apex around 37 weeks somewhat less well than the 2-knot and 3-knot models.

Figure 17: PFS visual fits of all 1, 2 and 3 knot models to the smooth spline hazards

Figure 18: TTP visual fits of all 1, 2 and 3 knot models to the smooth spline hazards

The spline models also provide superior statistical fit compared to the one-piece models in both arms, but particularly in the pembrolizumab arm. In terms of comparison between the AIC scores for the nine spline options, there is little to choose between any option for the SoC models. For the pembrolizumab+SoC arm, the 2 and 3-knot models had better fit than the 1-knot models. The hazard scale had the best fitting 1-knot model, fit was comparable among the 2-knot models and the hazard scale had a slightly worse fitting 3-knot model than the other two options. Overall, AIC did not provide a strong rationale to pick between models.

Table 10: AIC for spline models for PFS

	Pembrolizumab + SoC			**SoC**		
	knots=1	knots=2	knots=3	knots=1	knots=2	knots=3
hazard						
odds						
normal						

Table 11: AIC comparing one-piece with spline models

		Pembro		**SoC**	
	Model	**AIC**	**AIC Rank**	**AIC**	**AIC Rank**
One-piece	Exponential				
	Weibull				
	Log-normal				
	Log-logistic				
	Gompertz				
	Generalised gamma				

Confidential

Spline (hazard scale)	1-knot	[redacted]	[redacted]	[redacted]	[redacted]
	2-knot	[redacted]	[redacted]	[redacted]	[redacted]
	3-knot	[redacted]	[redacted]	[redacted]	[redacted]

We examined the 3-5 year treatment waned 5-year, 10-year, 15-year and 20-year extrapolations to make the final determination about which models to use as the base case for the spline-based analysis. Decisions on whether extrapolations were plausible or not were made referencing the two-piece curves that had previously been confirmed as plausible at the MSD UK advisory board (seven clinicians currently treating advanced cervical cancer in the NHS) and by the clinical experts at NICE ACM1 and ACM2. These original estimates were 7% OS at 20 years for the pembrolizumab+SoC arm and 1.3% OS in the SoC arm, although the clinical advisors noted that this was slightly optimistic and 0.5%-1% might be more plausible. After the committee's preferred 3-5 year treatment waning assumptions were applied, 20 year OS was 4.3% in the pembrolizumab arm in the original appraisal (37-week two-piece model).

The table below gives the landmark OS estimates generated by the economic model when all nine spline methods are implemented for PFS and TTP and when treatment waning 3-5 years is applied. It is notable that for the models on the hazard scale, all three options provide estimates close to those confirmed plausible at ACM2. The hazard scale options also provide low (1-knot), middle (3-knot) and high (2-knot) values to test in sensitivity analysis.

Table 12: Landmark OS estimates from the economic model using all spline options for PFS/TTP

Scale	Knot	Treatment	5-year survival	10-year survival	15-year survival	20-year survival
Normal	1	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	1	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Normal	2	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	2	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Normal	3	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	3	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Odds	1	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	1	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Odds	2	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]

Confidential

	2	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Odds	3	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	3	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Hazard	1	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	1	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Hazard	2	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	2	SOC	[redacted]	[redacted]	[redacted]	[redacted]
Hazard	3	Pem+SOC	[redacted]	[redacted]	[redacted]	[redacted]
	3	SOC	[redacted]	[redacted]	[redacted]	[redacted]

For any given number of knots, there were not big differences between the scales (hazard, odds and normal) in terms of statistical or visual fit but we noted the hazard models provided a reasonable range for sensitivity analysis that was close to estimates that had already been confirmed by experts. The hazard models were also those that were implemented in the model for ACM2. In the interests of parsimony, we implemented 1-knot, 2-knot and 3-knot hazard models in the economic model for both PFS and TTP and discarded the odds and normal options. We note that only one of these models (1-knot normal) produces more conservative results than are available in the hazard model suite; additionally, the difference is slight. The rest are less conservative.

As in the original submission, we selected the same model for TTP as PFS, based on the two datasets comprising very similar data but PFS having more events and being a primary trial outcome.

For the company base case, we selected the 3-knot hazard model as it produced estimates close to those previously considered and was the middle option. We examined 1-knot and 2-knot models as low/high sensitivity analysis.

3.3.2. Post Progression Survival

While the PFS and TTP curves at FA are largely just an extension of the IA1 data, there are some differences in the PPS curves between the two datasets.

It can be seen that survival time is longer in both arms than had been observed in IA1. This is likely because there is a positive correlation between TTP and PPS. In the original submission we explored this relationship to better understand the likely

Confidential

trajectory of OS (CS, Appendix Q; updated in the Appendix 1 of this report) and the FA PPS is confirmatory of that relationship (updated figures provided in the Appendix 1 of this document). The increase in PPS in both arms may be because the PPS cohort is now comprised more of people whose disease responded well to treatment before progressing (and therefore they are relatively less advanced within the PPS health state) and/or because the PPS cohort is now comprised more of patients whose disease is naturally slower progressing in general.

In addition to the PPS time in both arms lengthening, the separation between the curves is now more obvious and sustained than it was at IA1, although the HR is still not statistically significant (p=0.14). This is confirmatory of the NICE committee's preferences for separate curves in each arm.

Our interpretation of these data are the same as the clinical experts and NICE committee's; that longer PPS time observed in the pembrolizumab arm is a small but real effect attributable to greater magnitude of initial response to treatment prior to progression. Progression in KEYNOTE-826 was assessed from the greatest extent of response rather than from baseline so this makes sense.

As with the original submission, we fit parametric curves to the data for use in the economic model. Our criteria for model selection were as follows:-

1. Prioritization of one-piece model unless hazard function complex
2. Visual and statistical fit
3. Prioritise using the same model type between the arms unless there is a strong rationale not to, consistent with EAG advice detailed in the FAD
4. Mean life years longer in the pembrolizumab arm than SoC arm (curves don't cross), consistent with NICE committee's conclusions in the FAD

As with the original submission, the visual fit of the standard one-piece parametric models was good and we did not need to explore more flexible survival analyses. The log-normal and generalised gamma models had the lowest AIC and BIC among the options for the pembrolizumab combination arm and log-logistic and

Confidential

generalised gamma models had the lowest AIC and BIC among the options in the SoC arm.

We compared the projections to those that had been considered in the previous NICE ACMs and found that, given the greater observed survival time, median and mean PPS was understandably slightly longer than the 1.04 and 0.9 mean PPS life years produced by the previous independent generalised gamma models preferred by the committee. We selected the base case models conscious of the guidance we received from the EAG about using the same model in both arms unless there is a strong rationale not to. We compared pairs of curves and found that the log-normal and log-logistic curves crossed each other, leading to longer mean PPS life years in the SoC arm, which was contrary to the NICE committee's stated conclusions in the FAD. Where therefore excluded these models.

We selected the independently fitted generalised gamma curves as the most appropriate as they had among the lowest AIC, had the best visual fit to the data and were a pair where mean PPS was longer in the pembrolizumab arm/the curves did not cross. This is consistent with the committee's stated preferences in section 3.8 of the FAD (also independently fitted generalised gamma models with slightly longer PPS in the pembrolizumab arm).

Table 13 below summarises the rule-in and rule-out criteria above. Visual fit was assessed as amber if it was notably worse than the best fitting curve but not so poor that it could be directly ruled out.

Table 13: Rule-in and rule-out criteria applied to one-piece parametric curves

	Exponential	Weibull	Lognormal	Log-logistic	Gompertz	Generalised gamma
AIC/BIC for pembro among lowest	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AIC/BIC for SoC among lowest	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Visual fit for Pem+SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Confidential

Visual fit for SoC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mean PPS longer for Pem+SoC for pair	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 14: Statistical fit of parametric survival models fit to the PPS KM data for Pem+SoC and SoC in the CPS≥1 population of KEYNOTE-826 Final analysis

	Pem+SoC			SoC		
Model	**AIC**	**BIC**	**Average**	**AIC**	**BIC**	**Average**
Exponential	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Weibull	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Gompertz	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 19: PPS FA independent generalised gamma models

[redacted]

Figure 20: FA PPS Pem+SoC arm KM and one-piece extrapolations

[redacted]

Figure 21: FA PPS KM SoC arm and one-piece extrapolations

[redacted]

Consistent with the NICE committee's preferences, we have not updated/included the pooled PPS data.

3.4. Time on Treatment

Time on treatment on pembrolizumab was slightly longer than observed at IA1. The reason for this is that there were some patients who were censored in the original

Confidential

analysis and fuller follow-up data has revealed that there were some patients who missed cycles within the two-year window and therefore continued treatment after this point. In line with the marketing authorization and proposed use in the NHS, the maximum number of cycles received by the vast majority of patients in the study was 35, however, there were seven patients who received more than 35 cycles. Consistent with the committee's preferences from section 3.9 of the FAD, we examined removing these costs in sensitivity analysis but the difference was negligible as it affected so few patients. Overall, the incremental costs associated with pembrolizumab have increased in the updated analysis.

ToT for bevacizumab and platinum doublet chemotherapy have also been updated in the economic model but the differences between the IA1 and FA data are limited and do not affect the economic model's results.

Figure 22: Time on Treatment KM curves - Final Analysis

[redacted]

While the model retains the original functionality for estimating pembrolizumab treatment costs (complete KM curve adjusted by Relative Dose Intensity) we have updated the base case method to count the exact treatment cost. This is done by simply referencing a table of the number of patients receiving a dose of pembrolizumab at each patient-specific 3-weekly treatment cycle. We were able to be this precise as ToT is not linked to any other parameter in the model and the data were collected/reported like this in KEYNOTE-826. The result is very similar to the original method but avoids potential overestimation caused by censoring at the tail of the KM curve.

Figure 23: Exact number of patients who received pembrolizumab treatment at each patient-specific 3-weekly treatment cycle

[redacted]

3.5. Quality of life and utilities

EQ-5D data collected from the KEYNOTE-826 trial were analyses were conducted based on the final data cut. The total final analysis population with a CPS≥1 consisted of [redacted] patients, resulting in a combined total of [redacted] EQ-5D measurements. The population comprised of patients who were randomised (n = 548), received a study

Confidential

treatment (n = [redacted]), and completed at least one EQ-5D-5L questionnaire (n = [redacted]). The EQ-5D-5L data from KEYNOTE-826 were mapped onto the 3L scale using the algorithm developed by Hernandez-Alava et al. (2022) (5).

We provide three analyses; progression-based analyses calculated via both linear mixed effects and via naïve means and time-to-death utilities. Consistent with the NICE committee's preferences, we chose the same progression-based linear mixed effects model that had been used in the original appraisal as the base case (Table 15) and used the other two methods as sensitivity analyses (Table 16 and Table 17). We consider all to be potentially appropriate. We saw that PF utility increased over time in the supportive data analysis that we submitted during the original company submission and were unsure whether this was due to selection bias or the real effects of longer term PFS on HRQoL. The naïve means method, which weights by observation, better captures this than the linear mixed effects model, which effectively down-weights repeated observations in patients who have long survival. Time-to-death models also offer a potentially more nuanced way of capturing HRQoL than the progression-based approaches and are included for completeness. The methodology is described in a separate report. The difference between the three methods is not large and has minimal impact on the ICER.

Table 15: Base case – progression-based utilities for patients with CPS≥1 (KEYNOTE-826 final analysis; applied in the economic model)

Health state	Health state utility value	Lower bound	Upper bound
PF, no AEs	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]
Grade 3+ AE (disutility)	[redacted]	[redacted]	[redacted]

Table 16: Progression-based utilities (naïve means method) for patients with CPS≥1 (KEYNOTE-826 final analysis)

Health state	**Health state utility value**	**Lower bound**	**Upper bound**
PF	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]

Confidential

Table 17: Time-to-Death Utilities for patients with CPS≥1 (KEYNOTE-826 final analysis)

Health state	Health state utility value	Lower bound	Upper bound
Time to death >=360 days	[redacted]	[redacted]	[redacted]
Time to death 180-360 days	[redacted]	[redacted]	[redacted]
Time to death 90-180 days	[redacted]	[redacted]	[redacted]
Time to death 30-90 days	[redacted]	[redacted]	[redacted]
Time to death 0-30 days	[redacted]	[redacted]	[redacted]
Grade 3+ AE (disutility)	[redacted]	[redacted]	[redacted]

Methods for calculating and applying AE utility decrements remain the same as in Document B, Section B.3.4.5. Based on the updated analysis of KEYNOTE-826, the AE utility decrement was -[redacted] (Table 15). AE utility decrements are still not an important driver of the results and have negligible impact on the ICER.

