in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma [ID1609]

Produced by Kleijnen Systematic Reviews Ltd. in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University

Authors
Nigel Armstrong, Health Economist, KSR Ltd
Sabine Grimm, Health Economist, Maastricht UMC, the Netherlands
Bram Ramaekers, Health Economist, Maastricht UMC, the Netherlands
Marie Westwood, Reviews Manager, KSR Ltd
Annette Chalker, Systematic Reviewer, KSR Ltd
Mohammed Islam, Health Economist, KSR Ltd
Gill Worthy, Statistician, KSR Ltd
Shelley De Kock, Information Specialist, KSR Ltd
Manuela Joore, Health Economist, Maastricht UMC
Jos Kleijnen, Director, KSR Ltd, Professor of Systematic Reviews in Health Care, Maastricht University

Correspondence to
Nigel Armstrong, Kleijnen Systematic Reviews
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, United Kingdom
YO19 6FD

Date completed 17/03/2021

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number STA 13/30/12.

Declared competing interests of the authors

None.

Acknowledgements

Commercial in confidence (CiC) data are highlighted in blue throughout the report.

Academic in confidence (AiC) data are highlighted in yellow throughout the report.

Confidential comparator prices are highlighted in green throughout the report.

Any de-personalised data are highlighted in pink throughout the report.

Rider on responsibility for report

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

This report should be referenced as follows:

Armstrong N, Grimm S, Ramaekers B, Westwood M, Chalker A, Islam M, Worthy G, De Kock S, Joore MA, Kleijnen J. Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma [ID1609]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2021.

Contributions of authors

Nigel Armstrong acted as project lead, systematic reviewer and health economist on this assessment, critiqued the clinical effectiveness methods and evidence and the company's economic evaluation and contributed to the writing of the report. Sabine Grimm acted as health economic project lead, critiqued the company's economic evaluation and contributed to the writing of the report. Bram Ramaekers and Mohammed Islam acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Marie Westwood and Annette Chalker acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Gill Worthy acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Shelley De Kock critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AE	Adverse event
AIC	Akaike Information Criterion
ALT	Alanine aminotransferase
ASBI	Average Symptom Burden Index
AST	Aspartate aminotransferase
BCS	Best case scenario
BI	Budget impact
BIC	Bayesian information criterion
BICR	Blinded independent central review
BMJ	British Medical Journal
BMS	Bristol-Myers Squibb
BTS	British Thoracic Society
CADTH	Canadian Agency for Drugs and Technologies in Health
CE	Cost effectiveness
CEA	Cost effectiveness analysis
CEAC	Cost effectiveness acceptability curve
CHMP	Committee for Medicinal Products for Human Use
CI	Confidence interval
CNS	Central nervous system
CR	Complete response
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTR	Clinical trial results
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EMA	European Medicines Agency
eMIT	Electronic market information tool
EORTC	European Organisation for Research and Treatment of Cancer
EPAR	European Public Assessment Report
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-3L	European Quality of Life-5 Dimensions 3 levels
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
EUR	Erasmus University Rotterdam
FAS	Full analysis set
FAD	Final appraisal document
FDA	Food and Drug Administration
GHS	Global health status
HR	Hazard ratio
HRQoL	Health-related quality of life
HSUV	Health state utility value
HTA	Health technology assessment
IC	Investigator choice
ICD	International Statistical Classification of Diseases and Related Health Problems
ICER	Incremental cost effectiveness ratio
IHC	Immunohistochemistry
IQR	Interquartile range
ITC	Indirect treatment comparison
ITT	Intention to treat
IV	Intravenous
KM	Kaplan-Meier

KSR	Kleijnen Systematic Reviews
LCSS Meso	Lung Cancer Symptom Scale–Mesothelioma
LYs	Life years
LYG	Life years gained
MeSH	Medical subject headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MOS SF-36	Medical Outcomes Study Short Form Survey
MPM	Malignant pleural mesothelioma
MTA	Multiple technology appraisal
MTC	Mixed treatment comparison
NA	Not applicable
NCCN	National Comprehensive Cancer Network
NCRI	National Cancer Research Institute
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not reported
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PDC	Platinum doublet chemotherapy
PD-L1	Programmed death-ligand 1
PFS	Progression-free survival
PH	Proportional hazards
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
PRO	Patient-reported outcome
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSM	Partition survival model
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
Q3W	Every three weeks
QALY	Quality-adjusted life year
QLQ-C30	Quality of Life Questionnaire
QoL	Quality of life
RCT	Randomised controlled trial
RECIST	Response Evaluation Criteria in Solid Tumours
RR	Relative risk; Risk ratio
SAE	Serious adverse events
SC	Subcutaneous
ScHARR	School of Health and Related Research
SD	Standard deviation
SE	Standard error
SIGN	Scottish Intercollegiate Guidelines Network
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SmPC	Summary of product characteristics
SoC	Standard of care
STA	Single technology appraisal
STM	State transition model
TA	Technology assessment
TEAE	Treatment emergent adverse events

TRAE	Treatment-related adverse event
TSD	Technical Support Document
UK	United Kingdom
UMC	University Medical Centre
VAS	Visual analogue scale

Table of Contents

Abbreviations **3**

Table of Tables **8**

Table of Figures **10**

1. EXECUTIVE SUMMARY **11**

1.1 Overview of the ERG's key issues 11
1.2 Overview of key model outcomes 12
1.3 The decision problem: summary of the ERG's key issues 12
1.4 The clinical effectiveness evidence: summary of the ERG's key issues 13
1.5 The cost effectiveness evidence: summary of the ERG's key issues 14
1.6 Other key issues: summary of the ERG's view 19
1.7 Summary of the ERG's view 19

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM **21**

2.1 Population 25
2.2 Intervention 25
2.3 Comparators 26
2.4 Outcomes 27
2.5 Other relevant factors 27

3. CLINICAL EFFECTIVENESS **28**

3.1 Critique of the methods of review(s) 28
3.1.1 Searches 28
3.1.2 Inclusion criteria 30
3.1.3 Data extraction 32
3.1.4 Quality assessment 32
3.1.5 Evidence synthesis 32
3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these) 32
3.2.1 Design (including statistical analyses) of CheckMate-743 trial 32
3.2.2 Baseline characteristics of CheckMate-743 trial 36
3.2.3 Quality of CheckMate-743 trial 37
3.2.4 Results of CheckMate-743 trial 38
3.2.5 Subgroup analyses 49
3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison 52
3.4 Critique of the indirect comparison and/or multiple treatment comparison 52
3.5 Additional work on clinical effectiveness undertaken by the ERG 52
3.6 Conclusions of the clinical effectiveness section 52

4. COST EFFECTIVENESS **55**

4.1 ERG comment on company's review of cost effectiveness evidence 55
4.1.1 Searches performed for cost effectiveness section 55
4.1.2 Inclusion/exclusion criteria 57
4.1.3 Conclusions of the cost effectiveness review 57
4.2 Summary and critique of company's submitted economic evaluation by the ERG 58

4.2.1 NICE reference case checklist 58
4.2.2 Model structure 58
4.2.3 Population 60
4.2.4 Interventions and comparators 60
4.2.5 Perspective, time horizon and discounting 61
4.2.6 Treatment effectiveness and extrapolation 61
4.2.7 Adverse events 65
4.2.8 Health-related quality of life 65
4.2.9 Resources and costs 66

5. COST EFFECTIVENESS RESULTS 72

5.1 Company's cost effectiveness results 72
5.1.1 Company's subgroup analyses 74
5.2 Company's sensitivity and scenario analyses 76
5.3 Model validation and face validity check 77
5.3.1 Face validity assessment 77
5.3.2 Technical verification 77
5.3.3 Comparisons with other technology appraisals 77
5.3.4 Comparison with external data 77

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES 78

6.1 Exploratory and sensitivity analyses undertaken by the ERG 78
6.1.1 ERG base-case 78
6.1.2 ERG exploratory scenario analyses 79
6.1.3 ERG subgroup analyses 79
6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG 82
6.3 ERG's preferred assumptions 87
6.4 Conclusions of the cost effectiveness section 87

7. END OF LIFE 89

8. REFERENCES 90

Table of Tables

Table 1.1: Summary of key issues 11

Table 1.2: Key issue 1 Effectiveness and safety of expected nivolumab fixed dosing 12

Table 1.3: Key issue 2 Applicability of comparator 13

Table 1.4: Key issue 3 Immaturity of CheckMate-743 trial outcomes 13

Table 1.5: Key issue 4 Subsequent therapy 14

Table 1.6: Key issue 5 Subgroup effectiveness of nivolumab + ipilimumab according to PD-L1 status and histology 14

Table 1.7: Key issue 6 Model structure - the use of a PSM, without a STM approach to verify the results 15

Table 1.8: Key issue 7 Population – no subgroup cost effectiveness analyses presented 15

Table 1.9: Key issue 8 Intervention and comparators – two-year stopping rule may not be completely adhered to in trial 15

Table 1.10: Key issue 9 Treatment effectiveness and extrapolation – immaturity of the long-term PFS and OS data 16

Table 1.11: Key issue 10 Health-related quality of life – duration of utility benefits for nivolumab + ipilimumab 16

Table 1.12: Key issue 11 Resources and costs – estimation of time to treatment discontinuation 17

Table 1.13: Key issue 12 Resources and costs – uncertainty about subsequent treatments 17

Table 1.14: Key issue 13 Resources and costs – adverse events 18

Table 1.15: Key issue 14 Company's cost effectiveness results – proportion of (PF)LY accumulated beyond the observed data 18

Table 1.16: Summary of ERG's preferred assumptions and ICER 19

Table 2.1: Statement of the decision problem (as presented by the company) 21

Table 3.1: Data sources for the clinical effectiveness systematic review (as reported in CS and response to clarification) 28

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence 30

Table 3.3: CheckMate-743: study design 32

Table 3.4: CheckMate-743 statistical analyses 34

Table 3.5: CheckMate-743: baseline demographics (all randomised patients) 36

Table 3.6: Quality assessment of CheckMate-743 (NCT02899299) 37

Table 3.7: CheckMate-743: overall survival rates – all randomised patients 39

Table 3.8: Response rate per BICR 40

Table 3.9: CheckMate-743: safety summary – all treated patients 41

Table 3.10: CheckMate-743: treatment-related adverse events – all treated patients 42

Table 3.11: All-cause adverse events of grade 3 or 4 severity (≥1%) 43

Table 3.12: CheckMate-743: summary of deaths – all treated patients 46

Table 3.13: Subgroup analyses by PD-L1 status 50

Table 3.14: Subgroup analyses by histological subtype 51

Table 4.1: Data sources for the cost effectiveness systematic review 55

Table 4.2: NICE reference case checklist 58

Table 4.3: Health state utility values 66

Table 4.4: Costs per weekly cycle 68

Table 6.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 5.1) 80

Table 6.2: ERG base-case (deterministic unless indicated) 83

Table 6.3: Deterministic scenario analyses (conditional on ERG base-case) 84

Table 6.4: Probabilistic scenario analyses (conditional on ERG base-case, 1,000 iterations unless stated otherwise) 85

Table of Figures

Figure 3.1: CheckMate-743: Kaplan-Meier plot of overall survival (all randomised patients)..... 39

Figure 3.2: CheckMate-743: Kaplan-Meier plot of progression-free survival by blinded independent central review (all randomised patients) 40

Figure 3.3: EQ-5D-3L Utility Index: mean change from baseline scores by treatment group (patient-reported outcome analysis population) 47

Figure 3.4: EQ-5D VAS: mean change from baseline scores by treatment group (patient-reported outcome analysis population)..... 48

Figure 4.1: Model structure..... 59

Figure 5.1: CS base-case cost effectiveness plane..... 72

Figure 5.2: PSA convergence plot for ERG base-case 77

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. If possible, 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 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues related to the cost effectiveness. Other key issues are discussed in Section 1.6 while a summary in presented in Section 1.7.

