Pembrolizumab with lenvatinib for previously treated advanced, metastatic or recurrent endometrial cancer [ID3811]:

A Single Technology Appraisal

Produced by

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

Authors

Brian O'Toole[1]
Maxwell S Barnish[1]
Edward CF Wilson[1]
Justin Matthews[1]
Helen Coelho[1]
Naomi Shaw[1]
Melanie Powell[2]
Neil Ryan[3]
Emma Crosbie[4]
Alexandra Taylor[5]
Louise Crathorne[1]
G.J. Melendez-Torres[1]

[1] Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter

[2] Consultant Clinical Oncologist, St Bartholomew's Hospital, London

Produced by

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

[3] Specialty Registrar in Obstetrics and Gynaecology, Severn Deanery

[4] NIHR Advanced Fellow, Professor and Honorary Consultant in Gynaecological Oncology, Manchester

[5] Consultant in Clinical Oncology, The Royal Marsden NHS Foundation Trust

Correspondence to

Brian O'Toole
3.09 South Cloisters, St Luke's Campus, Heavitree Road, Exeter, EX1 2LU; b.otoole@exeter.ac.uk

Date completed 01/06/2022

Source of funding This report was commissioned by the NIHR Systematic Reviews Programme as project number 13/50/59.

Declared competing interests of the authors Authors have no interests to declare.

Acknowledgments The authors acknowledge the administrative support provided by Mrs Sue Whiffin and Ms Jenny Lowe (both PenTAG).

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

This report should be referenced as follows O'Toole B, Barnish MS, Wilson ECF, Coelho H, Shaw N, Powell M, Ryan N, Crosbie E, Taylor A, Crathorne L, Melendez-Torres GJ. Pembrolizumab with lenvatinib for previously treated advanced, metastatic or recurrent endometrial cancer [ID3811]: A Single Technology Appraisal. Peninsula Technology Assessment Group (PenTAG), 2022.

Author Contributions:

Brian O'Toole	Project lead. Lead for the ERG's appraisal of the economic evidence, drafted economic sections of the report, writing and editorial input

Author Contributions:	
Maxwell S. Barnish	Lead for the ERG's critical appraisal of the clinical evidence of the company submission, writing and editorial input
Edward CF Wilson	Critical appraisal of the economic evidence, checked and re-analysed the economic model, carried out ERG base case analyses and further scenario analyses, drafted economic sections of the report and editorial input
Justin Matthews	Critical appraisal of the clinical evidence and drafted sections of the report
Helen Coelho	Critical appraisal of the clinical evidence and drafted sections of the report
Naomi Shaw	Critical appraisal of the literature search strategies, conducted additional literature searching, and editorial review
Melanie Powell	Clinical advice
Neil Ryan	Clinical advice and review of draft report
Emma Crosbie	Clinical advice and review of draft report
Alexandra Taylor	Clinical advice
Louise Crathorne	Editorial review
G.J. Melendez-Torres	Critical appraisal of the clinical and economic evidence, editorial review

Table of Contents

List of tables

List of Figures

Abbreviations

AE	Adverse event
BNF	British National Formulary
CS	Company submission
CSR	Clinical study report
dMMR	Deficient mismatch repair
EC	Endometrial Cancer
ERG	Evidence Review Group
HRQoL	Health related quality of life
ICER	Incremental cost effectiveness ratio
KM	Kaplan-Meier
LY	Life year
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
OS	Overall survival
OWSA	One way sensitivity analysis
PD	Progressed disease
PD-1	Programmed cell death protein 1
PD-L1	Programmed death ligand 1
PEM+LEN	Pembrolizumab with lenvatinib
PFS	Progression free survival
pMMR	Proficient mismatch repair
PSSRU	Personal Social Services Research Unit
QALY	Quality adjusted life year
QoI	Quality of life
SLR	Systematic literature review
SmPC	Summary of product characteristics
ToT	Time on treatment
TPC	Treatment of physicians choice
TTD	Time to death
UK	United Kingdom

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the evidence review group (ERG) as being potentially important for decision-making. It also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 provides an overview of key model outcomes and the modelling assumptions that have the greatest effect on the ICER. Sections 1.3 to 1.6 explain the key issues in more detail, and Section 1.7 presents the preferred assumptions of the ERG. Background information on the condition, technology and evidence and information on non-key issues are in the main ERG report.

All issues identified represent the ERG's view, not the opinion of NICE.

1.1. Overview of the ERG's key issues

A brief overview of the key issues identified by the ERG in their appraisal of the company submission (CS) is provided in Table 1. Further detail of the issues is provided in Sections 1.3, 1.4, 1.5, and 1.6.

Table 1: Summary of key issues

ID[3811]	Summary of issues	Report sections
#1 and #2	Clinically distinct subgroups in the evidence base	Section 1.3, 1.5 and 4.2.3
#3	Uncertainty surrounding modelled OS	Section 1.5 and 4.2.6
#4	Uncertainty surrounding base case utility values (time to death approach)	Section 1.5 and 4.2.8
#5	Treatment waning	Section 1.5 and 4.2.6.3

Abbrevations: MMR, mismatch repair; OS, overall survival; ToT, time-on-treatment

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are outlined in Table 2.

Table 2: Key differences between the company's preferred assumptions and ERG's preferred assumptions

	Company's preferred assumption	ERG preferred assumption	Report Sections
Capping survival to ensure PFS ≤ OS	Blended approach	Hazards-based approach	Section 4.2.6.2 and 6.2.1
Proportion of patients receiving doxorubicin or paclitaxel in TPC	As observed in KEYNOTE-775 (74.5% received doxorubicin, 25.5% paclitaxel)	50/50 split between doxorubicin and paclitaxel	Section 4.2.3 and 6.2.2
Time on treatment	As observed in KEYNOTE-775	Capped to disease progression	Section 6.2.7
Health state utilities	Based on time to death	Based on health state (progression-free and progressed disease)	Section 4.2.8 and 6.2.6
Patient weight	70 kg	85 kg (plus associated increase in BSA)	Section 4.2.3 and 6.2.8
Patient age	63.5 years (median)	75 years	Section 4.2.3 and 6.2.9
OS for TPC	KM+Exponential	KM+Log-logistic	Section 4.2.6

Abbreviations: BSA, body surface area; ERG, evidence review group; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; TPC, treatment of physician's choice (control arm of KN775).

1.2. Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival) and quality of life in a quality-adjusted life year (QALY). An ICER is the ratio of the extra cost for every QALY gained.

Overall, the technology is modelled to affect QALYs by:

- Keeping a higher proportion of patients in the progression-free survival (PFS) and progressed disease (PD) health states, for longer. As PEM+LEN is modelled to delay progression and extend survival, patients accrue more QALYs and gain more life years (LYs) compared to those receiving doxorubicin or paclitaxel.
- Time spent in the PD health state and use of time-to-death to estimate utilities, as most of the PEM+LEN incremental QALY gain (73%) is accrued in the PD health state.

Overall, the technology is modelled to affect costs by:

- Drug acquisition costs, as PEM+LEN results in substantially higher costs compared to the comparator treatment arm (treatment of physician's choice, or TPC, of doxorubicin or paclitaxel). Drug costs are a key driver of incremental costs.
- Adverse event costs, end of life costs and subsequent treatment costs, as these are lower in the PEM+LEN arm (however the incremental cost difference between treatment arms is considered minor).

The modelling assumptions that have the greatest effect on the ICER are as follows:

- Based on scenario analysis conducted by the ERG, results are most sensitive to variation in OS extrapolation assumptions and treatment waning (see Section 6.2.10).
- Based on scenario analyses submitted by the company, the assumptions with the largest impact on the ICER were the discount rate for costs and benefits (1.5% for both), no dose reduction for lenvatinib (based on full dose of 20mg per week), health state utilities based on progression status (not time to death), basing the cost of doxorubicin on Caelyx® (branded liposomal/pegylated doxorubicin), and restricting time-on-treatment (ToT) to PFS. The company did not perform a sensitivity analysis on acquisition cost of pembrolizumab or lenvatinib. The results are moderately sensitive to these parameters.

1.3. The decision problem: summary of the ERG's key issues

The ERG reviewed the approach of the company to addressing the NICE decision problem for this appraisal and identified the following key issues for the committee's consideration

Key Issue 1: Clinically distinct subgroups in the evidence base

Report sections	Sections 2.4 and 3.2.3.1
Description of issue and why the ERG has identified it as important	The ERG noted that there were two clinically distinct subgroups in the population of the pivotal KEYNOTE-775 trial. Point estimate results suggested patients in the dMMR subgroup may have performed better than patients in the pMMR subgroup on both OS and PFS outcomes, although it should be noted that the study was not specifically powered to explore the impact of MMR status on survival outcomes and the follow-up period of KEYNOTE-775 was limited.
What alternative approach has the ERG suggested?	Clinical effectiveness results for dMMR and pMMR subgroups were provided. However, the company did not provide cost effectiveness subgroup results nor did the company model offer

Report sections	Sections 2.4 and 3.2.3.1
	the functionality to allow the ERG to implement sub-groups as a scenario analysis. The ERG recognised that the subgroups were not pre-defined in the NICE scope, but rather emerged from the clinical effectiveness results.
What is the expected effect on the cost-effectiveness estimates?	The expected impact on cost effectiveness of each subgroup remains unclear. However, due to improved OS in the dMMR subgroup, compared to the pMMR subgroup, the ICER for PEM+LEN is likely to be lower in this subgroup, all else remaining equal.
What additional evidence or analyses might help to resolve this key issue?	The provision of subgroup-specific cost effectiveness scenario analyses and the model functionality to produce these analyses would help resolve the uncertainty.

Abbreviations: dMMR, deficient mismatch repair; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; NICE, National Institute for Health and Care Excellence; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival; pMMR, proficient mismatch repair

1.4. The clinical effectiveness evidence: summary of the ERG's key issues

No clinical effectiveness key issues were identified.

1.5. The cost effectiveness evidence: summary of the ERG's key issues

Key Issue 2: Uncertainty surrounding the cost effectiveness of PEM+LEN within dMMR and pMMR subgroups

Report sections	Section 1.3 and 4.2.3
Description of issue and why the ERG has identified it as important	The company presented cost effectiveness results which were in alignment with the NICE final scope and the company's marketing authorisation i.e. for the treatment of advanced or recurrent endometrial carcinoma in adults who have disease progression on or following prior treatment with a platinum-containing therapy in any setting and who are not candidates for curative surgery or radiation. However, based on clinical expert opinion to the ERG, prognosis and treatment is likely to differ for patients based on MMR status. As part of KEYNOTE-775, the company conducted subgroup analyses for both the dMMR and pMMR patients, however cost effectiveness results were not presented. Given that overall OS for PEM+LEN varies depending on MMR status (as per the subgroup data outlined in Section 3.2.3.1), cost effectiveness results are expected to vary between subgroups. The company's base case analysis therefore does not explore the cost

Report sections	Section 1.3 and 4.2.3
	effectiveness of PEM+LEN in two clinically relevant subgroups, which represents an area of uncertainty for the ERG.
What alternative approach has the ERG suggested?	Cost effectiveness results presented for the dMMR and pMMR subgroups would have adequately addressed uncertainty. The ERG were unable to conduct subgroup analyses, due to time and data constraints.
What is the expected effect on the cost-effectiveness estimates?	The expected impact on cost effectiveness of each subgroup remains unclear. However, due to improved OS in the dMMR subgroup, compared to the pMMR subgroup, the ICER for PEM+LEN is likely to be lower in this subgroup, all else remaining equal.
What additional evidence or analyses might help to resolve this key issue?	Provision of cost effectiveness results for each subgroup would resolve this issue.

Abbreviations: dMMR, deficient mismatch repair; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; MMR, mismatch repair; NICE, National Institute for Health and Care Excellence; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; pMMR, proficient mismatch repair

Key Issue 3: Uncertainty surrounding extrapolation of OS

Report sections	Sections 4.2.6, 6.2 and 6.2.5
Description of issue and why the ERG has identified it as important	Based on clinical opinion to the ERG there was some concern surrounding the long term overall survival estimates modelled by the company, namely that the extrapolated curves lacked clinical plausibility and were too far apart. Specifically, five year OS in the PEM+LEN arm was considered optimistic whilst this was considered pessimistic in the doxorubicin or paclitaxel arm. The ERG considered that the company's base case extrapolation approach potentially biases the analysis in favour of PEM+LEN by overestimating life years, and underestimating life years in the doxorubicin or paclitaxel arm. Additionally, the ERG noted concerns surrounding the following: • The company's dismissal of alternative modelling approaches, including the use of restricted cubic splines (see Section 4.2.6 for commentary). • The use of the ECHO study as a means of validating OS in the doxorubicin or paclitaxel arm (see Section 4.2.6.4). • The ERG did not have access to the KEYNOTE-146 CSR, which introduced further uncertainty. The company included

Report sections	Sections 4.2.6, 6.2 and 6.2.5
	this trial in its submission, but is not the sponsor or owner of the CSR.
What alternative approach has the ERG suggested?	The ERG conducted additional scenario analyses using alternative parametric curves for OS extrapolation. For PEM+LEN, the KM + Weibull was used and for the doxorubicin or paclitaxel arm, the KM + Log logistic curve was used. The ERG also conducted a combined scenario analysis which used both of these alternative curves (see Section 6.2.5).
What is the expected effect on the cost-effectiveness estimates?	The impact of these changes caused the OS gap between treatment arms to narrow, thereby reducing the incremental LY gain in the PEM+LEN arm. Results were highly sensitive to these scenario analyses (see Section 6.2.5 for results).
What additional evidence or analyses might help to resolve this key issue?	Whilst the ERG acknowledged ECHO provided supplementary supportive evidence with respect to OS, the ERG identified several limitations with this study (see Section 3.3). Furthermore, in order to explore uncertainty surrounding OS extrapolation, the company could have also provided results using alternative modelling approaches including the use of restricted cubic splines.

Abbreviations: CSR, clinical study report; ERG, Evidence Review Group; KM, Kaplan-Meier; LY, life year; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib

Key Issue 4: Uncertainty surrounding the company's time to death utility approach

Report sections	Sections 4.2.8, and 6.2.6
Description of issue and why the ERG has identified it as important	In the base case analysis, the company used a TTD approach to derive utility values for modelled health states. The ERG considered that a more reasonable approach was to base utility values on progression status i.e. PF and PD. This approach is consistent with the company's model structure which includes progression-free and progressed disease as health states. Furthermore, the ERG noted that in the company's base case TTD approach, varying the PFS curve (whilst keeping OS unchanged) did not have an impact on QALYs, but did impact costs. This result appeared somewhat counter-intuitive. Based on scenario analysis provided by the company, results were sensitive to the estimation of utility values based on progression status.
What alternative approach has the ERG suggested?	The ERG preferred to base health state utility values on progression status. This preference forms part of the ERG base case.

Report sections	Sections 4.2.8, and 6.2.6
What is the expected effect on the cost-effectiveness estimates?	This scenario had a relatively small upward impact on the ICER.
What additional evidence or analyses might help to resolve this key issue?	Estimating health state utilities based on progression status mostly resolves this issue. However, longer-term QoL data would be helpful to validate modelled estimates.

Abbreviations: ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; OS, overall survival; PD, progressed disease; PF, progression-free; PFS, progression-free survival; QALY, quality-adjusted life year; QoL, quality of life; TTD, time-to-death

Key issue 5: Treatment waning

Report sections	Sections 4.2.6.3 and 6.2.3
Description of issue and why the ERG has identified it as important	The company's base case analysis assumes no waning of treatment effect i.e. after patients discontinue PEM+LEN the treatment effect is assumed to be maintained over time. Although the company provided some justification for not including treatment waning (see Section 4.2.6.3 and response to B.18 of the company's clarification response), the ERG considered there to be some uncertainty surrounding the maintenance of the PEM+LEN treatment effect. Clinical opinion to the ERG noted data on treatment waning are limited, however it may be reasonable to assume gradual waning once patients stop treatment.
What alternative approach has the ERG suggested?	The ERG conducted a scenario analysis which included a treatment waning effect in the PEM+LEN arm between years 2 and 5 (see Section 6.2.3 for details). The ERG did not include this scenario as part of its preferred base case due to the lack of data supporting this assumption. However, this scenario does highlight the sensitivity of results to the use of alternative treatment effect assumptions.
What is the expected effect on the cost-effectiveness estimates?	Results were highly sensitive to this scenario (see Section 6.2.10 for results).
What additional evidence or analyses might help to resolve this key issue?	Robust long-term treatment effectiveness data would help to resolve this uncertainty.

Abbreviations: dMMR, deficient mismatch repair; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; MMR, mismatch repair; NICE, National Institute for Health and Care Excellence; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; pMMR, proficient mismatch repair

1.6. Other issues: summary of the ERG's views

Issue 6: Time-on-treatment, percentage of patients receiving doxorubicin or paclitaxel, modelled baseline patient characteristics and approach to capping survival

Report sections	Sections 4.2.7, 4.2.4, 4.2.3.1, 6.2.7, 6.2.2, 6.2.8 and 6.2.9
Description of issue and why the ERG has identified it as important	ToT: In the base case analysis, the company modelled ToT independently for PEM+LEN i.e. a generalised gamma curve was used for both arms. The ERG considered that a more appropriate method is to cap ToT by PFS (for all treatments), as ToT should be coterminous with PFS. See section 4.2.7 for further discussion. Percentage of patients receiving doxorubicin or paclitaxel: In the base case analysis, the company assumed that 75% of patients would receive doxorubicin and 25% would receive paclitaxel. Based on clinical input to the ERG, a more even split (50/50) is likely to better represent clinical practice (see section 4.2.4 for further discussion). Modelled baseline patient characteristics: In the base case the company based patient weight and age on patient characteristics from KEYNOTE-775. Based on clinical input, patients in the UK are likely to be heavier and older than those in KEYNOTE-775 (see section 4.2.3 for further discussion). Capping of overall survival: The company used a 'hybrid' approach to capping overall survival to general population survival and PFS to OS. The ERG's preference is for the hazards-based approach as this generates more plausible estimates of survival (see Section 4.2.6.2 for further discussion).
What alternative approach has the ERG suggested?	ToT: The ERG conducted a scenario analysis which capped ToT by PFS (for all treatments). This has been included as part of the ERG's preferred base case. Percentage of patients receiving doxorubicin or paclitaxel: The ERG conducted scenario analyses which varied the proportion of patients receiving either doxorubicin or paclitaxel (see Section 6.2.2). The ERG preferred base case assumes that 50% of patients receive doxorubicin and 50% receive paclitaxel. Modelled patient baseline characteristics: The ERG has conducted a scenario analysis which increased mean patient weight to 85 kg (and BSA to 1.96 m^2) and patient age to 75 years. This has

Report sections	Sections 4.2.7, 4.2.4, 4.2.3.1, 6.2.7, 6.2.2, 6.2.8 and 6.2.9
	been included as part of the ERG's preferred base case. Capping of overall survival: In this scenario the ERG have used two alternative approaches to capping overall survival, the 'simple' approach and 'hazards' approach (see Section 6.2.1). The ERG preferred base case uses the hazards approach.
What is the expected effect on the cost-effectiveness estimates?	ToT capped by PFS: This caused the ICER for PEM+LEN to decrease (due to reduced drug costs). See Section 6.2.7. Percentage of patients receiving doxorubicin or paclitaxel (50/50): This scenario had minimal impact on the ICER. See Section 6.2.2. Modelled patient baseline characteristics: Altering patient age had a mild upward impact on the ICER, however increasing patient weight did not have a meaningful impact. See Section 6.2.8.
What additional evidence or analyses might help to resolve this key issue?	The additional analyses conducted by the ERG have addressed these issues.

Abbreviations: BSA, body surface area; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival; ToT, time-on-treatment

1.7. Summary of ERG's preferred assumptions and resulting ICER

The ERG's preferred base case results are presented below. Please note that the results include the PAS for pembrolizumab and list price for lenvatinib. [redacted] All of the ERG's analyses therefore include the latest PAS.

Table 3: ERG preferred assumptions (deterministic)

	ERG report section	Incremental cost	Incremental QALYs	ICER
Company's base case	5.1.1.1	[redacted]	1.75	[redacted]
ERG corrected company base case				
Drug costs corrected + additional PAS	6.1	[redacted]	1.75	[redacted]
ERG Preferred base case assumptions **(applied individually)**				

	ERG report section	Incremental cost	Incremental QALYs	ICER
Survival capped by hazards	6.2.1 and 6.2.10	[redacted]	1.77	[redacted]
50% of patients receive doxorubicin and 50% receive paclitaxel	6.2.2 and 6.2.10	[redacted]	1.75	[redacted]
ToT capped by PFS	6.2.7 and 6.2.10	[redacted]	1.76	[redacted]
Health state utilities based on progression status	6.2.6 and 6.2.10	[redacted]	1.59	[redacted]
Patient weight increased to 85 kg (and BSA to 1.96 m^2)	6.2.8 and 6.2.10	[redacted]	1.75	[redacted]
Patient age increased to 75 years	6.2.9 and 6.2.10	[redacted]	1.55	[redacted]
OS for doxorubicin or paclitaxel (KM +Log logistic)	6.2.5 and 6.2.10	[redacted]	1.31	[redacted]
Cumulative impact of ERG's preferences	1.7 and 6.3	[redacted]	1.07	[redacted]

Abbreviations: BSA, body surface area; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year; ToT, time-on-treatment

Table 4: ERG preferred assumptions (probabilistic)

	ERG report section	Incremental cost	Incremental QALYs	ICER
Company's base case	5.1.1.1	[redacted]	1.77	[redacted]
ERG corrected company base case				
Drug costs corrected + additional PAS	6.1	[redacted]	1.76	[redacted]
ERG Preferred base case assumptions (applied incrementally)				
Survival capped by hazards	6.2.1 and 6.2.10	[redacted]	1.77	[redacted]
50% of patients receive doxorubicin and 50% receive paclitaxel	6.2.2 and 6.2.10	[redacted]	1.75	[redacted]

	ERG report section	Incremental cost	Incremental QALYs	ICER
ToT capped by PFS	6.2.7 and 6.2.10	[redacted]	1.76	[redacted]
Health state utilities based on progression status	6.2.6 and 6.2.10	[redacted]	1.61	[redacted]
Patient weight increased to 85 kg (and BSA to 1.96 m^2)	6.2.8 and 6.2.10	[redacted]	1.76	[redacted]
Patient age increased to 75 years	6.2.9 and 6.2.10	[redacted]	1.56	[redacted]
OS for doxorubicin or paclitaxel (KM +Log logistic)	6.2.5 and 6.2.10	[redacted]	1.32	[redacted]
Cumulative impact of ERG's preferences	1.7 and 6.3	[redacted]	1.05	[redacted]

Abbreviations: BSA, body surface area; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year; ToT, time-on-treatment

Modelling errors identified and corrected by the ERG are described in Section 6.1. For further details of the exploratory and sensitivity analyses done by the ERG, see Section 6.2.