3.6. Cost and resource use

There has been no change to the cost categories, dosing schedules, unit costs or data sources (see to Document B, Section 3.5). The only inputs that were updated that specifically affect cost calculations are those related to the observed dosing data for pembrolizumab (see Section 3.4 section above for discussion of the new base case method), Relative Dose Intensity for treatment components and the duration of subsequent treatments based on the FA of KEYNOTE-826, as described below.

3.6.1. Relative Dose Intensity

As in the original CS, Relative Dose Intensity is used in combination with the ToT KM curves in the model to calculate treatment costs. These data have changed very slightly since the original appraisal and have been updated accordingly with an extremely minimal impact on the model's results.

Table 18: Relative Dose Intensity for Study Treatments

PEM+SoC			SoC		
Mean	**SD**	**N**	**Mean**	**SD**	**N**

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

3.6.2. Subsequent therapy

The proportion of patients receiving subsequent treatments available in the NHS remain the same as in the original appraisal. These were provided by the UK advisory board. Mean duration of treatment has been updated to reflect the FA from KEYNOTE-826. Although the observed data were somewhat higher than these figures, this is likely to be due to the international nature of the trial. Overall, we considered these estimates still the most appropriate to reflect UK clinical practice. We note that the therapy options are the same in both arms in the UK and are inexpensive.

Table 19: Subsequent treatments modelled

Subsequent treatment (KEYNOTE-826)	Pem + SoC		SoC	
	Proportion of patients	Mean treatment duration (days)	Proportion of patients	Mean treatment duration (days)
Paclitaxel	[redacted]	[redacted]	[redacted]	[redacted]
Doxorubicin	[redacted]	[redacted]	[redacted]	[redacted]
Fluorouracil	[redacted]	[redacted]	[redacted]	[redacted]
Cisplatin + Gemcitabine	[redacted]	[redacted]	[redacted]	[redacted]

3.7. Results

All results presented in this section include the commercial access agreement (CAA) currently in place for pembrolizumab; all other treatments are included at list prices.

The model's key base case assumptions are the following:-

- PFS and TTP using 3-knot splines
- Treatment waning from 3-5 years, consistent with the NICE committee's preferences

Confidential

- PPS using individual generalised gamma curves, consistent with the committee's preferences
- Progression-based utilities using the linear mixed effects model, consistent with the NICE committee's preferences
- Costs using exact dosing

3.7.1. Deterministic base case

The cost-effectiveness results for Pem+SoC versus SoC are presented in

Table 20. The results show that Pem+SoC is estimated to offer a substantial incremental health benefit compared with SoC, with an additional 2.68 mean life years (LYs) and [redacted] quality-adjusted life years (QALYs) per patient lifetime. This level of benefit supports the importance of Pem+SoC as a treatment for patients with persistent, recurrent or metastatic cervical cancer who would otherwise face a poor prognosis under highly limited treatment options. The ICERs are primarily driven by the cost of pembrolizumab and the improvement in (mainly) PFS and PPS, particularly among the subgroup of patients who responded well to treatment.

These ICERs should be considered in the context of Pem+SoC being an innovative, end-of-life technology that presents a step-wise improvement for patients with persistent, recurrent or metastatic cervical cancer.

Table 20: Base case ICER

Treatment	Totals per treatment arm			Incremental results			ICER (£/QALY)
	LYs	QALYs	Costs	LYs	QALYs	Costs	
SoC	2.65	[redacted]	[redacted]	2.68	[redacted]	[redacted]	[redacted]
Pem+SoC	5.33	[redacted]	[redacted]				
Key: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life year.							

3.7.2. Sensitivity analyses

3.7.2.1. Probabilistic sensitivity analysis (PSA)

Parameters were varied in PSA using appropriate distributions as outlined in the CS.

Table 21, Figure 24 and Figure 25 show the PSA results: the mean average outcomes

of the 5,000 probabilistic iterations result in an ICER of [redacted] per QALY, which is very similar to the base case analysis. The PSA indicated that pembrolizumab had an ICER of less than £50,000/QALY gained in 94.5% of iterations.

Table 21: Mean PSA results, Pem+SoC versus SoC

	Totals per treatment arm			Incremental results			ICER (£/QALY)
Treatment	**LYs**	**QALYs**	**costs**	**LYs**	**QALYs**	**Costs**	
SoC	2.51	[redacted]	[redacted]	2.60	[redacted]	[redacted]	[redacted]
Pem+SoC	5.11	[redacted]	[redacted]				
Key: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life year.							

Figure 24: PSA scatterplot, Pem+SoC versus SoC

[redacted]

Figure 25: Cost-effectiveness acceptability curve, Pem+SoC versus SoC

[redacted]

3.7.2.2. Deterministic sensitivity analysis

As established during the original appraisal, plausible variations in single parameters had a minimal influence on the cost-effectiveness results.

Table 22: One-way Sensitivity Analysis - most influential single parameters

[redacted]

3.7.3. Scenario analyses

The original Company Submission included a large number of scenario analyses that did not materially affect the ICER. Although the functionality remains in the model, we have omitted many of those in favour of focusing on areas of residual uncertainty that were identified in the FAD.

Table 23: Scenario analyses

Scenario label	Incremental costs	Incremental QALYs	Incremental LYs	ICER
MSD Base Case	[redacted]	[redacted]	2.68	[redacted]

PFS/TTP 1-knot spline	[redacted]	[redacted]	2.34	[redacted]
PFS/TTP 2-knot spline	[redacted]	[redacted]	3.00	[redacted]
PFS/TTP Two-piece gen-gamma	[redacted]	[redacted]	2.73	[redacted]
PFS/TTP Two-piece log-logistic	[redacted]	[redacted]	3.16	[redacted]
PFS/TTP Two-piece log-normal (46-weeks)	[redacted]	[redacted]	3.29	[redacted]
PPS Exponential Curves	[redacted]	[redacted]	2.43	[redacted]
Pembrolizumab costs KM+RDI method	[redacted]	[redacted]	2.68	[redacted]
Pembro costs beyond 35 cycles excluded	[redacted]	[redacted]	2.68	[redacted]
Treatment waning 5-7 years	[redacted]	[redacted]	3.26	[redacted]
Pooled PPS curve	[redacted]	[redacted]	2.26	[redacted]
Naïve mean utilities	[redacted]	[redacted]	2.68	[redacted]
Time to Death utilities	[redacted]	[redacted]	2.68	[redacted]

All scenario analyses are comfortably beneath the threshold that NICE typically consider cost-effective for life when the End of Life criteria are applied, indicating that there is minimal uncertainty that pembrolizumab is a cost-effective addition to standard care.

4. Discussion and Conclusion

The updated data from the clinical trial have strengthened clinical certainty about the effectiveness of pembrolizumab as an addition to standard care for this population of women with advanced cervical cancer.

The updated survival analyses strengthen the argument that flexible survival models for PFS and TTP are needed. The greater extent of follow-up has also narrowed the range of survival models that are suitable for use in the economic model. The longer PPS follow-up has also provided stronger evidence for differential PPS on

Confidential

pembrolizumab. The increased time on treatment on pembrolizumab has slightly increased incremental costs.

The updated economic model uses all the committee's preferred assumptions from the FAD and produces a base case ICER substantially below £50,000/QALY gained (ref section 3.13 of the FAD). It is not sensitive to plausible changes to its inputs and was cost-effective in ~95% of PSA iterations. In all reasonable scenario analyses examined by the company, the ICER for pembrolizumab combination was below that considered a cost-effective use of NHS resources when the End of Life criteria are applied. In section 3.13 of the FAD, the committee indicated that an ICER substantially below the threshold would result in a recommendation for baseline commissioning for pembrolizumab combination. In addition, it should be noted that the treatment waning assumptions used in the model are highly conservative and no evidence has yet emerged from any immunotherapy trial that supports when this effect might take place; a two-year relaxation of the year that they begin, an equally credible scenario, reduces the ICER by a substantial ~£5,000. Furthermore, no adjustment has been undertaken by the company to account for subsequent immunotherapy use in the control arm of the trial. It is possible that the true ICER lies below the one produced by the model using the current assumptions.

Overall, we conclude that pembrolizumab is a cost-effective addition to standard care for this underserved population.

Confidential

References

1. NICE. Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer [TA885] 2023 [Available from: https://www.nice.org.uk/guidance/TA885.
2. NICE. Final appraisal document: Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer 2023 [Available from: https://www.nice.org.uk/guidance/ta885/documents/final-appraisal-determination-document.
3. Colombo N, Dubot C, Lorusso D, Caceres MV, Hasegawa K, Shapira-Frommer R, et al. Pembrolizumab for Persistent, Recurrent, or Metastatic Cervical Cancer. N Engl J Med. 2021;385(20):1856-67.
4. Jackson C, Metcalfe P, Amdahl J, Warkentin MT, Sweeting M, Kunzmann K. Flexible Parametric Survival and Multi-State Models 2023 [Available from: https://cran.r-project.org/web/packages/flexsurv/flexsurv.pdf.
5. Hernandez-Alava M, Pudney S. Mapping between EQ-5D-3L and EQ-5D-5L: A survey experiment on the validity of multi-instrument data. Health Econ. 2022;31(6):923-39.
6. National Institute for Health and Care Excellence. Position statement on use of the EQ-5D-5L value set for England (updated October 2019) 2019 [Available from: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l.
7. van Hout B, Janssen MF, Feng YS, 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. 2012;15(5):708-15.

Confidential

Appendices:

Appendix 1: Association between time to progression and post-progression survival

Figure 1: Time to progression (TTP) and time from progression to death (PD2D) in KEYNOTE-826 by arm; PEM+SoC (top) and SoC (bottom).

Update on key data from Appendix Q, CS. These data are supportive of the hypothesis that, based on the evidence so far (patients who have had both a PFS->PD event and a PD->Death event, there is no negative relationship between TTP and PPS).

Key: evt.pd2d, binary indicator of event observed for pd2d (0 = no death event observed; 1 = death event observed); pd2d, (time from) progressed disease to death; PEM, pembrolizumab; SoC, standard of care; ttp, time to progression.

Confidential

Appendix 2: Health Utility Analysis Report

Introduction

Objectives

The purpose of this supplement report is to describe statistical models of the EuroQol EQ-5D-5L health utility, experienced by the CPS>=1 population in a Phase III study, KEYNOTE-826, according to UK value sets and methods accepted by NICE. The co-primary endpoints of the trial were progression-free survival (PFS), as assessed by investigator, and overall survival (OS). The results of this analysis will be used to inform inputs for the cost-effectiveness model.

(6, 7)Trial design

KEYNOTE-826, a double-blind trial, was conducted at 151 sites in 19 countries. Patients were randomly assigned in a 1:1 ratio to receive pembrolizumab (200 mg) or placebo every 3 weeks for up to 35 cycles. All the patients were to receive paclitaxel (175 mg per square meter of body-surface area) and the investigator's choice of cisplatin (50 mg per square meter) or carboplatin (area under the concentration–time curve, 5 mg per milliliter per minute) every 3 weeks. Patients could receive bevacizumab at a dose of 15 mg per kilogram of body weight every 3 weeks according to local practice at the investigator's discretion.

Trial site staff collected patient reported outcomes (PROs) from patients using an electronic tablet device at the beginning of each clinic visit and reported reasons for non-completion. Sites were contacted by the study sponsor in cases of missing PRO data.

In this study, the EuroQol EQ-5D-5L system was used to measure generic health status. It contains five health state dimensions: mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. Each dimension is rates on a 5-point scale from 1 (no problem) to 5 (extreme problem).

The analyses of the EQ-5D-5L utilities were based on the Final Analysis data-cut of 3 October 2022, using the CPS>=1 population. Subjects were analyzed in the treatment group allocated at randomization. The EQ-5D-5L population included a total of 545 subjects.

The EQ-5D-5L data collected in the trial can be converted to population-based utility valuations using published algorithms. For the analysis in this report, we followed the recommendation from NICE and estimated health state utilities based on mapping

EQ-5D-5L data collected in KEYNOTE-826 to EQ-5D-3L value set using the mapping function from Hernandez Alava 2022 [1,2].

Methods

All analyses was conducted in R version 4.2.2.

Descriptive analysis

The mean health utility is derived and presented according to whether disease has progressed.