Background information on the condition, technology and evidence and information on key as well as non-key issues are in the main ERG report, see Sections 2 (background), 3 (decision problem), 4 (clinical effectiveness) and 5 (cost effectiveness) for more details.

All issues identified represent the ERG's view, not the opinion of the National Institute for Health and Care Excellence (NICE).

1.1 Overview of the ERG's key issues

Table 1.1: Summary of key issues

ID1457	Summary of issue	Report sections
1	Effectiveness and safety of expected nivolumab fixed dosing	Section 2.2
2	Applicability of comparator to English NHS practice	Section 2.3
3	Immaturity of CheckMate-743 trial outcomes	Section 3.2.4
4	Subsequent therapy: difference between arms and applicability to English NHS practice	Section 3.2.4
5	Subgroup effectiveness of nivolumab + ipilimumab according to PD-L1 status and histology	Section 3.2.5
6	Model structure - the use of a PSM, without a STM approach to verify the results	Section 4.2.2
7	Population – no subgroup cost effectiveness analyses presented	Section 4.2.3
8	Intervention & comparators – two-year stopping rule may not be completely adhered to in trial	Section 4.2.4
9	Treatment effectiveness and extrapolation – immaturity of the long-term PFS and OS data	Section 4.2.6
10	Health-related quality of life – duration of utility benefits for nivolumab + ipilimumab	Section 4.2.8
11	Resources and costs – estimation of time to treatment discontinuation	Section 4.2.9
12	Resources and costs – uncertainty about subsequent treatments	Section 4.2.9
13	Resources and costs – adverse events	Section 4.2.9
14	Company's cost effectiveness results – proportion of PF LYs accumulated beyond the observed data	Section 5.1

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are matters of judgement relating to the effectiveness and relative effectiveness of nivolumab + ipilimumab versus platinum doublet chemotherapy (PDC) (regarding overall survival (OS) and progression-free survival (PFS)) and the long-term impact on health-related quality of life

(HRQoL). Further differences are in the estimation of costs regarding assumptions about time on treatment, subsequent treatments and adverse events (AEs).

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (OS) 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:

- Increased mean PFS (undiscounted time in the progression-free (PF) health state: 18.5 vs 10.5 months) and mean OS (undiscounted survival: 34.4 vs 20.6 months) compared with PDC.
- Increased health state utility values for the PF (0.74 vs 0.73) and PD (0.65 vs 0.58) health states compared with PDC.
- The PFS, OS and health state utility benefits are maintained for the whole duration of the time horizon (i.e. no waning of these treatment benefits).

Overall, the technology is modelled to affect costs by:

- its higher unit price than PDC prices
- cost-savings through delayed more severe health state costs and subsequent treatment costs
- potentially less costly subsequent treatments (uncertain) and potentially AEs (direction uncertain).

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

- treatment waning from five years onwards
- using the log-logistic distribution for estimating OS in the PDC arm
- using time to discontinuation (TTD) estimates with 100% dose intensity instead of the number of mean doses approach.

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

The ERG is reasonably satisfied that the population, which includes Eastern Cooperative Oncology Group (ECOG) performance status (PS) 0-1, as specified in the decision problem matches that in the CheckMate-743 trial and, although narrower than that in the scope, is appropriate. The ERG is also satisfied that this narrower population is consistent with the omission of best supportive care (BSC) as a comparator. The company have also provided evidence sufficient to support the exclusion of raltitrexed as a relevant comparator for the National Health Service (NHS) in England. This leaves two remaining key issues, shown in Tables 1.2 and 1.3.

Table 1.2: Key issue 1 Effectiveness and safety of expected nivolumab fixed dosing

Report section	Section 2.2
Description of issue and why the ERG has identified it as important	The effect of fixed dosing vs. weight-based dosing, as used in CheckMate-743, is uncertain.
What alternative approach has the ERG suggested?	The ERG requested evidence to support the relative efficacy and safety of the two dosing regimens. However, the evidence provided by the company lacked clarity or was not appropriate.
What is the expected effect on the cost effectiveness estimates?	Unknown

Report section	Section 2.2
What additional evidence or analyses might help to resolve this key issue?	The company could provide clarification regarding the analyses that they referred to in the response to clarification. There is also the possibility that further evidence exists that compares the two methods of dosing.

Table 1.3: Key issue 2 Applicability of comparator

Report section	Section 2.3
Description of issue and why the ERG has identified it as important	The extent to which the clinical judgments made as to investigator choice of PDC, i.e. carboplatin or cisplatin, in CheckMate-743 match those that would be made in English NHS practice is uncertain.
What alternative approach has the ERG suggested?	The ERG requested evidence as to the degree of consistency, to which the company responded by providing the proportion of patients in the UK who have received the two platinum-based treatments. However, because there appeared to be considerable variation between sources, the uncertainty remains unresolved.
What is the expected effect on the cost effectiveness estimates?	Unknown, with likely small impact on cost.
What additional evidence or analyses might help to resolve this key issue?	The ERG cannot conceive of a way to reduce the uncertainty and therefore this issue will probably subject to the application of judgment.

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

The CheckMate-743 trial is a relatively high-quality source of evidence to inform effect estimates for the outcomes listed in the scope for the comparison between nivolumab + ipilimumab and the most appropriate comparators, as explained in Section 1.3. However, there remain two key issues, as shown in Tables 1.4, 1.5 and 1.6.

Table 1.4: Key issue 3 Immaturity of CheckMate-743 trial outcomes

Report section	Section 3.2.4
Description of issue and why the ERG has identified it as important	The only results that have been presented are for an interim analysis with a database lock 3 April 2020.
What alternative approach has the ERG suggested?	The ERG asked for the results from a later data-cut, but the company stated that no further results were available and did not provide a date for their submission.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	The uncertainty in all outcomes especially OS and PFS and for subgroups (see Key issue 5) would be reduced considerably by the provision of updated results.

Table 1.5: Key issue 4 Subsequent therapy

Report section	Section 3.2.2
Description of issue and why the ERG has identified it as important	There was a difference in the number of patients taking each type of subsequent therapy between the nivolumab + ipilimumab and PDC arms of CheckMate-743 and, apparently, between the PDC arm and UK clinical experience.
What alternative approach has the ERG suggested?	The ERG requested evidence as to the effect that the differences described above may have and for the comparison with English NHS practice. However, the ERG could not validate the results regarding time survived on subsequent therapy or the nature of that subsequent therapy based on the reference provided.[1] With the FAC, the poster for that reference has now been provided to enable the ERG to validate the figures provided by the company. Nevertheless, the figures for percentage receiving each type of subsequent therapy received in UK clinical practice provided do appear to be quite different to those in the PDC arm of the CheckMate-743 trial.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	It is unlikely that the effect of any difference in subsequent therapy between the trial arms or between the PDC arm and English NHS practice can be estimated with any confidence.

Table 1.6: Key issue 5 Subgroup effectiveness of nivolumab + ipilimumab according to PD-L1 status and histology

Report section	Section 3.25
Description of issue and why the ERG has identified it as important	Subgroup analysis by both PD-L1 status and histology, which was included in the scope, reveals potential variation and in some cases 95% CIs that overlap the point of no difference for nivolumab + ipilimumab versus PDC for both OS and PFS. This is particularly the case for PD-L1<1% where for PFS there is little uncertainty (point estimate for HR greater than 1 and 95% CI does not include 1) that PDC is superior and for OS where there appears to be little difference between groups (95% CI includes 1).
What alternative approach has the ERG suggested?	No alternative approach has been suggested by the ERG other than to provide results from a later data-cut (see Key issue 4).
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	Given the current evidence, uncertainty would be reduced by considering analysing the decision problem using combined PD-L1 status and histology subgroups. It would also be reduced by submission of more complete results i.e. at a later data-cut.

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

The company's cost effectiveness model was well built and complied with the NICE reference case. The main critique points are modelling choices and assumptions. The overarching challenge was the immaturity of the data from CheckMate-743, which results in the ICER being very uncertain. A full

summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness results are presented in Section 5, the ERG's summary and detailed critique in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The key issues in the cost effectiveness evidence are discussed in Tables 1.7 to 1.15.

Table 1.7: Key issue 6 Model structure - the use of a PSM, without a STM approach to verify the results

Report section	Section 4.2.2
Description of issue and why the ERG has identified it as important	NICE TSD 19 recommended the use of state transition models (STMs) alongside partitioned survival models (PSMs) to verify the plausibility of PSMs extrapolations and explore key clinical uncertainties in the extrapolation period.
What alternative approach has the ERG suggested?	To develop a STM.
What is the expected effect on the cost effectiveness estimates?	Expected impact is unclear but might be substantial given the large proportion of outcomes that are accumulated beyond the observed data.
What additional evidence or analyses might help to resolve this key issue?	Develop a STM to validate the PSM results.

Table 1.8: Key issue 7 Population – no subgroup cost effectiveness analyses presented

Report section	Section 4.2.3
Description of issue and why the ERG has identified it as important	The company did not present subgroup cost effectiveness analyses despite relevant subgroups being listed in the scope, such as histologic subtype (epithelioid, sarcomatoid, biphasic) and level of PD-L1 expression. Cost effectiveness may differ in these subgroups.
What alternative approach has the ERG suggested?	Provide subgroup cost effectiveness analyses for subgroups in the scope.
What is the expected effect on the cost effectiveness estimates?	Unknown.
What additional evidence or analyses might help to resolve this key issue?	The presentation of those subgroup analyses.

Table 1.9: Key issue 8 Intervention and comparators – two-year stopping rule may not be completely adhered to in trial

Report section	Section 4.2.4
Description of issue and why the ERG has identified it as important	Two patients continued treatment with nivolumab + ipilimumab beyond 24 months, despite the protocol stipulating a 24-months stopping rule.

Report section	Section 4.2.4
What alternative approach has the ERG suggested?	If the proportion of patients continuing nivolumab + ipilimumab beyond 24 months increases or it is deemed unlikely to be adhered to in clinical practice: scenario analyses without the stopping rule in place.
What is the expected effect on the cost effectiveness estimates?	Unclear - may increase the ICER, but effectiveness may also change.
What additional evidence or analyses might help to resolve this key issue?	Provide proportions of patients continuing treatment with nivolumab + ipilimumab beyond 24 months and duration of continued treatment in future data cuts and analyses.

Table 1.10: Key issue 9 Treatment effectiveness and extrapolation – immaturity of the long-term PFS and OS data

Report section	Section 4.2.6
Description of issue and why the ERG has identified it as important	The majority of (PF)LY were accumulated beyond the observed data period (see section 5.1) and the validation of long-term PFS and OS using external data is limited, most importantly for nivolumab + ipilimumab. Moreover, the plausibility of assuming a continued treatment effect over the lifetime horizon of the model is unclear.
What alternative approach has the ERG suggested?	Alternative approaches to estimate PFS and OS as well as assumptions related to treatment waning are considered by the ERG. However, due to the immaturity of the data, using the April 2020 database lock of CheckMate-743 (minimum follow-up for all patients was 22.1 months; 23% and 15% of patients treated with nivolumab + ipilimumab and PDC, respectively were still alive at this point), it is unclear what approach is most plausible.
What is the expected effect on the cost effectiveness estimates?	Depending on the scenario, the impact can be substantial. This is also illustrated by the large majority of (PF)LY gains that are accumulated beyond the observed data period.
What additional evidence or analyses might help to resolve this key issue?	Using CheckMate-743 data with additional follow-up data.