2. INTRODUCTION AND BACKGROUND

2.1. Introduction

In this report, the Evidence Review Group (ERG) provides a review of the evidence submitted by Merck, Sharp and Dohme (MSD) in support of pembrolizumab with lenvatinib (PEM+LEN) for previously treated advanced, metastatic or recurrent endometrial cancer (EC).

2.2. Critique of the company's description of the underlying health problem

The ERG is broadly in agreement with the company's description of the underlying health problem. The company describe Stage III EC as advanced cancer that has spread outside the womb and Stage IV EC as cancer that has spread beyond the pelvis (womb, bowel or bladder). For clarity, the ERG refer to the BGCS Uterine Cancer Guideline Recommendations for Practice 2021[1] where Stage III and Stage IVA (spread to other areas of the pelvis) EC are described as advanced and Stage IVB (distal spread) as metastatic. These guidelines use internationally recognized FIGO and TNM staging methods.

The company provide information about the age of the population (i.e. highest incidence in people aged 75-79 years) but not about weight. The ERG noted that a large proportion of the population are overweight or obese.[2]

The ERG highlights the importance of separately considering the mismatch repair, or MMR, subgroups within the target population. Section B.2.4.2 of the CS provides an acknowledgment that the efficacy of PEM+LEN is expected to be greater for those with deficient MMR, or dMMR, EC (vs proficient MMR, or pMMR, EC), and clinical effectiveness data are provided according to MMR subgroups in Appendix P of the CS. However, these important subgroups are not highlighted in the CS from the outset, and not separately considered in the economic analyses. Clinical expert advice to the ERG confirms that dMMR tumours are generally (but not always) considered to have a better treatment response and prognosis than pMMR tumours, and most importantly are more likely to respond to immunotherapy. Recently, a clear difference in the treatment pathway has emerged for people with dMMR EC compared to those with pMMR EC: those with advanced or recurrent previously treated EC displaying dMMR are now able to access dostarlimab as monotherapy (NICE Technology Appraisal Guidance TA779).[3]

2.3. Critique of the company's overview of current service provision

The company provide a description of the clinical pathway for people with advanced, metastatic or recurrent EC (refer to B.1.3.2 and Appendix L in the company submission) alongside a diagram of the clinical pathway (refer to Figure 1, Section B.1.3.2 in the company submission). The ERG agree that this is largely consistent with the BGCS guidelines,[1] and accurate for the population in England and Wales, with the following exceptions:

- The description of service provision given by the company is mostly applicable to people with pMMR tumours. People with previously treated advanced or recurrent EC with MSI/dMMR may be responsive to immunotherapy monotherapy and can now be offered dostarlimab monotherapy (TA779).[3]
- Clinical expert advice to the ERG suggests that radiotherapy may sometimes be used in the advanced/recurrent setting for tumours not previously treated with radiotherapy (i.e. in the (neo)adjuvant setting).

2.4. Critique of company's definition of decision problem

The decision problem provided by the company (refer to B.1.1 in the company submission) is largely consistent with the NICE scope:

The company appropriately clarify that the target population are those who "have disease progression on or following prior treatment with a platinum-containing therapy" and that the population would be those who are "not candidates for curative surgery or radiation". The company also state that this can be in any setting, thereby opening up two positions for PEM+LEN: firstly, as a treatment option following platinum-containing chemotherapy provided in the advanced/recurrent setting; and secondly, for those with recurrent, advanced or metastatic cancer who had received platinum-based chemotherapy in the (neo)adjuvant setting. The ERG agree that these are appropriate treatment positions, but noted that for the latter positioning, re-challenge with platinum-containing doublet chemotherapy may be the first-choice treatment (for those receiving adjuvant platinum-based chemotherapy at least 12 months before), and would, therefore, be a useful comparator for this positioning. Whilst both carboplatin and doxorubicin are included as comparators, the key trial in the company submission does not use this doublet as a comparator.

The company has also narrowed the comparator in comparison with the NICE scope, with hormone therapy and best supportive care no longer considered. The ERG agreed that, even though best supportive care is not always limited to those not suitable for active treatment (occasionally people suitable for active treatment may choose best supportive care), the aims of this treatment differ from those of PEM+LEN and the exclusion of this comparator is, therefore, acceptable. The ERG also agree that hormone therapy is given with palliative intent in the recurrent/advanced population and it is reasonable to exclude such treatments as comparators. The ERG noted that the company listed cyclophosphamide as being a comparator in the NICE scope, whereas it was not listed in the NICE final scope document.[4]

The ERG agree that paclitaxel and doxorubicin are reasonable comparators and that both of these treatments are used in the UK setting and are considered to be equally effective. Following advice from clinical experts, the ERG noted that paclitaxel and doxorubicin are used in similar numbers of people with recurrent, advanced or metastatic endometrial cancer (rather than the preference towards doxorubicin in the KEYNOTE 775 data (see Section 3.2.2).

The ERG highlight that the treatment and population in the company's decision problem are aligned with the NICE scope. However, it is important to highlight that due to recent changes in the treatment pathway (those with EC displaying MSI-H/dMMR can now access dostarlimab monotherapy through the CDF (TA779)),[3] the treatment may bemore appropriate for people with pMMR EC than for those with dMMR EC. Clinical advice to the ERG indicated that immunohistochemistry was more accurate for identifying MMR status where available compared to MSI. Ideal comparators for PEM+LEN in this subgroup would, therefore, be immune checkpoint inhibitors as monotherapy.

The ERG noted that in certain cancers, use and licensing of PD-1/PD-L1 checkpoint inhibitor is conditioned on extent of PD-L1 expression. The company indicated that 'the treatment benefit of PEM+LEN compared with TPC was consistent across all the major subgroups tested in patients with advanced EC, including by histology' and explained that 'the regulatory license for this indication does not have a restriction based on the PD-L1 status'. Therefore the ERG did not consider PD-L1 expression any further. Clinical advice to the ERG indicated that MMR status is of much greater use in EC than PD-L1 expression.

Table 5: Summary of decision problem

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	People with advanced, metastatic or recurrent endometrial cancer, previously treated with platinum-based therapy who are not able to receive curative surgery or radiation	For the treatment of advanced or recurrent endometrial carcinoma in adults who have disease progression on or following prior treatment with a platinum-containing therapy in any setting and who are not candidates for curative surgery or radiation	Aligned to anticipated marketing authorisation	The ERG considered the decision problem addressed in the company submission was in alignment with the NICE scope. However, based on clinical expert opinion to the ERG, two clinically distinct subgroups exist within the overall population i.e. patients with dMMR and pMMR cancers. The company conducted subgroup analyses in these subgroups, however cost effectiveness results were not presented (See Section 4.2.3). The ERG noted that people with dMMR EC now have access to dostarlimab (TA779),[3] as monotherapy. Therefore, PEM+LEN may be most appropriately positioned for people with pMMR EC. Furthermore, whilst clinically appropriate, the positioning of PEM+LEN following platinum-based

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
				treatment in any setting (including following platinum-based chemotherapy in the (neo)adjuvant setting) creates questions about useful comparators (see below).
Intervention	Pembrolizumab with Lenvatinib	Pembrolizumab with lenvatinib (PEM+LEN)	N/A	The ERG agreed that the intervention is consistent with the NICE final scope.
Comparator(s)	Chemotherapy (including carboplatin and paclitaxel, paclitaxel monotherapy, doxorubicin monotherapy and carboplatin monotherapy Hormone therapy (such as medroxyprogesterone acetate and megestrol) Best supportive care	Chemotherapy (such as paclitaxel, carboplatin, doxorubicin)	Active comparators aligned with BGCS evidence-based recommendations and Company's consultation with clinical experts: • Cyclophosphamide is not used to treat advanced or recurrent EC. • Hormone therapy is only used if all other treatment options are exhausted or patients cannot tolerate further lines of chemotherapy and even then hormone therapy	The primary comparators were based on the physician's choice. This was assumed by the company to be doxorubicin or paclitaxel. The ERG considered these comparators to be reasonable (see Section 3.2.2.4 for further comment). However, for the dMMR subpopulation, the ideal comparators are likely to be immune checkpoint inhibitors as monotherapy. Nevertheless, the ERG acknowledge that such trials, using a population in England and Wales, would not be expected to be available, due to the recency of the availability

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
			has a palliative intent rather than an expectation of clinical response; this is not the target position for PEM+LEN • Best supportive care reserved for patients not fit for active treatment; this is not the target position for PEM+LEN	of dostarlimab through the CDF (TA779).[3] The ERG agreed with the exclusion of cyclophosphamide (but noted that this was not in the NICE final scope), best supportive care and hormone therapy as comparators. The ERG also noted that when PEM+LEN is positioned as first-line treatment in the advanced, metastatic or recurrent setting, a useful comparator is re-challenge with platinum-based chemotherapy (see p.62 and p.63).
Outcomes	Progression-free survival Overall survival Response rates Duration of response Adverse effects of treatment Health-related quality of life	As per the NICE final scope	N/A	The ERG agreed that the outcomes assessed and presented by the company were in line with the NICE final scope.
Economic analysis	The cost-effectiveness of treatments is expressed in terms of incremental cost per quality-adjusted life year	As per the NICE final scope	N/A	A cost utility analysis was provided by the company and results were presented as cost per QALY as appropriate. The time horizon used in

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	The time horizon for estimating cost-effectiveness was set at a lifetime horizon to sufficiently reflect any differences in costs or outcomes between the technologies being compared • Costs are considered from a NHS and Personal Social Services perspective • The availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account			the company's base case (40 years) was considered reasonable.
Subgroups	Not stated in final scope	No economic subgroup analyses were submitted by the company, which is consistent with the NICE final scope	N/A	PEM+LEN may be expected to perform better in people with dMMR EC but be most appropriately positioned for people with pMMR EC (those with dMMR EC now have access to dostarlimab as monotherapy).
Special considerations including issues related to equity or equality	None	None	N/A	The ERG did not identify any issues related to equity or equality.

Abbreviations: BGCS, British Gynaecological Cancer Society; dMMR, deficient mismatch repair; EC, endometrial cancer; ERG, Evidence Review Group; N/A, not applicable; NICE, National Institute for Health and Care Excellence; PEM+LEN, pembrolizumab with lenvatinib; pMMR, proficient mismatch repair; QALY, quality-adjusted life years

3. CLINICAL EFFECTIVENESS

The sections below discuss the evidence submitted by the company in support of the clinical effectiveness of pembrolizumab in combination with lenvatinib (PEM+LEN) for people with advanced, metastatic or recurrent endometrial cancer, previously treated with platinum-based therapy who are not able to receive curative surgery or radiation.

The ERG reviewed the details provided on:

- Methods implemented to identify, screen, extract data and assess the risk of bias in relevant evidence
- Clinical efficacy of PEM+LEN
- Safety profile of PEM+LEN
- Assessment of comparative clinical effectiveness of PEM+LEN against relevant comparators

A detailed description of an aspect of the CS is only provided where the ERG disagreed with the company's assessment or proposal, or where the ERG identified a particular area of concern that the ERG considered necessary to highlight for the Committee.

The ERG did not identify any clinical effectiveness key issues.

3.1. Critique of the methods of review(s)

The Company undertook two systematic literature reviews (SLRs) related to clinical effectiveness: an SLR of interventional evidence and one of real-world evidence (RWE).

3.1.1. Critique of the methods of the interventional evidence SLR

The SLR of interventional evidence was aimed at identifying randomised controlled trials (RCTs) or single-arm studies assessing the clinical effectiveness and safety of PEM+LEN, and comparator interventions, for recurrent or advanced cancer in people with disease progression on or following prior treatment with a platinum containing chemotherapy who were not candidates for curative surgery or radiation. The Company make clear in their inclusion criteria that their definition of advanced cancer is inclusive of stage IV metastatic disease (Appendix D.1.1.2, Table 5, in the Company submission), which is in line with the NICE scope.

The SLR of interventional evidence identified two relevant studies. One was a Phase III RCT (KEYNOTE-775),[5] relevant to the decision problem, and providing direct evidence on the clinical effectiveness and safety of PEM+LEN versus treatment of physician's choice (doxorubicin or paclitaxel monotherapy). A critique of the choice of comparator is in section 2.4. The other study was a single-arm Phase Ib/II study (KEYNOTE-146)[6] that was a precursor of KEYNOTE-775.

Overall, the ERG found this SLR to be of reasonable quality and likely to have identified all studies relevant to the Company's decision problem. A summary of the ERG's critique of the methods implemented in this SLR is presented in Table 6.

Table 6: Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix D.1.1.1, Tables 1-4	The searches of bibliographic databases are considered broadly appropriate, however, the ERG noted the following limitation: the Ovid Embase search strategy applied a filter excluding conference abstracts from search results. Database searches and manual searches of four conference proceedings may have mitigated this issue. The ERG conducted an additional search on Embase (reported in Appendix A) to check if any conference abstracts were missed by Company searches and is satisfied all relevant evidence has been identified.
Inclusion criteria	Appendix D.1.1.2, Table 5	The inclusion criteria were in line with the Company's decision problem.
Screening	Appendix D.1.1.2	Standard accepted methods
Data extraction	Appendix D.1.1.2.1, Table 6	Standard accepted methods
Tool for quality assessment of included study or studies	Appendix D.1.1.2.2	The Company state that RoB2 was used to assess KEYNOTE-775 (Appendix D.1.1.2.2 in the Company submission.). However, the described domains and summary assessments (low, unclear or high) do not correspond with RoB2. Following clarification, a

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		RoB2 assessment was provided to the ERG (see section 3.2.2.6). The company state that KEYNOTE-146 was assessed using the Newcastle-Ottawa Scale. This is an appropriate tool, but the ERG could not check this assessment because these results were not presented.
Evidence synthesis	Appendix D.1.4	The ERG agrees that NMA/unanchored ITC were not feasible

Abbreviations: CS, Company submission; ERG, Evidence Review Group; ITC, Indirect Treatment Comparison; NMA, network meta-analysis; RoB2, Cochrane Risk of Bias version 2

3.1.2. Critique of the methods of the real-world evidence SLR

The SLR of real-world evidence (RWE) had the objective of identifying observational and cross-sectional studies on clinical efficacy, safety, epidemiological burden and treatment for people with recurrent or advanced endometrial cancer. Overall, the ERG found this review to be of reasonable quality and the methods were likely to have identified the relevant observational and cross-sectional evidence available at the time the searches were conducted. However, the searches for this SLR were conducted in 2020 and are, therefore, out of date. The company were unable to provide updated results, pointing to the fact that the RWE does not form the primary basis of the clinical effectiveness evidence. Whilst the ERG agree that this is the case, the RWE is nevertheless important, and the lack of an up-to-date review risks bias in the choice of studies used to validate the data from KEYNOTE-775.

This SLR identified six retrospective cohort studies,[7-12] four providing data on re-challenge with platinum-based chemotherapy[9-12] and two relating to doxorubicin[7,8] and not considered further. It was not clearly stated why the doxorubicin studies were not used to validate data from the comparator arm of KEYNOTE-775, although presumably this was because of the composite nature of the comparator arm in KEYNOTE-775. Instead, an extra study was reported (the ECHO study,[13] Document B, Section B.2.9.3) and used to confirm/validate the survival data for the comparator arm in KEYNOTE-775. The ERG highlight that the ECHO study[13] was not identified through the SLR of RWE and the company have clarified that ECHO is a recently completed internal study and that the UK data have not yet been published. The inclusion of a

study identified through non-systematic methods introduces a risk of bias, particularly because the SLR of RWE is not up to date and could potentially have identified alternative relevant validation studies.

A summary of the ERG's critique of the methods implemented in this SLR is presented in Table 7.

Table 7: Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem (SLR of real world evidence)

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix D.1.2.1 and D.1.3.2, Tables 8-11	Searches were conducted in July 2020 and, therefore, it is not known if other relevant evidence has since become available. Searches did not include web searches for grey literature sources not included in bibliographic databases (e.g., UK cancer registries or reports derived from electronic health records).
Inclusion criteria	Appendix D.1.3.3, Table 12	The inclusion criteria were in line with the Company's decision problem. The inclusion criteria table in the CS (Appendix D.1.3.3, Table 12) states that subgroups of interest were "disease stage, line of therapy, treatment setting, risk factors for progression" but does not list MMR status as a subgroup of interest here. The ERG noted that subgroup data based on MMR status are of particular interest in this population, particularly with regards immunologic treatments (see section 2.4 for details).
Screening and selection	Appendix D.1.3.3 and D.1.3.6	An additional study (the ECHO study[13]) was described in Document B, Section B.2.9.3 and used to confirm/validate the survival data for the comparator arm in KEYNOTE-775. This is a recent study by the Company and was not

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		identified by the SLR of RWE, as it is not published.
Data extraction	Appendix D.1.3.5	Standard accepted methods
Tool for quality assessment of included study or studies	Appendix D.1.3.5	Studies were assessed using the ROBINS-I.[14] The ERG noted that the ROBINS-I is best suited to evaluating non-randomised comparative studies. The ERG could not check how the ROBINS-I was applied to the included retrospective cohort studies because the assessments were not provided.

Abbreviations: CS, Company submission; ERG, Evidence Review Group; MMR, mismatch repair; ROBINS-I, Risk Of Bias In Non-randomised Studies - of Interventions; RWE, real-world evidence; SLR, systematic literature review

3.2. Critique of trials of the technology of interest, the company's analysis and interpretation (and any standard meta-analyses of these)

3.2.1. Studies included in the clinical effectiveness review

The company presented evidence from one pivotal Phase III trial of PEM+LEN against physician's choice (typically doxorubicin or paclitaxel monotherapy) in KEYNOTE-775.[5,15] This trial informs the company's economic model. One supportive Phase 1b/II dose-finding trial of PEM+LEN, KEYNOTE-146,[6] was used to validate model extrapolations. Limited information about KEYNOTE-146 was included in the CS. The ERG asked the company at the clarification call if further information about this study was available. The company indicated that only limited information was available as KEYNOTE-146 was not conducted by the submitting company and that CSR or further methodological information was available. Subsequently, the ERG identified that a protocol had been published as an appendix to a published results paper from the study.[6] Therefore, the ERG used information from the published protocol to provide additional information regarding the study methods.

A summary of the clinical evidence included in the CS is presented in Table 8 .

Table 8: Clinical evidence included in the CS

Study name and acronym	Study design	Population	Intervention	Comparator	Study type
KEYNOTE-775	Multi-centre, open-label, randomised Phase III trial	People with advanced (including metastatic) or recurrent EC who have disease progression following prior systematic therapy with platinum chemotherapy, and are not candidates for curative surgery or radiation.	PEM+LEN	Physician's choice, typically doxorubicin or paclitaxel monotherapy	Phase III
KEYNOTE-146	Multi-centre, open-label, single-assignment Phase Ib/II basket trial	Phase Ib: people with selected tumour types who have progressed after treatment with approved therapies or for whom there are no standard effective therapies available. Phase II: people with metastatic selected solid tumour types who have received 0−2 prior lines of systemic therapy.	PEM+LEN	None	Phase Ib/II

Abbreviations: CS, company submission; EC, endometrial cancer; PEM+LEN, pembrolizumab with lenvatinib

3.2.2. Description and critique of the design of the studies

3.2.2.1. Design of the studies

The company's primary evidence for the combination of PEM+LEN comes from the KEYNOTE-775 study,[5,15,16] which was a global multi-centre, open-label, randomised Phase III trial of PEM+LEN against physician's choice, typically doxorubicin or paclitaxel in advanced or recurrent EC. This was the only trial of PEM+LEN that was used in the company economic model.

The study compromised a 28-day screening period followed by a period of treatment and finally a period of efficacy follow-up. Patients were enrolled using random assignment in a 1:1 ratio into one of two treatment arms: pembrolizumab 200 mg administered via IV every 3 weeks (Q3W) up to 35 cycles, plus lenvatinib 20 mg every day (QD); or treatment of physician's choice of either doxorubicin 60 mg/m2 Q3W up to a maximum cumulative dose of 500 mg/m2 or paclitaxel 80 mg/m2 every week (QW) on a 28 day cycle, 3 weeks on and 1 week off. The efficacy follow-up period was measured from the day after the end of treatment visit and continued for the duration of each patient's lifetime, or until the data cutoff date for the primary OS analysis if the participant was still alive.

A summary of the methodology of KEYNOTE-775 is provided in Table 9.