Mixed effects regression by progression status

We explore a series of mixed effects regressions with the objective of identifying the most parsimonious statistical model of health utility as it may be explained in this patient population by progressive disease status (progression-free vs. progressive disease), treatment assignment (pembrolizumab vs. placebo) and experience of adverse events. Since one patient could have multiple utility measures within the same health state, mixed linear effects models with random intercept were used for this analysis to account for within-subject correlation.

The variables and models are described in Table 124 and Table 225.

Table 124. Variable description

Variable (variable name)	Variable description
Utility	EQ-5D-3L utility based on mapping EQ-5D-5L data to EQ-5D-3L value set using the mapping function from Hernandez Alava (2022)
Progression Status (PFINVFL)	Progression-free (No_PD) vs. progressive disease (W_PD), according to RECIST, version 1.1, as based on investigator's assessment. An "Unknown" category was created for records measured with unknown progression status.
AE (G35AEFL (w/o Grade3+ AE))	Indicator for an EQ-5D-5L score measured during grade 3+ AEs; with G35AEFL set to 0 for 'Without (w/o) Grade 3+ AE' and set to 1 when 'During Grade 3+ AE'. In regression model with G35AEFL, the intercept is associated with "w/o Grade3+ AE".
TRT01P	Indicator for treatment group

Table 225. Models description

Model	Specification*
M1	$Utility_{ij} = \beta_{0i} + \beta_1 AE_{ij} + \beta_2 PFSINVFL_{ij} + e_{ij}$
M2	$Utility_{ij} = \beta_{0i} + \beta_1 AE_{ij} + \beta_2 PFSINVFL_{ij} + \beta_3 TRT01P_{ij} + e_{ij}$

Confidential

M3	$Utility_{ij} = \beta_{0i} + \beta_1 AE_{ij} + \beta_2 PFSINVFL_{ij} + \beta_3 TRT01P_{ij} + \beta_4 (PFINVFL_{ij} * TRT01P_{ij}) + e_{ij}$
*i denotes individual and j denotes time when the EQ-5D-5L measures was taken. β_{0i} is the intercept term, representing the baseline utility level β_1, β_2, β_3, and β_4 are the regression coefficients corresponding to the AE, progression status, treatment, and their interaction term, respectively, representing the effects of these variables on utility outcomes. e_{ij} is the residual term	

The best fitting model was identified by considering AIC and BIC model fit statistics.

Missing data

There are 32 records with unknown 'PFINVFL' [progression status], thus we treat them as missing value for "progression status". Analyses were performed on a complete case basis.

Results

Descriptive analysis results

Table 326 presents descriptive analysis or naïve values based on progression status (i.e., with and without disease progression), based on the PFINVFL variable.

Table 326. Descriptive means and standard deviations (SD) for utility values by progression status

Health state	No of records	Mean utility	Standard deviation (SD)
Progression free	[redacted]	[redacted]	[redacted]
Progressed disease	[redacted]	[redacted]	[redacted]

Mixed effects regression by progression status and time to death

We present multiple regression models for health utility by progression status and time to death category. The following analyses were conducted with 7228 EQ-5D records from 545 subjects in the CPS>=1 population. The results were reported based on the UK algorithm with EQ-5D-5L mapped to EQ-5D-3L using Hernandez Alava 2022 method [2]. The numbers used here to identify each model do not necessarily correspond to those used in any separate report.

Summary fit statistics for the mixed effect regressions by health state are provided in Table 427. Model 1 had the lowest AIC and may be considered the more conservative approach of the analyses under review.

Table 427: Statistical model fits by health state

Model number	Terms	Number of fixed effect parameters	AIC	BIC
M1	Progression status + Experiencing grade 3+AEs	[redacted]	[redacted]	[redacted]
M2	Progression status + Experiencing grade 3+AEs + Treatment indicator	[redacted]	[redacted]	[redacted]
M3	Progression status + Experiencing grade 3+AEs + Treatment group indicator + Interaction term	[redacted]	[redacted]	[redacted]

Model 1. Utility by Progression Status, accounting for Grade 3+ AE

Parameter values rounded to three decimal places for the Model 1, are given in **Table 528** along with the variance-covariance matrix in **Table 629**.

Table 528. Parameters for the fixed effect

Parameter	Description	Coefficients	p-value
(Intercept)		[redacted]	[redacted]
G35AEFLDuring Grade3 + AEs	With Grade 3+ AEs	[redacted]	[redacted]
PFINVFLW_PD	Progression status-progressed disease	[redacted]	[redacted]
Key: AE, adverse event.			

Table 629. Variance-covariance matrix

	(Intercept)	G35AEFLDuring Grade3 + AEs	PFINVFLW_PD
(Intercept)	[redacted]	[redacted]	[redacted]
G35AEFLDuring Grade3 + AEs	[redacted]	[redacted]	[redacted]
PFINVFLW_PD	[redacted]	[redacted]	[redacted]

Model 2. Utility by Progression Status, accounting for Grade 3+ AE and Treatment indicator

Confidential

Parameter values rounded to three decimal places for the Model 2, are given in **Table 730**. The additional covariate for treatment assignment was not statistically significant (Wald P=0.556).

Table 730. Parameters for the fixed effect

Parameter	Description	Coefficients	p-value
(Intercept)		[redacted]	[redacted]
G35AEFLDuring Grade3 + AEs	With Grade 3+ AEs	[redacted]	[redacted]
PFINVFLW_PD	Progression status-progressed disease	[redacted]	[redacted]
TRT01P	Treatment group indicator: Pembrolizumab + Chemotherapy	[redacted]	[redacted]
Key: AE, adverse event.			

Model 3. Utility by Progression Status, accounting for Grade 3+ AE and Treatment indicator and interaction term

Parameter values rounded to three decimal places for the Model 3, are given in Table 8. The additional covariates for treatment assignment and interaction were not statistically significant.

Table 831. Parameters for the fixed effect

Parameter	Description	Coefficients	p-value
(Intercept)		[redacted]	[redacted]
G35AEFLDuring Grade3 + AEs	With Grade 3+ AEs	[redacted]	[redacted]
PFINVFLW_PD	Progression status-progressed disease	[redacted]	[redacted]
TRT01P	Treatment group indicator: Pembrolizumab + Chemotherapy	[redacted]	[redacted]
PFINVFL:TRT01P	Interaction between progression status and treatment group	[redacted]	[redacted]
Key: AE, adverse event.			

Recommendation

The recommended model is **M1** based on model comparison from AIC/ BIC statiticis. This is a mixed effects regression by progression status with grade 3+AEs. This is a

Confidential

conservative model: it considers the grade 3+ AEs (significant factor) and does not consider the impact of treatment indicator (as it is found to be not significant in models M2 and M3).

Discussion

In summary, the recommended base case for utility is mixed effects regression by progression status with grade 3+AEs (model 1). The fixed effects table for Model 1 indicated that progression status and experiencing grade 3+ AEs had significant effects on health utility. Patients with progressive disease had a lower utility compared to those with progression-free disease, with an estimated coefficient of -[redacted]. Experiencing grade 3+ AEs also negatively impacted health utility, with an estimated coefficient of [redacted]

There is a limitation of the analysis based on trial-collected utility data, which is collected for up to one year or end of treatment, whichever comes first, as well as at time of discontinuation and at the 30-day post-treatment discontinuation follow-up visit. No further utility data were collected in the trial. Therefore, the utility data for progressive disease state is limited and it is usually collected right after progression.

Additional analyses were also considered using time to death approach (data on file). Time to death is calculated as the time between the health utility observation and the time of death and recorded in categories: <30, 30-89, 90-179, 180-359, ≥360 days until death. The ≥360 days category includes observations made ≥360 days from a censoring event for overall survival; otherwise, if overall survival is censored, observations are recorded as 'unknown' time to death (Hatswell 2014) [3]. Time-to-death categories were further classified as <180 and >=180 days until death to increase the stability of estimation. One important limitation of the time-to-death utility approach is that the records measured within 360 days from OS censoring date cannot be assigned to a time-to-death category due to the uncertain date of death.

Also, it has been suggested combining time to death and health state-based formulations, could potentially lead to overfitting (Hatswell 2021) [4] and hence, not performed. Hence, the recommended model utility is mixed effects regression by progression status is a conservative selection for base case analysis.

Although not preferred by the committee, we updated the Time to Death analysis using the same method outlined in the CS and implemented the results in the economic model for use in scenario analysis.

Confidential

References

1. 2022 NICE methods guide. NICE health technology evaluations: the manual. Process and methods [PMG36] Published: 31 January 2022
2. Mónica Hernández Alava, Steve Pudney, Allan Wailoo. Estimating the relationship between EQ-5D-5L and EQ-5D-3L: results from an English Population Study. (2020) Pharmacoeconomics; November 30 2022, 9p.
3. Hatswell AJ, Pennington B, Pericleous L, Rowen D, Lebmeier M, Lee D. Patient-reported utilities in advanced or metastatic melanoma, including analysis of utilities by time to death. Health and quality of life outcomes. 2014 Dec;12(1):1-9.
4. Hatswell AJ, Bullement A, Schlichting M, Bharmal M. What is the Impact of the Analysis Method Used for Health State Utility Values on QALYs in Oncology? A Simulation Study Comparing Progression-Based and Time-to-Death Approaches. Applied health economics and health policy. 2021 May;19:389-401.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Rapid review

Pembrolizumab in combination with platinum-based chemotherapy for treating recurrent, persistent or metastatic cervical cancer [TA885]

Clarification questions

July 2023

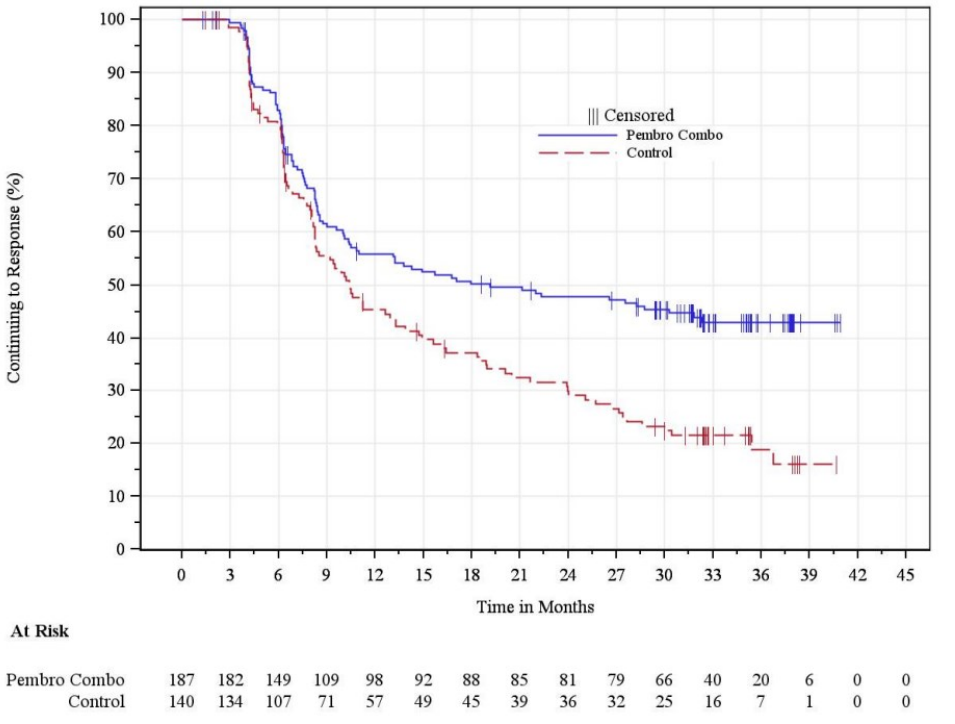
File name	Version	Contains confidential information	Date
ID6279 TA885 pembro clarification questions 20230811 Redacted.docx	V1	Yes	11th August 2023

Section A: Clarification on effectiveness data

A1. Please provide the Final analysis (FA) data from the KEYNOTE 826 trial for the remaining outcomes that were reported in the original company submission: Duration of response; health related quality of life outcomes; subgroup analyses for progression free survival (PFS) and overall survival (OS)

Duration of Response

Figure 1: Duration of response based on investigator assessment per RECIST 1.1for confirmed response (CPS ≥ 1 population)



The median duration of response (DoR) at the final analysis was 19.2 months and 10.4 months in the pembrolizumab combination and placebo combination arm, respectively. The median DoR for the pembrolizumab arm increased by 1.2 months compared with the first interim analysis. There was no change between the analyses for the placebo combination arm. The percentage of subjects with a response lasting equal to or greater than 24 months was 48% and 30% in the pembrolizumab combination and placebo combination arms, respectively. These results are consistent with previous analysis provided.