Table 1.11: Key issue 10 Health-related quality of life – duration of utility benefits for nivolumab + ipilimumab

Report section	Section 4.2.8
Description of issue and why the ERG has identified it as important	The treatment dependent utilities, used in the CS base-case, result in utility benefits for nivolumab + ipilimumab compared to PDC. This is 0.004 and 0.072 for the PF and PD health states. In the CS base-case, these utility benefits are maintained for the whole duration of the time horizon. The plausibility of this assumption can be debated. Although the company's responses to clarification question B12 were informative and seemed to indicate that there might be a utility benefit when patients are off treatment (clarification response Tables 26 and 27), the duration/extrapolation of the utility benefit is unclear.

Report section	Section 4.2.8
What alternative approach has the ERG suggested?	Not assuming that the utility benefits are maintained for the whole duration of the time horizon.
What is the expected effect on the cost effectiveness estimates?	The ERG adjustment using the treatment dependent utilities (with the nivolumab + ipilimumab utility benefit) up to three years and treatment independent utilities afterwards increased the ICER by ~£2,700 (when applied to the company's corrected base-case).
What additional evidence or analyses might help to resolve this key issue?	It might be informative for the company to explore the time point until which the utility benefits are maintained in CheckMate-743.

Table 1.12: Key issue 11 Resources and costs – estimation of time to treatment discontinuation

Report section	Section 4.2.9
Description of issue and why the ERG has identified it as important	Using number of mean doses to estimate time on treatment in the model may be biased due to right-censoring. Treatment cost is a major driver of cost effectiveness in this model.
What alternative approach has the ERG suggested?	Use parametric survival analysis based on TTD data from CheckMate-743: differential distributions could be used (e.g. the best-fitting generalised gamma for the nivolumab + ipilimumab arm and Gompertz for the pemetrexed + cisplatin arm as reported in Appendix K). The stopping rule for nivolumab + ipilimumab can be included by discontinuing all patients still on treatment at 24 months. Missed and delayed doses can be reflected for both arms using dose intensity as informed by CheckMate-743. No stopping rule will then be required for the pemetrexed + cisplatin arm.
What is the expected effect on the cost effectiveness estimates?	This will likely increase the ICER. The magnitude of the effect is unknown as this is depending on dose intensity.
What additional evidence or analyses might help to resolve this key issue?	Nothing further.

Table 1.13: Key issue 12 Resources and costs – uncertainty about subsequent treatments

Report section	Section 4.2.9
Description of issue and why the ERG has identified it as important	Uncertainty about proportion of patients using subsequent treatments, the mix of treatments used and the duration of subsequent treatments.
What alternative approach has the ERG suggested?	Enable in the model differential treatment durations for each treatment arm to enable further scenario analysis.
What is the expected effect on the cost effectiveness estimates?	This may increase the ICER if there is evidence for longer subsequent treatment duration in the nivolumab + ipilimumab arm than in the PDC arm, but this is currently unclear. The impact is likely small.
What additional evidence or analyses might help to resolve this key issue?	Provide CheckMate-743 analyses of subsequent treatment proportions of use, mix of treatments and duration of subsequent treatment if possible. Explore Waterhouse et al data for differential second-line treatment duration by first-line treatment (if available). Explore expert

Report section	Section 4.2.9
	opinion on subsequent treatment proportions of use, mix of treatments and duration of subsequent treatments.

Table 1.14: Key issue 13 Resources and costs – adverse events

Report section	Section 4.2.7 and Section 4.2.9
Description of issue and why the ERG has identified it as important	The exclusion of many adverse events from the model may introduce bias in favour of nivolumab + ipilimumab.
What alternative approach has the ERG suggested?	Provide cost effectiveness analyses with all-causality (treatment-emergent) adverse events instead of only treatment-related adverse events and change the restriction on the incidence to >1% instead of >2%.
What is the expected effect on the cost effectiveness estimates?	ICER will likely increase, but the impact is likely not large.
What additional evidence or analyses might help to resolve this key issue?	Provide Supplementary Table S.6.6.2 of the CheckMate-743 CSR.

Table 1.15: Key issue 14 Company's cost effectiveness results – proportion of (PF)LY accumulated beyond the observed data

Report section	Section 5.1
Description of issue and why the ERG has identified it as important	The proportion of (PF)LY accumulated beyond the observed data is substantially larger for nivolumab + ipilimumab than for PDC. Moreover, considering the increments, approximately [redacted] of the LYs are gained beyond the observed data period for nivolumab + ipilimumab compared with PDC while this is even larger (approximately [redacted] for PFLY. While the company's response to clarification questions B5 and B8 give some indication about the plausibility of the long-term extrapolations, the finding that the large majority of gains are accumulated beyond the observed data period and hence additional explanation of the mechanism by which the model generated these differences as well as a justification for why they are plausible based upon available evidence is warranted (as requested but not provided in the company's response to clarification question B17). This includes verifying the plausibility of the partitioned survival model extrapolations.
What alternative approach has the ERG suggested?	Providing additional explanation of the mechanism by which the model generated the differences as well as a justification for why they are plausible based upon available evidence is warranted. This includes verifying the plausibility of the partitioned survival model extrapolations.
What is the expected effect on the cost effectiveness estimates?	The expected impact is unclear but is potentially substantial.
What additional evidence or analyses	See suggestions above, as well as using CheckMate-743 data with additional follow-up data.

Report section	Section 5.1
might help to resolve this key issue?	

1.6 Other key issues: summary of the ERG's view

There are no other key issues.

1.7 Summary of the ERG's view

The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was £111,898 per QALY gained. The probabilistic ERG base-case analyses indicated cost effectiveness probabilities of 0%, 0% and 0% at willingness to pay thresholds of £20,000, £30,000 and £50,000 per QALY gained. The most influential adjustments were implementing treatment waning from five years onwards and using the log-logistic distribution for estimating OS in the PDC arm. The ICER increased most in the scenario analysis using TTD estimates with 100% dose intensity instead of the number of mean doses approach. Since dose intensity was likely lower in the trial, this may be regarded as a the upper bound of the ICER using alternative scenarios on time on treatment.

There is large remaining uncertainty about the effectiveness and relative effectiveness of nivolumab + ipilimumab versus PDC, which can be at least partly resolved with future analyses of CheckMate-743 data. In view of the immaturity of the CheckMate-743 study it was not possible for the ERG to quantify all uncertainty now. Further data cuts could potentially result in additional survival gains for the nivolumab + ipilimumab arm. However, it is currently questionable whether nivolumab + ipilimumab can be cost effective compared to PDC.

Table 1.16: Summary of ERG's preferred assumptions and ICER

Technologies	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Company's corrected base-case			
Nivolumab + ipilimumab vs Pemetrexed + cisplatin or carboplatin	54,417	0.702	77,531
Matter of judgement 1: do not use piecewise approach (key issue 9)			
Nivolumab + ipilimumab vs Pemetrexed + cisplatin or carboplatin	54,579	0.719	75,867
Matter of judgement 2: use log-logistic distributions for OS in both treatment arms (using piecewise) (key issue 9)			
Nivolumab + ipilimumab vs Pemetrexed + cisplatin or carboplatin	53,269	0.576	92,413
Matter of judgement 3: implement treatment waning from 5 years onwards (key issue 9)			
Nivolumab + ipilimumab vs Pemetrexed + cisplatin or carboplatin	52,988	0.443	119,543
Matter of judgement 4: change to treatment-independent utilities from 3 years onwards (key issue 10)			
Nivolumab + ipilimumab vs Pemetrexed + cisplatin or carboplatin	54,417	0.678	80,206
ERG base-case (Changes 1-4)			

Nivolumab + ipilimumab vs Pemetrexed + cisplatin or carboplatin	53,327	0.476	112,005
ERG base-case probabilistic (5,000 runs)			
Nivolumab + ipilimumab vs Pemetrexed + cisplatin or carboplatin	53,076	0.612	111,898

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

Table 2.1: Statement of the decision problem (as presented by the company)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
Population	Adults with untreated unresectable MPM.	As per scope.	Not applicable.	The inclusion criteria reported for CheckMate-743 (Table 7 of the CS) specify patients with ECOG PS 0-1. The company also did not provide comparison with BSC on the basis that BSC would only be indicated if PS>1.
Intervention	Nivolumab with ipilimumab.	As per scope.	Not applicable	The intervention is in line with the NICE scope, although dosing in CheckMate-743 was by weight. This is different to the cost effectiveness analysis, which employed a flat dosage of 360 mg every 3 weeks and was stated to align with the anticipated EMA licence.
Comparator(s)	Pemetrexed with cisplatin Raltitrexed with cisplatin (for people for whom treatment with pemetrexed is unsuitable) Pemetrexed with carboplatin (for people for whom treatment with cisplatin is unsuitable)	Pemetrexed with cisplatin or carboplatin (referred to as PDC)	In CheckMate-743, participants were randomised 1:1 to either open-label nivolumab + ipilimumab or pemetrexed + cisplatin or carboplatin. The choice of cisplatin or carboplatin was the investigator's choice, and the use of cisplatin was preferred; however, carboplatin was used at the	The choice of cisplatin or carboplatin may indicate clinically identifiable subgroups and the applicability to the English NHS of the choice of carboplatin or cisplatin as

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	Best supportive care		discretion of the investigator, and switching from cisplatin to carboplatin and vice versa were allowed if reported in the case report form. Raltitrexed is not approved for use in the UK for the first-line treatment of MPM and is not used in the NHS according to UK registry data (see Section B.1.3.4.1), as well as the UK clinical experts we have consulted (Appendix N) and the scope consultation comments from the British Thoracic Oncology Group. For these reasons, BMS have not included raltitrexed as a comparator in this submission. Best supportive care is also not included as a comparator in this submission. This is because first-line systematic anticancer therapies are only used in patients with good PS (0-1), in accordance with BTS guidelines. In line with clinical practice and the NICE recommendation for pemetrexed, the eligibility criteria of CheckMate-743 only included patients with an ECOG PS of 0-1. According to the UK clinical experts we have consulted (Appendix N) and the scope consultation comments from the British Thoracic Oncology Group, best supportive care is not an appropriate	observed in CheckMate-743 is questionable. It is unclear the extent to which raltitrexed is part of current standard of care. If the population is broader than ECOG PS 0-1, BSC should also be considered as a comparator.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
			comparator because this technology relates to a particular group of fit patients for whom best supportive care would not be deemed acceptable or ethical unless specifically requested by the patient.	
Outcomes	Overall survival Progression-free survival Response rate Adverse effects of treatment Health-related quality of life	As per the scope	Not applicable	The outcomes reported are in line with the NICE scope.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and Personal Social Services perspective.	This was not included in Table 1 of the CS.	Not applicable.	The economic analysis is in line with the NICE reference case.
Subgroups to be considered	Histologic subtype (epithelioid, sarcomatoid, biphasic)	Histology: epithelioid and non-epithelioid	Clinical efficacy data are presented for the prespecified subgroup analyses in CheckMate-743, which included	Response was not reported by histological subtype and no statistical test of

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG comment
	Level of PD-L1 expression	PD-L1 expression: ≥ 1% or < 1%	histology and PD-L1 expression subgroups as per the scope.	difference was reported by PD-L1 expression. No cost effectiveness analyses were presented for the subgroups.
Special considerations including issues related to equity or equality	None	The company are not aware of specific equality issues for this appraisal. However, MPM is a preventable, occupational-related disease caused by asbestos exposure. BMS wish to highlight that MPM incidence rates vary across England, with higher rates in areas of heavy industry (e.g. the northeast and southern England). Also, as MPM is a rare cancer, patients may be referred in the NHS to a limited number of specialist mesothelioma multidisciplinary teams, which may require patients to travel long distances from their homes for appointments if they live in a rural setting. Patients with MPM are often older and diagnosed at a late stage of the disease. Consequently, they can be too frail to travel for treatment, which may limit their treatment options.		No comment.