Table 9: Summary of KEYNOTE-775 trial methodology

Trial name	**KEYNOTE-775 (NCT03517449)**
Location	International, multi-centre trial with 167 sites across 21 countries, including nine sites in the United Kingdom (other sites were located in Argentina, Australia, Brazil, Canada, Colombia, France, Germany, Ireland, Israel, Italy, Japan, Korea, Mexico, New Zealand, Poland, Russia, Spain, Taiwan, Turkey and US)
Trial design	Multi-centre, randomised, open-label, Phase III study
Method of randomisation	Patients were randomised in a 1:1 ratio to receive PEM+LEN or TPC. Randomisation followed a predefined randomisation scheme based on the following stratification factors: MMR status (pMMR or dMMR) ECOG performance status (0 or 1) Geographic region (Region 1: Europe, USA, Canada, Australia, New Zealand, and Israel or Region 2: rest of the world) Prior history of pelvic radiation (yes or no)

Trial name	KEYNOTE-775 (NCT03517449)
	First, patients were stratified according to MMR status. Patients within the pMMR stratum were further stratified according to ECOG performance status, geographic region, and prior history of pelvic radiation. A total of 9 strata were used for the study.
Eligibility criteria for patients	Key inclusion criteria: Female patients who were ≥18 years of age Histologically confirmed EC Documented evidence of advanced, recurrent, or metastatic EC Radiographic evidence of disease progression after 1 prior systemic, platinum-based chemotherapy regimen for EC (participants may have received up to 1 additional line of platinum-based chemotherapy if given in the neoadjuvant or adjuvant treatment setting) [a] Provided a fresh or archival tumour sample for determination of MMR status Had at least 1 measurable target lesion according to RECIST 1.1, including a non-nodal target lesion ≥1 cm in the longest diameter and lymph node lesion that measured ≥1.5 cm in the short axis ECOG Performance Status of 0 or 1 within 7 days of starting treatment Adequately controlled blood pressure with or without antihypertensive medications (defined as ≤150/90 mm Hg at screening) Key exclusion criteria: Had carcinosarcoma (malignant mixed Müllerian tumour), endometrial leiomyosarcoma and endometrial stromal sarcomas Had central nervous system metastases, unless they have completed local therapy and have discontinued the use of corticosteroids for this indication for at least 4 weeks before starting treatment in this study Had gastrointestinal malabsorption, gastrointestinal anastomosis, or any other condition that might affect the absorption of lenvatinib Had a pre-existing Grade ≥3 gastrointestinal or non-gastrointestinal fistula Had significant cardiovascular impairment within 12 months of the first dose of study drug Full eligibility criteria are provided in Appendix N.1.
Trial drugs and method of administration	Intervention (n=411) Pembrolizumab (200 mg administered intravenously, every 3 weeks on Day 1 of a 21-day cycle; 35 doses maximum) plus lenvatinib (20 mg taken orally once daily) Comparator (n=416) Doxorubicin (60 mg/m2 administered intravenously, every 3 weeks on Day 1 of a 21-day cycle) or paclitaxel (80 mg/m2 administered intravenously, every week on Days 1, 8 and 15 of a 28-day cycle)[a] Participants continued to receive study treatment until disease progression was confirmed by BICR, development of unacceptable toxicity, withdrawal

Trial name	KEYNOTE-775 (NCT03517449)
	of consent, receipt of 35 administrations of pembrolizumab (approximately 2 years), or a lifetime cumulative dose of 500 mg/m² of doxorubicin Pembrolizumab and doxorubicin/paclitaxel were administered in the clinic by qualified site personnel, whilst lenvatinib was dispensed to patients for oral self-administration
Permitted and disallowed concomitant medication	Permitted concomitant medications: Hormone replacement therapy Thyroid hormone suppressive therapy Adjuvant hormonal therapy for history of definitively treated breast cancer Anticoagulants including low molecular weight heparin, warfarin, anti-Xa agents Anti-inflammatory agents Bisphosphonates or denosumab Antihypertensive therapy (including additional antihypertensive treatment as appropriate if blood pressure increases once the participant is enrolled) Palliative radiotherapy to non-target bone metastases or brain lesions may be permitted after consultation Disallowed concomitant medications: Concurrent anticancer therapies such as chemotherapy, targeted therapies, hormonal therapy directed at EC, radiotherapy, antitumour interventions, or cancer immunotherapy Other concurrent investigational drugs Live vaccines Systemic glucocorticoids for any purpose other than to modulate symptoms from an AR that is suspected to have immunologic aetiology
Primary endpoints	PFS, defined as the time from date of randomisation to the date of the first documentation of disease progression, as determined by BICR per RECIST 1.1, or death from any cause, whichever occurs first OS, defined as the time from date of randomisation to date of death from any cause
Key secondary endpoints	ORR, defined as the proportion of patients who have best overall response of either CR or PR, as determined by BICR per RECIST 1.1 HRQL, assessed using the global score of the EORTC QLQ-C30 The EORTC QLQ-C30 physical functioning score, EORTC QLQ EN24 urological symptoms score and the EQ-5D-5L VAS score were included as exploratory endpoints Incidence of treatment emergent AEs, SAEs, and immune-related AEs Proportion of patients discontinuing study treatment due to treatment emergent AEs Time to treatment failure due to toxicity, defined as the time from the date of randomisation to the date that a participant discontinues the study treatment due to treatment-emergent AEs[b]

Trial name	KEYNOTE-775 (NCT03517449)
	Plasma concentration vs. time, clearance and AUC for lenvatinib[b]
Subgroup analysis	Pre-specified subgroup analyses were performed in the all-comer population for PFS, OS and ORR. The subgroup analyses were conducted using the same methods described for the primary efficacy endpoints and were based on the following baseline demographic and disease characteristics: Age (<65, ≥65 years) Race (White, Asian, other) Region (Region 1, Region 2) MMR status (pMMR, dMMR) ECOG status (0, 1) Prior history of pelvic radiation (yes, no) Histology (endometrioid, non-endometrioid) Prior lines of therapy (1, 2, ≥3)

Key: AE: adverse event; AUC, area under the curve; BICR: Blinded Independent Central Review; CBR: clinical benefit rate; CR: complete response; CSR: clinical study report; DCR : disease control rate; dMMR: deficient mismatch repair; DOR: duration of response; ECOG : Eastern Cooperative Oncology Group; EORTC, European Organisation for the Research and Treatment of Cancer; HRQL: health-related quality of life; LVEF: Left ventricular ejection fraction; MMR: mismatch repair; ORR: overall response rate; OS: overall survival; PD: progressive disease; PEM+LEN, pembrolizumab with lenvatinib; PFS: progression free survival; PFS2: progression free survival on next line therapy; pMMR: proficient mismatch repair; PR: partial response; RECIST: Response Evaluation Criteria in Solid Tumours; SAE: serious adverse event; SD: standard deviation; TTR: time to response; VAS, visual analogue scale.

Notes: [a], There was no restriction regarding prior hormonal therapy; [b], These endpoints have not been presented as part of this submission but are available in the CSR.

Source: CS, Table 4, pp.18-21, based on KEYNOTE-775 Clinical Study Report[15]

The company also presented supplementary evidence for the clinical effectiveness of PEM+LEN from the KEYNOTE-146 study.[6], which was a single-arm phase Ib/II trial of PEM+LEN. While limited information on this study was available in the CS, a study protocol was available as an appendix to the study results paper.[6] The ERG requested further information on KEYNOTE-146 via NICE, but the provision of this information was refused, inhibiting a full critique of this study which informs the model validation.

3.2.2.2. Population

In KEYNOTE-775, eligible participants were adult females aged at least 18 years with documented evidence of advanced, recurrent or metastatic endometrial cancer with completely resected Stage IB (tumours at least 4 cm) to Stage IIIA, who had an ECOG performance status 0-1 and who were able to receive cisplatin-based chemotherapy. Patients with carcinosarcoma

or sarcoma were excluded. Detailed inclusion and exclusion criteria were provided in the CS (Appendix N.1). These overall appeared reasonably aligned with the NICE scope and company decision problem.

There were a total of 827 participants, of whom 411 were randomised to the PEM+LEN arm and 416 were randomised to the comparator arm. The study recruited from 167 sites across 21 countries globally. Nine of the study sites were located in the United Kingdom (UK), no specific breakdown was provided for England and Wales, the UK nations for which this appraisal is applicable. Clinical advice to the ERG indicated that treatment pathways are unlikely to differ substantially between countries, but that the trial profile may underestimate the age and weight of patients encountered in routine clinical practice in the UK.

Baseline characteristics for KEYNOTE-775 are provided below as Table 10.

Table 10: Baseline characteristics for KEYNOTE-775

Characteristic	PEM+LEN (n=411)	TPC (n=416)
Sex, n (%)		
Female	411 (100)	416 (100)
Age in years, n (%)		
<65	206 (50.1)	204 (49.0)
≥65	205 (49.9)	212 (51.0)
Mean (SD)	63.2 (9.1)	63.8 (9.2)
Median (min, max)	64.0 (30, 82)	65.0 (35, 86)
Race, n (%)		
Asian	85 (20.7)	92 (22.1)
Black or African American	17 (4.1)	14 (3.4)
White	261 (63.5)	246 (59.1)
Other	12 (2.7)	20 (4.8)
Age in years at initial diagnosis, n (%)		
<65	253 (61.6)	255 (61.3)
≥65	158 (38.4)	161 (38.7)
Mean (SD)	61.3 (9.1)	61.5 (9.3)
Median (min, max)	62.4 (30, 81)	62.1 (27, 84)
Region,[a] n (%)		
Region 1	234 (56.9)	240 (57.7)

Characteristic	PEM+LEN (n=411)	TPC (n=416)
Region 2	177 (43.1)	176 (42.3)
MMR Status, n (%)		
pMMR	346 (84.2)	351 (84.4)
dMMR	65 (15.8)	65 (15.6)
ECOG, n (%)		
0	246 (59.9)	241 (57.9)
1	164 (39.9)	175 (42.1)
3	1 (0.2)	0 (0.0)
Prior history of pelvic radiation, n (%)		
Yes	168 (40.9)	173 (41.6)
No	243 (59.1)	243 (58.4)
Elapsed time in years from initial diagnosis		
Mean (SD)	2.4 (2.4)	2.9 (2.8)
Median (min, max)	1.7 (0, 21)	2.1 (0, 26)
Histology of initial diagnosis, n (%)		
Clear cell carcinoma	30 (7.3)	17 (4.1)
Endometrioid carcinoma	83 (20.2)	103 (24.8)
Endometrioid carcinoma with squamous differentiation	7 (1.7)	7 (1.7)
High grade endometrioid carcinoma	94 (22.9)	90 (21.6)
High grade mucinous carcinoma	0 (0.0)	1 (0.2)
High grade serous carcinoma	65 (15.8)	65 (15.6)
Low grade endometrioid carcinoma	59 (14.4)	54 (13.0)
Low grade mucinous carcinoma	1 (0.2)	0 (0.0)
Mixed	22 (5.4)	16 (3.8)
Neuroendocrine	2 (0.5)	0 (0.0)
Serous carcinoma	38 (9.2)	50 (12.0)
Unclassified	0 (0.0)	3 (0.7)
Undifferentiated histology	4 (1.0)	3 (0.7)
Other	6 (1.5)	7 (1.7)
FIGO stage at initial diagnosis, n (%)		
I	10 (2.4)	11 (2.6)

Characteristic	PEM+LEN (n=411)	TPC (n=416)
IA	54 (13.1)	64 (15.4)
IB	47 (11.4)	64 (15.4)
II	32 (7.8)	26 (6.3)
III	5 (1.2)	8 (1.9)
IIIA	28 (6.8)	33 (7.9)
IIIB	11 (2.7)	11 (2.6)
IIIC	30 (7.3)	24 (5.8)
IIIC1	17 (4.1)	25 (6.0)
IIIC2	27 (6.6)	27 (6.5)
IV	27 (6.6)	26 (6.3)
IVA	7 (1.7)	8 (1.9)
IVB	116 (28.2)	89 (21.4)
Brain metastasis,[c] n (%)		
Yes	2 (0.5)	2 (0.5)
No	409 (99.5)	414 (99.5)
Bone metastasis,[c] n (%)		
Yes	39 (9.5)	33 (7.9)
No	372 (90.5)	383 (92.1)
Liver metastasis,[c] n (%)		
Yes	101 (24.6)	98 (23.6)
No	310 (75.4)	318 (76.4)
Lung metastasis,[c] n (%)		
Yes	164 (39.9)	152 (36.5)
No	247 (60.1)	264 (63.5)
Intra-abdominal metastasis,[b,c] n (%)		
Yes	164 (39.9)	166 (39.9)
No	247 (60.1)	250 (60.1)
Lymph node metastasis,[c] n (%)		
Yes	224 (54.5)	225 (54.1)
No	187 (45.5)	191 (45.9)

Key: PEM+LEN, pembrolizumab with lenvatinib; TPC: treatment of physician's choice.
Notes: [a], Region 1: Europe, USA, Canada, Australia, New Zealand, Israel; Region 2: Rest of World; [b], Includes reported locations of colon, abdominal cavity, omentum, small intestine, peritoneal cavity, and

Characteristic	PEM+LEN (n=411)	TPC (n=416)
peritoneum. Does not include lymph nodes or other organs; [c], Lesion location as determined by investigator review.		

Source: CS, Table 5, pp. 21-23, based on KEYNOTE-775 CSR.[15]

In KEYNOTE-146, the participant profile was subdivided between two phases. In Phase Ib, participants were people with selected tumour types who have progressed after treatment with approved therapies or for whom there are no standard effective therapies available. In Phase II, participants were people with metastatic selected solid tumour types who had received up to two prior lines of systemic therapy. It is stated that 6 separate cohorts were enrolled into Phase II based on tumour location. A total of 125 participants were enrolled, of whom one was excluded due to leiomyosarcoma. Of the remaining 124 participants, nine were first line and 115 were second line. KEYNOTE-146 included up to 25 study sites from the United States and the European Union. There were no UK sites in KEYNOTE-146, which may limit generalisability to a UK decision-making context. Baseline characteristics for KEYNOTE-146 are shown in Table 11 below.

Table 11: Baseline characteristics for KEYNOTE-146

	Previously treated EC[a]			**All EC**
Characteristic	**MSS/pMMR (n = 94)**	**MSI-H/dMMR (n =11)**	**Total[b] (n = 108)**	**(n = 124)**
Age, years				
Mean	65.4	62.4	65.1	65.3
SD	7.42	9.45	7.60	7.83
Race, n (%)				
White	81 (86.2)	9 (81.8)	93 (86.1)	108 (87.1)
Black or African American	6 (6.4)	0	6 (5.6)	7 (5.6)
Asian	4 (4.3)	1 (9.1)	5 (4.6)	5 (4.0)
American Indian or Alaskan native	1 (1.1)	0	1 (0.9)	1 (0.8)
Native Hawaiian or other pacific islander	0	1 (9.1)	1 (0.9)	1 (0.8)
Other	2 (2.1)	0	2 (1.9)	2 (1.6)
ECOG PS, n (%)				
0	49 (52.1)	1 (9.1)	53 (49.1)	62 (50.0)

	Previously treated EC[a]			**All EC**
Characteristic	**MSS/pMMR (n = 94)**	**MSI-H/dMMR (n =11)**	**Total[b] (n = 108)**	**(n = 124)**
1	45 (47.9)	10 (90.9)	55 (50.9)	62 (50.0)
Histologic subtype, n (%)				
Endometrioid adenocarcinoma	46 (48.9)	8 (72.7)	55 (50.9)	67 (54.0)
FIGO grade I	10 (10.6)	2 (18.2)	12 (11.1)	15 (12.1)
FIGO grade 2	15 (16.0)	4 (36.4)	19 (17.6)	22 (17.7)
FIGO grade 3	21 (22.3)	2 (18.2)	24 (22.2)	30 (24.2)
Serous adenocarcinoma	33 (35.1)	0	35 (32.4)	39 (31.5)
Clear-cell adenocarcinoma	5 (5.3)	1 (9.1)	6 (5.6)	6 (4.8)
Dedifferentiated/ undifferentiated carcinoma	0	1 (9.1)	1 (0.9)	1 (0.8)
Adenocarcinoma, not otherwise specified	1 (1.1)	0	1 (0.9)	1 (0.8)
Other	1 (0.8)	1 (9.1)	10 (9.3)	10 (8.1)
PD-L1 status, n (%)				
Positive	46 (48.9)	7 (63.6)	53 (49.1)	60 (48.4)
Negative	39 (41.5)	4 (36.4)	43 (39.8)	52 (41.9)
Not available	9 (9.6)	0	12 (11.1)	12 (9.7)
Prior treatment regiments for EC, n (%)				
0	0	0	0	9 (7.3)
1	48 (51.1)	7 (63.6)	57 (52.8)	60 (48.4)
2	36 (38.3)	3 (27.3)	40 (37.0)	43 (34.7)
≥3	10 (10.6)	1 (9.1)	11 (10.2)	12 (9.7)
Prior treatment, n (%)				
Bevacizumab	5 (5.3)	1 (9.1)	6 (5.6)	7 (5.6)
Platinum + taxane combination (with or without other anticancer medication	92 (97.9)	11 (100.0)	106 (98.1)	113 (91.1)
Other anticancer combinations	9 (9.6)	1 (9.1	11 (10.2)	12 (9.7)

	Previously treated EC[a]			All EC
Characteristic	MSS/pMMR (n = 94)	MSI-H/dMMR (n =11)	Total[b] (n = 108)	(n = 124)
Monotherapy	33 (35.1)	3 (27.3)	36 (33.3)	37 (29.8)
Prior history of/ current hypertension, n (%)				
Yes	60 (63.8)	9 (81.8)	71 (65.7)	79 (63.7)

Key: CPS: combined positive score; dMMR: deficient mismatch-repair; EC: endometrial carcinoma; ECOG PS: Eastern Cooperative Oncology Group performance status; FIGO: International Federation of Gynecology and Obstetrics; MSI-H: microsatellite instability high; MSI/MMR: microsatellite instability/mismatch repair; PD-L1: programmed death-ligand 1; pMMR: mismatch-repair proficient; SD: standard deviation.

Notes: [a], Enrolled before July 1, 2018; [b], Three patients had an unknown MSI/MMR tumour status; c, Predominantly mixed histology; d, PD-L1 status is positive if CPS is ≥ 1 and negative if CPS is < 1; e, The majority of patients received therapies in the adjuvant or metastatic setting; 9 patients received therapy in the neoadjuvant setting; the setting for 2 patients was unknown; f, Patients may be counted in multiple categories.

Source: CS Appendix O.2, Table 55, based on Makker et al.[6,17]

3.2.2.3. Intervention

The intervention in KEYNOTE-775[15,16] was pembrolizumab 200 mg administered via IV every 3 weeks (Q3W) up to 35 cycles, plus lenvatinib 20 mg every day (QD). Phase Ib of KEYNOTE-146[6] sought to determine the maximum tolerated dose. Dosing began at the full dose of both PEM+LEN due to well-established safety profiles and non-overlapping mechanisms of action, with lower dose levels being explored as necessary based on observed toxicity. Phase Ib began with Dose Level 1; lenvatinib 24 mg/day orally and pembrolizumab 200 mg every 3 weeks IV were administered to participants with selected solid tumors on a 21-day treatment cycle. Two dose de-escalation steps were included: Dose Level 2 (lenvatinib 20 mg/day orally + pembrolizumab 200 mg Q3W, IV) and Dose Level 3 (lenvatinib 14 mg/day orally + pembrolizumab 200 mg Q3W, IV). In Phase II, following confirmation of the maximum tolerated dose, treatment proceeded at that dose.

3.2.2.4. Comparator

The comparator arm in KEYNOTE-775[15,16] was treatment of physician's choice of either doxorubicin 60 mg/m2 Q3W up to a maximum cumulative dose of 500 mg/m2 or paclitaxel 80 mg/m2 every week (QW) on a 28-day cycle, 3 weeks on and 1 week off. There was no comparator arm in KEYNOTE-146.

3.2.2.5. Outcomes

The outcomes covered in the KEYNOTE-775[15,16] study were summarised in the CS section B.2.3.1.

The primary efficacy outcome measures for this study were:

- Progression-free survival, defined as the time from date of randomisation to the date of the first documentation of disease progression, as determined by BICR per RECIST 1.1, or death from any cause, whichever occurs first.
- Overall survival, defined as the time from date of randomisation to date of death from any cause.

The secondary efficacy outcome measures for this study were:

- Overall response rate, defined as the proportion of patients who have best overall response of either complete respond or partial response, as determined by BICR per RECIST 1.1.
- Health-related quality of life, assessed using the global score of the EORTC QLQ-C30.

Exploratory endpoints for this study were:

- Health-related quality of life, assessed using The EORTC QLQ-C30 physical functioning score, EORTC QLQ EN24 urological symptoms score and the EQ-5D-5L VAS score.

Safety outcome measures for this study were:

- Incidence of treatment emergent AEs, SAEs, and immune-related AEs
- Proportion of patients discontinuing study treatment due to treatment emergent AEs
- Time to treatment failure due to toxicity, defined as the time from the date of randomisation to the date that a participant discontinues the study treatment due to treatment-emergent AEs
- Plasma concentration vs. time, clearance and AUC for lenvatinib

The ERG considered that the outcomes presented in KEYNOTE-775[15,16] generally encompassed the outcomes from the NICE scope. Data were presented for duration of response, although it was not included in the list of outcomes in the company methods.

The ERG also noted that EQ-5D-5L was only an exploratory endpoint in KEYNOTE-775, despite being a key health-related quality of life outcome for health technology appraisals. EQ-5D-5L scores were mapped to EQ-5D-3L using the Van Hout algorithm.[18].

Information regarding outcomes is not presented in a clear list for KEYNOTE-146.[6] It seems that in addition to safety outcomes, the key effectiveness outcomes in this study were overall response rates and duration of response.

3.2.2.6. Critical appraisal of the design of the studies

Following clarification, the Company provided RoB2 assessments for KEYNOTE-775 (for PFS and OS).[16] The Company's broad RoB2 judgements are provided in Table 12, alongside ERG comments. The Company also provided the ERG with more detailed (item-by-item) RoB2 judgements, which the ERG mostly agreed with (minor disagreements did not alter the domain judgements provided in Table 12).

Table 12: Summary of the RoB2 assessments for KEYNOTE-775

Bias domain	Company RoB2 assessment		ERG Comment
	Makker et al (2022) PFS	Makker et al (2022) OS	
1. Bias arising from the randomisation process	Lower risk of bias	Lower risk of bias	Agree with domain judgements
2. Bias due to deviations from intended interventions	Some concerns	Some concerns	Low risk of bias (PFS and OS)
3. Bias due to missing outcome data	Lower risk of bias	Lower risk of bias	Agree with domain judgements
4. Bias in measurement of the outcome	Lower risk of bias	Lower risk of bias	Agree with domain judgements
5. Bias in selection of the reported result	Lower risk of bias	Lower risk of bias	Agree with domain judgements
Overall bias	Some concerns	Some concerns	Low risk of bias (PFS and OS)

Abbreviations: ERG, Evidence Review Group; RoB2, Cochrane Risk of Bias version 2; OS, overall survival; PFS, progression-free survival

The ERG agreed with the Company that the primary risk of bias in KEYNOTE-775[16] was from the open-label study design and the resultant lack of blinding of those delivering and undergoing treatment. The ERG agreed that there was no evidence to suggest that this lack of blinding led

to trial-contextual issues that would impact upon the delivery of the interventions (item 2.3 in the item-by-item RoB2 assessment supplied to the ERG was rated as 'probable no') or have a substantial impact on OS or PFS. However, according to the RoB2 judgement algorithm, this should lead to a domain 2 (and therefore overall bias) judgement of 'low' rather than 'some concerns' (see Table 12). However, if a 'no information' judgement had been given to item 2.3 (this was not the case, but would have been reasonable), domain 2 and overall bias judgements of 'some concerns' would be appropriate.