Health related Quality of Life Outcomes

Final analysis data is consistent with the IA1 (Document B section B.2.6.2). The between group difference in least-squares mean change from baseline at Week 30 stayed the same – 1.69 (95% CI: -1.80, 5.18; p = 0.3414) (Table 1).

Compared with patients in the placebo group, more patients treated with pembrolizumab reported improved and stable patient reported outcome scores (1, 2), demonstrating the treatment with pembrolizumab did not adversely impact the HRQL compared with current treatments available for patients with recurrent, persistent or metastatic cervical cancer.

Table 1: Analysis of Change from Baseline in EQ-5D-5L VAS Score to Week 30 (CPS ≥1 Full Analysis Set)

	Baseline		Week 30		Change from Baseline at Week 30	
Treatment	N	Mean (SD)	N	Mean (SD)	N	LS Mean (95% CI) [a]
Pembrolizumab combination	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Control	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pairwise comparison					Difference in LS Means 95% CI)	p-Value
Pembrolizumab + chemotherapy vs. Placebo + chemotherapy					[redacted]	[redacted]

[a] Based on a cLDA model with the PRO scores as the response variable with covariates for treatment by study visit interaction, stratification factors metastatic at diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS <1, CPS 1 to <10, CPS ≥10).
For baseline and Week 30, N is the number of participants in each treatment group with non-missing assessments at the specific time point; for change from baseline, N is the number of participants in the analysis population in each treatment group.
P-value is based on two-sided t test.
Database Cutoff Date: 03OCT2022

Mean change from baseline and 95% CI for EQ-5D-5L VAS for the pembrolizumab and placebo groups are presented in Figure 2. Compared to IA1 a slightly larger improvement in mean change was seen in the pembrolizumab group compared with patients treated with placebo over time. While we have not captured this formally in the economic model it is plausible that it is an additional real effect in favour of pembrolizumab; the alive cohort is comprised of an ever greater percentage of Complete Responders over time within the trial period.

Figure 2: Empirical Mean Change from Baseline and 95% CI for the EQ-5D-5L VAS Score Over Time by Treatment Group (CPS ≥1 Full Analysis Set)

Subgroup analyses for PFS

The forest plots for PFS (based on investigator assessment per RECIST 1.1) by subgroup factors for the CPS ≥ 1 population are presented in Figure 3. The FA data is consistent with the IA1 data cut. The benefit of pembrolizumab was demonstrated across all patients when compared with placebo. The HRs in the pembrolizumab group were less than 1 in all pre-specified subgroups analysed within the CPS ≥ 1 population, and the 95% CIs for all subgroups overlapped with that of the overall population. The clinical benefit of pembrolizumab compared with placebo was also generally consistent across various pre-specified subgroups.

Figure 3: Forest Plot of PFS survival hazard ratio by subgroup factors based on investigator assessment per RECIST 1.1 (CPS ≥ 1 population)

Subgroup analyses for OS

The forest plots for OS (based on investigator assessment per RECIST 1.1) by subgroup factors for the CPS ≥ 1 population are presented in Figure 4. Similarly, to the PFS, the OS FA subgroup data is consistent with the IA1.

Figure 4: Forest Plot of OS survival hazard ratio by subgroup factors (CPS ≥ 1 population)

A2. Priority Question: In addition to the information provided in Table 19 of the rapid review submission, please provide FA data on all subsequent treatments observed in the KEYNOTE 826 trial (including immunotherapies) categorized by treatment arm, along with the corresponding duration of these treatments (the EAG recognises that these data may not reflect UK practice).

Table 2 shows, as at the time of the data cut off, the majority of patients in both arms did not receive any subsequent treatment, 70.7% and 60.8% for pembrolizumab combination and the control arm, respectively. The treatments listed are not all reflective of the UK practice although there is no defined standard of care in the NHS. This was validated by two clinicians MSD consulted since receiving the Clarification Questions.

The control arm had a small number of patients who received the following immunotherapy treatments pembrolizumab (n=12), atezolizumab (n=2), dostarlimab (n=1) and AK 104 (n=1, phase II study) in the second line (24 patients had immunotherapy at all later lines). The most common second line treatments received by patients in the control arm, either alone or in combination, were carboplatin (n=32), paclitaxel (n=25), gemcitabine (n=23), cisplatin (n=14) and bevacizumab (n=13). The range of therapies used indicates there is variation in the subsequent treatments.

Table 2: Duration and Utilization of Subsequent Systemic Oncologic Therapies in days - Participants Who Completed or Discontinued from Study Treatment - Participants with CPS ≥1 - (APaT)

	Pembrolizumab Combo (N=256)		Control (N=263)	
	n (%)	Mean Duration (SE)	n (%)	Mean Duration (SE)
Participants who did not receive subsequent oncologic therapies	[redacted]		[redacted]	
Participants who died	[redacted]		[redacted]	
Participants who are alive	[redacted]		[redacted]	
Participants who received subsequent oncologic therapies	[redacted]		[redacted]	
2nd line	[redacted]		[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
3rd line	[redacted]		[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
4th line	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
5th line	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
6th line	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
7th line	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8th line	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Every participant is counted once for each applicable row and column.
Percentages are based on the number of participants who completed or discontinued study treatment.
If a treatment reflects a combination of regimens, the duration is derived using the first initiation date of any treatment within the regimen and the last date at which any treatment within the regimen is observed.
Database Cutoff Date: 03OCT2022

Section B: Clarification on cost-effectiveness data

B1. Priority Question: A key rationale for using a state-transition model was the comparative maturity of PFS data relative to OS. The available OS data

from KEYNOTE 826 is however now much more mature and it is less clear that a state-transition model represents the most appropriate modelling approach.

a) **Please justify the use a state-transition model over a partition survival model.**

MSD's understanding was that the Rapid Review would focus on the updated clinical data and their impact on the cost-effectiveness results with as few variations as possible versus the evidence already considered by the committee. We acknowledge that the data are now more complete, potentially making a Partitioned Survival Model (PSM) more relevant. However, the availability of the final data cut strengthens the reliability of all survival extrapolations, irrespective of the model structure. While it does not make the STM unsuitable, we acknowledge the interest in the PSM approach. We have therefore provided the data requested, although consider it a suboptimal structural sensitivity analysis.

MSD UK continues to consider the State Transition Model (STM) appropriate for characterising outcomes in advanced cervical cancer. The availability of the Final Analysis does not diminish the clinical justifications supporting the STM:-

- Clinicians at ACM1 supported the structural link between PFS and OS
- PFS was a good surrogate for OS in GOG240
- TTP and PPS are not negatively correlated
- Non-cancer progression-related mortality e.g. death relating to age or comorbidities is not an important factor determining OS in this population
- No effective subsequent treatments are available in either arm that might differentially affect PPS

Close to 90% of incremental QALYs in the company's base case are obtained within the PFS health state. Importantly, since PFS is modelled exactly the same way in both an STM and a PSM, the only difference should lie in PPS. This means adopting a plausibly parametrised PSM should not be expected to result in cost-ineffective ICERs.

Given that PFS will be modelled the same way, there are only two ways in which a PSM could increase the ICER vs the STM. The first is if the PSM calculated that PPS time was significantly shorter in the pembrolizumab arm than the SoC arm. The second is if the structural choice was made for the extrapolation of OS to cap the extrapolation of PFS, thus altering incremental PFS QALYs.

In section 3.8 of the FAD, the committee concluded that patients on pembrolizumab would have “at least a modest benefit in post-progression survival compared with placebo”. This was based on the evidence from GOG240 (3), the testimony of clinical experts at ACM1 and the evidence from the Interim Analysis of KEYNOTE-826, which has been strengthened in the Final Analysis (see section 2.7 of the Company’s Rapid Review submission). It is important that any economic model used for decision-making reflects this conclusion.

OS and PFS are fairly close at the end of the KM curves in the Final Analysis indicating that the significant majority of alive patients are PF. It can be seen from the KM curves by responder status that many are still in a state of Complete Response (see section 2.5 figure 4 of the Company’s Rapid Review submission). It is obvious from visual inspection of the KM curves that, due to OS data too immature to capture the trend that has become obvious in PFS, most logical PSM extrapolations will eventually result in an unrealistic crossing of OS and PFS curves. A decision will have to be made about whether the PFS (now also OS) health state continues along the OS curve or the PFS curve. Either situation is, of course, suboptimal, especially when the intersection will likely take place while a significant proportion of patients remain alive and therefore a large influence on the ICER is to be expected. In either case, there will be no patients within the PSM’s Progressed Disease health state after the curves cross and therefore no patients are assumed to be progressing before they die. This is clearly contrary to the natural history of advanced cervical cancer. The STM approach does not face this issue, avoiding the need to make such an explicit assumption.

The committee stated that the most appropriate modelling approach may change when more data is available. It is worth clarifying that despite the extended follow-up provided by the final analysis of KEYNOTE-826, the principal limitation of using the

partitioned survival analysis model (PSM) on this dataset (OS not being mature enough) still persists.

In a PSM there are no explicit transition probabilities, rather, health state occupancy is calculated using simple arithmetic relationships between PFS and OS extrapolations. This can cause a problem when survival extrapolations cross one another as the health state occupancy will sum to more than 1. To overcome this the modeller must either reject well-fitting models with crossing curves or choose one of either the PFS or OS extrapolation to be used for both outcomes. If this is employed OS and PFS are both modelled using the same extrapolation, PFS or OS. In either case there will be zero patients in the PD health state from the time of the crossing onwards, which is likely unrealistic.

Crossing is not normally a big problem in PSMs as it typically takes place past the point in model-time where one choice or the other meaningfully influences the outcomes of the model. In this dataset, however, the curves of the best-fitting PFS and OS curves cross when ~15% of patients (in the base case) are still PF in the Pem+SoC arm. The crossing of curves affects far fewer patients in the SoC arm and therefore the approach to curve crossing often does not meaningfully influence SoC outcomes. Having zero patients in the PD health state is highly unrealistic in advanced cervical cancer; the clinicians at ACM1 explained that it is progression of disease that causes death in this indication and that this relatively young population seldom die from other causes.

The reason for the curve crossing in the pembrolizumab arm is that not enough OS events have yet been observed to fully describe the expected trend in OS that is evident from the PFS curve, which now has a very noticeable plateau driven largely by durable Complete Responders (please see the data we provided on PFS/OS by responder status in section 2.5 of the Rapid Review submission).

The curve-crossing issue makes the PSM structure suboptimal in this particular dataset and therefore the STM should still be considered as the primary analysis. However, when the PSM is used we believe that in the event of the curves crossing both outcomes should be modelled using the PFS extrapolations. This is because the PFS dataset is far more complete and is able to describe the expected trajectory

among the remaining cohort better than the OS curve, where the trend in the last year of the trial is still being influenced by higher event rates among patients who haven't responded well to immunotherapy.

b) **Please provide parametric extrapolations of OS using KEYNOTE-826 FA data, including relevant diagnostics and fit statistics. At a minimum, this should include all standard one-piece models but may also include more flexible approaches that the company feels are appropriate e.g. two-piece models or splines. Please state clearly the company's preferred extrapolation of OS considering relevant diagnostics and fit statistics as well the clinical plausibility of projections.**

We have fitted one-piece and spline options to the OS data as requested. OS extrapolations were validated the using the clinical opinion that had been provided at the MSD UK advisory board (comprising seven UK clinicians currently treating advanced cervical cancer in the NHS) and by experts at NICE ACM1 and ACM2 to determine longer term plausibility. Furthermore, two individual 30-minute interviews with UK NHS clinical oncology consultants were conducted in August 2023. These validation interviews included OS FA KM extrapolations, expected survival on SoC and review of the subsequent treatments received by this trial's patients. Briefly, the two clinical experts were consulted to understand the appropriateness of using one-piece or spline models, and to discuss the plausibility of long-term projections. This included survival estimates for both trial arms and mean predicted life years generated from each model. They suggested that all extrapolations where mean survival was > 3 years were not appropriate, and that the two-knot or three-knot spline models may provide the most plausible longer term projections for both arms.

<u>Overall survival – One-piece models</u>

It can be seen in the graph below (Figure 5 and Figure 6) that no one-piece models are able to capture the complex hazard function for the OS for the Pem+SoC arm, providing poor visual fit. A similar although less pronounced situation is observed in the standard of care arm.