Based on Table 1 of the CS.[2]

BMS = Bristol-Myers Squibb; CS = company submission; BTS = British Thoracic Society; ECOG = Eastern Cooperative Oncology Group; MPM = malignant pleural mesothelioma; NHS = National Health Service; PD-L1 = programmed death-ligand 1; PS = performance status; QALY = quality-adjusted life year; UK = United Kingdom.

2.1 Population

The decision problem specified in the scope defines the population as adults with untreated, unresectable malignant pleural mesothelioma (MPM). The inclusion criteria reported for CheckMate-743 (Table 7 of the company submission (CS)) specify patients with ECOG PS 0-1. The company have also not provided a comparison with BSC, as argued in Table 1, because *"...first-line systematic anticancer therapies are only used in patients with good PS (0-1), in accordance with BTS guidelines."*[2]

A marketing authorisation application has been filed in Europe for nivolumab in combination with ipilimumab for the first-line treatment of adult patients with unresectable MPM. Regulatory approval and marketing authorisation are expected in [redacted].[2]

ERG comment: On the one hand, the company state that the decision problem population is as per scope and no further qualification is mentioned in the request for marketing authorisation. On the other hand, ECOG PS 0-1 is an entry criterion for the pivotal trial and the reason for excluding BSC as comparator. The ERG therefore requested clarification.[3] The company confirmed that evidence presented in the submission was only for patients with ECOG PS 0-1.[4] They did mention that many patients might have unrecorded ECOG PS, but the ERG would argue that this does not imply that status would be unknowable to the treating clinician.

2.2 Intervention

The intervention is nivolumab with ipilimumab, as per scope. It is expected to be given by intravenous infusion of 360 mg nivolumab every three weeks + 1 mg/kg ipilimumab every six weeks.[2] A 2-year treatment stopping rule is expected to be applied in clinical practice to the nivolumab + ipilimumab regimen, which is consistent with the CheckMate-743 clinical trial design.[2]

ERG comment: Nivolumab dosing in the trial was according to weight, but the cost effectiveness analysis employed a flat nivolumab dosage of 360 mg every three weeks, which was stated to align with the anticipated licence. The ERG therefore requested evidence that this difference in dosing will have no effect on effectiveness, quality of life or safety.[3] The company cited a conference presentation the purpose of which was to show that the fixed licensed dose would produce both efficacy and safety outcomes that were similar to those observed with weight-based dosing in the trial.[4] However, the pharmacokinetic analysis showed that a large difference was observed with Cmax1 peak serum concentration after the first dose, i.e. 67.4% higher with 360 mg Q3W.[5] This was reported to not be a problem because it was *"~82% below the median Cmaxss (peak serum concentration at steady state) when administered as NIVO 10 mg/kg Q2W, a dosing regimen previously demonstrated to be safe and well tolerated*". Although this does provide some reassurance regarding safety, a judgment of safety/tolerance is not a substitute for actual AE rates at the given fixed dose. The presentation also stated that: *"...efficacy and safety were evaluated by characterising the relationships between simulations of NIVO exposure and OS or grade ≥ 2 immune-mediated adverse events (grade 2+ IMAEs), respectively, using the multivariate Cox proportional-hazard model"*. However, it is not clear to the ERG precisely how outcomes could be estimated for a fixed dose without evidence from patients who received that dosing regimen. Subgroup analyses by weight were also provided, but again these do not show the effect of patients receiving a lower or higher dose, as would have been the case if dosing had been weight-based. Therefore, uncertainty remains regarding the effectiveness and safety of the expected licensed dose of nivolumab. The potential implications of this are discussed further in Section 4.2.4 and form the basis of Key Issue 8.

Despite the company stating that a 2-year stopping rule had been applied in the CheckMate-743 trial, Figure 10 in Appendix K appears to show two patients still on treatment at 25 months.[6]

2.3 Comparators

The NICE scope listed the following four comparators:

- Pemetrexed with cisplatin
- Raltitrexed with cisplatin (for people for whom treatment with pemetrexed is unsuitable)
- Pemetrexed with carboplatin (for people for whom treatment with cisplatin is unsuitable)
- Best supportive care

The company only included one comparator, which was a combination of pemetrexed plus either cisplatin or carboplatin, referred to as PDC, i.e. they chose not to separate into two comparators on the basis that which one was received in CheckMate-743 was according to investigator choice (IC). The clinical study report (CSR) states*: "The use of cisplatin was preferred; however, carboplatin may be used at the discretion of the investigator."*[7]

ERG comment: The British Thoracic Society (BTS) guideline recommends carboplatin only*: "Where cisplatin is contraindicated, or has adverse risk,"* (p.i2).[8] This might imply clinically identifiable subgroups and thus that the most appropriate way of estimating the effectiveness and cost effectiveness would be by such subgroups. However, the ERG recognises that such analyses may not be required if the choice of comparator in CheckMate-743 was made in a way that is consistent with English NHS practice and that the proportion of those that would receive each treatment is approximately that which would be observed in the English NHS. The ERG also acknowledges that subgroup analysis by cisplatin or carboplatin would be hindered by the fact that the choice of cisplatin or carboplatin was at the discretion of the clinician and not part of the randomisation. Therefore, the intention to treat (ITT) analysis of the control group as a whole is the most appropriate one. The ERG therefore asked for reassurance of the applicability of CheckMate-743 to English NHS practice to which the company responded by providing a set of estimates of the percentages of United Kingdom (UK) patients treated with either carboplatin or cisplatin.[4] Although the company seemed to believe that these estimates validated the results of CheckMate-743, the percentage of patients who had received carboplatin or cisplatin varied between sources:

- UK National Mesothelioma Audit 2020: of patients treated with chemotherapy, pemetrexed with carboplatin was the most common regimen used (48%), followed by pemetrexed with cisplatin (20%), i.e. about 42% of those who received PDC[9].
- Real-world treatment data from the CAS registry in England from January 2013-December 2017 (3,159 unresected patients received first-line SACT): of patients treated with PDC, [redacted].[1]
- EU cross-sectional study including a smaller cohort of 248 UK patients: of patients treated with PDC, [redacted].[10]

Therefore, a key issue remains given the continued uncertainty regarding the applicability of the comparator used in CheckMate-743.

As stated in Section 2.1, the ERG also requested clarification on the applicability of BSC as comparator.[3] The company responded, as described in Section 2.1, that the index population of the evidence submission is those with ECOG PS 0-1, which would seem to eliminate BSC as a comparator.[4]

In support of the omission of raltitrexed as a comparator, the CS stated that: *"The BTS guidelines state that pemetrexed can be replaced with raltitrexed and cisplatin can be replaced with carboplatin as alternatives; however, in clinical practice, raltitrexed is not used in the UK NHS."* The main reference given for treatment patterns is the 2016-2018 UK National Mesothelioma Audit; this report does not describe the chemotherapy regimens received by patients who did not receive pemetrexed with carboplatin or cisplatin (32%) and does not mention raltitrexed.[9] The company did provide expert opinion in Appendix N that raltitrexed is not used, but this is only from two clinicians.[6] The ERG therefore requested that the company either provide further evidence that raltitrexed is not currently used in the UK NHS or include raltitrexed as a comparator.[3] The company responded by providing two sources of data:[4]

- Real-world treatment data from the CAS registry in England from January 2013-December 2017: no recorded use of raltitrexed during the study period.[1]
- Real-world cross-sectional study on treatment patterns in Europe. In the UK in 2019, [redacted] received combination treatment with off-label raltitrexed.[10]

The ERG is therefore satisfied that raltitrexed can reasonably be omitted as a comparator.

2.4 Outcomes

The outcomes are as per scope:[2]

- OS
- PFS
- Response rate
- Adverse effects of treatment

2.5 Other relevant factors

There are none.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

A clinical systematic literature review (SLR) was performed in October 2020 according to NICE requirements to identify studies relevant to nivolumab + ipilimumab for the treatment of previously untreated unresectable MPM in adults.[2]

3.1.1 Searches

Appendix D of Document C of the CS details a SLR conducted to identify randomised and non-randomised trials evaluating the efficacy and safety of first-line, second-line and later treatments for adults with MPM. The last search was undertaken on 5 October 2020. There were no date limits.[6] A language limit was reported but this did not appear to be applied at the searching stage. A summary of the sources searched is provided in Table 3.1.

Table 3.1: Data sources for the clinical effectiveness systematic review (as reported in CS and response to clarification)

	Resource	Host/source	Date ranges	Dates searched
Electronic databases	Embase	Embase.com	From inception	5.10.21
	MEDLINE	Embase.com		
	MEDLINE In-Process and Ahead of Print	PubMed		
	CENTRAL	Wiley		
	CDSR	Wiley		
Conference proceedings	American Society of Clinical Oncology (ASCO)	https://meetinglibrary.asco.org/results/(Keywords:"Mesothelioma");page=0	2018-2020	October 2020
	European Society for Medical Oncology (ESMO)	https://www.sciencedirect.com/search?qs=mesothelioma&pub=Annals%20of%20Oncology&cid=321639&years=2020&lastSelectedFacet=years	2018-2020	
	American Association for Cancer Research (AACR)	https://www.aacr.org/professionals/meetings/previous-aacr-meetings/previous-aacr-meetings-2018/	2018	
		https://www.aacr.org/professionals/meetings/previous-aacr-meetings/previous-aacr-meetings-2019/ https://cancerres.aacrjournals.org/content/79/13_Supplement	2019	
		https://www.aacr.org/professionals/meetings/previous-aacr-meetings/previous-aacr-meetings-2018/	2020	

	Resource	Host/source	Date ranges	Dates searched
	International Society for Pharmacoeconomics and Outcomes Research (ISPOR)	https://www.ispor.org/heor-resources/presentations-database/search	2018-2020	
	World Conference on Lung Cancer (WCLC)	https://wclc2018.iaslc.org/wp-content/uploads/2018/09/WCLC2018-Abstract-Book_vF-LR-REV-SEPT-25-2018.pdf	2018	
		https://wclc2019.iaslc.org/wp-content/uploads/2019/08/WCLC2019-Abstract-Book_web-friendly.pdf	2019	
	European Lung Cancer Congress (ELCC)	https://www.jto.org/issue/S1556-0864(18)X0004-5	2018	
		https://www.sciencedirect.com/journal/annals-of-oncology/vol/30/suppl/S2?page=3#article-201	2019	
	International Mesothelioma Interest Group (IMIG)	Not searched		
Additional resources				
National Institute for Health and Care Excellence (NICE)				
Scottish Medicines Consortium (SMC)				
Institute for Quality and Efficiency in Health Care (IQWiG)				
Haute Autorité de Santé (HAS)				
Canadian Agency for Drugs and Technologies in Health (CADTH)				
Pharmaceutical Benefits Advisory Committee (PBAC)				
Food and Drug Administration (FDA)				
European Medicines Agency (EMA)				
Gesellschaft der Epidemiologischen Krebsregister in Deutschland (GEKID)				
Belgian Cancer Registry				
Dutch Cancer Registry				
Italian Association of Cancer Registries (ITACAN)				
Red Española de Registros de Cáncer (REDECAN)				
Nordic Cancer Registry (NORDCAN)				
Surveillance, Epidemiology, and End Results Program (SEER)				
National Lung Cancer Audit annual report				

ERG comment:

- A range of databases and conference proceedings were searched as well as health technology assessment (HTA) agencies, regulatory agencies and registries. The CS provided sufficient details for the ERG to appraise the literature searches.
- The update searches which were reported, were well-conducted and documented making them transparent and reproducible.
- Databases were searched from inception to the search date.
- A restriction to English language publications was reported but this did not appear to be a searching restriction.
- Study design filters were appropriately used although not referenced in the CS. Upon clarification it was explained that they were based on clinical effectiveness filters from a number of sources including Scottish Intercollegiate Guidelines Network (SIGN), British Medical Journal (BMJ) Best Practice and Canadian Agency for Drugs and Technologies in Health (CADTH). They appeared sufficient to find both randomised and non-randomised study designs.
- Cochrane Library searches for observational studies and real-world evidence reported use of a filter and the use of filters is not recommended in Cochrane Library databases which are study design specific.[11] However, as the results for the Cochrane Library observational studies search had the same number of hits as the Cochrane Library search for controlled evidence and the flowchart does not combine these two searches, it is likely that the Cochrane Library search for observational studies and real-world evidence was incorrectly reported.