The company did not provide Newcastle-Ottawa scale assessment results for KEYNOTE-146.[6] The ERG was therefore unable to comment on the Company's risk of bias assessment of this study.

3.2.3. Description and critique of the results of the studies

3.2.3.1. Clinical effectiveness results

Progression-free survival

In KEYNOTE-775, median PFS was significantly improved with PEM+LEN compared with TPC; 7.2 and 3.8 months respectively, HR 0.56 (95% CI 0.47, 0.66; p< 0.0001). For KEYNOTE-146, median PFS was reported in the CS in (Table 11). For the PEM+LEN arm this was 7.4 months.

Overall survival

In KEYNOTE-775, Median OS was significantly longer in the PEM+LEN group compared with the control group; 18.3 and 11.4 months respectively, HR 0.62 (95% CI 0.51, 0.75; p< 0.0001) at interim assessment time point 1. For KEYNOTE-146, median OS was reported in the CS (Table 11). For the PEM+LEN arm this was 17.7 months.

Response rates

In KEYNOTE-775, overall response rate was 31.9% (95% CI 27.4, 36.6) in the PEM+LEN group compared to 14.7% (95% CI 11.4, 18.4) in the control arm, with an estimated difference of 17.2% (95% CI 11.5, 22.9%, p<0.0001). In KEYNOTE-146, overall response rate was 39.8% for pre-treated endometrial cancer patients.

Duration of response

Among patients achieving a response, the median duration of response was 14.4 months (range: 1.6, 23.7) in the PEM+LEN group compared to 5.7 months (range: 0.0, 24.2) for the

control group. In KEYNOTE-146, median duration of response was 22.9 months (95% CI: 10.2, not estimable) for pre-treated endometrial cancer participants.

Health-related quality of life

As described above (Section 3.2.2.5), quality of life in KEYNOTE-775 was assessed using several different scores. The scores from the EQ-5D-5L visual analogue scale results are presented here, as they informed the company's economic model, though in the trial this was only an exploratory endpoint. On this measure, both groups improved significantly over the 12-week follow-up period (PEM+LEN mean change -4.44, 95% CI -6.43, -2.46; control mean change -6.79, 95% CI -8.98, -4.60). However, there was no statistically significant difference between the two arms in terms of the extent of improvement over the 12-week period (difference in least squares mean change from baseline 2.35, 95% CI -0.44, 5,14, [redacted]). No health-related quality of life data were presented from KEYNOTE-146 in the CS, as this outcome was not assessed in this trial.

Subgroup analyses

Subgroup analyses were presented in the CS examining the differential effectiveness of PEM+LEN by age, race, region, MMR status, ECOG performance status, prior history of pelvic radiation, histology (endometrioid vs non-endometrioid) and prior lines of therapy (CS, Appendix E). Those considered by the ERG to be of greatest importance were MMR status and region. The region subgroup analysis divided the world into two regions: Region 1 being Europe, USA, Canada, Australia, New Zealand, and Israel; Region 2 being the rest of the world. There was no significant difference in OS or PFS between Regions 1 and 2 (CS Appendix E, Figure 5). The grouping is fairly broad and includes heterogeneous health systems and is therefore limited in its applicability to assessing the generalizability of the findings to a UK decision-making context.

There was a statistically significant difference in favour of PEM+LEN on both PFS and OS for both the pMMR and dMMR subgroups. However, the effect in favour of PEM+LEN was stronger in the dMMR subgroup for both PFS (dMMR HR 0.36, 95% CI 0.23-0.57; pMMR HR 0.60, 95% CI 0.50-0.72) and OS (dMMR HR 0.37, 95% CI 0.22-0.62; pMMR HR 0.68, 95% CI 0.56, 0.84). The trial was not powered specifically to explore differences between sub-groups, so these findings should be regarded as exploratory.

These findings were also evidence in median survival times, measured in months. Median PFS was 6.6 (95% CI 5.6, 7.4) months in the PEM+LEN group and 3.8 (95% CI 3.6, 5.0) months in

the control arm in the pMMR population and 10.7 (5.6, NR) months in the PEM+LEN group and 3.7 (3.1, 4.4) months in the control group in the dMMR population. Median OS was 17.4 (95% CI 14.2, 19.9) months in the PEM+LEN group and 12.0 (95% CI 10.8, 13.3) in the control group in the pMMR population and not reached in the PEM+LEN group and 8.6 (95% CI 5.5, 12.9) months in the control group in the dMMR population.

Finally, differences were also in evidence in survival proportions. Six-month OS was 82.9 (78.5, 86.5) in the PEM+LEN group and 77.9 (73.1, 81.9) in the control group in the pMMR population and 80.0 (68.1, 87.9) in the PEM+LEN group and 61.7 (48.5, 72.5) in the control group in the dMMR population. Twelve-month OS was 27.6 (22.5, 32.8) in the PEM+LEN group and 13.1 (8.9, 18.3) in the control group in the pMMR population and 67.2 (54.2, 77.2) in the PEM+LEN group and 39.1 (26.7, 51.3) in the control group in the dMMR population.

Adverse effects

Adverse events from KEYNOTE-775 were reported in section B.2.10 and Table 54, Appendix R.3 of the CS. These data were supplemented with adverse events data from KEYNOTE-146 (CS Appendix F).

Adverse effects: KEYNOTE-775

Adverse events in KEYNOTE-775 were provided for the safety population (n=406 for PEM+LEN; n=388 for TPC). Incidences of Grade 3 to 5 adverse events (AEs; 88.9% versus 72.7%), Grade 3 to 5 drug-related AEs (77.8% versus 59.0%), serious adverse events (SAEs; 52.7% versus 30.4%) and drug-related SAEs (52.7% versus 30.4%) were higher for treatment with PEM+LEN group compared with TPC. Likewise, dose interruptions (69.2% versus 27.1%), reductions (66.5% versus 12.9%) and discontinuations (33.0% versus 8%) due to AEs occurred more in the PEN+LEN arm than the TPC arm (CS B.2.10.1.2, Table 14), and discontinuations due to AEs were higher for lenvatinib (30.8%) than for pembrolizumab (18.7%).

Duration of exposure was longer in the PEM+LEN arm than the TPC arm; median (range) duration of exposure in days was 231.0 (1.0-817.0) for PEM+LEN and 104.5 (1.0-785.0) for TPC (see CS B.2.10.1.1, Table 13). Following adjustment for duration of exposure (see Table 13), lower rates of Grade 3 to 5 AEs and deaths were evident in the PEM+LEN arm compared with the TPC arm and SAEs were more similar between the two groups. Dose modifications, interruptions and reductions due to AEs remained higher in the PEM+LEN arm (Table 13).

There were more interruptions and discontinuations of treatment with LEN than PEM (see Table 15, section B.2.10.1.2 in the CS).

Table 13: Exposure-adjusted adverse event summary (KEYNOTE-775)

AE, event count rate (events/100 person-months)	PEM+LEN (n=406)	TPC (n=388)
Total exposure in person-months	3919.48	1765.17
One or more AE	9091 (231.94)	4526 (256.41)
No AE	1 (0.03)	2 (0.11)
Drug-related AEs	5221 (133.21)	2703 (153.13)
Toxicity grade 3-5 AEs	1216 (31.02)	861 (48.78)
Toxicity grade 3-5 drug-related AE	726 (18.52)	609 (34.50)
SAEs	398 (10.15)	178 (10.08)
Treatment-related SAEs	202 (5.15)	72 (4.08)
Dose modification due to and AE	1486 (37.91)	328 (18.58)
Dose interruption due to an AE	830 (21.18)	203 (11.50)
Dose reduction due to AE	594 (15.16)	84 (4.76)
Deaths	23 (0.59)	19 (1.08)
Deaths due to AEs	6 (0.15)	8 (0.45)
Discontinuations due to AEs	196 (5.00)	41 (2.32)
Discontinuation due to treatment-related AEs	156 (3.98)	31 (1.76)
Discontinuation due to SAE	95 (2.42)	15 (0.85)
Discontinuation due to a treatment-related SAE	64 (1.63)	8 (0.45)

Abbrevations: AE, adverse event; PEM+LEN, pembrolizumab with lenvatinib; SAE, serious adverse event TPC, treatment of physician's choice. Source: adapted from CS B.2.10.1.2, Table 15

The ERG agree that the specific adverse events from KEYNOTE-775 were consistent with what would be expected for the study treatments (see CS B.2.10.1.3, Table 16). Serious adverse events (SAEs) were also higher with PEM+LEN than with TPC (52.7% vs 30.4%; see CS B.2.10.1.7, Table 22).

The Company state that after adjusting for duration of exposure, most specific adverse events were lower or similar in the PEM+LEN arm compared with the TPC arm. The ERG noted that whilst this statement is not incorrect, there were some adverse events that remained higher with PEM+LEN than with TPC, including endocrine disorders, diarrhoea, decreased weight and appetite, hypertension and musculoskeletal and connective tissue disorders (CS Appendix R.3, Table 54).

AEs of special interest were provided in Tables 17 and 18, section B.2.10.1.3 of the CS, with the most common AEs of special interest in the PEM+LEN arm being hypothyroidism (57.6%), hyperthyroidism (11.6%), colitis (4.7%), skin reactions (3.2%) and infusion reactions (3.0%).

Treatment-related adverse events (TRAEs) are shown in Table 14, with hypertension, hypothyroidism, diarrhoea, nausea, and decreased appetite all reported in ≥ 30% of the PEM+LEN arm and nausea, anaemia, neutropenia and alopecia reported in ≥ 30% of the TPC arm. Grade 3 to 5 TRAEs (77.8% vs. 59.0%) and treatment-related SAEs (33.3% vs. 14.2%) were higher in the PEM+LEN than the TPC arm. There was a higher incidence of discontinuations due to TRAEs in the PEM+LEN compared with the TPC arm; discontinuations due to TRAEs were higher for lenvatinib than pembrolizumab (22.7% vs. 9.9%).

Table 14: Summary of treatment-related AEs (incidence ≥ 10% in one or more arms) in the KEYNOTE-775 trial (safety all-comer population)

AE, n (%)	PEM+LEN (n=406)	TPC (n=388)
One or more treatment-related AE	395 (97.3)	364 (93.8)
Hypertension	248 (61.1)	4 (1.0)
Hypothyroidism	221 (54.4)	0 (0.0)
Diarrhoea	171 (42.1)	42 (10.8)
Nausea	158 (38.9)	157 (40.5)
Decreased appetite	149 (36.7)	64 (16.5)
Fatigue	113 (27.8)	92 (23.7)
Proteinuria	102 (25.1)	4 (1.0)
Vomiting	99 (24.4)	59 (15.2)
Weight decreased	90 (22.2)	7 (1.8)
Arthralgia	84 (20.7)	17 (4.4)
Palmar-plantar erythrodysaesthesia syndrome	84 (20.7)	3 (0.8)

AE, n (%)	PEM+LEN (n=406)	TPC (n=388)
Dysphonia	76 (18.7)	2 (0.5)
Asthenia	75 (18.5)	76 (19.6)
Stomatitis	70 (17.2)	46 (11.9)
Alanine aminotransferase increased	63 (15.5)	14 (3.6)
Anaemia	58 (14.3)	150 (38.7)
Aspartate aminotransferase increased	58 (14.3)	12 (3.1)
Myalgia	54 (13.3)	13 (3.4)
Headache	53 (13.1)	14 (3.6)
Rash	47 (11.6)	6 (1.5)
Mucosal inflammation	45 (11.1)	35 (9.0)
Platelet count decreased	43 (10.6)	20 (5.2)
Blood thyroid stimulating hormone increased	40 (9.9)	1 (0.3)
Hyperthyroidism	39 (9.6)	1 (0.3)
Hypomagnesaemia	38 (9.4)	12 (3.1)
Constipation	36 (8.9)	51 (13.1)
Dry mouth	33 (8.1)	9 (2.3)
Dysgeusia	32 (7.9)	26 (6.7)
Lipase increased	32 (7.9)	2 (0.5)
Thrombocytopenia	31 (7.6)	22 (5.7)
Abdominal pain	30 (7.4)	13 (3.4)
Abdominal pain upper	28 (6.9)	28 (6.9)
Pruritus	27 (6.7)	7 (1.8)
Blood alkaline phosphatase increased	26 (6.4)	5 (1.3)
Pyrexia	26 (6.4)	26 (6.4)
Epistaxis	25 (6.2)	7 (1.8)
Hypertriglyceridaemia	24 (5.9)	1 (0.3)
Neutropenia	22 (5.4)	127 (32.7)
Blood creatinine increased	21 (5.2)	2 (0.5)
Leukopenia	20 (4.9)	47 (12.1)
Alopecia	17 (4.2)	117 (30.2)
Neutrophil count decreased	17 (4.2)	93 (24.0)
Lymphopenia	15 (3.7)	26 (6.7)

AE, n (%)	PEM+LEN (n=406)	TPC (n=388)
White blood cell count decreased	15 (3.7)	58 (14.9)
Lymphocyte count decreased	10 (2.5)	22 (5.7)
Neuropathy peripheral	8 (2.0)	21 (5.4)
Febrile neutropenia	1 (0.2)	21 (5.4)

Abbreviations: AE: adverse event; PEM+LEN, pembrolizumab with lenvatinib; TPC: treatment of physician's choice.
Source: CS B.2.10.1.4, Table 19

The most frequently reported Grade 3 to 5 TRAEs were hypertension, decreased weight, decreased appetite, and diarrhoea in the PEM+LEN arm and neutropenia, decreased neutrophil count, anaemia, decreased white blood cell count, leukopenia, and febrile neutropenia in the TPC arm (see CS, B.2.10.1.6, Table 21). The most frequently reported treatment-related SAEs (incidence ≥1%) for PEM+LEN were hypertension, colitis, decreased appetite, vomiting, diarrhoea, pyrexia, and acute kidney injury and the most frequently reported treatment-related SAEs for TPC were febrile neutropenia, neutropenia, and anaemia.

The Company state that deaths due to AEs were similar in the two trial arms and the ERG agrees with this: there were six deaths due to AEs in the PEM+LEN arm and eight in the TPC arm. The six deaths related to AEs in the PEM+LEN arm were considered to be treatment-related: one death was considered to be related to both pembrolizumab and lenvatinib (due to multiorgan dysfunction syndrome), three deaths were considered to be related to lenvatinib (one each due to cerebrovascular accident, right ventricular dysfunction and myelodysplastic syndrome), and one death was considered to be related to pembrolizumab (due to colitis). The eight deaths related to AEs in the TPC arm were considered to be related to doxorubicin (two due to pneumonia, and one each due to aspiration, pulmonary embolism, cardiogenic shock, toxic cardiomyopathy, cardiac failure, and sepsis).

Adverse events in KEYNOTE-146

Safety data from KEYNOTE-146 were provided in Appendix F of the CS (see CS Appendix F, Table 18 for a summary of all AEs up until Jan 10 2019 and CS Appendix F, Table 20 for a summary of all AEs up until Aug 18 2020). Table 15 provides a summary of TRAEs from this study. The ERG agrees that the data presented in Appendix F of the CS were broadly consistent with the safety data from KEYNOTE-775. The Company did not provide the CSR for

KEYNOTE-146. The ERG was unable, therefore, to verify the safety data provided for this study against the CSR.

Table 15: Overview of treatment-related adverse events in the KEYNOTE-146 trial (August 18, 2020)

Parameter, n (%)	Previously treated EC[a] (n = 108)
Patients with any treatment related AEs	104 (96.3)[a]
Patients with treatment related AEs leading to study-drug discontinuation[b]	23 (21.3)
Both lenvatinib and pembrolizumab	9 (8.3)
Lenvatinib[c]	19 (17.6)
Pembrolizumab[d]	17 (15.7)
Patients with treatment related AEs leading to study-drug dose reduction of lenvatinib	73 (67.6)
Patients with treatment related AEs leading to study-drug interruption[b]	80 (74.1)
Both lenvatinib and pembrolizumab	34 (31.5)
Lenvatinib[c]	77 (71.3)
Pembrolizumab[d]	47 (43.5)

Abbreviations: AE: adverse event; EC: endometrial cancer.

Notes: a, 94 (87.0%) and 10 (9.3%) patients experienced Grade ≤3 and Grade ≥4 treatment related AEs, respectively; b, Drug action taken is for lenvatinib or/and pembrolizumab; c, Drug action taken for lenvatinib, regardless of action taken for pembrolizumab; d, Drug action taken for pembrolizumab regardless of action taken for Lenvatinib.

Source: CS Apppendix F, Table 19

3.3. Additional clinical evidence submitted

No indirect treatment comparison or standard meta-analyses were presented. The ERG considered these decisions to be generally appropriate in light of the presence of relevant head-to-head data. The ERG agreed that it was not feasible to conduct a network meta-analysis due to the lack of connecting nodes. The ERG considered it could have been feasible to construct MAIC(s) between PEM+LEN and those comparators not trialled head to head e.g. paclitaxel monotherapy, but also noted significant uncertainty and limitations associated with bringing together data from a wide range of sources.

In considering the feasibility of conducting an indirect treatment comparison, the company discussed the Endometrial Cancer Health Outcomes – Europe (ECHO) study.[13] This is unpublished and as such was not identified through the SLR nor its methods and results included in the main clinical effectiveness section. This was a retrospective, multicentre chart review real-world evidence study evaluating treatment patterns and clinical outcomes in advanced or recurrent EC patients previously treated with systemic therapy. Data from the UK cohort were presented. This comprised [redacted] eligible patients aged at least 18 years at the time of advanced or recurrent endometrial cancer diagnosis, who were not considered a candidate for curative-intent surgery, did not participate in any other endometrial cancer-related clinical trials during treatment and who had a known medical history from the date of advanced or recurrent endometrial cancer diagnosis. Eligible patients also did not have any prior malignancy active within the past three years, except from locally curable cancers that had been cured.

The majority of patients [redacted] were treated with [redacted] with the remainder [redacted] receiving [redacted] The data were re-weighted to exclude treatments considered to be investigative. Clinical advice to the ERG indicated that the proportion of patients with clear cell histology ([redacted] data provided in clarification response) was far in excess of what would be expected in a UK setting. The median OS from the start of second-line systemic therapy was [redacted] The company noted that the survival outcome in ECHO was comparable to the control group in the pivotal KEYNOTE-775 trial. The ERG noted that there was very limited methodological information available about the ECHO study.[13]

3.4. Additional work on clinical effectiveness undertaken by the ERG

The ERG undertook additional searches (see Appendix A) to identify any evidence that had not been identified by the company. The ERG identified a small number of conference abstracts that had not been included by the company. However, the company did not provide a full list of excluded articles from the full-text screen of the interventional SLR, so the ERG was unable to comment on whether the company identified and excluded these abstracts or did not identify them. The ERG considered that the additional abstracts, while potentially eligible for the interventional SLR, did not provide additional data that would enhance the already identified clinical effectiveness evidence base.

3.5. Conclusions of the clinical effectiveness section

The ERG considered that the company's SLRs had identified the relevant evidence related to PEM+LEN and key comparators, except that the ECHO study on comparator treatments could not be identified through the SLR as it was unpublished. The ERG considered that the pivotal KEYNOTE-775 trial covered the relevant outcomes in the NICE final scope.[4] The ERG considered that generally the company's SLR and included trial were adequately described, although certain information was not described in sufficient detail. The ERG considered that the KEYNOTE-146 study which served as supplementary clinical evidence for model validation purposes was not well described in the CS, but further information was available through a published protocol identified by the ERG. The ERG requested further information on KEYNOTE-146 via NICE, but the provision of this information was refused, inhibiting a full critique of this study which informs the model validation. The ERG also considered that the unpublished ECHO study,[13] which also served for model validation purposes, was not described in adequate detail.

There was one pivotal clinical trial that informed the base case economic model – KEYNOTE-775. This was a multi-centre, open-label, randomised Phase III trial comparing PEM+LEN with treatment of physician's choice (paclitaxel or doxorubicin) for people with advanced, metastatic or recurrent endometrial cancer, previously treated with platinum-based therapy who are not able to receive curative surgery or radiation. The ERG was satisfied that all relevant studies were identified and that the pivotal KEYNOTE-775 trial was generally of high quality and well reported. The ERG was satisfied that the company's decision to not conduct an ITC was appropriate given the existence of a suitable directly comparative trial.

The ERG was satisfied that there was evidence for a statistically significant benefit in the KEYNOTE-775 trial for both OS and PFS for patients on PEM+LEN compared to patients on physician's choice of doxorubicin or paclitaxel. In the subgroup results, the ERG noted that the benefit of PEM+LEN, while statistically significant in both the pMMR and dMMR subgroups, was consistently greater in the dMMR subgroup.

The ERG considered that there were no key issues in the clinical effectiveness evidence base.

4. COST-EFFECTIVENESS

4.1. ERG comment on company's review of cost-effectiveness evidence

The company conducted SLRs to identify existing cost-effectiveness evidence, health related quality of life (HRQoL) evidence, and cost and resource use evidence of PEM+LEN and comparator treatments, in adult patients (aged 18 years and older) with endometrial cancer limited to recurrent (Stage I and II), Stage III/IV, metastatic, irrespective of line of therapy.

In Appendix G, the company stated that an initial search was conducted on May 6, 2019, which included studies relevant to advanced/metastatic (stage III and IV) endometrial cancer between 1999 and 2019. An updated search was conducted on January 6 and November 8, 2021 which expanded the inclusion criteria to include recurrent early stage (stage I and II) endometrial cancer patients in addition to advanced/metastatic patients.