The statistical fit for one-piece models versus the more flexible spline-based approaches is compared via the AIC and BIC than one-piece models for both arms

(Table 3 and Table 4). The spline models have significantly lower AICs than the one-piece models in the pembro arm whereas there is a single one-piece model (the log-logistic) that has a comparable AIC to the best fitting spline based approaches for the SoC arm.

The evidence on visual and statistical fit suggests that one-piece curves are unsuitable for modelling the Pem+SoC arm. These extrapolations were validated with 2 consultant clinical oncologists. Both of them stated that the log-logistic one-piece model overestimated expected long-term overall survival rates on SoC and that all one-piece models are likely to underestimate long term survival in the Pem+SoC arm, given the positive outcomes observed so far and their experience of using pembrolizumab in other settings. The one-piece log-logistic model for the SoC arm predicted 20-year OS of 1.9%, which was deemed too high. This value was also higher than the 1.5% value that experts had confirmed was “optimistic” at the MSD UK advisory board. Some one-piece models produce long term estimates for SoC that are similar to the spline based approaches, albeit with worse AIC.

Figure 5: OS FA vs One-piece models for pembrolizumab arm

Figure 6: OS FA vs One-piece models for SoC arm

Table 3: AIC comparing one-piece

		Pembro		SoC	
	Model	**AIC**	**AIC Rank**	**AIC**	**AIC Rank**
One-piece	Exponential	[redacted]	[redacted]	[redacted]	[redacted]
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
	Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]

Table 4: BIC comparing one-piece

		Pembro		SoC	
	Model	**BIC**	**BIC Rank**	**BIC**	**BIC Rank**
One-piece	Exponential	[redacted]	[redacted]	[redacted]	[redacted]
	Weibull	[redacted]	[redacted]	[redacted]	[redacted]
	Log-normal	[redacted]	[redacted]	[redacted]	[redacted]
	Log-logistic	[redacted]	[redacted]	[redacted]	[redacted]
	Gompertz	[redacted]	[redacted]	[redacted]	[redacted]
	Generalised gamma	[redacted]	[redacted]	[redacted]	[redacted]

Overall survival – Splines

We followed the same standard methodology for fitting spline models to the data as in the ACM2 submission; 1, 2 and 3 knot spline models on the normal, hazard and odds scales were fitted, providing nine options for each KM curve. The locations of the knots were predetermined by the R package (flexsurvspline) at the relevant standard event quantiles as at ACM2 (4).

To assess which curves to implement in the economic model we examined:-

1. Visual fit to the KM and smooth hazard curves

2. Statistical fit assessed via AIC

3. Clinical plausibility of long-term survival vs. estimates confirmed as plausible at the UK advisory board, NICE ACM meetings and recent discussions with clinicians

The spline models provide significantly improved visual fit to the data over one-piece models, particularly in the pembrolizumab arm (Figure 7 and Figure 8).

Figure 7: OS FA vs Spline models for pembrolizumab arm*

*Spline 2 knots and 3 knots extrapolations are overlapping

Figure 8: OS FA vs Splines models for SoC arm

Table 5: AIC for spline models for OS

	Pembrolizumab + SoC			**SoC**		
	knots=1	knots=2	knots=3	knots=1	knots=2	knots=3
hazard	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
odds	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
normal	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 9: OS visual fits of all 1, 2 and 3 knot models to the smooth spline hazards

OS: Pembrolizumab + chemotherapy

OS: Placebo + chemotherapy

[redacted]

There were not meaningful differences between the odds, hazard and normal scales in terms of AIC or predicted survival. We therefore restricted analyses to the hazard scale for simplicity.

The company's preferred projections for OS

Our preferred projection of OS for both arms is the 3-knot hazard spline (while acknowledging there is not much meaningful differentiation between any of the 2-knot and 3-knot models). These models have among the lowest AIC, the best visual fit to the KM curve, the best visual fit to the smooth-spline hazards over time and produce OS estimates within the range that has already been confirmed as plausible by clinical experts at ACM1 and by separate recent discussion with two UK clinicians.

For the SoC arm, the only one-piece alternative based on similar AIC would be the one-piece log-logistic curve. However, this curve has a very long tail and produces OS of 1.9% at 20 years and consequently mean life years of 3.00, which are above the upper limit of plausibility confirmed by clinicians at the MSD advisory board and those consulted since (see section c) for more details).

The 2-knot spline is a reasonable alternative to the 3-knot spline but has a slightly longer tail, producing 2.76 mean LYGs vs. 2.66, which was the same as the LYGs produced by the base case STM.

For the Pem+SoC arm, all one-piece models have much higher AICs and poorer visual fits than the 2-knot and 3-knot spline alternatives.

The proportion of people alive at various landmark timepoints is influenced by whether PFS patients should continue along the PFS or OS curve. Even in the more optimistic scenario that PFS/OS continues along the PFS curve, the model still produces slightly more conservative estimates than the STM estimates that have already been confirmed as plausible by clinical experts at ACM1 and at the MSD UK

advisory board. 20-year OS is 3% in the Pem+SoC arm under these assumptions (PSM with 3-knot splines for both PFS and OS and following PFS after the curves cross). 20-year OS was 4.3% in the original appraisal and, despite the change of modelling approach from piecewise to splines in the Rapid Review resubmission, OS was also 4.3% in the base case analysis we submitted. If it has been confirmed that 20-year OS is not zero in the SoC arm then 3% (or indeed 4.3%) does not seem unreasonable, given the proportion of durable complete responders in the Pem+SoC arm of the trial. In the scenario where PFS follows the OS curve after the crossover, 20-year OS is just 1.5%, which would be quite pessimistic given it is not zero (0.75% in the base case) in the SoC arm and bevacizumab is available in both arms.

Beyond absolute survival estimates, there is a further clinical justification for wanting to use a more flexible modelling approach for OS. It is clear from the KM data by responder status (section 2.5 of the Rapid Review Submission) and the extreme turn in the shape of the hazard function in the PFS curve that this cohort (particularly the pembrolizumab arm) is becoming increasingly comprised of patients who have responded very well to treatment. This trend is not yet as starkly visible in the Pem+SoC OS curve but, given the likelihood of OS events slowing down and being more representative of the rates towards the tail of the observed KM curve, it makes good clinical sense to 'up-weight' the tail by using a flexible approach such as splines over a one-piece model when making projections.

c) **Please present scenario analyses incorporating the extrapolated OS data within the economic model, thereby updating the model structure to a partition survival model. For simplicity, the company may wish to provide this as a separate model rather than using a single model that switches between the alternative model structures.**

All results presented in this section include the commercial access agreement (CAA) currently in place for pembrolizumab; all other treatments are included at list prices.

The PSM has been built into the existing economic model using switches. The base case assumptions from the company's Rapid Review submission remain consistent in both models. Namely:-

- PFS using 3-knot splines (for detail please see the Rapid Review submission; it is very clear that no one-piece models can be considered appropriate for the Pem+SoC arm)
- Treatment waning from 3-5 years, consistent with the NICE committee's preferences
- Progression-based utilities using the linear mixed effects model, consistent with the NICE committee's preferences
- Costs using exact dosing

Partitioned survival model curve selection criteria and structural choices

Crossing of curves

As noted in our response to part a), OS and PFS crossing is a significant issue when implementing a PSM in this dataset, particularly in the Pem+SoC arm. We suggest that in the PSM structure, PFS/OS should follow PFS in the base case rather than OS. This is because the extrapolations are more reliable due to the maturity of the data. This is particularly important in light of the large difference in hazard rate between response groups within the trial.

Importantly, curve crossing should not be used as a justification to specify a significantly more pessimistic PFS curve for Pem+SoC that avoids this issue but does not fit the PFS data.

Curve selection criteria

Our principal curve selection criteria for PFS are the same as in the STM and the same range of scenarios are considered appropriate.

The curve selection criteria for OS are the following:-

- Amongst lowest AIC (predominantly splines)
- Good visual fit (predominantly splines)

- Produces extrapolations within ranges that have been agreed as clinically plausible by the clinicians at ACM1, the clinicians at the MSD advisory board and those consulted since

- Does not produce negative incremental post-progression survival

The model in the original MSD submission produced an overall survival estimate of 1.5% at 20 years for the SoC arm, which clinicians at the MSD advisory board agreed was ‘optimistic’. We have therefore set this as the upper bound of what could be considered plausible. The revised MSD STM based on the final analysis produced OS of 0.75% at 20 years, which may be plausible, given the same clinicians agreed that there would be a rare patient who would respond very will and could be alive at this time i.e. OS would not be zero at 20 years.

In the FAD, the NICE committee agreed that, based on the evidence from GOG240 and KEYNOTE-826, the testimony of clinicians and the agreement of the CDF clinical lead, that there would be an incremental PPS gain for Pem+SoC (Section 3.8). The PPS data MSD has presented from the final analysis are stronger than those shown at ACM1 and ACM2 and therefore reinforce this. We therefore rejected models which produced negative incremental PPS gains.

We consulted two NHS clinicians treating advanced cervical cancer and discussed what assumptions might be unrealistic in their experience. They stated that mean life years of >3 was unrealistic for SoC and that very few patients would be alive at 20 years but survival would not be zero.

Table 6 presents a variety of scenarios around modelling OS, in all instances the 3-knot spline is used to model PFS for both arms, in line with the model selection process we outlined in the rapid review submission. All of the scenarios using a PSM suffer from the problem of curves cross and result in negative incremental PPS QALYs estimates, which are implausible. Nevertheless, they consistently result in ICERs below the End of Life threshold. While we do not believe that the PSM should be used as the principal basis for decision-making, it contributes some structural sensitivity analysis supporting the conclusion that pembrolizumab is a cost-effective addition to SoC in this indication

Table 6: PSM model results and comparison with STM

OS Model type	PFS or OS cap	20y Pem+ SoC OS	20y SoC OS	LYs SoC	LYs Pem+ SoC	Curves do not cross	Positive inc. PPS	ICER
3k spline	PFS	[redacted]	[redacted]	2.70	4.91	No	No	[redacted]
2k spline	PFS	[redacted]	[redacted]	2.78	4.94	No	No	[redacted]
3k spline	OS	[redacted]	[redacted]	2.66	4.56	No	No	[redacted]
2k spline	OS	[redacted]	[redacted]	2.76	4.56	No	No	[redacted]
One-piece loglog	OS	[redacted]	[redacted]	3.00	4.88	No	No	[redacted]
One-piece loglog	PFS	[redacted]	[redacted]	3.00	4.97	No	No	[redacted]
STM base case	Not needed	[redacted]	[redacted]	2.65	5.33	Yes	Yes	[redacted]

We examined other PSM curve choices and it is possible to specify a model that produces positive PPS if non-flexible parametric models are used for PFS. We present them here as they were analyses of interest to the EAG during the original appraisal process but highlight that they provide an extremely poor fit to the PFS data (section 2.2 of the Company's Rapid Review submission) and are therefore not suitable for decision making. Nevertheless, the relevant ICERs are still below the threshold.

Table 7: Results using best fitting one-piece PFS curves (OS 3-knot splines)

PFS Model type	PFS or OS cap	20y Pem+ SoC OS	20y SoC OS	LYs SoC	LYs Pem+ SoC	Curves do not cross	Positive inc PPS	ICER
Gen-gamma	Not needed*	[redacted]	[redacted]	4.561	2.664	Yes*	Yes	[redacted]
Log-logistic	Not needed*	[redacted]	[redacted]	2.664	4.561	Yes*	Yes	[redacted]

*crossing occurs very late in the model when few patients are alive

It is possible to specify a PSM with a cost-ineffective ICER when the modeller makes the following selections:

- Some of the poorer fitting (AIC and visual inspection) one-piece OS curves for Pem+SoC or 1-knot spline AND

- PFS/OS continues along OS curve after curves cross

Given the array of issues that would be associated with such selections (suboptimal model fit, pessimistic projections compared to clinical expectation, prioritisation of the less mature dataset, curves crossing with zero PD patients and negative incremental PPS), we do not consider such analyses informative.

Overall we conclude firstly that the STM is still the most appropriate model for decision making as it produces survival estimates that have been validated by clinicians at ACM1 and since, positive PPS benefit for pembrolizumab, includes patients in the PD state throughout the model, is faithful to the natural history of disease and the pattern of response to treatment in KEYNOTE-826 and has been shown to be robust to all reasonable scenario analyses. Secondly, we conclude that the model's overall conclusion is extremely robust; in all reasonable STM and PSM scenarios, pembrolizumab remains a cost-effective addition to standard care in this population.