3.1.2 Inclusion criteria

The eligibility criteria used in the search strategy for randomised controlled trials (RCTs) and non-RCTs is presented in Table 3.2. All inclusion screening was performed by two independent reviewers, followed by a quality check by a third independent reviewer.[6]

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence

	Description	Justification
Inclusion criteria		
Population	Gender: Any Race: Any Ethnicity: Any Disease: Malignant pleural mesothelioma	Consistent with scope.
Interventions	Doxorubicin Picoplatin Oxaliplatin Raltitrexed Cyclophosphamide Pemetrexed Carboplatin Gemcitabine Vinorelbine Fluorouracil Vinblastine	Unclear given the decision problem excluded all but the comparator in the company trial, CheckMate-743.

	Description	Justification
Inclusion criteria		
	Pemetrexed + cisplatin or carboplatin Erlotinib Bevacizumab Cisplatin Navelbine Platinum Topotecan Liposomal doxorubicin Irinotecan Mitomycin Paclitaxel Adriamycin Nivolumab + ipilimumab Pembrolizumab Best supportive care Active symptom control	
Outcomes	Overall survival Progression-free survival Disease control rate Duration of response Post progression survival Duration of therapy Overall response rate Adverse effects Study withdrawals/discontinuations Time-to-treatment discontinuation	Reported as outcomes to extract, rather than to include.
Study design	RCT: parallel group (triple/double blind) RCT: cross-over (triple/double blind) RCT: post hoc and open-label extension RCTs: Unblinded Pooled studies of RCTs Non-randomised controlled trials Cohort studies (retrospective observational) Cohort studies (prospective observational) Single-arm studies Literature reviews/systematic reviews/meta-analysis/relevant general reviews	Unclear why non-RCTs were included given that company trial, CheckMate-743, which is an RCT, was the only one included.

	Description	Justification
Inclusion criteria		
Language restrictions	English language only	Not reported.
Exclusion criteria: None reported.		
Source: Table 1 of Appendix D.[6] RCT = randomised controlled trial		

3.1.3 Data extraction

All data extraction was performed by two independent reviewers, followed by a quality check by a third independent reviewer.[6]

3.1.4 Quality assessment

The critical appraisal of randomised studies was conducted using the NICE checklist as recommended in the NICE STA manufacturer's template.[12]

3.1.5 Evidence synthesis

Because only one RCT was included, there was no synthesis.[2]

ERG comment: The systematic review appears to have been largely well conducted with the inclusion of more studies than are required given that the submission relies solely upon evidence from the CheckMate-743 trial. This trial was considered to be the most appropriate evidence, assuming that PDC is the only relevant comparator (see Section 2.3), because it provides a direct comparison between nivolumab + ipilimumab and PDC.

3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these)

3.2.1 Design (including statistical analyses) of CheckMate-743 trial

The CheckMate-743 trial is an international, multicentre, randomised, open-label, active-controlled phase 3 trial (See Table 3.3).[2] The population of the CheckMate-743 trial included individuals who had a histological diagnosis of MPM, had advanced unresectable disease that was not amenable to therapy, had available pathological samples for centralised programmed death-ligand 1 (PD-L1) immunohistochemistry (IHC) testing, ECOG PS 0-1, and could have had prior palliative radiotherapy.[2] The trial locations comprised of 103 sites, with six of these sites being based in the UK, however, additional locations were not further identified.[2] The intervention in the CheckMate-743 trial was nivolumab 3 mg/kg Q2W and ipilimumab 1 mg/kg Q3W for up to two years. The comparator was cisplatin or carboplatin + pemetrexed Q3W for six cycles. The use of cisplatin or carboplatin was based on the investigator's choice and, thus, are not treated as separate comparators. Statistical analyses are shown in Table 3.4.

Table 3.3: CheckMate-743: study design

Study	CheckMate-743 (NCT02899299)
Study Design (n)	International, multicentre, randomised, open-label, active-controlled phase 3 trial (n=605)
Population	• Males and females aged ≥18 years. • Histological diagnosis of MPM; determination of epithelioid vs. non-epithelioid histology.

Study	CheckMate-743 (NCT02899299)
	• Patients with advanced unresectable disease that is not amenable to therapy with curative intent (surgery with or without chemotherapy). • Available (archival and/or fresh) pathological samples for centralised PD-L1 IHC testing. • Prior palliative radiotherapy is acceptable; however, ≥14 days must have passed prior to first treatment, and all signs of toxicity must have remitted. Prior prophylactic radiotherapy to a pleurodesis drainage tract or biopsy site is allowed. • ECOG PS 0-1. • Measurable disease is defined as: – Mesothelioma tumour thickness perpendicular to the chest wall or mediastinum that can be measured in up to 2 positions at 3 separate levels on transverse cuts of computed tomography scan (cuts must be ≥10 mm apart), for a total of up to 6 measurements. Each single tumour measurement must be ≥10 mm to qualify as measurable disease and contribute to the sum that defines the pleural measurement. – Non-pleural metastatic target lesions measured unidimensional as per RECIST v1.1 criteria. Patients who present without pleural lesions that can be considered measurable but with metastatic lesions meeting criteria for target lesion by RECIST v1.1 criteria may be considered for inclusion after consultation with the Medical Monitor. As of 3 April 2020, database lock, 713 patients enrolled included: 605 patients randomised to each treatment arm: • 303 patients in the Nivolumab + ipilimumab arm • 302 patients in the PDC arm
Intervention	Nivolumab 3 mg/kg Q2W + ipilimumab 1 mg/kg Q6W for up to 2 years, n = 303
Comparator	PDC – pemetrexed 500 mg/m^2 plus cisplatin 75 mg/m^2 or carboplatin (AUC of 5 mg/mL/minute), on day 1 of a 21-day cycle for 6 cycles
Reported outcomes specified in the decision problem	Measures of disease severity and symptom control: • OS • ORR • PFS HRQoL: • EQ-5D-3L • VAS • LCSS-Meso Safety outcomes: • AEs
All other reported outcomes	• Disease control rate (DCR) • Composite correlation of PD-L1 • Time to response (TTR) • Duration of response (DOR) • Eastern Cooperative Oncology Group performance status (ECOG PS)
Duration of study and follow-up	CheckMate-743 is ongoing. At the latest database cut of 3 April 2020 after 419 observed events, the median follow up was 29.7 months. Most of the patients received around 90% or more of planned doses. The median duration of patients in the nivolumab + ipilimumab arm was longer than patients in the PDC arm.

Study	CheckMate-743 (NCT02899299)
	The maximum duration of treatment per protocol was 24 months for nivolumab + ipilimumab and 6 cycles of PDC. A final primary OS analysis will be performed when 473 deaths have occurred. Estimated date for primary completion is April 2021and study final completion date is April 2022.
Countries	103 sites in Australia, New Zealand, Europe, Asia, North America, and South America (6 sites in the UK)

Source: Adapted from Table 6 and Table 7 of the CS [2]
AE = adverse event; BICR = blinded independent central review; DCR = disease control rate; ECOG = Eastern Cooperative Oncology Group; MPM = malignant pleural mesothelioma; ORR = objective response rate; OS = overall survival; PD-L1 = programmed death-ligand 1; PFS = progression-free survival; Q2W = every 2 weeks; Q3W = every 3 weeks; Q6W = every 6 weeks. AUC = area under the curve; CTLA-4 = cytotoxic T-lymphocyte-associated protein 4; DOR = duration of response; ECOG PS = Eastern Cooperative Oncology Group performance status; IHC = immunohistochemistry; LCSS-Meso = Lung Cancer Symptom Scale–Mesothelioma; PD 1 = programmed death-1; PDC = platinum-based doublet chemotherapy; PD-L2 = programmed death-ligand 2; RECIST = Response Evaluation Criteria in Solid Tumours; TTR = time to response; UK = United Kingdom.

Follow-up visit 1 = 30 days from the last dose ± 7 days or coincides with the date of discontinuation (± 7 days) if date of discontinuation is > 35 days after last dose. Follow-ups visit 2 = 90 days (± 7 days) from follow-up visit 1.

Table 3.4: CheckMate-743 statistical analyses

Study	CheckMate-743 (NCT02899299)
Hypothesis objectives	Evaluate and compare the OS of nivolumab + ipilimumab vs. pemetrexed + cisplatin or carboplatin as first-line treatment in patients with unresectable MPM
Statistical analysis	OS was analysed between the treatment groups at the interim and final analyses by utilising a stratified log-rank test. Stratified factors observed were histology and sex of patients. An O'Brien and Fleming α-spending function was used to determine the nominal significance levels for the interim and final analyses. The stratified hazard ratio between the treatment groups was to be introduced along with 100*(1-α) % CI (adjusted for interim). A two-sided P value was accounted for the analysis of the OS. OS was to be assessed by utilizing KM techniques. A two-sided 95% CI for median OS in each treatment group was to be computed via the log-log transformation method. OS rates at fixed time points (e.g. six months, depending on the minimum follow-up) were to be introduced alongside their associated 95% CIs. These estimates were derived from the KM estimates and relating CIs were determined on Greenwood formula for variation derivation and on log-log transformation applied on the survivor function. The status of patients who are controlled in the OS KM investigation was arranged for every treatment groups utilising the accompanying classifications: • On study (on treatment, in follow up) • Off study (lost to follow up, withdrawn consent, never treated) The influence of baseline and demographic characteristics on the treatment effect among all randomised patients was also to be explored for specific subgroups, including age, sex, race, ECOG PS, histology, and PD-L1. • Principal analyses of PFS and ORR were based on the BICR evaluation. No formal testing of the secondary objectives was done. Results were descriptive. PFS was estimated using the KM methodology and analysed similarly to OS. Response and disease