Table 16: Summary of ERG's critique of the methods implemented by the company to identify cost-effectiveness evidence

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix G.1.1 and Tables 25, 28.	The searches of bibliographic databases and grey literature sources are considered broadly appropriate.
Inclusion criteria	Appendix G.1.2 and Table 21	The company excluded studies prior to 1999, studies reporting clinical data only, simple costing studies, those studies that did not report model outputs and studies which included non-pharmaceutical interventions. The ERG considered the company's inclusion criteria to be broadly reasonable.
Screening	Appendix G.1.4	Studies (titles and abstracts) were independently assessed by two reviewers using the basic selection criteria. Eligible studies were screened at full text stage by two independent reviewers and any discrepancies were reconciled by a 3rd independent reviewer. The ERG considered the company's screening methods to be broadly reasonable.
Data extraction	Appendix G.1.4	The company state that data extraction was conducted systematically based on a predefined data extraction template in

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
		line with standards required for HTA purposes. This appeared reasonable.
QA of included studies	Appendix G.1.5	QA was completed using the Drummond checklist, as recommended by NICE. The ERG considers the QA to be appropriate.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; QA, quality assessment

Table 17: Summary of ERG's critique of the methods implemented by the company to identify health related quality of life

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix H.1.1 and Appendix G.1.1	The searches of bibliographic databases and grey literature sources are considered broadly appropriate.
Inclusion criteria	Appendix H.1.2 and Table 30	The inclusion criteria as outlined in Table 30 were considered to be mostly reasonable. The company stated that studies reporting HRQoL values only were excluded i.e. those reporting HRQoL scores without utility or disutility values.
Screening	Appendix H.1.4	The company stated that '*the same selection process as described in Appendix G.1.3 was used for the SLR conducted for utilities.*' The ERG assumed that the company used the same screening strategy as per G.1.4, which is considered appropriate.
Data extraction	Appendix H.1.4	It appeared that the company used the same data extraction approach as per G.1.4, which is considered appropriate.
QA of included studies	Appendix H.1.5	QA was completed using the Drummond checklist, as recommended by NICE. The ERG considers the QA to be appropriate.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; QA, quality assessment

Table 18: Summary of ERG's critique of the methods implemented by the company to identify healthcare resource use and costs

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Searches	Appendix I.1 and Table 37	The searches of bibliographic databases and grey literature sources are considered broadly appropriate.
Inclusion criteria	Appendix I and Table 35	The search for cost and healthcare resource use studies was restricted to those in a US and UK setting. The ERG considered this restriction to be reasonable.
Screening	Appendix I and I.1	The same independent two reviewer screening approach appeared to have been used for costs and healthcare resource use studies, as for economic evaluations. The ERG considered this to be reasonable.
Data extraction	Appendix I.1	The company identified 3 studies which were considered generalisable to the UK, however these were not used in the appraisal, as reporting of data were considered too limited. See section 4.2.9 for further commentary on the sources and modelled inputs used by the company for costs and healthcare resource use.
QA of included studies	Appendix I.1	Not mentioned by the company.

Abbreviations: CS, Company Submission; ERG, Evidence Review Group; HRQoL, health-related quality of life; QA, quality assessment

4.2. Summary and critique of company's submitted economic evaluation by the ERG

4.2.1. NICE reference case checklist

Table 19: NICE reference case checklist

Attribute	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	QALYs were used as appropriate, which captured the health benefit to patients. The company did not include carer disutility.

Attribute	Reference case	ERG comment on company's submission
Perspective on costs	NHS and PSS	NHS and PSS as appropriate.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	The company submitted a cost utility analysis and presented pairwise results.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	A 40-year time horizon was used in the company's base case which was considered to be a lifetime horizon. The ERG considered this to be appropriate.
Synthesis of evidence on health effects	Based on systematic review	Clinical data used in the economic model was derived from the pivotal KEYNOTE-775 study. The ERG noted that median OS and PFS were reached. KEYNOTE-775 data were used to estimate modelled OS and PFS outcomes for both the intervention arm (PEM+LEN) and the comparator treatment arm (doxorubicin or paclitaxel) treatment. Information from KEYNOTE-146 and ECHO was provided as suporting data by the company to validate OS.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs were used as appropriate.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	QoL data were captured directly from patients in the KEYNOTE-775 study using the EQ-5D-5L. These values were then mapped to EQ-5D-3L values.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Utility values were estimated according to time to death. The company used a linear mixed effects regression model which was fitted to HRQoL data from KEYNOTE-775.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	There were no equity concerns.

Attribute	Reference case	ERG comment on company's submission
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Resource use and costs were based on NHS reference costs and the 2019/20 PSSRU, as appropriate. The company also used prior NICE appraisals for ovarian, cervical and uterine cancers to estimate resource frequency.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Costs and QALYs were discounted at 3.5% as appropriate. However estimates of life years were discounted at 0%. The ERG did not consider this to be appropriate.

Abbreviations: EQ-5D-3L, EuroQol 5 dimension 3 level; EQ-5D-5L, EuroQol 5 dimension 5 level; ERG, Evidence Review Group; HRQoL, health-related quality of life; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival; PSS, Personal Social Services; PSSRU, Personal Social Services Research Unit; QALY, quality-adjusted life year; TA, technology appraisal

4.2.2. Model structure

The model is a partitioned survival model (PSM; CS document B, section B.3.2.2) which is a common structure for modelling late stage cancer. Patients are defined as residing in one of three health states: progression-free (PF), progressed (PD) or dead, and the cycle length is 1 week. The advantage of partitioned survival models is that they are relatively simple and straightforward to implement, based on extrapolations of overall and progression-free survival curves from a clinical trial. Overall, the ERG considered the model structure to be appropriate for decision making.

As a general note, the key disadvantage of these models is that most implementations tend to draw on only one source of evidence (typically the key Phase III study for the product in question) for both baseline prognosis and treatment effect. NICE guidelines state that evidence on outcomes should be obtained from systematic review and meta-analysis provided there are sufficient relevant and valid data (NICE 2013; sections 5.2.2 and 5.2.8). As such it would have been preferable for the company to make use of a meta-analysis of appropriate data for baseline prognosis and/or treatment effect rather than the single KEYNOTE-775 study.

4.2.3. Population

The patient population within the company's economic analysis is adults who have disease progression on or following prior treatment with a platinum-containing therapy in any setting and

who are not candidates for curative surgery or radiation. The ERG noted several points of uncertainty surrounding company's positioning and population.

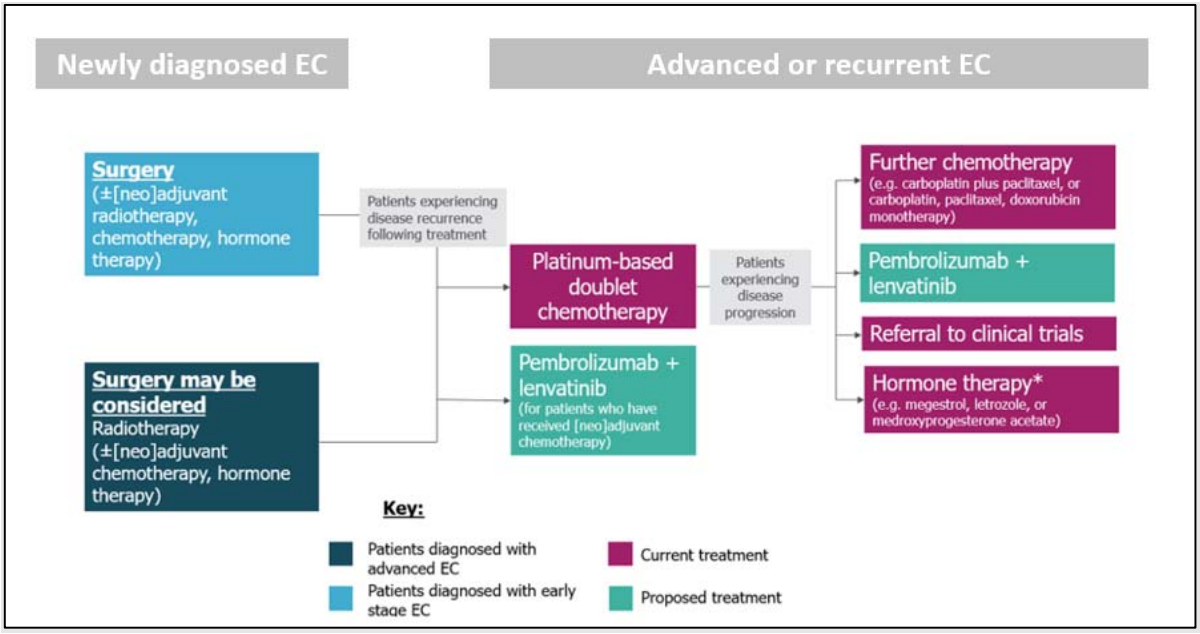
4.2.3.1. Modelled patient baseline characteristics

Modelled patient characteristics were based on patients from KEYNOTE-775. Within this pivotal study, average patient weight was 70.5 kg and median patient age was 63.5 years. The ERG noted KEYNOTE-775 was a multi-centre study, therefore patient characteristics used in the model were not specifically from a UK cohort. Due to generalisability concerns, the ERG asked clinical experts to comment on the appropriateness of the patient baseline characteristics. Based on responses, UK patients are likely to be heavier and older than the company's baseline characteristics. It was highlighted that endometrial cancer is most common amongst obese patients and that it usually affects older women i.e. between the ages of 75 and 79 years. In order to explore this uncertainty with respect to impact on cost effectiveness, the ERG conducted scenario analyses which used a higher weight and age (see Section 6.2.8 and 6.2.9). Whilst these scenario analyses did not have a meaningful effect on the ICER, they were included in the ERG's preferred base case as they were considered to better reflect UK patients.

4.2.3.2. Positioning

As noted in Section 2.4 and in Figure 1 below, the company appear to be positioning PEM+LEN as a treatment option following platinum-containing chemotherapy provided in the advanced/recurrent setting and secondly for those with recurrent, advanced or metastatic cancer who had received platinum-based chemotherapy in the (neo)adjuvant setting. The ERG noted that for the latter positioning, the most appropriate comparator is re-challenge with platinum-containing doublet chemotherapy and that the treatments provided in KEYNOTE-775 were primarily doxorubicin or paclitaxel (and not specifically platinum re-challenge). The company did provide a scenario analysis which assumed a proportion of patients received carboplatin in combination with paclitaxel as re-challenge (see p.65).

Figure 1: Positioning of PEM+LEN



Abbreviation: EC, endometrial cancer

4.2.3.3. Lack of subgroup analyses

Whilst the ERG noted the population to be consistent with the NICE final scope[4] and KEYNOTE-775 trial population, the appropriateness of assessing cost effectiveness in potentially clinically relevant subgroups was not explored. As noted in Section 3.2.3.1, the company conducted clinical subgroup analysis in dMMR and pMMR patients, however no cost effectiveness results were provided. During clarification with the company (B19), the ERG asked for further rationale as to why an economic analysis was not conducted based on MMR subgroups. The company stated that *'The indication covers the overall patient population irrespective of MMR status. As such the cost-effectiveness analyses focus on the overall patient population and such claim was not made by MSD as it would deviate from the final scope issued by NICE.'* Whilst the ERG agree that the overall population covered by the indication is in alignment with the NICE final scope, clinical opinion to the ERG suggested that prognosis and treatment options provided to patients may vary depending on MMR status.

As outlined in Section 4.2.4, clinical opinion to the ERG suggested that monotherapy immunotherapy treatments are currently used in patients with dMMR (due to high response to treatment). However, monotherapy appears to have limited efficacy in patients with pMMR. Clinical opinion to the ERG included interest in using PEM+LEN within the pMMR subgroup, as

the dual combination of PEM+LEN is likely to be more effective than single agent use. The ERG considered the lack of cost effectiveness results for MMR subgroups to be an area of uncertainty, particularly the lack of results for pMMR patients, as PEM+LEN is most likely to be used in this subgroup in practice.

4.2.4. Interventions and comparators

Pembrolizumab plus lenvatinib (dual therapy) is compared with a composite comparator of doxorubicin or paclitaxel, which the company state is reflective of physician's choice or TPC. This represents a blended comparator whereby the company assumed that 74.5% of patients received doxorubicin and 25.5% received paclitaxel. Based on clinical input to the ERG, doxorubicin or paclitaxel were considered appropriate comparators and were likely to be displaced by PEM+LEN, however choice of treatment varied, with most experts indicating that paclitaxel is used more than doxorubicin. In order to explore uncertainty surrounding the proportion of patients receiving doxorubicin or paclitaxel, the ERG conducted scenario analyses which varied proportions (see Section 6.2.2). Based on clinical input received, the ERG's preferred base case assumed 50% of patients received doxorubicin and 50% received paclitaxel.

Initially, the ERG had some concerns surrounding the company's assumption that doxorubicin and paclitaxel were comparable in terms of efficacy. Based on clinical input to the ERG, doxorubicin and paclitaxel were likely to be similarly '*effective or ineffective*', however choice between paclitaxel and doxorubicin would be based on the side-effect profile (cardiac vs renal).

The ERG noted that hormone therapy was not considered as an appropriate comparator within the company's economic analysis. The company justified the exclusion of hormone therapy on the basis that it is only used *'if all other treatment options are exhausted or patients cannot tolerate further lines of chemotherapy'*. Clinical opinion to the ERG, noted that hormone therapy is primarily given as a palliative treatment and therefore agreed with the company's decision to exclude it. The ERG are aware of a recent NICE appraisal for endometrial carcinoma dostarlimab (TA779),[3] which was recommended for patients with recurrent or advanced dMMR/MSIH EC who have progressed on or after platinum-based chemotherapy. Given that this is relatively recent guidance, published in March 2022, the ERG considered the exclusion of dostarlimab to be appropriate. Furthermore, dostarlimab is recommended for a subgroup of patients, which is narrower than the population for which pembrolizumab is indicated.

As noted on p. 70 of the CS (document B), the company conducted a scenario analysis, which compared PEM+LEN to a mixed chemotherapy arm. Treatment costs for this comparator were based on a weighted approach which used data from ECHO i.e. the mixed chemotherapy arm was assumed to consist of [redacted] % paclitaxel, [redacted] % doxorubicin, [redacted] % carboplatin, and [redacted] % carboplatin plus paclitaxel [as re-challenge]. Results were not sensitive to this analysis. The ERG noted that this analysis was subject to several simplifying assumptions, namely that the mixed chemotherapy arm was assumed to have equivalent efficacy to that of the TPC arm in KEYNOTE-775.

4.2.5. Perspective, time horizon and discounting

All costs and outcomes were estimated from a NHS and PSS perspective as appropriate. The time horizon used in the analysis was 40 years, which was considered by the company to be a lifetime horizon. The ERG noted that a 40 year time horizon had been used previously several ovarian cancer appraisals including niraparib (TA528)[19] and (TA673),[20] and a 30 year time horizon in others including rucaparib (TA611)[21] and olaparib (TA620).[22] Within the recent appraisal of dostarlimab (TA779),[3] for the treatment of patients with recurrent or advanced dMMR/MSI-H endometrial cancer, a 40 year time was used. Overall the ERG considered the time horizon was sufficiently long to adequately capture the differences in costs and outcomes between treatments and was broadly in line with appraisals for similar conditions. Furthermore, the company provided a scenario analysis which reduced the time horizon to 30 years, however this did not have a significant impact on results.

With respect to discounting used in the model, both costs and QALYs were discounted at 3.5%, in line with NICE guidance. However, estimates of life years were not discounted. NICE guidance states that "the same annual discount rate should be used for both costs and benefits (currently 3.5%)." [NICE 2013, paragraph 5.6.1].[23] The ERG conducted an analysis which discouted life years at 3.5%. This was included as part of the ERG's preferred base case (see Section 6.3).

4.2.6. Treatment effectiveness and extrapolation

4.2.6.1. Critique of general modelling approach

The clinical data used to model PFS, OS and time-on-treatment (ToT) were taken from the phase III KEYNOTE-775 study (data cut 26 October 2020). Due to the lack of long-term data

from this trial (and to assess the cost effectiveness of PEM+LEN over a lifetime horizon), the company extrapolated OS and PFS beyond the clinical trial's last follow-up.

The company considered several modelling approaches, including a 'one-piece' approach (standard parametric) and 'two-piece' (piecewise KM followed by parametric) approach independently fitted to each treatment arm. As noted in Appendix P of the CS, propotional hazards-based methods were not considered to be a viable means to estimate OS, citing Schoenfeld residuals and proportional hazard plots. The ERG agreed that a proportional hazards modelling approach was not justifiable on the basis of log-cumulative hazard plots (Document B Appendix, figures 26 and 30). The company considered modelling OS or PFS using a one-piece parametric survival curve; however, the company considered these fits implausible/inappropriate (sections B.3.3.3 and B.3.3.4). The ERG agreed that the one-piece approach appeared not to fit the hazard function well, most notably for OS (CS document B, figures 13 and 14).

In the base case analysis the company opted for the two-piece approach to estimate both OS and PFS in the trial period of both arms of KEYNOTE-775. The company argued (document B section 3.3.3.1; clarifications to B9, B11, B13) that the two-piece approach provided a good visual fit (adequate internal validity) and plausible extrapolated survival (adequate external validity).

The CS supplied plots of the modelled hazards for the one- and two-piece approaches in CS document B (figures 13, 14, 17, 18) for OS but not PFS. Within these figures was a curve labelled 'smooth spline estimate'. This was clarified by the company (responding to clarification questions B10 and B12) as a representation of the hazard function using many-knot (31) basis splines, as opposed to a flexible modelling approach of the type outlined in TSD 21 (e.g. restricted cubic spline). The ERG hereon terms the former the empirical hazard function. 'Zoomed-in' versions of some of these plots were supplied in clarification response section C. PFS hazard function plots for the two-piece approach were not provided in the CS but obtained in clarification response to B12 (figures 9 and 10).

The 'two-piece' approach used by the company uses an initial nonparametric (KM) fit followed by standard parametric fits to later data points. The ERG acknowledges that the two-piece or 'Liverpool approach' has been used in previous appraisals and is one possibility outlined in TSDs 14 and 21.[24,25] A criticism of the two-piece approach is that placement of the breakpoint (between KM and parametric) can be arbitrary. In the CS, the results of a Chow test to more

objectively set the breakpoint were cited, though these were only supplied after clarification (B6).

The ERG noted a range of issues with the two-piece approach. When comparing the two-piece fitted hazard to the empirical hazard (labelled 'smooth spline fit'), a failure to track the hazard function closely is apparent in all fitted two-piece models see (e.g. clarification response figure 8). Sudden changes in hazard at the breakpoint, mentioned in TSD21[25] as potentially implausible and a drawback of the two-piece approach, are apparent with some parametric model choices in the two-piece approach (see e.g. doc B figs 17 and 18). A 26-week breakpoint was selected for OS and 10-week breakpoint for PFS on the basis of Chow test results, visual inspection of the hazard function and a preference for earlier breakpoints, thereby providing more data for parametric fitting in the second piece (doc B p80). Taking the Chow test at face value, the plots supplied at clarification (B6) do not appear to clearly support the 10-week breakpoint selection for PFS (clarification figs 4 and 5), nor the 26-week OS breakpoint in the TPC arm (clarification fig 3). Moreover, the ERG believes the Chow test to be an invalid approach because it is inappropriate to use a 'test statistic surface' to determine relevant breakpoints, and in the event Chow test statistics revealed a range of plausible breakpoints.

The company chose not to use other flexible fitting approaches outlined in TSD 21[25] in the CS, and declined to do so for clarification (see e.g. clarification responses B9, B11). It is not possible therefore to assess any improvement in fit over the two-piece approach, nor the plausibility of any extrapolations under an improved fit. The ERG recommends restricted cubic splines are applied and assessed as these are the best combination of flexibility and generalisability given the hazard functions in evidence.

The company introduced further constraints to survival modelling in the form of capping to ensure that PFS and ToT never exceed OS, and to ensure OS is capped to general mortality (further discussed in section 4.2.6.2).

Validation of extrapolations

The company appeared to have presented its selected fits to clinicians who indicated that they were plausible (CS document B section 3.3.3.1). The CS indicates that 'All participants were more comfortable predicting plausible extrapolations for the TPC arm, given their experience in treating patients with chemotherapy regimens in this treatment area'. In the TPC arm clinicians to the company favoured certain extrapolations (the 'bottom group of curves in figure 16'

representing Weibull, Gompertz and exponential; see clarification response C2). It is not clear to the ERG from the CS that clinicians *selected* from the extrapolating model(s) in the PEM+LEN arm, though it appears they accepted the company's choice (log-logistic) was plausible.

The company attempted to validate the extrapolated curves with the use of longer-term information. The company marshalled information from other studies (KEYNOTE-146 for PEM+LEN and ECHO for TPC) to validate its extrapolations. This aspect is discussed in more detail in section 4.2.6.4 . For the TPC arm the ERG found reporting to be inadequate for the purpose, and there were marked discrepancies in patient characteristics between the ECHO and KEYNOTE-775 TPC arm. For the PEM+LEN arm, the supporting study (KEYNOTE-146) was comparable in many ways to the KEYNOTE-775 arm, though some information remained unavailable (e.g. time since diagnosis).

Specific issues with base case extrapolations

Turning to the choice of parametric model under the two-piece approach (but noting the ERG's preference for a restricted cubic spline approach as discussed above), the ERG disagreed with the company base-case choice for OS in the TPC arm.

For OS, the empirical hazard function declines at later follow-up times in both TPC and PEM+LEN arms (clarification figs 7 and 8). The company selected a model to track this decline in the PEM+LEN arm (log-logistic selected), but not so the in the TPC arm (exponential selected). This is depicted in Figure 2 below. The company argued (CS document B, p. 85) that 'the hazards in the PEM+LEN arm have a strong decreasing trend after 26 weeks, which does not occur in the TPC arm'. However, the ERG noted a decline in hazards from about 60 weeks in the TPC arm, albeit delayed compared to PEM+LEN and with less precision, and in this context questions the selection of a uniform hazard (exponential model).