PSM deterministic results:-

Table 98: PSM ICER

Treatment	Totals per treatment arm			Incremental results			ICER (£/QALY)
	LYs	QALYs	Costs	LYs	QALYs	Costs	
SoC	2.67	[redacted]	[redacted]	2.21	[redacted]	[redacted]	[redacted]
Pem+SoC	4.91	[redacted]	[redacted]				
Key: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life year.							

Table 109 : PSM model – Mean PSA results, Pem+SoC versus SoC

	Totals per treatment arm			Incremental results			ICER (£/QALY)
Treatment	LYs	QALYs	costs	LYs	QALYs	Costs	
SoC	2.81	[redacted]	[redacted]	2.35	[redacted]	[redacted]	[redacted]
Pem+SoC	5.17	[redacted]	[redacted]				
Key: ICER, incremental cost-effectiveness ratio; LY, life year; QALY, quality-adjusted life year.							

*slight non-linearity possibly caused by curve crossing issue

Figure 10: PSM scatterplot, Pem+SoC versus SoC

Figure 11: PSM Cost-effectiveness acceptability curve, Pem+SoC versus SoC

B2. Priority Question: The company's updated base case analysis predicts a mean life expectancy of 2.65 years on standard of care. This is significantly longer than 2 years. Do the company consider that the end-of-life criteria are met?

The committee, as part of the discussions for TA885 (4), decided the end of life criteria had been met as the survival is 'normally less than 24 months' for people having standard care. This was based upon the median overall survival in the chemotherapy and bevacizumab arms of GOG240 (13.3 and 16.8 months), the percentage of people in the control arm of KEYNOTE-826 who were alive at 24 months (39%), median OS of 1.375 years and the testimony of NHS clinicians at ACM1. Mean Life Years (LYs) on SoC was not below 2 years in any of the analyses previously considered by the committee (range 2.06 – 2.51 years, ACD response Table 5). MSD considers that little has changed and that the end of life criteria are still met.

MSD believes this is the last TA that will apply the end of life criteria. Since its inception in 2009, there has been ambiguity around the meaning of 'normally less than 24 months', particularly which statistical measure of 'normal' should be applied: mean or median or some consideration of both. We note that the prognosis for the majority of patients is often better reflected by the median than the mean, which can be heavily skewed by strongly positive outcomes in a small number of patients with good response.

In this trial, median survival is 1.375 years (16.5 months), while mean is unknown exactly and dependent on survival extrapolations. Given NICE guidance is developed in language to support clinicians, commissioners and patients, we believe the lay audience would understand 'short life expectancy, normally less than 24 months' as 'the prognosis for most people' rather than a mean average dragged upwards by outliers. We consider the statement from a clinician on this topic in the appeal hearing for NICE TA788, which had similar LY estimates to this appraisal, to be the most compelling and relevant definition of 'normal':-

"He stated that bladder cancer patients in his clinic often ask him "how long they have got left?". He stated that his response was that 12 to 18 months is a reasonable estimate of life expectancy and anyone who told patients that they would normally expect to survive two years would be misleading them."

This statement would appear highly relevant in the context of this Technology Appraisal. The end-of-life appeal point was upheld in the TA788 hearing (5).

Following receipt of this clarification question, MSD also consulted with two clinicians currently treating advanced cervical cancer in the NHS, asking the question "how long would you normally expect a patient in the KN826 disease setting to survive?" and received the following responses:-

"For the patients eligible for bevacizumab my expectation would be around 12-15 months; but if a patient can tolerate only sub-optimal treatment – 9 months and if none of the treatment can be tolerated – 3-6 months."

"I would expect median overall survival for these patients to be around 9-12 months. The prognosis depends on a patients age and health status, and de novo vs recurrent metastatic cervical cancer."

These estimates show that the clinicians we spoke to understood this to be an end-of-life setting.

References:

1. Colombo N, Dubot C, Lorusso D, Caceres MV, Hasegawa K, Shapira-Frommer R, et al. Pembrolizumab for Persistent, Recurrent, or Metastatic Cervical Cancer. N Engl J Med. 2021;385(20):1856-67.
2. Merck Sharp & Dohme Corp (MSD). A Phase 3 Randomized, Double-Blind, Placebo-Controlled Trial of Pembrolizumab (MK-3475) Plus Chemotherapy Versus Chemotherapy Plus Placebo for the First-Line Treatment of Persistent, Recurrent, or Metastatic Cervical Cancer (KEYNOTE-826). 2023.
3. Tewari KS, Sill MW, Penson RT, Huang H, Ramondetta LM, Landrum LM, et al. Bevacizumab for advanced cervical cancer: final overall survival and adverse event analysis of a randomised, controlled, open-label, phase 3 trial (Gynecologic Oncology Group 240). The Lancet. 2017;390(10103):1654-63.
4. NICE. TA885 Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer 2023 [Available from: https://www.nice.org.uk/guidance/ta885/chapter/3-Committee-discussion#end-of-life.

5. TA788 Avelumab for maintenance treatment of locally advanced or metastatic urothelial cancer after platinum-based chemotherapy [Internet]. 2022. Available from: https://www.nice.org.uk/guidance/ta788.

Single Technology Appraisal (STA)

Pembrolizumab with chemotherapy for treating recurrent, persistent or metastatic cervical cancer (Rapid Review of TA885) [ID6279]

Produced by CRD and CHE Technology Assessment Group, University of York, Heslington, York, YO10 5DD

Authors Matthew Walton, Research Fellow, Centre for Reviews and Dissemination, University of York

Mark Rodgers, Research Fellow, Centre for Reviews and Dissemination, University of York

Helen Fulbright, Research Fellow in Information Science/Information Specialist, Centre for Reviews and Dissemination, University of York

Alison Eastwood, Professor of Research, Centre for Reviews and Dissemination, University of York

Robert Hodgson, Senior Research Fellow, Centre for Reviews and Dissemination, University of York

Correspondence to Robert Hodgson, Senior Research Fellow, Centre for Reviews and Dissemination, University of York

Date completed 12/09/2023

Source of funding

This report was commissioned by the NIHR Evidence Synthesis as project number NIHR136086.

Declared competing interests of the authors

None

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

Contributions of authors

Matthew Walton and Robert Hodgson critiqued the company's model and submitted economic evidence, and co-authored Sections 1, 3, and 4 of the report. Robert Hodgson took overall responsibility for the cost effectiveness sections. Helen Fulbright provided information support. Mark Rodgers critiqued the clinical evidence and wrote Section 2 of the report. Alison Eastwood provided advice, commented on drafts of the report and took overall responsibility for the clinical effectiveness sections.

Note on the text

All commercial-in-confidence (CIC) data have been highlighted in blue and underlined, all academic-in-confidence (AIC) data are highlighted in yellow and underlined.

Contents

1 Introduction and Overview

In this report, the external assessment group (EAG) has reviewed the company submission for the Rapid Review of TA885 which presents updated clinical and cost-effectiveness results using the KEYNOTE-826 trial Final Analysis (FA) data cut. The Final Appraisal Document for TA885, Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer, was published on 29 March 2023.

This overview provides a brief summary of the key issues identified by the EAG as being important for decision-making.

The EAG had no substantive concerns with the implementation of the FA data cut from KEYNOTE-826 in accordance with the committee's preferences as set out in the Final Appraisal Document (FAD) in TA885. Whilst a number of parameters relating to the trial data have been updated, they do not indicate any need for a departure from the committees preferred assumptions established in TA885.

However, the EAG identified three primary uncertainties with a potentially significant impact on the cost-effectiveness of pembrolizumab, and which may affect the interpretation of the cost-effectiveness estimates presented by the company.

1.1 Key Issue 1: Ongoing immaturity of OS data in KEYNOTE-826

In the TA885 FAD, the committee considered the continued use of a state transition model to be potentially inappropriate when final KEYNOTE-826 data became available. The EAG therefore requested that the company incorporate the FA data cut into a partitioned survival model structure for this rapid review. However, OS remained too immature to generate extrapolations with plausible long-term hazard rates with respect to established PFS trends. The assumptions necessary to maintain the PSM structure (i.e. many patients die immediately upon progression) given this immaturity are likely to be overly conservative. The cost-effectiveness results produced under these assumptions remain relatively robust. The model structure is therefore unlikely to be a key driver of uncertainty in this appraisal, however, the immaturity of the FA data cut means much of the benefit of pembrolizumab remains within the extrapolated portion of the survival curves, and is therefore subject to continuing uncertainty. This issue is discussed in Sections 2.2 and 3.2.

1.2 Key Issue 2: Generalisability of KEYNOTE-826 outcomes

Long-term PFS and OS outcomes on SoC predicted by the company's model remain much better than are likely to be expected in practice. This was highlighted as a concern by the company's clinical advisers, who estimated median survival of 9 to 12 months in this population, where median OS in

KEYNOTE-826 was 16.5 months. The EAG is concerned that this may be indicative of a trial population selected for greater fitness and propensity to a more durable treatment response, and therefore may overestimate the proportion of patients able to achieve such durable responses on either treatment in practice. This is likely to overestimate the cost-effectiveness of pembrolizumab in practice. This issue is discussed in Section 3.2.

1.3 Key Issue 3: Applicability of the end of life criteria

The committee considered that the end of life (EoL) criteria could be applied in TA885 despite mean OS on SoC being in excess of two years (2.06 years). With the KEYNOTE-826 FA data cut, SoC now offers mean survival of 2.65 years, well above the usual EoL criteria threshold. Furthermore, whilst the present appraisal is to be run using pre-2022 methods, the EAG is uncomfortable applying EoL given that NICE moved away from this approach in recognition of the lack of evidence that society places additional value on treatments at the end of life. This issue is discussed in Section **Error! Reference source not found.**.

2 Description and critique of new clinical evidence

This summary refers to the updated clinical results from KEYNOTE-826 Final Analysis (FA), and their relationship to the Interim Analysis (IA1) data previously considered by the NICE committee. For the EAG's broader commentary on the KEYNOTE-826 trial, please refer to the Evidence Review Group's Report for TA885 Pembrolizumab with chemotherapy for treating recurrent, persistent or metastatic cervical cancer.[1]

The company submission for the Rapid Review of TA885 reported the following endpoints from the final planned analysis of KEYNOTE-826:

- Progression free survival (PFS)
- Time to progression
- Overall survival (OS)
- Response rate
- Subsequent treatments
- Post progression survival (PPS)
- Adverse events

In response to the EAG's request for clarification, the company also provided FA data from KEYNOTE 826 for the remaining outcomes that were reported in the original company submission for TA885:

- Duration of response

- Health related quality of life outcomes
- Subgroup analyses for PFS and OS

2.1 Progression free survival (PFS) and time to progression (TTP)

Table 1 shows that the final analysis PFS data from the company submission are similar to the interim analysis data presented in the original appraisal (ID3798). Final analysis Kaplan-Meier estimates for PFS and TTP (which differs from PFS in that deaths are considered censors rather than events) are presented in the company submission.

The original company submission for TA885 cited the GOG-240 trial[2] which reported median PFS of 6.0 months for cisplatin plus paclitaxel chemotherapy and 8.2 months with the addition of bevacizumab. The median PFS for SoC patients in KEYNOTE-826 (63% of whom received bevacizumab) is 8.2 months (95% CI 6.3 to 8.5), suggesting that PFS in the SoC arm of KEYNOTE-826 is towards the upper end of that previously observed.

Table 1 PFS in the KEYNOTE-826 trial (CPS ≥1 population)

	Interim analysis		**Final analysis**	
	Pembrolizumab + chemotherapy ± bevacizumab (n = 273)	**Placebo + chemotherapy ± bevacizumab (n = 275)**	**Pembrolizumab + chemotherapy ± bevacizumab (n = 273)**	**Placebo + chemotherapy ± bevacizumab (n = 275)**
Number of events, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Median PFS, months (95% CI, months) [a]	10.4 (9.7, 12.3)	8.2 (6.3, 8.5)	10.5 (9.7, 12.3)	8.2 (6.3, 8.5)
PFS HR (95% CI) [b]	0.62 (0.50, 0.77)		0.58 (0.47, 0.71)	
p-value [c]	< 0.0001		<0.0001	
6-month PFS, % (95% CI)	[redacted]	[redacted]	81.5 (76.2, 85.7)	67.1 (61.0, 72.4)
12-month PFS, % (95% CI)	45.5 (39.2, 51.5)	34.1 (28.3, 40.0)	45.6 (39.3, 51.6)	33.7 (27.9, 39.5)
18-month PFS, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
24-month PFS, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]

Key: CI, confidence interval; CPS, combined positive score; FIGO, International Federation of Gynaecology and Obstetrics; HR, hazard ratio; PD-L1, Programmed death-ligand 1; PFS, progression-free survival.
Notes: [a] From product-limit (Kaplan–Meier) method for censored data. [b] Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10). [C] One-sided p-value based on log-rank test stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10).