Study	CheckMate-743 (NCT02899299)
	control rate estimates were presented along with their exact two-sided 95% CIs by Clopper and Pearson. • DOR was to be estimated using the KM product limit method. CIs for secondary endpoints were at the two-sided 95% level. • Safety: Descriptive statistics of safety were presented using MedDRA version 22.1 and NCI-CTCAE version 4.0. All on-study AEs, drug-related AEs, SAEs, drug-related SAEs, IMAEs, and select AEs were tabulated using worst grade per NCI-CTCAE version 4.0 criteria by system organ class and preferred term. Frequency, management, and resolution of IMAEs and select AEs were analysed. Patient-reported outcome analyses: Continuous data were described using descriptive statistics. Categorical data were summarised using counts and percentages, for which "missing" was used when applicable. Where relevant, significance testing was two-sided at the 0.05 level, with no adjustment for multiplicity.
Sample size, power calculation	For the OS primary endpoint, a general two-sided alpha (type 1 error rate) was set at 0.05. 605 patients were randomized with 1:1 proportion to two treatment arms. 473 OS events were required for the final analysis. The sample size was determined to compare OS between nivolumab + ipilimumab (Arm A) versus pemetrexed + cisplatin or carboplatin regime (Arm B). One conventional interim analysis was performed for OS at 403 OS events. Key parameters for the primary analysis were as per the following: • Targeted power: 90% • Target hazard ratio: 0.72 • 0-6 months: 1 • 6-34 months: 0.767 • After 34 months: 0.002 • Alpha: 0.05, two-sided (0.03 at interim; 0.041 at final analyses) • Sample size: 606 • Target number of events: 473 • Expected number of events for interim analysis: 403 (85% of target) • Duration (monthly accrual rate = 34 patients): 56 months
Date management and patient withdrawals	OS was censored on the last date a patient was known to be alive. For PFS, patients who died with no reported progression were considered to have progressed on the date of death. Patients who did not progress or die were censored on the date of their last evaluable tumour assessment. Patients who did not have any on-study tumour assessments and did not die were censored on their date of randomisation. Patients who had palliative local therapy or initiated anticancer therapy without a prior reported progression were censored on the date of their last evaluable tumour assessment on or before the initiation of subsequent anticancer therapy or palliative local therapy. For DOR, patients who did not progress or die were censored on the date of their last evaluable tumour assessment. Patients who started subsequent therapy without a prior reported progression were censored at the last evaluable tumour assessments before initiation of the subsequent anticancer therapy. Patients who died without a reported prior progression were considered to have progressed on the date of their death. For patients who neither progressed nor died, DOR was censored on the date of their last evaluable tumour assessment.

Study	CheckMate-743 (NCT02899299)
Missing data	Patients who remained lost to follow-up then the last recognised alive date was determined by an investigator was reported and accounted in the patient's clinical records.

Source: Adapted from Table 9 of the CS.[2]
AE = adverse event; BICR = blinded independent central review; CI = confidence interval; DOR = duration of response; ECOG PS = Eastern Cooperative Oncology Group performance status; IMAE = immune-mediated adverse event; KM = Kaplan-Meier; MedDRA = Medical Dictionary for Regulatory Activities; MPM = malignant pleural mesothelioma; NCI-CTCAE = Common Terminology Criteria for Adverse Events; ORR = objective response rate; OS = overall survival; PD-L1 = programmed death-ligand 1; PFS = progression-free survival; SAE = serious adverse event.

3.2.2 Baseline characteristics of CheckMate-743 trial

The baseline characteristics of the CheckMate-743 trial are presented in Table 3.5. The participants in the trial were randomised on a 1:1 basis.[2] The median age of the randomised participants was 69.0 years. The majority of participants were white and male and at baseline had an advanced disease stage.[2] Almost all participants had quantifiable PD-L1 expression, with 77.0% at ≥1% and 23% <1%.[2] According to UK clinical experts, the trial population was representative of a treatment naïve MPM population in England.[2]

Table 3.5: CheckMate-743: baseline demographics (all randomised patients)

	Nivolumab + ipilimumab (n = 303)	PDC (n = 302)	Total (n= 605)
Age, median (IQR), years	69 (65-75)	69 (62-75)	69 (64-75)
Male, N (%)	234 (77)	233 (77)	467 (77)
ECOG performance status, N (%)			
0	114 (38)	128 (42)	242 (40)
1	189 (62)	173 (57)	362 (60)
Disease stage at study entry			
I	12 (4)	20 (7)	32 (5)
II	23 (8)	22 (7)	45 (7)
III	103 (34)	106 (35)	209 (35)
IV	160 (53)	149 (49)	309 (51)
Unknown	5 (2)	5 (2)	10 (2)
Smoking status, N (%)			
Never	127 (42)	122 (40)	249 (41)
Current/former	173 (57)	171 (57)	344 (57)
Histology,[a] N (%)			
Epithelioid	229 (76)	227 (75)	456 (75)
Non-epithelioid[b]	74 (24)	75 (25)	149 (25)
Prior radiotherapy, %	10	9	9
PD-L1 quantifiable at baseline,[c] N	289	297	586
< 1%,[d] N (%)	57 (20)	78 (26)	135 (23)
≥ 1%,[d] N (%)	232 (80)	219 (74)	451 (77)

	Nivolumab + ipilimumab (n = 303)	PDC (n = 302)	Total (n= 605)
Sources: Table 8 CS.[2] ECOG = Eastern Cooperative Oncology Group; IHC = immunohistochemistry; IQR = interquartile range; PDC = platinum-based doublet chemotherapy; PD-L1 = programmed death-ligand 1. [a] Based on case report form source. [b] Included 47% sarcomatoid and 53% mixed/other in the nivolumab + ipilimumab arm and 48% and 52%, respectively, in the chemotherapy arm. [c] Determined by the PD-L1 IHC 28-8 pharmDx assay (Dako). [d] Based on PD-L1 quantifiable at baseline, 95% and 98% of patients in the nivolumab + ipilimumab and chemotherapy arms, respectively.			

ERG comment: Baseline characteristics seemed to be similar between the arms in the trial. However, the CheckMate-743 had just 38 patients from the UK, which was 6.3% of total patients randomised. The company was therefore requested to provide evidence of generalisability to the UK in response to the clarification letter.[4] They stated that the clinicians the company consulted for this appraisal considered this evidence in addition to the baseline characteristics of the trials to indicate generalisability to English NHS practice.

3.2.3 Quality of CheckMate-743 trial

The critical appraisal of RCTs was conducted utilising the NICE checklist. The quality assessment of the CheckMate-743 trial is presented in Table 3.6. It was unclear how many reviewers were involved in the quality assessment.

Table 3.6: Quality assessment of CheckMate-743 (NCT02899299)

	Company appraisal	ERG appraisal
Was randomisation carried out appropriately?	Yes/No	Yes
Was the concealment of treatment allocation adequate?	No – open-label trial	Yes
Were the groups similar at the outset of the study in terms of prognostic factors?	Yes	Yes
Were the care providers, participants, and outcome assessors blind to treatment allocation?	No – open-label trial	No
Were there any unexpected imbalances in dropouts between groups?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	No

	Company appraisal	ERG appraisal
Was randomisation carried out appropriately?	Yes/No	Yes
Did the analysis include an ITT analysis? **If so, was this appropriate and were appropriate methods used to account for missing data?**	Yes	Yes
Did the authors of the study publication declare any conflicts of interest?	Yes	Yes
Does the trial reflect routine clinical practice in England?	Yes	Unsure
Sources: Table 12, CS.[2] ITT = intention to treat.		

ERG comment: The ERG agrees with the quality assessment. The CheckMate-743 trial was a high-quality study in some respects, the major flaw being in the lack of blinding and questionable applicability to clinical practice in the NHS in England given the small number of UK patients and possible variation in judgement as to whether carboplatin or cisplatin prescribed (see also Section 2.3).

3.2.4 Results of CheckMate-743 trial

The results presented in the CS were reported to be from an interim analysis with a database lock of 3 April 2020.

3.2.4.1 Overall survival

OS was the primary endpoint of the CheckMate-743 trial and was defined at the time of randomisation to the date of death from any cause. According to the CS, a statistically significant benefit was observed for patients who were treated with nivolumab + ipilimumab when compared to patients treated with PDC.[2] The company noted that treatment with nivolumab + ipilimumab reduced the risk of death by 26% when compared to PDC (hazard ratio (HR), 0.74; 96.6% confidence interval (CI), 0.60 to 0.91; stratified log-rank $P = 0.0020$).[2] Those treated with nivolumab + ipilimumab were noted to have a median OS of 18.1 months (95% CI: 16.8 to 21.4 months), whereas the those treated with PDC had a median OS of 14.1 months (95% CI: 12.4 to 16.2 months). The OS rates for all randomised patients are depicted below in Table 3.7. The company notes that additional follow-up will demonstrate a long-term, durable benefit with dual immunotherapy with nivolumab + ipilimumab.[2]

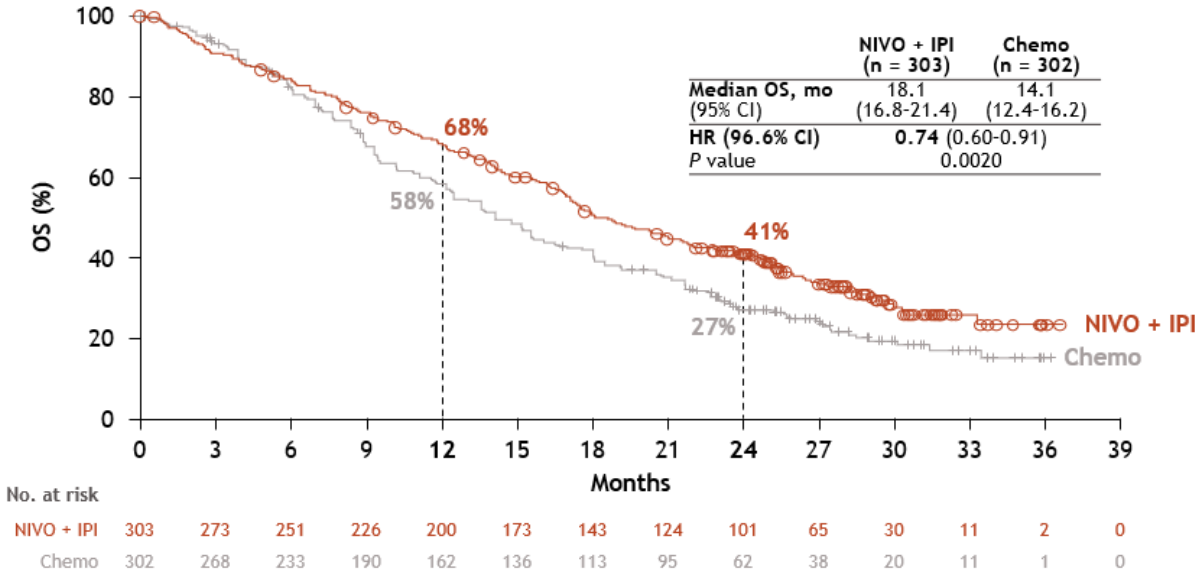
The ERG requested results of formal statistical analyses for the comparison of the OS-related outcomes, as presented in Table 3.7, to which the company responded that OS was compared in two randomised arms via a two-sided, long-rank test stratified by histology and gender at the interim analysis cut-off only.[4] In the response to clarification, the company also noted that at the time of the prespecified interim analysis, the median follow-up for OS 29.7 months (interquartile range (IQR): 26.7 to 32.9), with a minimum of follow-up of 22.1 months.[4]

Table 3.7: CheckMate-743: overall survival rates – all randomised patients

Median overall survival (95% CI)	Nivolumab + ipilimumab (n=303)	PDC (n=302)
6 months	84.0 (79.4-87.7)	82.2 (77.3-86.2)
12 months	67.9 (62.3-72.8)	57.7 (51.7-63.2)
18 months	50.5 (44.7-56.1)	40.6 (34.8-46.3)
24 months	40.8 (35.1-46.5)	27.0 (21.9-32.4)

Sources: Table 14, CS[2]
CI= confidence interval; PDC= platinum-based doublet chemotherapy.
Note: Based on Kaplan-Meier estimates.