Extrapolated OS (up to 10 years) for ERG and company curve fit selections is shown in

Figure 3. In the TPC arm, higher survival is predicted in the longer term by the ERG choice (log-logistic) than the company (exponential). The ERG choice responds to clinicians advising the ERG: there was some variation in responses, however on balance long-term OS extrapolation under the company choice in the TPC arm was considered an underestimate. The estimates of OS within and beyond the trial period at 1, 2, 5 and 10 years are shown in Table 20.

Based on the supplied survival curves for PFS (CS document B figures 21 and 22), there is less divergence between models and the ERG has not altered the company's base case choices as these were viewed to be reasonable. The ERG noticed (but could not explain) the relatively jagged form of the empirical hazard supplied for PFS in the TPC arm (clarification response figure 10).

Figure 2: Company and ERG base-case model choices (after 26-week breakpoint) with empirical hazard (black line) for OS (ERG-constructed figure)

Abbreviations: CS, company submission; ERG, Evidence Review Group; KN, KEYNOTE (trial); OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; TPC, treating physician's choice

Source: company-supplied survival data and parameter estimates

Figure 3. Fitted OS models with CS and ERG selections, extrapolated to 10 years

[redacted]

Abbreviations: CS, company submission; ERG, Evidence Review Group; KN, KEYNOTE (trial); OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; TPC, treating physician's choice

Table 20: Survival estimates from the main trial (KEYNOTE-775) and base case extrapolations by ERG and company

Years		1	2	5	10
PEM+LEN	KEYNOTE-775	[redacted]	[redacted]	-	-
	ERG/CS model (log-logistic)	[redacted]	[redacted]	[redacted]	[redacted]
	KEYNOTE-146	[redacted]	[redacted]	[redacted]	-
TPC	KEYNOTE-775	[redacted]	[redacted]	-	-

Years		1	2	5	10
	CS base case (exponential)	[redacted]	[redacted]	[redacted]	[redacted]
	ERG base case (log-logistic)	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: CS, company submission; ERG, Evidence Review Group; PEM+LEN, pembrolizumab with lenvatinib; TPC, treating physician's choice

4.2.6.2. Capping of OS and PFS

As OS, PFS and ToT are modelled independently, the spreadsheet model implements capping to ensure that PFS and ToT never exceed OS. Furthermore OS is capped to general population mortality (CS document B, Section 3.2.2, p. 66). This is implemented in a 'hybrid' fashion. For example, overall survival at time t is set to be the minumum of the predicted overall survival at time t from the chosen OS model, and the overall survival at t-1 multiplied by 1-hazard of death in the general population:

$OS(t) = \min[OS_{pred}(t), OS(t-1)*(1-h_{pop}(t)]$

where $OS_{pred}(t)$ is the overall survival at time t predicted by the chosen survival function, and $h_{pop}(t)$ is the hazard of death in the general popuation at time t (i.e. for the age and gender of the subject patient).

Likewise, PFS is calculated in the same manner:

$PFS(t) = \min[PFS_{pred}(t), PFS(t-1)*(1-h_{OS}(t)]$

where $PFS_{pred}(t)$ is the PFS at t predicted by the chosen function, and $h_{OS}(t)$ is the hazard of overall survival at t, after adjusting for overall population mortality.

This hybrid approach is somewhat inconsistent. For example, in the latter case it mixes together the 'stock' of PFS survival as predicted by the chosen survival function, which is itself a function of the 'flows' of hazards predicted purely by the chosen survival function and the 'flows' of hazards capped for OS and population mortality. A simpler approach would be to cap OS at the minimum of OSpred and general population mortality, and to cap PFS at the minimum of PFS_{pred} and OS:

$OS(t) = \min[OS_{pred}(t), OS_{pop}(t)]$

$PFS(t) = \min[PFS_{pred}(t), OS(t)]$

where $OS_{pop}(t)$ is general population survival at t. We describe this as the 'simple' approach.

Alternatively, perhaps more plausibly, it could be argued that the *risk* (hazard) of death in patients with advanced EC each period should be the greater of that predicted by the chosen survival function and that experienced in the general population:

$OS(t) = OS(t-1) * (1 - \max[h_{OS,pred}(t), h_{pop}(t)])$

And likewise the hazard of progression or death should be the greater of that predicted by the chosen PFS survival function and that for overall survival (and by definition from the equation above, greater than the hazard of death for the general population):

$PFS(t) = PFS(t-1) * (1 - \max[h_{PFS,pred}(t), h_{OS,pred}(t), h_{pop}(t)])$

We describe this as the 'hazards' approach.

It should be noted that these alternative approaches were applied to both the PEN+LEN arm and the TPC (and mixed chemotherapy arms), as part of ERG scenario analyses. The impact of these alternative approaches is explored in the ERG's scenario analyses (see Section 6.2.10).

4.2.6.3. Treatment waning

The company did not include treatment waning in their economic model, on the basis of precendent and stated this was 'validated by long-term KN-146 data'. The appraisals described in the CS (document B, Table 24) to demonstrate non-applicability of treatment waning were for PARP inhibitors in ovarian cancer. During clarification, the ERG asked the company to provide additional rationale for excluding exploration of a waning in treatment effect. In clarification response B18, the company cited two additional pembrolizumab appraisals which did not use a treatment waning assumption (TA531:[26] untreated PD-L1-positive metastatic non-small-cell lung and TA357:[27] advanced melanoma after disease progression with ipilimumab). The company stated that as longer-term immunotherapeutic effects were demonstrated after stopping treatment in these appraisals, it could be expected that PEM+LEN would offer a sustained treatment effect. The ERG noted that it may not be appropriate to assume that PEM+LEN would mirror the treatment effect seen in these appraisals, on the basis that there would be differences across patient populations with respect to baseline characteristics, drug mechanisms, disease

type and treatments received. Furthermore, based on a review of dostarlimab TA779,[3] the ERG noted that treatment waning was considered as part of the company's base case.

In order to validate the company's decision to exlcude treatment waning, the ERG sought clinical expert opinion. Responses to the ERG were somewhat mixed and noted there to be a lack of data surrounding waning of effect. However on balance clinicians considered that after stopping treatment with PEM+LEN, there may be gradual waning. It was also noted that there would be patients who will relapse/experience disease progression. In order to explore uncertainty surrounding treatment waning, the ERG conducted a scenario analysis which included a treatment waning assumption. Results were highly sensitive to this analysis (see Section 6.2.10).

The company also stated (clarification B18) that treatment waning was not explored on the basis that long-term OS data from KEYNOTE-146 showed a durable and sustained treatment effect beyond the 2-year treatment period with PEM+LEN i.e. sustained OS in the form of a plateau (with 30% of patients alive at 5 years). However, the ERG noted that the latter part of the KM curve shown (CS document B, figure 9) is still subject to considerable censoring with small numbers at risk after 28 months, and confidence intervals may be wide. The confidence intervals were not shown and the company informed the ERG that the underlying data were not available (clarification response B7). Furthermore, based on clinical opinion to the ERG, 30% survival at 5 years is likely higher than in UK clinical practice. On the other hand in case of sufficient support for the notion that some patients are cured (best demonstrated through sufficient maturity and precision in survival curves), survival modelling may need to incorporate a cure fraction (TSD21 section 3.6).[25]

4.2.6.4. Validation of extrapolations

The population characteristics of the KEYNOTE-146 study, used to validate extrapolation in the KEYNOTE-775 PEM+LEN arm, are shown in Table 21, derived by the ERG from the CS. The ERG interpretation is that many characteristics are well-matched when available, but on the other hand the ERG observes the CS statement (Document B p.37) that 'KEYNOTE-146 and KEYNOTE-775 are heterogeneous in terms of study design and population'. Information is sometimes limited (see also section 3.2.1), and the ERG noted in particular that the time since diagnosis has not been supplied for KEYNOTE-146 (despite a request in clarification A13) and that the distribution of stages is only available for endometrioid cancers. This significantly limits the value of KEYNOTE-146 for validation of extrapolations.

Table 21: Comparison of baseline characteristics for PEM+LEN arms of KEYNOTE-775 and KEYNOTE-146 (EC subgroup).

Baseline characteristics		**KEYNOTE-775: PEM+LEN arm**	**KEYNOTE-146: endometrial cancer subgroup**	
			Previously treated	All
Age (mean)		63.2	65.1	65.3
Ethnicity (%)	Asian	20.7	4.6	4.0
	Black	4.1	5.6	5.6
	White	63.5	86.1	87.1
	Other [a]	2.7	3.7	3.2
ECOG status (%)	0	59.9	49.1	50
	1	39.9	50.9	50
	2	-	-	-
	3	0.2	-	-
MMR (%)	pMMR	84.2	87	-
	dMMR	15.8	13	-
PD-L1 (%)	+	-	49.1	48.4
	-	-	39.8	41.9
	N/A	-	11.1	9.7
FIGO grading (%) [b,c]	I	27	11.1	12.1
	II	7.8	17.6	17.7
	III	29	22.2	24.2
	IV	36	-	-
Histology (%)[d]	Clear cell	7.3	5.6	4.8
	Endometrioid [2]	59.2	50.9	54
	Serous [1]	25	32.4	31.5
Time since diagnosis (mean, years)		2.4	-	-
Prior treatment with (%)	monotherapy	-	33.3	29.8
	Platinum+taxane	-	98.1	91.1
Patients		Advanced recurrent or metastatic EC, progression after 1 prior systemic platinum-based chemo	No more than 2 previous systemic therapies	

Abbreviations: dMMR, deficient mismatch repair; EC, endometrial cancer; ECOG, Eastern Cooperative Oncology Group; FIGO, International Federation of Gynecology and Obstetrics; MMR, mismatch repair; N/A, not applicable; PD-L1, programmed death-ligand 1; PEM+LEN, pembrolizumab with lenvatinib; pMMR, proficient mismatch repair

Sources : Doc B Table 5 , Doc B Appendices Table 55

Notes:[a] pooled over small categories for 'Other' by ERG; [b] pooled over subcategories by ERG; [c] in KN146 for endometrioid cancers only – see clarification response A14; [d] pooled over subcategories by ERG. 1: serous + high-grade serous ; 2: endometrioid + endometrioid with squamous differentiation + high grade endometrioid + low grade endometrioid

The population characteristics of the ECHO study,[13] used to validate extrapolation of the KEYNOTE-775 TPC arm, are shown in Table 22, derived by the ERG from the CS. There are numerous characteristics that differ markedly between the cohorts, including [redacted]. There are further discrepancies e.g. proportion of dMMR patients.

The ERG has further concerns about the use of ECHO :

- limited information presented by the company to support its suitability for validation, and no study report, no protocol, no peer-reviewed publication available;
- offers only a short extrapolation period of [redacted] (updated from [redacted] in the CS – see clarification A15) even though it represents standard treatment; and
- [redacted] of patients in ECHO are on doxorubicin or paclitaxel.

A clinician advising the ERG indicated that in the UK population performance status was roughly in the proportions (0=10%, 1=50%, 2=30%, 3=10%), MMR status in the proportions (dMMR=30%, pMMR=70%) and in the relapsed setting histology of (endometrioid=40%, serous=40%, clear-cell=15-20% and mucinous= <5%). Comparing the supporting study (ECHO, KEYNOTE-146) characteristics to the routine UK population, the ERG noted:

- low proportion of performance status (as measured by ECOG) grades 2 or 3 in ECHO (compared to 30% and 10% respectively in the UK);
- a smaller proportion of MMRd (around [redacted]) in ECHO and a larger proportion (around [redacted]) in KN-146 compared to the UK (approximately 30%); and
- differences in the proportions of serous or endometrioid cancers in both studies compared to the UK.

The company further informs of the following difference (CS document B, p. 83): 'ECHO included some patients who received investigational treatments not routinely available in UK clinical practice as a subsequent treatment (such as PD-1/PD-L1 and VEGF/VEGFR inhibitors).

In order to obtain the fullest information for extrapolating the TPC arm, the ERG recommends extending the search for RWE on survival to include web searches for grey literature sources not included in bibliographic databases (e.g., UK cancer registries or reports derived from electronic health records). Searches for the company's RWE SLR were conducted in July 2020 and updating these searches may also identify additional evidence. Paclitaxel or doxorubicin in combination or alone could be informative. The ERG noted that the dostarlimab appraisal (TA779)[3] provided results on what may be a relevant cohort, but the information is confidential.

Should further external sources for validation be obtained, the ERG recommends carrying out a comparison of the population characteristics of any extrapolating studies to those of the target population (UK clinical practice), and consideration of adjusted extrapolations by standardising to the target population.

Table 22: Comparison of baseline characteristics for TPC arm of KEYNOTE-775 and RWE study ECHO.

Baseline characteristics		**KEYNOTE-775 : doxorubicin or paclitaxel arm**	**ECHO**
Age in years at initial diagnosis (mean)		61.5	[redacted]
Ethnicity (%)	White	59.1	[redacted]
	Black or African/ Caribbean-origin	3.4	[redacted]
	Middle Eastern/ North-African	-	[redacted]
	Asian	22.1	[redacted]
	Other	4.8	[redacted]
MMR Status (%)	dMMR	15.6	[redacted]
	pMMR	84.4	[redacted]
MSI status (%)	MSI-H/dMMR	-	[redacted]
	Non-MSI-H/pMMR	-	[redacted]
	Mixed	-	[redacted]
ECOG at recurrent or advanced diagnosis (%)	0	57.9	[redacted]
	1	42.1	[redacted]

Baseline characteristics		KEYNOTE-775 : doxorubicin or paclitaxel arm	ECHO
	2	-	[redacted]
	3	0.0	[redacted]
Radiation (%)		41.6	[redacted]
Histology [a] (%)	Clear cell	4.1	[redacted]
	Endometrioid [2]	61.1	[redacted]
	Serous [1]	27.6	[redacted]
	Mucinous [3]	0.2	[redacted]
Elapsed time in years from initial diagnosis (years, mean)		2.9	[redacted] [e]
Staging at initial diagnosis [b] (%)	I	33.4	[redacted]
	II	6.3	[redacted]
	III	30.7	[redacted]
	IV	29.6	[redacted]
Metastatic site(s) at diagnosis (%)		-	[redacted]
	Liver metastasis	23.6	[redacted]
	Distant [c] lymph node(s)	54.1	[redacted]
	Lung metastasis	36.5	[redacted]
	Bone metastasis	7.9	[redacted]
	Brain metastasis	0.5	[redacted]
	Pancreas	-	[redacted]
	Kidney	-	[redacted]
Treatment with doxorubicin or paclitaxel [d]	Yes	100%	[redacted]

Abbreviations: dMMR, deficient mismatch repair; EC, endometrial cancer; ECOG, Eastern Cooperative Oncology Group; MMR, mismatch repair; MSI, microsatellite instability; MSI-H, microsatellite instability-high; pMMR, proficient mismatch repair; RWE, real-world evidence; TPC, treating physician's choice

Sources: Doc B table 5; clarification question response A11

Notes:[a] pooled over subcategories by ERG : 1: serous + high-grade serous ; 2: endometrioid + endometrioid with squamous differentiation + high grade endometrioid + low grade endometrioid; 3: low-grade mucinous + high-grade mucinous; [b] pooled over subcategories by ERG; [c] described as 'distant' in ECHO but not in KN775; [d] inferred from text; [e] ERG conversion from presumed reported months

4.2.7. Time-on-treatment and stopping rules

For ToT, the company investigated one-piece parametric survival models and applied stopping rules, derived from dosage/cycle limits for doxorubicin and pembrolizumab, to each selected

model. The company's base case analysis included a 24 month stopping rule for pembrolizumab and assumed that treatment with doxorubicin would be limited to a maximum lifetime cumulative dose of 500 mg/m². Based on clinical expert opinion to the ERG, these assumptions were considered to be reasonable. For completeness, the company conducted scenario analyses which assumed no maximum dosing rule for doxorubicin and which assumed a maximum duration of 6 months for paclitaxel. Results were not sensitve to these analyses.

Long term drug acquisition costs for lenvatinib, pembrolizumab and TPC were modelled using a generalised gamma curve. The ERG noted that pembrolizumab and lenvatinib were modelled seperately to account for the costs associated with each treatment, however ToT for doxorubicin and paclitaxel were not modelled separately, but rather as a sigle arm 'TPC' (see **Figure 4**). The company stated that this was due to the short duration of treatment and low costs associated with the treatments.

For pembrolizimab, lenvatinib and TPC, the company stated that the generalised gamma provided a plausible fit to the observed data from KEYNOTE-775. The company provided some scenario analyses to test the impact of alternative ToT assumptions on the ICER, which included the the use of an alternative parametric function (Weibull) for both arms, the assumption that ToT cannot exceed PFS (in both arms) and estimating ToT based on KM data, see Table 52, Section B.3.8.3 of the CS. The ICER was not sensitive to these analyses. The ERG considered that the most appropriate method of estimating drug costs was to cap ToT by PFS, as ToT should be largely coterminous with PFS i.e. progression would often trigger a change or desistance in treatment. The ERG conducted a scenario analysis based on this approach (and considered this as part of the ERG preferred base case). Results were not especially sensitive to this (see Section 6.2.7 and 6.2.10).

Figure 4: ToT extrapolation used in the company's base case

Abbreviations: KN, KEYNOTE (trial); OS, overall survival; LEN, lenvatinib; PEM, pembrolizumab; PEM+LEN, pembrolizumab with lenvatinib; ToT, time-on-treatment; TPC, treating physician's choice

4.2.8. Health-related quality of life

Quality of life data collected directly from patients in KEYNOTE-775 were used to derive health state utilities in the model. In KEYNOTE-775, patients were given the EQ-5D-5L to complete on day 1 of each cycle, for the equivalent of four cycle lengths and the end of treatment visit. On CS document B p. 101, the company stated that completion of HRQoL questionnaires following the end of treatment visit i.e. post treatment discontinuation, was not mandatory. The ERG noted that the EQ-5D-5L values were mapped to EQ-5D-3L values using the Van Hout cross walk method (as per NICE's position statement). In the base case analysis the company opted to use a time to death (TTD) approach to derive base case utilities (utilities presented in Table 23 below), as opposed to a progression status approach whereby values are presented based on whether patients are progression-free (PF) or have progressed diseased (PD).

Table 23: Time to death utility values used in the company's base case analysis

Time to death	Mean utility value
≥360 days	[redacted]
270 - 359 days	[redacted]
180 - 269 days	[redacted]
90 - 179 days	[redacted]
30 - 89 days	[redacted]
<30 days	[redacted]

To estimate TTD utility values for the six time-based modelled health states, the company used a linear mixed effects regression model which was fitted to HRQoL data from KEYNOTE-775. Explanatory variables were dummy variables for time to death less than 30 days, 30-90, 90-180, 270-360, greater than 360 and absence of AEs (grade 3 and above). (Note 180-270 days to death is therefore the default.)

Health state utilities were estimated assuming no AEs. The manufacturer did not use the estimated coefficient to estimate disutility of AEs from the model, opting instead to estimate a sum of disutilities for each grade 3/4 AE individually, weighted for the probability and duration to estimate a QALY penalty per cycle. The ERG assessed that this was reasonable to account for the duration of adverse events.

On p.102 of the CS, the company justified the TTD approach to estimating health state utility on the basis that it captures *'the decrease in utility as patients move closer to death, driven by the underlying impact of the disease over time, removing the dependence on clinical assessment of progression status.'* The company further stated that this approach has been used in previous oncology appraisals including TA531[26] and TA357.[27]. Whilst the ERG acknowledged that time to death had been used previously, the approach does not adequately account for progression status i.e. utilities based on time to death rather than progression status divorced health related quality of life from disease status in the model. Furthermore, the ERG noted that changing the PFS curve whilst holding OS the same made no difference to QALYs gained (only costs). This appeared somewhat counterintuitive.

Based on a review of the dostarlimab (TA779)[3] committee papers, the company used a regression equation which estimated utility based on time to death, but also included progression status as a covariate. Utility values were therefore estimated for pre progression (> 5 cycles from death and ≤5 cycles from death) and post progression (> 5 cycles from death and

≤5 cycles from death). The ERG considered this approach, which also captured progression status, to be more appropriate.

The company did conduct a scenario analysis whereby health state utilities were estimated based on progression status. Using this approach, the mean utility value for PF was estimated to be [redacted] and the mean utility for PD was [redacted]. Results were sensitive to this analysis, resulting in a moderate upward impact in the ICER (See Section 5.2.3). Overall, the ERG preferred utility estimation according to progression status, therefore this approach has been used in the ERG's preferred base case (see Section 6.2.6).

4.2.9. Resources and costs

The company's model included drug acquisition costs, administration and monitoring costs, adverse event costs, subsequent treatment costs and end of life care costs.

4.2.9.1. Drug acquisition costs

Drug costs were included for the intervention (PEM+LEN) and comparator treatment arms (doxorubicin or paclitaxel). The dosing regimen for PEM+LEN was based on the EMA and MHRA marketing authorisation and the KEYNOTE-775 protocol. For pembrolizumab, patients received 200mg every 3 weeks and for lenvatinib patients received 20mg every day. For doxorubicin and paclitaxel, dosing was based on KEYNOTE-775 protocol (see table 42, p.112 of the CS). Based on clinical opinion to the ERG, the dosing schedule used appeared to be appropriate.

Unit costs were derived from MIMS and the drugs and pharmaceutical electronic tool kit (eMIT). The ERG noted some uncertainty surrouding costs and sources for several drugs i.e. the incorrect cost appeared to have been used for medroxyprogesterone and doxorubicin. The company were asked to comment on these during the clarification stage and acknowleged the incorrect costs had been used. The company confirmed that when the correct prices were used for these treatments, this had minimal impact on the ICER. The ERG's preferred base case uses the correct prices for these treatments (see Section 6.1).