2.2 Overall survival (OS)

In the original company submission for TA885 reported the KEYNOTE IA1 data cut in which OS data were immature, with median OS yet to have been reached for the pembrolizumab arm in the

presented interim analysis. In the final analysis, median OS has been reached in both pembrolizumab and placebo arms: 28.6 months (95% CI: 22.1, 38.0) vs 16.5 months (95% CI: 14.5, 20.0).

Table 2 shows that the interim and final analysis OS data are similar. Final analysis Kaplan-Meier estimates for OS are presented in the company submission.

GOG-240 trial[2] reported a median OS of 13.3 months for cisplatin plus paclitaxel chemotherapy and 16.8 months with the addition of bevacizumab. The median OS for SoC patients in KEYNOTE-826 (63% of whom received bevacizumab) is 16.5 months (95% CI 14.5 to 20.0), suggesting that OS in the SoC arm of KEYNOTE-826 was towards the upper end of that previously observed.

In response to a clarification question from the EAG, the company consulted with two clinicians currently treating advanced cervical cancer in the NHS, asking the question "how long would you normally expect a patient in the KN826 disease setting to survive?" and received the following responses:-

"For the patients eligible for bevacizumab my expectation would be around 12-15 months; but if a patient can tolerate only sub-optimal treatment – 9 months and if none of the treatment can be tolerated – 3-6 months."

"I would expect median overall survival for these patients to be around 9-12 months. The prognosis depends on a patients age and health status, and de novo vs recurrent metastatic cervical cancer."

Consequently, it appears that survival on SoC in KEYNOTE-826 is longer than would be expected in NHS practice. This could potentially be attributable to the subsequent treatments received by patients in this study arm (see section 2.4), though as noted in section 2.1, PFS (which would not be influenced by subsequent treatments) is also longer than observed in GOG-240.

Table 2 OS in the KEYNOTE-826 trial (CPS ≥1 population)

	Interim analysis		Final analysis	
	Pembrolizumab + chemotherapy ± bevacizumab (n = 273)	**Placebo + chemotherapy ± bevacizumab (n = 275)**	**Pembrolizumab + chemotherapy ± bevacizumab (n = 273)**	**Placebo + chemotherapy ± bevacizumab (n = 275)**
Number of events, n (%)	[redacted]	[redacted]	[redacted]	[redacted]
Median OS, months (95% CI, months) [a]	NR ([redacted])	16.3 ([redacted])	28.6 (22.1, 38.0)	16.5 (14.5, 20.0)
OS HR (95% CI) [b]	0.64 (0.50, 0.81)		0.60 (0.49, 0.74)	
p-value [c]	0.0001		<0.0001	
6-month OS, % (95% CI)	[redacted]	[redacted]	91.9 (88.0, 94.6)	85.5 (80.7, 89.1)
12-month OS, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
18-month OS, % (95% CI)	[redacted]	[redacted]	[redacted]	[redacted]
24-month OS, % (95% CI)	53.0 (46.0, 59.4)	41.7 (34.9, 48.2)	53.5 (47.4, 59.2)	39.4 (33.6, 45.2)

Key: CI, confidence interval; CPS, combined positive score; FIGO, International Federation of Gynaecology and Obstetrics; HR, hazard ratio; PD-L1, Programmed death-ligand 1; PFS, progression-free survival.
Notes: [a] From product-limit (Kaplan–Meier) method for censored data. [b] Based on Cox regression model with Efron's method of tie handling with treatment as a covariate stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10). [C] One-sided p-value based on log-rank test stratified by metastatic at initial diagnosis (FIGO [2009] stage IVB) (yes or no), bevacizumab use (yes or no) and PD-L1 status (CPS < 1, CPS 1 to < 10, CPS >=10).

2.3 Response rate

Table 3 shows the reported response rate data from the interim and final analyses of KEYNOTE-826. As for other endpoints, the response rates in the pembrolizumab and comparator arms are similar for the two analyses. The company submission presents Kaplan Meier estimates of PFS and OS separated by response category. As in the original appraisal, a patient's response status is highly prognostic of both OS and PFS, with poorer outcomes observed for each decrease in response category.

Table 3 Confirmed objective response based on investigator assessment per RECIST 1.1 (CPS ≥ 1 population)

	Interim analysis		Final analysis	
	Pembrolizumab + chemotherapy ± bevacizumab (n = 273)	**Placebo + chemotherapy ± bevacizumab (n = 275)**	**Pembrolizumab + chemotherapy ± bevacizumab (n = 273)**	**Placebo + chemotherapy ± bevacizumab (n = 275)**
Number of confirmed OR	[redacted]	[redacted]	[redacted]	[redacted]
ORR, % (95% CI)	68.1 (62.2, 73.6)	50.2 (44.1, 56.2)	[redacted]	[redacted]
CR, n (%)	62 (22.7)	36 (13.1)	[redacted]	[redacted]
PR, n (%)	124 (45.4)	102 (37.1)	[redacted]	[redacted]
SD, n (%)	58 (21.2)	88 (32.0)	[redacted]	[redacted]
PD, n (%)	9 (3.3)	29 (10.5)	9 (3.3)	29 (10.5)
Not evaluable n (%)	NR	NR	1 (0.4)	2 (0.7)
No assessment n (%)	NR	NR	19 (7.0)	18 (6.5)
Difference in percentage pembrolizumab group versus placebo group	[redacted]		17.6 (NR)	
p-value	[redacted]		NR	

Key: CI, confidence interval; CPS, combined positive score; OR, objective response; ORR, objective response rate; RECIST 1.1, response evaluation criteria in solid tumours version 1.1.

2.4 Subsequent treatments

The company submission reported subsequent oncologic treatments received by KEYNOTE-826 participants. Notably [redacted] patients in the SoC arm and [redacted] in the pembrolizumab arm had subsequent immunotherapies. The company suggested that since no immunotherapies are currently available for advanced cervical cancer in the UK, the KEYNOTE-826 data is likely to overestimate outcomes (particularly post-progression survival) in the SoC arm versus what would be seen in the UK setting.

2.5 Post progression survival (PPS)

The submission reported median PPS of [redacted] months (95% CI: [redacted]) and [redacted] months (95% CI: [redacted]) for the pembrolizumab combination and SoC arms, respectively. The associated Kaplan-Meier curves were presented in the submission.

2.6 Adverse events

Table 4 summarises the number of adverse events (AEs) and serious adverse events (SAEs), both overall and related to study drugs for the interim and final analyses of KEYNOTE-826. Observed AEs and SAEs were almost identical between analyses.

Table 4 Summary of adverse events (APaT population) – interim and final analyses

	Interim analysis				Final analysis			
	Pembrolizumab + chemotherapy ± bevacizumab (n = 307)		Placebo + chemotherapy ± bevacizumab (n = 309)		Pembrolizumab + chemotherapy ± bevacizumab (n = 307)		Placebo + chemotherapy ± bevacizumab (n = 309)	
	Any Grade	Grade 3–5	Any Grade	Grade 3–5	Any Grade	Grade 3–5	Any Grade	Grade 3–5
No. of patients, n (%)								
≥ 1 AE	305 (99.3)	251 (81.8)	307 (99.4)	232 (75.1)	305 (99.3)	253 (82.4)	307 (99.4)	233 (75.4)
≥ 1 drug-related AE	298 (97.1)	210 (68.4)	300 (97.1)	198 (64.1)	298 (97.1)	212 (69.1)	300 (97.1)	201 (65.0)
SAEs	[redacted] (49.8)	–	[redacted] (42.4)	–	157 (51.1)	–	132 (42.7)	–
Drug-related SAEs	[redacted]	–	[redacted]	–	94 (30.6)	–	73 (23.6)	
Death due to drug-related AEs	2 (0.7)	–	4 (1.3)	–	2 (0.7)	–	4 (1.3)	–
Discontinued any drug due to an AE	NR	–	NR	–	125 (40.7)	–	91 (29.4)	–
Key: AE, adverse event; APaT, all patients as treated; SAE, serious adverse event.								

2.7 Duration of response

In response to a request from the EAG, the company provided data on duration of response (DoR) from the final analysis. The median DoR was 19.2 months and 10.4 months in the pembrolizumab combination and SoC arms, respectively. The median DoR for the pembrolizumab arm increased by 1.2 months compared with the first interim analysis, with no change between the analyses for the SoC arm.

2.8 Health related quality of life outcomes

As in the interim analysis, the between group differences in least-squares mean change from baseline over the same period did not significantly differ between the pembrolizumab and SoC arms ([redacted]).

No additional data were reported for the principal pre-specified HRQoL measure for the trial (EORTC QLQ-C30 global score).

2.9 Subgroup analyses for PFS and OS

On the request of the EAG, the company provided PFS and OS data for the following subgroups: Metastatic at initial diagnosis (yes/no); bevacizumab use (yes/no); age (< 65 years/ ≥ 65 years); race (white/ non-white); and ECOG performance status (0/1).

The pattern of subgroup effects was very similar to those observed in the interim analyses and discussed in section 3.2.3.1 of the ERG report for TA885 (i.e. hazard ratios comparing pembrolizumab to placebo were less than 1 in all subgroups, and were consistent with the overall hazard ratio. However, the 95% CI for patients who were metastatic at initial diagnosis and the 95% CI for patients aged ≥65 years intersected the line of "null effect" for both PFS and OS).

Points for critique

The key clinical uncertainty addressed by the company submission relates to OS, with median OS having been reached for the pembrolizumab arm after a further 17 months of follow-up. These data significantly favoured pembrolizumab over SoC.

All reported endpoints from the final analyses were similar to the interim analyses: effect sizes were generally similar in magnitude with greater precision due to the additional observed events.

3 Description and critique of new economic evidence

3.1 Model structure

The economic model presented alongside the company submission maintained the state transition model (STM) structure adopted in TA885. A key rationale for adopting this structure over a partitioned survival model (PSM) was the comparative maturity of PFS data compared to OS. Overall survival data are now significantly more mature. The EAG noted the committee's conclusion in the Final Appraisal Determination for TA885, that '*although the company's model may be adequate for decision making with the data currently available, when further data becomes available from KEYNOTE-826, the most appropriate modelling approach may change*'. In light of this, the EAG requested that the company explore the plausibility of a PSM approach using the FA data cut.

In their clarification response, the company reiterated their preference for a STM but also presented a PSM which integrated the FA results. The company noted that the best-fitting extrapolations of these data continue to result in the crossing of OS and PFS curves when around 15% of patients are predicted to remain progression free in the Pem + SoC arm. Indeed, all reasonable extrapolations result in the crossing of OS and PFS curves. This is because while a plateau in PFS is well developed, it is only starting to become emergent on OS, and thus hazard rates adopted by most parametric models naturally result in the intersection of OS and PFS extrapolations.

This requires a judgment to be made as to whether to cap PFS at OS, or assume the hazard trends observed on PFS better represent long-term outcomes, and thus assume OS follows PFS beyond the point the curves cross (i.e. patients die upon progression). In the presented PSM base case, the company assume the latter, using PFS curves to model OS when these curves intersect. However, this

means that the committee's conclusion that patients on pembrolizumab would have 'at least a modest benefit in post-progression survival compared with placebo' is not reflected in the model.

Points for critique

The EAG acknowledges the company's concerns regarding the limitations of the final data cut. The results support the assumption that a substantial proportion of patients in the KEYNOTE-826 trial achieve a long-term reduction in PFS hazards which is yet to be reflected in the less mature OS data. The EAG considers this analysis an informative demonstration of the immaturity of OS and the relative robustness of cost-effectiveness outcomes to the structural implementation of the FA results. The EAG agrees that the PSM analysis is likely to fail to capture post-progression survival outcomes.

3.2 Survival analysis

Consistent with the STM model structure, the company base-case analysis applies parametric extrapolations of TTP, PFS and PPS to inform health state occupancy. In the alternative PSM structure presented by the company in their clarification response, OS was modelled directly in place of PPS. Each outcome was informed by data from the KEYNOTE-826 trial, which was the primary data source for the economic analysis. All model inputs from the KEYNOTE-826 trial are based on the FA data set.