Figure 3.1: CheckMate-743: Kaplan-Meier plot of overall survival (all randomised patients)

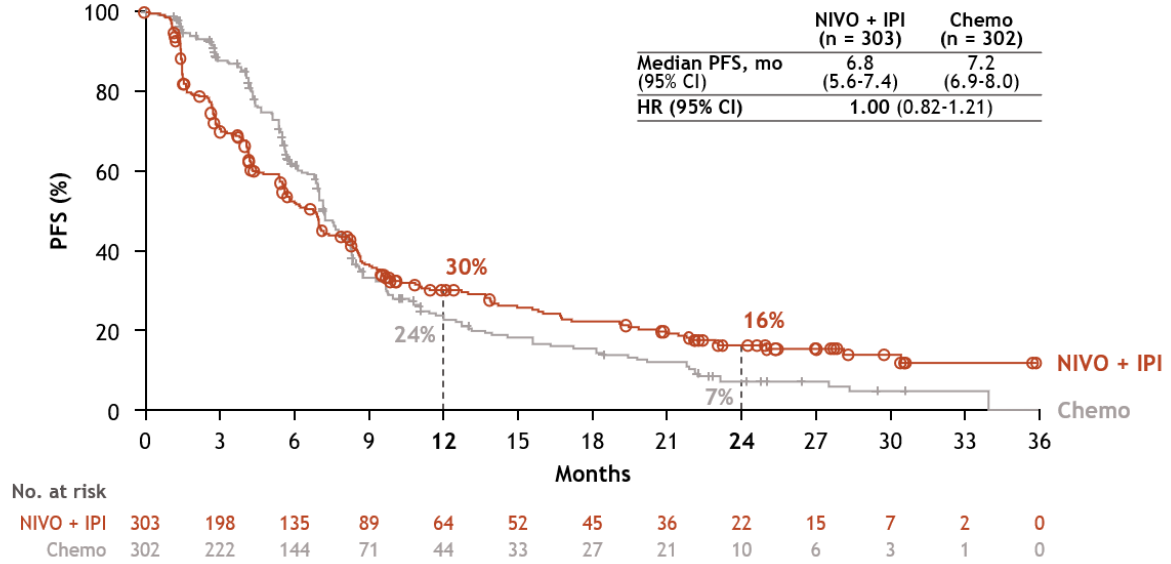


CI = confidence interval; HR = hazard ratio; IPI = ipilimumab; NIVO = nivolumab; OS = overall survival.
Notes: Subsequent systemic therapy was received by 44% of patients in the NIVO + IPI arm and 41% in the chemo arm; subsequent immunotherapy was received by 3% and 20%, and subsequent chemotherapy by 43% and 32%, respectively.
Chemo in figure refers to platinum-based doublet chemotherapy.
Source: Figure 12 from the CS.[2]

3.2.4.2 Progression-free survival

At the point of the interim analysis, 85.1% of patients in both arms had experienced a progression event according to the BICR assessment.[2] However, there was no statistically significant difference in PFS between patients treated with nivolumab + ipilimumab and patients treated with PDC.[2] The median PFS for patients treated with nivolumab + ipilimumab was 6.8 months (95% CI: 5.6 to 7.4 months), whereas the median PFS in the PDC group was 7.2 months (95% CI: 6.9 to 8.0 months).[2]

Figure 3.2: CheckMate-743: Kaplan-Meier plot of progression-free survival by blinded independent central review (all randomised patients)



CI = confidence interval; HR = hazard ratio; IPI = ipilimumab; NIVO = nivolumab; PFS = progression-free survival.
Notes: Per adapted mRECIST for pleural mesothelioma lesions and/or RECIST v1.1 for non-pleural lesions.
Chemo in figure refers to platinum-based doublet chemotherapy.
Source: Baas [13] Based on Figure 13 of the CS [2]

3.2.4.3 Objective response

In the CheckMate-743 trial, the nivolumab + ipilimumab and PDC arms displayed similar objective response rate (ORR) according to the BICR.[2] Patients in the nivolumab + ipilimumab arm had an ORR per BICR of 39.6% (95% CI: 34.1to 45.4%), whereas those in the PDC arm had an ORR of 42.7% (95% CI: 37.1 to 48.5%).[2] The company noted that a BOR of CR was observed in 5 (1.7%) patients in the nivolumab + ipilimumab group, while this was not observed in any patients in the PDC arm.[2]

ORR per BICR was noted to be similar in both treatment arms in patients with PD-L1-positive tumours.[2] When the ERG requested further clarification regarding the use of any formal statistical analyses for the comparison of all response outcomes, the company reiterated that results were descriptive.[4] In the response to clarification, the company provided additional analyses of the response outcomes with 95% CIs presented in Table 3.8.

Table 3.8: Response rate per BICR

Outcome	Nivolumab + ipilimumab (n=303)	PDC (n=302)
ORR per BICR[a]		
ORR,[b] n (% [95% CI])	120/303 39.6 (34.1-45.4)	129/302 42.7 (37.1-48.5)
Median TTR, months	2.7	2.5
DOR (95% CI), months[c]	11.0 (8.1-16.5)	6.7 (5.3-7.1)
Best overall response (BOR), n (% [95% CI])		

Outcome	Nivolumab + ipilimumab (n=303)	PDC (n=302)
CR	5/303 (1.7 [0.5-3.8])	0 (0)
PR	115/303 38.0 (32.5-43.7)	129/302 42.7 (37.1-48.5)
Stable disease	112/303 37.0 (31.5-42.7)	125/302 41.4 (35.8-47.2)
Progressive disease	55/303 18.2 (14.0-23.0)	14/302 4.6 (2.6-7.7)
DCR (95% CI), % (CR+PR+SD)	76.6 (71.4-81.2)	85.1 (80.6-88.9)

BICR = blinded independent central review; CI = confidence interval; CR = complete response; DCR = disease control rate; DOR = duration of response; HR = hazard ratio; ORR = objective response rate; OS = overall survival; PDC = platinum-based doublet chemotherapy; PFS = progression-free survival; PR = partial response; RECIST = Response Evaluation Criteria in Solid Tumours; SD = stable disease; TTR = time to response.

[a] Per adapted modified RECIST for pleural mesothelioma lesions and/or RECIST v1.1 for non-pleural lesions.

[b] 95% CI Clopper and Pearson Method.

[c] Kaplan-Meier estimates.

Source: Table 7 in the response to clarification.[4]

ERG comment: At the time of the interim analysis with a database lock of 3 April 2020, a statistically significant OS benefit was observed for patients who were treated with nivolumab + ipilimumab when compared to patients treated with PDC.[2] However, there was no statistically significant difference in PFS and results for ORR were similar.[2] The ERG did request results from a more recent data cut, but the company replied: "*As CheckMate-743 met its primary endpoint at the 3 April 2020 database lock, this analysis was considered the final analysis. However, follow up of CM-743 is ongoing and additional data cuts are expected, likely in Q2/Q3 2021 (TBC). As the timing of the analysis is event driven, there is uncertainty on the exact timing of future database locks*."[4] Although it is unlikely that the results will change the interpretation that nivolumab + ipilimumab is more effective in terms of OS, the precise size of the difference might be important particularly in determining if cost effective. The interpretation of PFS may change, given that progression data are incomplete.

3.2.4.4 Adverse events

The CS noted that the frequencies of all-cause AEs and treatment-related adverse events (TRAEs) were similar between treatment groups (see Table 3.9).[2]

Table 3.9: CheckMate-743: safety summary – all treated patients

	Nivolumab + ipilimumab (n=300)		PDC (n=284)	
Safety parameters, n (%)	Any grade	Grade 3-4	Any grade	Grade 3-4
All-causality SAEs	164 (54.7)	103 (34.3)	72 (25.4)	54 (19.0)
Treatment-related SAEs	64 (21.3)	46 (15.3)	22 (7.7)	17 (6.0)
All-causality AEs leading to discontinuation	88 (29.3)	59 (19.7)	58 (20.4)	28 (9.9)

Safety parameters, n (%)	Nivolumab + ipilimumab (n=300)		PDC (n=284)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Treatment-related AEs leading to discontinuation	69 (23.0)	45 (15.0)	45 (15.8)	21 (7.4)
All-causality AEs	299 (99.7)	159 (53.0)	277 (97.5)	121 (42.6)
Treatment-related AEs	240 (80.0)	91 (30.3)	233 (82.0)	91 (32.0)

AE = adverse event; MedDRA = Medical Dictionary for Regulatory Activities; PDC = platinum-based doublet chemotherapy; SAE = severe adverse event.
Note: Definitions of events were based on MedDRA version 22.1; Common Terminology Criteria version 4.0. Includes events reported between first dose and 30 days after the last dose of study drug, unless otherwise indicated.
Source: Table 17 of the CS [2]

The most commonly reported TRAEs with nivolumab + ipilimumab were diarrhoea and pruritus.[2] For patients who were treated with PDC, the most commonly experienced TRAEs were nausea, anaemia, and neutropenia, as presented in Table 3.10.[2] The company noted that most of the treatment-related select AEs and most IMAEs had resolved at the time of the database lock, with the exception of endocrine-related events.[2] The reported median time to resolution ranged from 0.14 to 12.14 weeks for select AEs and 0.14 to 17.14 weeks for IMAEs.[2]

Table 3.10: CheckMate-743: treatment-related adverse events – all treated patients

Safety parameters, n (%)	Nivolumab + ipilimumab[a] (n=300)		PDC[b] (n=284)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
TRAEs leading to discontinuation of any component of the regimen[c]	69 (23.0)	45 (15.0)	45 (15.8)	21 (7.4)
Serious TRAEs[c]	64 (21.3)	46 (15.3)	22 (7.7)	17 (6.0)
Any TRAE[c]	240 (80.0)	91 (30.3)	233 (82.0)	91 (32.0)
≥ 15% of patients in any treatment group				
Diarrhoea	62 (20.7)	10 (3.3)	21 (7.4)	2 (0.7)
Pruritus	49 (16.3)	3 (1.0)	1 (0.4)	0
Fatigue	41 (13.7)	3 (1.0)	55 (19.4)	5 (1.8)
Nausea	30 (10.0)	1 (0.3)	104 (36.6)	7 (2.5)
Decreased appetite	29 (9.7)	2 (0.7)	50 (17.6)	2 (0.7)
Asthenia	25 (8.3)	0	44 (15.5)	12 (4.2)
Anaemia	6 (2.0)	1 (0.3)	102 (35.9)	32 (11.3)
Neutropenia	2 (0.7)	2 (0.7)	71 (25.0)	43 (15.1)
Treatment-related select AEs				
Endocrine	52 (17.3)	4 (1.3)	0	0
Gastrointestinal	66 (22.0)	16 (5.3)	23 (8.1)	3 (1.1)
Hepatic	36 (12.0)	16 (5.3)	6 (2.1)	0
Pulmonary	20 (6.7)	2 (0.7)	0	0
Renal	15 (5.0)	4 (1.3)	19 (6.7)	1 (0.4)

Safety parameters, n (%)	Nivolumab + ipilimumab[a] (n=300)		PDC[b] (n=284)	
	Any grade	Grade 3-4	Any grade	Grade 3-4
Skin	108 (36.0)	9 (3.0)	28 (9.9)	1 (0.4)
Hypersensitivity/infusion reactions	36 (12.0)	4 (1.3)	7 (2.5)	0

AE = adverse event; PDC = platinum-based doublet chemotherapy; TRAE = treatment-related adverse event.
Note: Person-years of exposure: nivolumab + ipilimumab, 220.3; chemotherapy, 94.5. Nivolumab + ipilimumab dosages were nivolumab 3 mg/kg every 2 weeks + ipilimumab 1 mg/kg every 6 weeks.
[a] Median (interquartile range) doses for treated patients: nivolumab, 12.0 (5.0-23.5); ipilimumab, 4.0 (2.0-7.0).
[b] Median (interquartile range) doses for treated patients: pemetrexed, 6.0 (4.0-6.0); cisplatin 5.0 (3.0-6.0); carboplatin 6.0 (4.0-6.0).
[c] Includes events reported between first dose and 30 days after last dose of study drug.
Source: Table 18 of the CS[2]

In the request to clarification, the company also provided a table reporting treatment emergent i.e. all-cause AEs of grade 3 or 4 severity (≥1%).[4] This is shown in Table 3.11.