The ERG noted that for lenvatinib, the company used the relative dosing intensity (RDI) from the KEYNOTE-775 study, which was estimated to be [redacted]. The company's base case approach therefore assumed that a proportion of patients do not remain on lenvatinib 20mg, but experience dose reduction over time (dropping to 14mg, 10mg, 8mg and 4mg). The ERG noted that the cost of lenvatinib 10mg and 4mg is equivalent (£1,437). Overall, the ERG considered

that the use of dosing data from KEYNOTE-775 may be appropriate (if reflective of clinical practice). For completeness, the ERG sought further clinical expert opinion in order to determine whether dose reduction in clinical practice is likely. Based on feedback received, most patients are likely to receive dose reduction with lenvatinib (approximately 66%). As an exploratory analysis, the ERG conducted a scenario analysis which assumed no dose reduction for lenvatinib. Results were not overly sensitive to this (see Section 6.2.10).

Finally, the company has excluded pre-medication costs for paclitaxel for simplicity i.e. as per the SmPC for paclitaxel patients should receive steroids, antihistamines and H2-receptor antagonists. The company further noted that this is a conservative asssumption as this underestimates the costs of TPC. Due to the relatively small costs associated with these pre-medications, the ERG did not consider this to be an issue and found the company's approach to be reasonable.

4.2.9.2. Health state, monitoring and administration costs

Disease management costs (including monitoring costs) were included in the model and estimated for each health state i.e. PF or PD (see Table 44 on p.117 of the CS for a full list). Health state costs were calculated according to time spent in each state and specific healthcare resources used in that state. The company derived healthcare resource use estimates from previous NICE TA's including TA620[22] and ID1547,[28] which the ERG considered to be reasonable. The total weekly cost (cost per model cycle) associated with the PF and PD health states was estimated to be £43.06 and £35.08 respectively. Unit costs taken from the Personal Social Services Research Unit (PSSRU) and 2019/20 NHS reference costs as appropriate. In order to explore the impact of health state costs on the ICER, the ERG varied the cost per model cycle in the PF and PD health states by +/- 50%, however this did not have a meaningful impact on the ICER. With respect to administration costs, the company assumed no administration cost for lenvatinib, on the basis that it it an oral treatment. For pembrolizumab, paclitaxel and doxorubicin, the cost of intravenous administration were sourced from 2019/20 NHS reference costs (see Table 43, p.115 of the CS). The ERG considered the company's handling of administration costs to be reasonable.

4.2.9.3. Subsequent treatment costs

The model incorporated subsequent treatment costs (see Table 24 for subsequent treatments and proportions used in the base case). Subsequent treatments were based on those given in KEYNOTE-775 (excluding treatments that are not provided in the UK setting) and were

modelled as a one-off cost, applied at point of treatment discontinuation. Overall, based on clinical opinion to the ERG, the list of subsequent treatments and proportions used by the company were largely appropriate. For completeness the company conducted a scenario analysis which assumed subsequent treatments to be reflective of those received by patients in ECHO (see Table 46, p. 123 of the CS). However results were not sensitive to this analysis.

Unit costs for subsequent treatments were derived from eMIT and MIMS, which are considered to be appropriate sources (see Table 47, on p. 124 of the CS for full list of subsequent treatment costs). The ERG noted that MIMS provided a range of prices for bevacizumab i.e. £205.55 to 242.66 for 100mg/4ml and £810.10 to 924.40 for 400mg/16ml, and that the company used the cheapest price in their analysis (without providing justification). However, the ERG did not consider this to be an issue as bevacizumab is included at 0%.

Table 24: Modelled subsequent treatments

Subsequent treatments	After PEM+LEN	After TPC
Paclitaxel	[redacted]	[redacted]
Doxorubicin	[redacted]	[redacted]
Carboplatin	[redacted]	[redacted]
Gemcitabine	[redacted]	[redacted]
Cisplatin	[redacted]	[redacted]
Pembrolizumab	[redacted]	[redacted]
Bevacizumab	[redacted]	[redacted]
Lenvatinib	[redacted]	[redacted]
Hormone therapy	[redacted]	[redacted]

Abbreviations: PEM+LEN, pembrolizumab with lenvatinib; TPC, treating physician's choice

The company stated that PD-L1 regiments are currently not available in the UK for use as subsequent treatment, however clinical opinion to the ERG noted that recently there had been some immunotherapy use during the Covid 19 pandemic, particularly nivolumab. As per advice from NHS England regarding the use of interim treatment options during the Covid 19 pandemic, dostarlimab has recently displaced nivolumab as a viable subsequent treatment option for patients with previously treated advanced or recurrent endometrial cancer with high microsatellite instability or mismatch repair deficiency.

4.2.9.4. Adverse event costs

The manufacturer estimated the cost of AEs based on incidence, recurrence and duration of Grade 3+ AEs that were observed in more than 5% of patients in KEYNOTE-775. The ERG noted that the AE cost per cycle in the TPC arm was around 4 times higher than in the PEM+LEN arm (£42.19 vs £10.74). This is driven primarily by the incidences of neutropenia and febrile neutropenia in TPC versus a higher incidence of hypertension in PEM+LEN. This is reasonable as hypertension is less costly to treat than (febrile) neutropenia.

Unit costs for each adverse event were mostly taken from NHS reference costs 2019/20, as appropriate, however serveral costs including hypokalaemia and proteinuria were assumed to be £0. Adverse events were not considerd to be a key driver of cost effectiveness within this appraisal. Based on one-way sensitivity analysis conducted by the company which excluded AE costs from the model, the ICER increased by approximately 1%.

4.2.9.5. End of life costs

End of life costs were applied as a one off cost when a patient entered the death health state. The company derived the cost from a published study by Georghiou et al. (2014),[29] which was a Nuffield Trust report that explored care costs towards the end of life. This was estimated to be £6,015, however the company inflated this to the current year using PSSRU inflation indices, resulting in a cost of £6,520.55. The ERG identified various end of life costs in the report i.e. hospital care costs for those patients diagnosed with cancer in the final 2 years of life (£4,580), however the ERG were not able to identify the £6,015 figure from the report. Therefore there is some uncertainty as to what this cost consists of i.e. care setting (hospital or hospice) and resource use involved.

Based on a review of olaparib TA598,[30] end of life costs were derived from an alternative source i.e. Guest et al. (2006),[31] which assessed palliative care treatment patterns and associated costs of healthcare resource use for specific advanced cancer patients in the UK. The cost was reported to be £7,638.51. Whilst the ERG noted some variation in end of life case costs depending on the source used, overall end of life care costs were not considered to be a key driver of cost effectiveness. Varying the cost by +/- 50% did not have a meaningful impact on the ICER.

5. COST-EFFECTIVENESS RESULTS

5.1. Company's cost-effectiveness results

5.1.1.1. Base case results

The company submitted base case results are in Table 25. [redacted]. It should be noted that these results do not include the PAS for lenvatinib. CMU prices were also not included (see the cPAS Appendix for results relevant to decision making).

Based on the company's base case analysis, PEM+LEN resulted a deterministic and probabilistic ICER of [redacted] and [redacted] respectively, compared to doxorubicin or paclitaxel. The ERG noted the primary driver of incremental costs to be the drug acquisition costs associated with pembrolizumab and lenvatinib (in the progression-free health state), whilst the incremental QALY gain was primarily driven by an increase in life years i.e. due to the company's OS extrapolation approach patients receiving PEM+LEN lived longer and therefore accrued more QALYs than those in the comparator arm.

Table 25: Company base case results (with pembrolizumab PAS)

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
Company deterministic base case					
PEM+LEN	[redacted]	[redacted]	-	-	-
Doxorubicin or paclitaxel	[redacted]	[redacted]	[redacted]	1.75	[redacted]
Company probabilistic base case					
PEM+LEN	[redacted]	[redacted]			
Doxorubicin or paclitaxel	[redacted]	[redacted]	[redacted]	1.77	[redacted]

Abbreviations: PAS, patient access scheme; PEM+LEN, pembrolizumab with lenvatinib; QALYs, quality adjusted life years.

5.2. Company's sensitivity analyses

5.2.1. One-way sensitivity analysis

The company provided one-way sensitivity analyses which tested several clinical, QoL and cost variables (see Section B.3.8.2 of the CS for results). The company stated that parameters were

varied by the upper and lower limits of the confidence intervals reported in Appendix P. Results were most sensitive to variation in OS, utility and ToT parameters. Overall, the ERG considered the OWSA to be of limited use for decision making, as the relevant confidential price discounts for lenvatinib (and CMU prices for other treatments) were not incorporated. Furthermore, all the OWSAs considered are based on the deterministic results, not the probabilistic. This yields a biased estimate of the expected incremental costs and outcomes.

5.2.2. Probabilistic sensitivity analysis

The company conducted probabilistic sensitivity analysis which varied multiple model parameters simultaneously over 1000 iterations (see Table 25, for the company's base case probabilistic results). The ERG repeated the PSA multiple times generating a coefficient of variation of the ICER of approximately 0.5%, suggesting 1000 simulations are sufficient to minimise Monte Carlo error (a general rule of >2% implies insufficient simulations). Results were also presented in the form of a scatter plot and CEAC (see p. 131 and p. 132 of the CS). Based on the CEAC results (list prices), PEM+LEN had a [redacted] and [redacted] probability of being cost effective at a willingness to pay threshold of £30,000 and £50,000 respectively. Overall, the ERG considered the company's handling of the PSA to be appropriate and did not identify any errors.

As a general observation, on p. 131 of the CS, the company stated agreement between probabilistic analysis and deterministic analysis as evidence of robustness of the model, *"Therefore, the outcomes from the cost-effectiveness model are considered robust to uncertainty from parameter distributions."* The ERG do not consider this statement to be true, because the agreement between probabilistic and deterministic analyses depends on the degree of 'non-linearity' in the model and not robustness.

5.2.3. Scenario analyses

The company conducted a range of scenario analyses whereby alternative assumptions were used in the model (for the full list see p. 169, Appendix Q). Table 26 below presents five scenarios which had the largest impact on the company's base case ICER. Overall, results were not especially sensitive to changes in key model assumptions, however it should be noted that these results include the PAS for pembrolizumab (and list price for lenvatinib and comparator treatments).

Table 26: Company scenario analyses

Parameter	ICER	% change from company base case
Discount rate (1.5% for both costs and utilities)	[redacted]	[redacted]
Lenvatinib weekly dosing (full 20mg)	[redacted]	[redacted]
Health state utilities based on progression status	[redacted]	[redacted]
Use Caelyx® cost for doxorubicin	[redacted]	[redacted]
ToT cannot exceed PFS (both arms)	[redacted]	[redacted]

Abbreviations: ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; ToT, time-on-treatment

5.3. Model validation and face validity check

The company stated that the model was quality assured by the economists who constructed the model and an external economist (not involved in the model's construction) reviewed the technical implemetation of calculations and coding. A checklist was used to document the list of inconsistencies and errors. Overall, the ERG considered the company's model to be valid i.e. no major coding errors were identified. However, several minor errors were found and amended by the ERG (see Section 6.1).

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

The ERG identified a number of limitations within the company's base case and has explored the impact of parameter values, and assumptions, which the ERG believes are more plausible. This section is organised as follows:

- Section 6.1 outlines the errors identified by the ERG in the company's model.
- Section 6.2 details a series of scenario analyses exploring the robustness of the cost-effectiveness results to specific assumptions and additional uncertainties identified by the ERG. These analyses were conducted within the ERG corrected company base case analysis.
- Results for all scenario analyses are presented in Section 6.2.10.
- In Section 6.3, the ERG base case is presented based on a combination of the exploratory analyses presented in Section 6.2.

6.1. ERG corrections and adjustments to the company's base case model

The ERG noted a number of minor errors and typographical errors in the company's submitted model. These were:

- Typographical error in cost of doxorubicin (£20.20 corrected to £20.02)
- Error in cost of medroxyprogesterone (£1.84 corrected to £58.67)
- Life years discounting default set to 0% not 3.5%.

The company submitted a revised version of the model with the typographical errors corrected. These made no material difference to the results (Table 27). The ERG edited the default discount rate for life years to 3.5%. [redacted] The results below include this revised PAS for pembrolizumab and list price for lenvatinib.

Table 27: ERG-corrected company base case results

	Discounted costs	Discounted QALYs	Incremental discounted costs	Incremental discounted QALYs	Cost per QALY gained
ERG corrected company deterministic base case					
PEM+LEN	[redacted]	[redacted]	-	-	-
Doxorubicin or paclitaxel	[redacted]	[redacted]	[redacted]	1.75	[redacted]
ERG corrected company probabilistic base case					
PEM+LEN	[redacted]	[redacted]	-	-	-
Doxorubicin or paclitaxel	[redacted]	[redacted]	[redacted]	1.76	[redacted]

Abbreviations: ERG, Evidence Review Group; PEM+LEN, pembrolizumab with lenvatinib; QALY, quality-adjusted life years

6.2. Exploratory and sensitivity analyses undertaken by the ERG

6.2.1. Survival function capping

As described in Section 4.2.6.2, the company used a 'hybrid' approach to capping overall survival to general population survival and PFS to OS (

Figure 5). In this scenario the ERG have explored two alternative approaches to capping overall survival, a ‘simple’ approach (Figure 6) and a ‘hazards’ approach (

Figure 7). The ERG's preference is for the hazards-based approach as this generates more plausible estimates of survival. For example, this ensures the hazard of death in the patient population increases in line with that of the general population at older ages, avoiding a plateau of mortality under the simple approach. Results were insensitive to this adjustment under the company's base case, but may be sensitive to this under alternative survival functions. See Section 6.2.10 for results.

Figure 5: Company's base case approach to capping survival

Abbreviations: GenPopOS, general population overall survival; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival

Figure 6: Survival function capped via 'simple' approach

Abbreviations: GenPopOS, general population overall survival; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival

Figure 7: Survival function capped by hazards ('hazards' approach)

Abbreviations: GenPopOS, general population overall survival; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival

6.2.2. Comparator weighting

For these scenarios, the ERG explored the impact of altering the proportion of patients receiving doxorubicin or paclitaxel. Scenario a) assumed that 100% patients receive doxorubicin, scenario b) assumed that 100% of patients receive paclitaxel and scenario c) assumed that 50% of patients will receive doxorubicin and 50% will receive paclitaxel. Based on clinical advice received, the ERG have opted to use scenario c) as part of the ERG preferred base case. Results were not sensitive to these analyses. See Section 6.2.10 for results.

6.2.3. Treatment waning

As noted in Section 4.2.6.3, there is some uncertainty surrounding the long-term treatment effect of PEM+LEN. For this scenario the ERG implemented a waning treatment effect between years 2 and 5. This was implemented by substituting the hazard of OS and PFS in the PEM+LEN arm for a weighted average of the predicted OS and PFS in the PEM+LEN and TPC arms, with the weight increasing linearly between years 2 and 5, such that by year 5, the hazard in the PEM+LEN arm was equal to the hazard in the TPC arm. Under some model extrapolations, the predicted hazard at later time points in PEM+LEN exceeded that in the TPC arm. Thus to prevent a 'treatment waxing' effect, the hazard was set at the maximum of the

predicted hazard in PEM+LEN and the weighted average. Results were highly sensitive to this analysis (and to the start and stop timings of the waning). See Section 6.2.10 for results.

6.2.4. No dose reduction for lenvatinib

In the base case analysis, the dose for lenvatinib was based on the dosing observed in the KEYNOTE-775 study i.e. a dose reduction was observed, with relative dose intensity estimated to be [redacted]. The ERG has asked clinical experts to comment on whether dose reduction (as witnessed in KEYNOTE-775 is likely to occur in clinical practice. Based on the response received, most patients are likely to receive dose reduction with lenvatinib (approximately [redacted]). However, in order to explore uncertainty surrounding lenvatinib dosing, the ERG have conducted a scenario analysis which assumes no dose reduction i.e. patients receive 20mg weekly. Results were mildly sensitive to the analysis. See Section 6.2.10 for results.

6.2.5. OS extrapolation

The ERG conducted three scenario analyses to examine the impact of alternative OS modelling assumptions on the ICER. These are as follows:

- **Extrapolate OS using a one-piece model:** The company did not provide an analysis using a one-piece model to extrapolate OS in the PEN+LEN arm, on the basis that this modelling approach produced implausible OS estimates, when compared to longer term data in KEYNOTE-146 (See Section 4.2.6 for further discussion). Given that the company provided scenario analysis using a one-piece modelling approach for PFS (in both treatment arms), the ERG considered that for consistency it would be useful to have a scenario analysis which estimated OS based on this alternative modelling approach. For this scenario, OS in both arms was extrapolated using the best fitting curves based on AIC and BIC (Log-Normal for PEM+LEN and Log-Logistic for the TPC arm). Results were highly sensitive to the analysis. See Section 6.2.10 for results.
- **Two-piece modelling approach using alternative parametric distribution for OS in the PEM+LEN arm:** Given that the Weibull was the best fitting curve (based on AIC/BIC scores), the ERG explored the impact of KM+Weibull in place of the company base case of KM+Log-Logistic. During clarification (B5), the company was asked to explain why the Weibull was not used in the base case to extrapolate OS in the PEN+LEN arm. The company acknowledged, that whilst the Weibull (and exponential) curves provided a good statistical fit, they provided an insufficient fit to decreasing hazards in KEYNOTE-775 (as

per Fig 17 in the CS) and long term KM data from KEYNOTE-146. The company further stated the extrapolated OS estimates from these models were clinically implausible and underestimated long-term survival. Results were highly sensitive to the analysis. See Section 6.2.10 for results.

- **Two-piece modelling approach using alternative parametric distribution for OS in the TPC arm:** As noted in Section 4.2.6.1, clinical opinion to the ERG noted that modelled OS in the TPC arm was considered to be underestimated. In this scenario analysis the ERG explored the impact of using KM+Log-Logistic in place of the company base case of KM+Exponential. This analysis has been included as part of the ERG's preferred base case. Results were moderately sensitive to the analysis. See Section 6.2.10 for results.
- As an exploratory analysis, the ERG tested the impact of reducing the OS gap between the PEM+LEN arm and TPC arm. This scenario analysis combines the prior two options. It should be noted that this analysis is considered to be highly exploratory. Results were highly sensitive to this analysis. See Section 6.2.10 for results.

6.2.6. Health state utilities based on progression status

As noted in Section 4.2.8, the ERG identified several concerns surrounding the appropriateness of using time to death utilities within the base case. For this scenario, health state utilities were estimated on a health state basis i.e. progression-free, progression and dead, in place of the proximity to death approach. This analysis has been considered as part of the ERG's preferred base case. The results were mildly sensitive to the analysis. See Section 6.2.10 for results.

6.2.7. ToT capped by PFS

In the company base case, ToT was modelled independently from health state, allowing patients to continue treatment post progression. As noted in section 4.2.7, the ERG considered that ToT is more appropriately estimated by capping ToT by PFS. This analysis has been considered as part of the ERG's base case. The results were mildly sensitive to the analysis. See Section 6.2.10 for results.

6.2.8. Patient weight increased to 85kg / Body Surface Area to 1.96m^2

According to the clinical advice received by the ERG, patients enrolled in the clinical trial were of a lower mean weight than those typically seen in UK clinical practice (company base case: 70kg). The ERG therefore conducted a scenario analysis at a patient mass of 85kg.

In the company's model, dosing of paclitaxel, doxorubicin, gemcitabine and cisplatin is set according to body surface area (company base case: 1.77m^2). Bevacizumab is dosed according the body mass (company base case: 70kg). Body surface area is a function of height and weight, for which a number of alternative formulae exist.[32] However, the company's model does not explicitly link the two. Therefore, the ERG conducted a scenario analysis setting patient mass to 85kg and body surface area to 1.96m^2. (In 2019, the average height of a woman in England was 162cm.[33] Using the average weight of a patient in KM-775 of 70kg, the Mosteller formula[34] generates a BSA equal to the company base case of 1.77m^2. Using the same formula with a mass of 85kg yields an estimated BSA of 1.96m^2.) The results were insensitive to the analysis. See section 6.2.10 for results.

6.2.9. Patient age increased to 75

According to clinical advice received by the ERG, patients enrolled in the clinical trial were of a lower age than those typically seen in UK clinical practice. Supporting evidence was provided to the ERG, which highlighted that incidence rates for uterine cancer are highest amongst females aged 75-79 years.[35] For this scenario the ERG explored set the mean age of patients to be 75 (which was also used in the ERG's preferred base case). This had a relatively minor upward impact on the ICER. See Section 6.2.10 for results.

6.2.10. Impact on the ICER of additional clinical and economic analyses undertaken by the ERG

The ERG made the changes described in Sections 6.2.1 to 6.2.9. Each change has been made individually. The results of the ERG's exploratory analyses are provided in Table 28 and Table 29 below. All results are based on the updated pembrolizumab PAS and lenvatinib list price.