3.2.1 Approach to extrapolations

The company considered a range of extrapolations including one piece, two piece, and spline-based models. Evaluation of alternative extrapolation approaches and curve selection was conducted considering visual fit to the observed data, statistical fit (Akaike information criterion (AIC) and Bayesian information criterion (BIC)) as well as the clinical plausibility of long-term survival estimates. As per committee preferences in TA885 all extrapolations of TTP, PFS and OS assumed a phased waning of the treatment effect between 3 and 5 years.

3.2.2 TTP and PFS

Following evaluation of standard one-piece parametric models, the company concluded that a single piece model (a parametric distribution fitted to the whole KM curve) had poor visual fit to the observed data and was unable to appropriately capture the complexity of the hazard functions for TTP and PFS.

The company therefore considered more flexible modelling approaches, including two-piece and spline-based approaches. Consistent with TA885, the company explored a range of two-piece models in which transition probabilities up to 37 weeks were modelled directly using observed TTP and PFS KM data, followed by the use of parametric models fitted to the remaining data. Using the criteria outlined above, the log-logistic model was selected as the most appropriate curve of the simple

parametric models. Scenario analysis was also presented using the generalised gamma (considered the 2nd choice curve). The company further explored a range of spline-based models including 1, 2 and 3 knot spline models on the normal, hazard and odds scales. The company selected a 3-knot spline model on the hazard scale as their preferred extrapolation and applied this in their base case analysis. The rationale for selecting a spline model in favour of a two-piece model was not described. Scenario analysis, however, explored two-piece alternatives and demonstrated that a 3-knot spline model is a more conservative choice (i.e. generated a higher ICER).

3.2.3 PPS

The company base case used standard one-piece parametric models to extrapolate PPS data from KEYNOTE-826, more flexible models were considered unnecessary. The company's base case uses generalised gamma curves fitted to each treatment arm independently.

3.2.4 OS

Extrapolations of OS were provided following a request from the EAG at points for clarification (PFC B1). These are subsequently used in the alternative PSM model structure, which was also provided at clarification. The company's approach to extrapolating OS adopted a similar approach to that used to evaluate TTP, PFS and PPS extrapolations but also considered the plausibility of post-progression survival predictions as an additional selection criterion. The company explored both one-piece and spline-based models. Two-piece models appear not to have been considered. The EAG is not overly concerned by this omission as spline and two-piece models represent similarly flexible approaches. The company did not consider one-piece models appropriate to model OS, noting poor visual and statistical fit. Clinical expert advice also considered that all one-piece models were likely to underestimate survival on Pem+SoC. The company's preferred extrapolation was based on the 3-knot spline model using the hazards scale.

Points for critique

The EAG is generally satisfied with the company's approach to extrapolating the observed survival data for use in the PSM. In TA885, concerns were raised about the suitability of flexible parametric models, highlighting limited evidence to support a sustained plateau in PFS hazards. The longer follow-up in the FA analysis provides stronger support for the plateau in PFS, though this remains largely absent from the observed OS data. The EAG concurs with the company's assessment that one-piece parametric models inadequately capture the complex hazard trends observed in the Pem+SoC arm and considers that the company-preferred base-case spline models for both PFS and OS offer good visual and statistical fit to the observed data. The EAG, however, makes the following observations.

Firstly, the clinical plausibility of the selected extrapolations depends upon additional assumptions made regarding the waning of the treatment effect. The landmark estimates of OS predicted by the model (Tables 9 and 12 of the CS) are therefore conditional on a treatment waning effect also being applied. In the absence of this treatment waning, predicted OS is substantially higher. For example, 5-, 10- and 20-year survival using the company's preferred extrapolations without waning are [redacted] respectively. These are highly optimistic predictions of OS driven by a sharp decline in modelled PFS events. The clinical plausibility of these predictions is important because these extrapolations determine the proportion of patients that transition to SoC hazards following the waning period, which, as outlined below, also suggests sharply declining hazards and very long survival.

Secondly the company's preferred extrapolations of PFS also results in highly optimistic predictions of OS on SoC, generating 10-year survival predictions of [redacted] This is acknowledged by the company as clinically unrealistic and was highlighted as a concern by the company's clinical advisory panel. This is important because hazards modelled for the SoC arm are ultimately applied to both treatment arms following the application of effect waning. It is unclear whether these long tails predicted by the spline (and two-piece) models are clinically realistic and may result in OS being overestimated in both the Pem+SoC and SoC arms of the model.

Thirdly, the optimistic estimates of PFS and OS on SoC are not wholly attributable to the extrapolation approach. The observed PFS and OS data from the KEYNOTE-826 trial is relatively mature, and thus these predictions simply reflect the observed data. Outcomes for patients in the SoC arm appear to be better than those in clinical practice. Indeed, the company highlights that median survival is substantially longer in KEYNOTE-826 than that estimated by their clinical experts (16.5 months vs. 9 to 12 months predicted by clinical experts). The EAG is concerned that this indicates that the KEYNOTE-826 trial recruited a highly optimised population and may overrepresent the proportion of patients who are able to achieve durable responses in practice (on either treatment). This would tend to favour Pem+SoC and is a critical concern because so much of the modelled benefit is as a result of a proportion of patients achieving sustained survival benefits.

In summary, while the final analysis has helped resolve some uncertainties regarding the benefits of Pem+SoC much of this benefit remains within the extrapolated portion of the survival curves and as such remains uncertain.

3.3 Other updated model inputs

3.3.1 Time on treatment

Time on treatment outcomes were updated using the FA data cut. Mean time on treatment was slightly longer on pembrolizumab than at IA1, which results in a small increase in incremental costs. The EAG is satisfied that the company's implementation of the updated data cut is appropriate.

3.3.2 Health-related quality of life

The company analysed the final EQ-5D-5L data collected from KEYNOTE-826, which were mapped onto the 3L scale using the Hernandez-Alava *et al*. algorithm. The company's base-case approach adopts the progression-based linear mixed effects model preferred by the company in TA885, and generates utilities with only negligible differences to those accepted in that appraisal. The EAG is satisfied that the company's preferred utility set based on the FA data cut is appropriate.

3.3.3 Resource use

The company have made no material changes to the costs adopted in the model with the exception of those inputs affected by the updated data cut, namely relative dose intensity (RDI) and the duration of treatment with subsequent therapies. These updates had a minimal impact on total costs. The EAG is satisfied that the company's implementation of the updated data is appropriate.

3.4 End of life

In the analyses considered by the committee in TA885, mean overall survival on the standard of care was 2.06 years. Given the influence of the long tail on mean survival, and that 58.3% patients had died at 24 months, the committee concluded that on balance the end of life (EoL) criteria could be applied on the basis of the TA788 appeal decision, in which the committee were obliged to consider other measures of life expectancy.

In the company's base-case analysis incorporating the final data cut from KEYNOTE-826, SoC generated mean discounted life years (LYs) of 2.65, i.e. ~32 months. This is considerably longer than mean OS in the previous appraisal, and for consistency with previous appraisals, and the usual interpretation of life expectancy and QALY gain, should not be considered EoL.

Median OS was 16.5 months in KEYNOTE-826. It is unclear whether it could be argued that survival is 'normally less than 24 months', when 40% of patients remain alive in the model at two years. Given the significant extension to OS on SoC in the latest data cut, the EoL criteria may not continue to be applicable in the present appraisal.

As discussed in Section 3.2, survival on SoC appeared to be substantially longer in KEYNOTE-826 than would be expected in practice, which indicates the trial population is healthier than might be expected in practice. This may mean the trial overestimates real-world OS on SoC to the extent that

the EoL criteria may be applicable. However, in such a scenario, current cost-effectiveness estimates should no longer be interpreted as presented. The outcomes on the Pem+SoC arm are unlikely to be representative of those achieved in practice, as long-term (i.e. post-waning) PFS hazards are based on those achieved on SoC. Whether or not the KEYNOTE-826 trial can sufficiently represent achievable outcomes in the NHS population presents a potentially significant uncertainty with regards to both the interpretation of cost-effectiveness outcomes, and the application of the EoL criteria.

Furthermore, the EAG is uncomfortable with applying EoL despite the present appraisal running on pre-2022 methods. The current NICE methods were developed in recognition of evidence that society does not place additional value on treatments at the end of life. It is therefore unclear whether it is appropriate to make decisions on the basis of a £50,000 willingness-to-pay threshold, particularly when the EoL criteria are not strictly met.

4 Updated economic model

The company's cost-effectiveness model adopts all of the assumptions agreed by the committee in the TA885 FAD, updating only the inputs derived directly from KEYNOTE-826, as discussed over the previous sections. Note that all results presented in this document are inclusive only of the currently approved patient access scheme (PAS) for pembrolizumab. A confidential appendix is provided which presents results inclusive of all available confidential commercial arrangements.

The company states the key assumptions of the company's base case as follows:

- PFS and TTP extrapolated using 3-knot spline models
- Treatment waning between 3-5 years
- PPS using individual generalised gamma curves
- Progression-based utilities using the linear mixed effects model
- Costs using exact dosing

These assumptions are consistent with those preferred by the committee in TA885.

4.1 Results of company base-case analysis

4.1.1 Deterministic analysis

The cost-effectiveness results for the company's base-case analysis are presented in Table 5.

Table 5 Company base case (deterministic)

Treatment	Totals per treatment arm			Incremental results			ICER (£/QALY)
	Costs	LYs	QALYs	Costs	LYs	QALYs	
Company base case							

SoC	[redacted]	2.65	[redacted]				
Pem+SoC	[redacted]	5.33	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

4.1.2 Probabilistic sensitivity analysis

The company's probabilistic sensitivity analysis (PSA) results are presented in Table 6 using 5,000 model iterations. Pembrolizumab had a [redacted] probability of being the most cost-effective treatment option at a willingness-to-pay threshold of £50,000 per QALY gained. This drops to [redacted] at a threshold of £30,000.

Table 6 Company base case (probabilistic)

Treatment	Totals per treatment arm			Incremental results			ICER (£/QALY)
	Costs	LYs	QALYs	Costs	LYs	QALYs	
Company base case							
SoC	[redacted]	2.51	[redacted]				
Pem+SoC	[redacted]	5.11	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 1 Cost-effectiveness plane (company base case)
[REDACTED]

4.1.3 Scenario analyses

In their submission and clarification response, the company present a range of scenario analyses which explore the impact of alternative extrapolations, waning assumptions, utility sets, and the imposition of a PSM structure upon cost-effectiveness estimates. A selection of the results presented across the two submission documents are reproduced in

Table 7. These results include only the currently available PAS discount for pembrolizumab, and are replicated inclusive of all commercial arrangements available to the NHS in the confidential appendix to the report.

Table 7 Company scenario analyses (deterministic)

Treatment	Totals per treatment arm			Incremental results			ICER (£/QALY)
	Costs	LYs	QALYs	Costs	LYs	QALYs	
Company base case							
SoC	[redacted]	2.65	[redacted]				
Pem+SoC	[redacted]	5.33	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS/TTP 1-knot spline							
SoC	[redacted]	2.50	[redacted]				
Pem+SoC	[redacted]	4.84	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PFS/TTP 2-knot spline							
SoC	[redacted]	2.73	[redacted]				
Pem+SoC	[redacted]	5.73	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PPS modelled using exponential curve							
SoC	[redacted]	2.65	[redacted]				
Pem+SoC	[redacted]	5.01	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment waning applied between 5-7 years							
SoC	[redacted]	2.65	[redacted]				
Pem+SoC	[redacted]	5.91	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pooled PPS curve							
SoC	[redacted]	2.46	[redacted]				
Pem+SoC	[redacted]	4.72	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
PSM structure (OS 3-knot spline)							
SoC	[redacted]	2.70	[redacted]				
Pem+SoC	[redacted]	4.91	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

5 References

1. National Institute for Health and Care Excellence. *Pembrolizumab plus chemotherapy with or without bevacizumab for persistent, recurrent or metastatic cervical cancer [TA885]*: NICE; 2023. Available from: https://www.nice.org.uk/guidance/ta885/

2. Tewari KS, Sill MW, Penson RT, Huang H, Ramondetta LM, Landrum LM, et al. Bevacizumab for advanced cervical cancer: final overall survival and adverse event analysis of a randomised, controlled, open-label, phase 3 trial (Gynecologic Oncology Group 240). *Lancet* 2017;**390**:1654-63.