Table 28: ERG's exploratory analyses (deterministic)

Exploratory scenarios	Section in ERG report	Incremental discounted costs	Incremental discounted QALYs	ICER £/QALY	+/- company base case
Company base case	5.1.1.1	[redacted]	1.75	[redacted]	---
Approach to capping the survival function					
a. Simple capping method	4.2.6.2 and 6.2.1	[redacted]	1.82	[redacted]	[redacted]
b. Hazards capping method	4.2.6.2 and 6.2.1	[redacted]	1.77	[redacted]	[redacted]
Comparator					

Exploratory scenarios	Section in ERG report	Incremental discounted costs	Incremental discounted QALYs	ICER £/QALY	+/- company base case
a. 100% of patients receive doxorubicin	4.2.4 and 6.2.2	[redacted]	1.75	[redacted]	[redacted]
b. 100% of patients receive paclitaxel	4.2.4 and 6.2.2	[redacted]	1.75	[redacted]	[redacted]
c. 50% of pts receive doxorubicin and 50% receive paclitaxel	4.2.4 and 6.2.2	[redacted]	1.75	[redacted]	[redacted]
Treatment waning assumed for PEM+LEN (waning from year 2 to 5)	4.2.6.3 and 6.2.3	[redacted]	0.56	[redacted]	[redacted]
No dose reduction for lenvatinib (20mg weekly assumed to be maintained)	4.2.9.1 and 6.2.4	[redacted]	1.75	[redacted]	[redacted]
Overall survival					
a. OS extrapolated using best-fitting one-piece model for both treatment arms (Log-Normal curve used for the PEM+LEN arm and Log-logistic curve used for the doxorubicin or paclitaxel arm)	4.2.6 and 6.2.5	[redacted]	0.85	[redacted]	[redacted]
b. OS for PEM+LEN (KM+Weibull)	4.2.6 and 6.2.5	[redacted]	0.81	[redacted]	[redacted]
c. OS for doxorubicin or paclitaxel (KM +Log logistic)	4.2.6 and 6.2.5	[redacted]	1.31	[redacted]	[redacted]
d. (b) & (c) combined	6.2.5	[redacted]	0.37	[redacted]	[redacted]
Health state utilities based on progression status	4.2.8 and 6.2.6	[redacted]	1.59	[redacted]	[redacted]

Exploratory scenarios	Section in ERG report	Incremental discounted costs	Incremental discounted QALYs	ICER £/QALY	+/- company base case
ToT capped by PFS	4.2.7 and 6.2.7	[redacted]	1.76	[redacted]	[redacted]
Patient weight increased to 85 kg (and BSA to 1.96 m^2)	4.2.3 and 6.2.8	[redacted]	1.75	[redacted]	[redacted]
Patient age increased to 75 years	4.2.3 and 6.2.9	[redacted]	1.55	[redacted]	[redacted]

Abbreviations: BSA, body surface area; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival; QALY, quality adjusted life year; ToT, time-on-treatment

Table 29: ERG's exploratory analyses (probabilistic)

Exploratory scenarios	Section in ERG report	Incremental discounted costs	Incremental discounted QALYs	ICER £/QALY	+/- company base case
Company base case	5.1.1.1	[redacted]	1.77	[redacted]	---
Approach to capping the survival function					
a. Simple capping method	4.2.6.2 and 6.2.1	[redacted]	1.83	[redacted]	[redacted]
b. Hazards capping method	4.2.6.2 and 6.2.1	[redacted]	1.77	[redacted]	[redacted]
Comparator					
a. 100% of patients receive doxorubicin	4.2.4 and 6.2.2	[redacted]	1.77	[redacted]	[redacted]
b. 100% of patients receive paclitaxel	4.2.4 and 6.2.2	[redacted]	1.76	[redacted]	[redacted]
c. 50% of pts receive doxorubicin and 50% receive paclitaxel	4.2.4 and 6.2.2	[redacted]	1.75	[redacted]	[redacted]
Treatment waning assumed for PEM+LEN (waning from year 2 to 5)	4.2.6.3 and 6.2.3	[redacted]	0.57	[redacted]	[redacted]
No dose reduction for lenvatinib (20mg weekly assumed to be maintained)	4.2.9.1 and 6.2.4	[redacted]	1.77	[redacted]	[redacted]
Overall survival					

Exploratory scenarios	Section in ERG report	Incremental discounted costs	Incremental discounted QALYs	ICER £/QALY	+/- company base case
a. OS extrapolated using best-fitting one-piece model for both treatment arms (Log-Normal curve used for the PEM+LEN arm and Log-logistic curve used for the doxorubicin or paclitaxel arm)	4.2.6 and 6.2.5	[redacted]	0.86	[redacted]	[redacted]
b. OS for PEM+LEN (KM+Weibull)	4.2.6 and 6.2.5	[redacted]	0.85	[redacted]	[redacted]
c. OS for doxorubicin or paclitaxel (KM +Log logistic)	4.2.6 and 6.2.5	[redacted]	1.32	[redacted]	[redacted]
d. b) & (c) combined	6.2.5	[redacted]	0.38	[redacted]	[redacted]
Health state utilities based on progression status	4.2.8 and 6.2.6	[redacted]	1.61	[redacted]	[redacted]
ToT capped by PFS	4.2.7 and 6.2.7	[redacted]	1.76	[redacted]	[redacted]
Patient weight increased to 85 kg (and BSA to 1.96 m^2)	4.2.3 and 6.2.8	[redacted]	1.76	[redacted]	[redacted]
Patient age increased to 75 years	4.2.3 and 6.2.9	[redacted]	1.56	[redacted]	[redacted]

Abbreviations: BSA, body surface area; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; OS, overall survival; PEM+LEN, pembrolizumab with lenvatinib; PFS, progression-free survival; QALY, quality adjusted life year; ToT, time-on-treatment

6.3. ERG's preferred assumptions

The ERG's preferred base case results are presented below. All results are based on the updated pembrolizumab PAS and lenvatinib list price.

Table 30: ERG preferred assumptions (deterministic)

	ERG report section	Incremental cost	Incremental QALYs	ICER
Company's base case	5.1.1.1	[redacted]	1.75	[redacted]
ERG corrected company base case				
Drug costs corrected + additional PAS	6.1	[redacted]	1.75	[redacted]
ERG Preferred base case assumptions (applied individually)				
Survival capped by hazards	6.2.1 and 6.2.10	[redacted]	1.77	[redacted]
50% of patients receive doxorubicin and 50% receive paclitaxel	6.2.2 and 6.2.10	[redacted]	1.75	[redacted]
ToT capped by PFS	6.2.7 and 6.2.10	[redacted]	1.76	[redacted]
Health state utilities based on progression status	6.2.6 and 6.2.10	[redacted]	1.76	[redacted]
Patient weight increased to 85 kg (and BSA to 1.96 m^2)	6.2.8 and 6.2.10	[redacted]	1.75	[redacted]
Patient age increased to 75 years	6.2.9 and 6.2.10	[redacted]	1.55	[redacted]
OS for doxorubicin or paclitaxel (KM +Log logistic)	6.2.5 and 6.2.10	[redacted]	1.31	[redacted]
Cumulative impact of ERG's preferences	1.7 and 6.3	[redacted]	1.07	[redacted]

Abbreviations: BSA, body surface area; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year; ToT, time-on-treatment

Table 31: ERG preferred assumptions (probabilistic)

	ERG report section	Incremental cost	Incremental QALYs	ICER
Company's base case	5.1.1.1	[redacted]	1.77	[redacted]
ERG corrected company base case				
Drug costs corrected + additional PAS	6.1	[redacted]	1.76	[redacted]
ERG Preferred base case assumptions (applied incrementally)				
Survival capped by hazards	6.2.1 and 6.2.10	[redacted]	1.77	[redacted]
50% of patients receive doxorubicin and 50% receive paclitaxel	6.2.2 and 6.2.10	[redacted]	1.75	[redacted]

	ERG report section	Incremental cost	Incremental QALYs	ICER
ToT capped by PFS	6.2.7 and 6.2.10	[redacted]	1.76	[redacted]
Health state utilities based on progression status	6.2.6 and 6.2.10	[redacted]	1.61	[redacted]
Patient weight increased to 85 kg (and BSA to 1.96 m^2)	6.2.8 and 6.2.10	[redacted]	1.76	[redacted]
Patient age increased to 75 years	6.2.9 and 6.2.10	[redacted]	1.56	[redacted]
OS for doxorubicin or paclitaxel (KM +Log logistic)	6.2.5 and 6.2.10	[redacted]	1.32	[redacted]
Cumulative impact of ERG's preferences	1.7 and 6.3	[redacted]	1.05	[redacted]

Abbreviations: BSA, body surface area; ERG, Evidence Review Group; ICER, incremental cost-effectiveness ratio; KM, Kaplan-Meier; OS, overall survival; PFS, progression-free survival; QALY, quality adjusted life year; ToT, time-on-treatment

6.4. Conclusions of the cost-effectiveness section

Overall, the company's model was of good quality. The company's base case yielded a deterministic ICER of [redacted] and a probabilistic ICER of [redacted] (based on the original pembrolizumab PAS as per Section 5.1.1.1). The ERG disagreed with a number of the company's base case assumptions, most of which had a minor impact on the ICER with the exception of the overall survival function for the comparator arm (TPC).

Based on clinical opinion to the ERG, the modelled long-term survival gap between PEM+LEN and TPC appeared to lack clinical plausibility. In particular, modelled OS for the TPC arm was considered to underestimate the proportion of patients alive at 5 years. In order to estimate more clinically plausible overall survival estimates, the ERG opted to use an alternative parametric curve for extrapolation (see Sections 4.2.6 and 6.2.5). In isolation, this increased the deterministic ICER by [redacted] to [redacted] (Table 28) and increased the probabilistic ICER by [redacted] to [redacted] (Table 29).

It should be noted that the results are much more sensitive to the survival function selected for the PEM+LEN arm; when KM+Weibull is assigned, the ICER increased by [redacted] to [redacted] (deterministic) and by [redacted] to [redacted] (probabilistic). Furthermore, the results are highly sensitive to a number of other scenarios the ERG explored, in particular the impact of treatment waning, which increased the deterministic ICER by [redacted] to [redacted] (Table 28) and increased the probabilistic ICER by [redacted] to [redacted] (Table 29).

However, the cumulative impact of the ERG's preferred scenario yields a deterministic ICER of [redacted] and a probabilistic ICER of [redacted].

7. END OF LIFE

The company provided several data sources to support the application of NICE end of life criteria. The ERG noted NICE end of life criteria to be as follows;

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months, and;
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment.

As noted in Section B.2.13.4 of the CS, the company refer to results from KEYNOTE-775 which reported median survival (for patients receiving current standard of care) to be 11.4 months. In the company's base case, mean survival in TPC as estimated by the company's model was [redacted] months, though the ERG's base case estimated this as [redacted] years in both deterministic and probabilistic analyses . The company further outlined survival results from ECHO, a retrospective UK chart review (note: full study details were not available to the ERG and were stated to be on file). In this study, median survival was reported to be [redacted] months for standard of care. Based on survival data from these sources, the ERG agreed that that life expectancy for the patient population under review could be plausibly less than 24 months.

Furthermore, based on overall survival data from KEYNOTE-775, median overall survival was significantly longer in the PEM+LEN group compared with the control group; 18.3 and 11.4 months respectively (demonstrating an extension of life of approximately 6.9 months).

The ERG sought further clinical opinion to determine whether end of life criteria would be met if separate subgroups were to be considered i.e. according to dMMR and pMMR status. Clinical opinion noted that average life expectancy is likely to be less than 24 months for each subpopulation, PEM+LEN would result in an extension of life of at least an additional 3 months and that patient numbers are sufficiently small.

Based on the evidence provided by the company and clinical opinion received, the ERG considered that it may be appropriate to consider NICE end of life criteria for this appraisal, though the choice of extrapolation in the TPC arm is potentially dispositive.

References

1. Morrison J, Balega J, Buckley L, Clamp A, Crosbie E, Drew Y, et al. BGCS uterine cancer guidelines: Recommendations for practice, 2021: British Gynaecological Cancer Society. Available from: https://www.bgcs.org.uk/wp-content/uploads/2021/11/British-Gynaecological-Cancer-Society-v13-for-website-with-figure1.pdf.
2. Crosbie EJ, Kitson SJ, McAlpine JN, Mukhopadhyay A, Powell ME, Singh N. Endometrial cancer. Lancet. 2022;399(10333):1412-28.
3. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Dostarlimab for previously treated advanced or recurrent endometrial cancer with high microsatellite instability or mismatch repair deficiency [TA779], 2022. Available from: https://www.nice.org.uk/guidance/ta779.
4. National Institute for Health and Care Excellence (NICE). Pembrolizumab with lenvatinib for previously treated advanced, metastatic or recurrent endometrial cancer. Final scope, 2021. Available from: https://www.nice.org.uk/guidance/gid-ta10692/documents/final-scope.
5. Makker V, Colombo N, Casado Herráez A, Santin A, Colomba E, Miller D, et al. A multicenter, open-label, randomized, phase 3 study to compare the efficacy and safety of lenvatinib in combination with pembrolizumab vs treatment of physician's choice in patients with advanced endometrial cancer. Gynecologic Oncology. 2021;162:S4.
6. Makker V, Taylor MH, Aghajanian C, Oaknin A, Mier J, Cohn AL, et al. Lenvatinib plus pembrolizumab in patients with advanced endometrial cancer. Journal of Clinical Oncology. 2020;38(26):2981-92.
7. Julius JM, Tanyi JL, Nogueras-Gonzalez GM, Watkins JL, Coleman RL, Wolf JK, et al. Evaluation of pegylated liposomal doxorubicin dose on the adverse drug event profile and outcomes in treatment of recurrent endometrial cancer. International Journal of Gynecological Cancer. 2013;23(2):348-54.
8. Moreira E, Paulino E, Ingles Garces Á H, Fontes Dias MS, Saramago M, de Moraes Lino da Silva F, et al. Efficacy of doxorubicin after progression on carboplatin and paclitaxel in advanced or recurrent endometrial cancer: a retrospective analysis of patients treated at the Brazilian National Cancer Institute (INCA). Medical Oncology. 2018;35(3):20.
9. Nagao S, Nishio S, Michimae H, Tanabe H, Okada S, Otsuki T, et al. Applicability of the concept of "platinum sensitivity" to recurrent endometrial cancer: the SGSG-012/GOTIC-004/Intergroup study. Gynecologic Oncology. 2013;131(3):567-73.
10. Rubinstein M, Halpenny D, Makker V, Grisham RN, Aghajanian C, Cadoo K. Retreatment with carboplatin and paclitaxel for recurrent endometrial cancer: A retrospective study of the Memorial Sloan Kettering Cancer Center experience. Gynecologic Oncology Reports. 2019;28:120-3.
11. Sovak MA, Dupont J, Hensley ML, Ishill N, Gerst S, Abu-Rustum N, et al. Paclitaxel and carboplatin in the treatment of advanced or recurrent endometrial cancer: a large retrospective study. International Journal of Gynecological Cancer. 2007;17(1):197-203.
12. Ueda Y, Miyake T, Egawa-Takata T, Miyatake T, Matsuzaki S, Yokoyama T, et al. Second-line chemotherapy for advanced or recurrent endometrial carcinoma previously treated with paclitaxel and carboplatin, with or without epirubicin. Cancer Chemotherapy and Pharmacology. 2011;67(4):829-35.
13. Endometrial Cancer Health Outcomes. Europe (ECHO EU5): A retrospective, multicenter chart review study evaluating treatment patterns and clinical outcomes in advanced or recurrent EC (aEC) patients previously treated with systemic therapy in United Kingdom. Unpublished. 2021.

14. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.
15. Merck Sharp Dohme. A multicenter, open-label, randomized, phase 3 trial to compare the efficacy and safety of lenvatinib in combination with pembrolizumab versus treatment of physician's choice in participants with advanced endometrial cancer. Clinical Study Report. 2021.
16. Makker V, Colombo N, Casado Herráez A, Santin AD, Colomba E, Miller DS, et al. Lenvatinib plus pembrolizumab for advanced endometrial cancer. New England Journal of Medicine. 2022;386(5):437-48.
17. Makker V, Aghajanian C, Cohn AL, Romeo M, Bratos R, Brose MS, et al. Lenvatinib and pembrolizumab in advanced endometrial carcinoma (EC): long-term efficacy and safety update from a phase 1b/2 study. Journal for ImmunoTherapy of Cancer. 2021;9(354):A381.
18. Van Hout B, Janssen MF, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health. 2012;15(5):708-15.
19. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Niraparib for maintenance treatment of relapsed, platinum-sensitive ovarian, fallopian tube and peritoneal cancer [TA528], 2018. Available from: https://www.nice.org.uk/guidance/TA528.
20. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Niraparib for maintenance treatment of advanced ovarian, fallopian tube and peritoneal cancer after response to first-line platinum-based chemotherapy [TA673], 2021. Available from: https://www.nice.org.uk/guidance/TA673.
21. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Rucaparib for maintenance treatment of relapsed platinum-sensitive ovarian, fallopian tube or peritoneal cancer [TA611], 2019. Available from: https://www.nice.org.uk/guidance/ta611.
22. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Olaparib for maintenance treatment of relapsed platinum-sensitive ovarian, fallopian tube or peritoneal cancer [TA620], 2020. Available from: https://www.nice.org.uk/guidance/ta620.
23. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. Processes and methods, 2013. Available from: https://www.nice.org.uk/process/pmg9.
24. NICE DSU. NICE DSU Technical Support Document 14: Survival analysis for economic evaluations alongside clinical trials - extrapolation with patient level data, 2013: Decision Support Unit, ScHARR, University of Sheffield. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf.
25. NICE DSU. NICE DSU Technical Support Document 21: Flexible methods for survival analysis, 2020: Decision Support Unit, ScHARR, University of Sheffield. Available from: http://nicedsu.org.uk/wp-content/uploads/2020/11/NICE-DSU-Flex-Surv-TSD-21_Final_alt_text.pdf.
26. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer [TA531], 2018. Available from: https://www.nice.org.uk/guidance/ta531.
27. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Pembrolizumab for treating advanced melanoma after disease progression with ipilimumab [TA357], 2017. Available from: https://www.nice.org.uk/guidance/ta357.

28. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Avelumab with axitinib for untreated advanced or metastatic renal cell carcinoma [TA645], 2020. Available from: https://www.nice.org.uk/guidance/ta645/.
29. Georghiou T, Bardsley M. Exploring the cost of care at the end of life, 2014: Nuffield Trust. Available from: https://www.nuffieldtrust.org.uk/files/2017-01/end-of-life-care-web-final.pdf.
30. National Institute for Health and Care Excellence (NICE). Technology appraisal guidance. Olaparib for maintenance treatment of BRCA mutation-positive advanced ovarian, fallopian tube or peritoneal cancer after response to first-line platinum-based chemotherapy [TA598], 2019. Available from: https://www.nice.org.uk/guidance/ta598.
31. Guest JF, Ruiz FJ, Greener MJ, Trotman IF. Palliative care treatment patterns and associated costs of healthcare resource use for specific advanced cancer patients in the UK. European Journal of Cancer. 2006;15(1):65-73.
32. Calculator.net. Body surface calculator, 2022. Available from: https://www.calculator.net/body-surface-area-calculator.html.
33. NHS Digital. Health Survey for England 2019. Overweight and obesity in adults and children, 2020. Available from: https://files.digital.nhs.uk/9D/4195D5/HSE19-Overweight-obesity-rep.pdf.
34. Mosteller RD. Simplified calculation of body-surface area. New England Journal of Medicine. 1987;317(17):1098.
35. Cancer Research UK. Uterine cancer incidence statistics, 2021. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/uterine-cancer/incidence.

Appendix A: Additional searches conducted by ERG

Additional searches conducted by ERG

The ERG conducted an additional search to test the impact of the exclusion of conference abstracts in Ovid Embase in the SLR of interventional evidence. This search retrieved 454 results, and these were single screened by the Information Specialist, with 38 records selected for further consideration. Two reviewers independently screened the 38 records. The ERG considered that the additional abstracts, while potentially eligible for the interventional SLR, did not provide additional data that would enhance the already identified clinical effectiveness evidence base.

The search strategy for Ovid Embase is provided below:

Embase <1974 to 2022 May 06>

1	exp *endometrium carcinoma/	14269
2	exp *endometrium cancer/	32580
3	((endometrium or endometrial) adj1 (cancer* or carcinoma* or tumo?r* or neoplasm*)).ti.	24628
4	((endometrium or endometr.ial) adj1 adenocarcinoma*).ti.	1711
5	((endometrium or endometrial) adj1 (metastasis or metastatic*)).ti.	246
6	or/1-5	35843
7	Clinical Trial/	1033747
8	Randomised Controlled Trial/	707273
9	controlled clinical trial/	465536
10	multicenter study/	322509
11	Phase 3 clinical trial/	60337
12	Phase 4 clinical trial/	4742
13	exp RANDOMISATION/	94020
14	Single Blind Procedure/	46022
15	Double Blind Procedure/	194618
16	Crossover Procedure/	70225
17	PLACEBO/	379874
18	randomi?ed controlled trial$.tw.	284401
19	rct.tw.	46580
20	(random$ adj2 allocat$).tw.	49830
21	single blind$.tw.	28771
22	double blind$.tw.	229762
23	((treble or triple) adj blind$).tw.	1547
24	placebo$.tw.	342325
25	Prospective Study/	763144
26	(single-arm or single arm).tw.	22768
27	(Phase II or Phase 2).tw.	148543
28	Phase 2 clinical trial/	96477
29	or/7-28	2762428

30 Case Study/ 85034
31 case report.tw. 484899
32 letter/ 1147956
33 Editorial.pt. 725072
34 Letter.pt. 1222404
35 Note.pt. 892557
36 or/30-35 3398463
37 29 not 36 2629509
38 6 and 37 3605
39 exp endometrium carcinoma/ 22335
40 exp endometrium cancer/ 55826
41 ((endometrium or endometrial) adj3 (cancer* or carcinoma* or tumo?r* or neoplasm*)).ti,ab. 46815
42 ((endometrium or endometrial) adj3 adenocarcinoma*).ti,ab. 6030
43 ((endometrium or endometrial) adj3 (metastasis or metastatic*)).ti,ab. 1566
44 or/39-43 66646
45 Clinical Trial/ 1033747
46 Randomised Controlled Trial/ 707273
47 controlled clinical trial/ 465536
48 multicenter study/ 322509
49 Phase 3 clinical trial/ 60337
50 Phase 4 clinical trial/ 4742
51 exp RANDOMISATION/ 94020
52 Single Blind Procedure/ 46022
53 Double Blind Procedure/ 194618
54 Crossover Procedure/ 70225
55 PLACEBO/ 379874
56 randomi?ed controlled trial$.tw. 284401
57 rct.tw. 46580
58 (random$ adj2 allocat$).tw. 49830
59 single blind$.tw. 28771
60 double blind$.tw. 229762
61 ((treble or triple) adj blind$).tw. 1547
62 placebo$.tw. 342325
63 Prospective Study/ 763144
64 (single-arm or single arm).tw. 22768
65 (Phase II or Phase 2).tw. 148543
66 Phase 2 clinical trial/ 96477
67 or/45-66 2762428
68 Case Study/ 85034
69 case report.tw. 484899
70 abstract report/ or letter/ 1237405
71 Conference proceeding.pt. 0
72 Conference abstract.pt. 4389743
73 Editorial.pt. 725072
74 Letter.pt. 1222404
75 Note.pt. 892557
76 or/68-75 7738096
77 67 not 76 1998071
78 44 and 77 6279
79 38 not 78 1227

80	(advanced or recurrent or metastatic or inoperable or irresectable or unresectable or resistant or progressive).ti,ab.	2479879
81	79 and 80	454