Table 3.11: All-cause adverse events of grade 3 or 4 severity (≥1%)

Event, n (%)	Nivolumab + ipilimumab (n=300)	PDC (n=284)
	Grade 3-4	Grade 3-4
Total subjects with an event	159 (53.0)	121 (42.6)
≥ 1% of patients in any treatment group		
General disorders and administration site conditions	**29 (9.7)**	**30 (10.6)**
Fatigue	9 (3.0)	5 (1.8)
Pyrexia	4 (1.3)	2 (0.7)
Asthenia	4 (1.3)	12 (4.2)
Oedema peripheral	0	0
Non-cardiac chest pain	(1.7)	1 (0.4)
Chest pain	4 (1.3)	3 (1.1)
Pain	0	2 (0.7)
Malaise	2 (0.7)	0
Mucosal inflammation	0	2 (0.7)
Peripheral swelling	0	
General physical health deterioration	0	3 (1.1)
Gastrointestinal disorders	**28 (9.3)**	**21 (7.4)**
Diarrhoea	12 (4.0)	2 (0.7)
Nausea	2 (0.7)	7 (2.5)
Constipation	1 (0.3)	2 (0.7)
Vomiting	0	6 (2.1)
Abdominal pain	2 (0.7)	2 (0.7)

Event, n (%)	Nivolumab + ipilimumab (n=300) Grade 3-4	PDC (n=284) Grade 3-4
Colitis	7 (2.3)	1 (0.4)
Respiratory disorders	**27 (9.0)**	**17 (6.0)**
Dyspnoea	7 (2.3)	9 (3.2) 0
Cough	2 (0.7)	0
Pleural effusion	3 (1.0)	2 (0.7)
Pneumonitis	3 (1.0)	0
Hiccups	0	0
Pulmonary embolism	3 (1.0)	3 (1.1)
Skin and tissue disorders	**12 (4.0)**	**1 (0.4)**
Pruritus	3 (1.0)	0
Rash	3 (1.0)	0
Rash maculo-papular	2 (0.7)	0
Dry skin	0	0
Infections and infestations	**25 (8.3)**	**12 (4.2)**
Nasopharyngitis	1 (0.3)	0
Pneumonia	8 (2.7)	5 (1.8)
Lower respiratory tract infection	3 (1.0)	1 (0.4)
Metabolism and nutrition disorders	**22 (7.3)**	**21 (7.4)**
Decreased appetite	3 (1.0)	4 (1.4)
Hypoalbuminaemia	1 (0.3)	2 (0.7)
Hyponatraemia	5 (1.7)	4 (1.4)
Dehydration	3 (1.0)	2 (0.7)
Hypokalaemia	0	3 (1.1)
Musculoskeletal and connective tissue disorders	**13 (4.3)**	**2 (0.7)**
Arthralgia	3 (1.0)	0
Myalgia	0	0
Back pain	2 (0.7)	1 (0.4)
Pain in extremity	0	0
Musculoskeletal pain	2 (0.7)	0
Investigations	**32 (10.7)**	**9 (3.2)**
Blood creatinine increased	1 (0.3)	0

Event, n (%)	Nivolumab + ipilimumab (n=300) Grade 3-4	PDC (n=284) Grade 3-4
Lipase increased	16 (5.3)	1 (0.4)
Amylase increased	9 (3.0)	1 (0.4)
Alanine aminotransferase increased	6 (2.0)	0
Blood alkaline phosphatase increased	2 (0.7)	0
Weight decreased	0	1 (0.4)
Aspartate aminotransferase increased	5 (1.7)	0
Nervous system disorders	**15 (5.0)**	**2 (0.7)**
Headache	0	0
Dizziness	0	0
Dysgeusia	0	0
Syncope	4 (1.3)	1 (0.4)
Blood and lymphatic system disorders	**18 (6.0)**	**84 (29.6)**
Anaemia	8 (2.7)	39 (13.7)
Neutropenia	3 (1.0)	45 (15.8)
Thrombocytopenia	2 (0.7)	11 (3.9)
Leukopenia	0	8 (2.8)
Pancytopenia	0	5 (1.8)
Febrile neutropenia	0	3 (1.1)
Endocrine disorders	5 (1.7)	0
Hypothyroidism	0	0
Hypopituitarism	3 (1.0)	0
Psychiatric disorders	**2 (0.7)**	**2 (0.7)**
Insomnia	0	0
Anxiety	0	0
Injury, poisoning and procedural complications	**7 (2.3)**	**2 (0.7)**
Infusion related reaction	4 (1.3)	0
Neoplasms benign, malignant and unspecified (including cysts and polyps)	**12 (4.0)**	**5 (1.8)**
Malignant neoplasm progression	9 (3.0)	5 (1.8)
Cardiac disorders	**10 (3.3)**	**5 (1.8)**
Atrial fibrillation	2 (0.7)	3 (1.1)

Event, n (%)	Nivolumab + ipilimumab (n=300) Grade 3-4	PDC (n=284) Grade 3-4
Vascular disorders	12 (4.0)	1 (0.4)
Hypertension	6 (2.0)	1 (0.4)
Hepatobiliary disorders	**17 (5.7)**	**0**
Hepatic function abnormal	5 (1.7)	0
Immune-mediated hepatitis	3 (1.0)	0
Renal and urinary disorders	**9 (3.0)**	**2 (0.7)**
Acute kidney injury	5 (1.7)	0

Included events reported between the first dose of study drug and 30 days after the last dose of study drug.
Source: Table 11 from the response to clarification [4]

At the time of the database lock, 198 (66%) patients who were treated with nivolumab + ipilimumab had died, while 212 (75%) patients who received PDC had died.[2] In both treatment arms, disease progression was the most common cause of death.[2]

Table 3.12 shows the main causes of death.

Table 3.12: CheckMate-743: summary of deaths – all treated patients

Safety parameters, n (%)	Nivolumab + ipilimumab[a] (n=300)	Chemotherapy[b] (n=284)
Number of patients who died	198 (66.0)	212 (74.6)
Within 30 days of last dose	28 (9.3)	14 (4.9)
Within 100 days of last dose	55 (18.3)	50 (17.6)
Primary reason for death		
Disease	183 (61.0)	199 (70.1)
Study drug toxicity	3 (1.0)[c]	1 (0.4)[d]
Unknown	3 (1.0)	2 (0.7)
Other	9 (3.0)	10 (3.5)

Note: Person-years of exposure: nivolumab + ipilimumab, 220.3; chemotherapy, 94.5. Nivolumab + ipilimumab dosages were nivolumab 3 mg/kg every 2 weeks + ipilimumab 1 mg/kg every 6 weeks.
[a] Median (interquartile range) doses for treated patients: nivolumab, 12.0 (5.0-23.5); ipilimumab, 4.0 (2.0-7.0).
[b] Median (interquartile range) doses for treated patients: pemetrexed, 6.0 (4.0-6.0); cisplatin 5.0 (3.0-6.0); carboplatin 6.0 (4.0-6.0).
[c] 3 deaths due to nivolumab + ipilimumab: pneumonitis, encephalitis, and acute heart failure.
[d] 1 death due to chemotherapy: myelosuppression.
Source: Table 20 of the CS.[2]

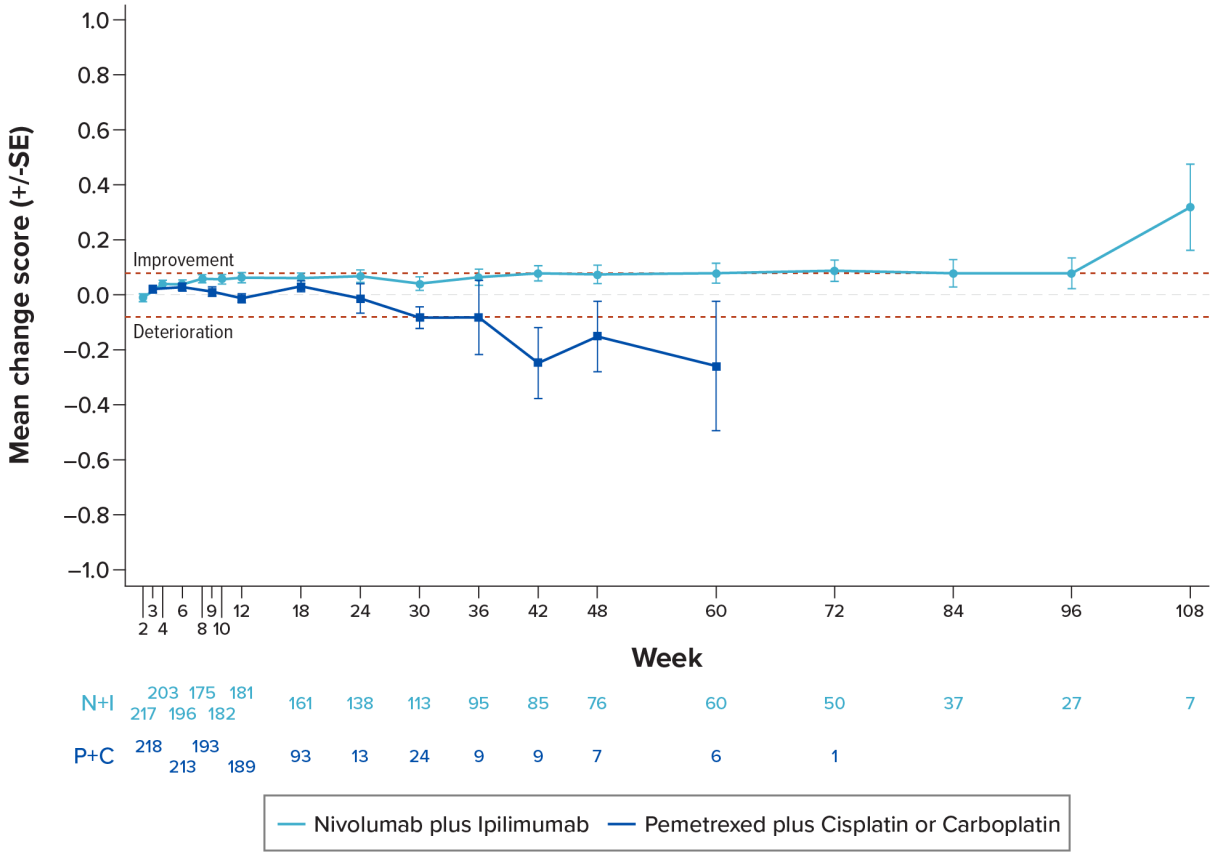
ERG comment: The rate of SAEs, treatment-related SAEs and AEs leading to discontinuation was considerably higher in patients treated with nivolumab + ipilimumab than for those in the PDC group. Although the rate of death was higher in the PDC group, three patients were reported to have died due to study drug toxicity in the nivolumab + ipilimumab arm: one of these died because of pneumonitis, for which there were also three Grade 3-4 events, respiratory tract infections also being more common

in the nivolumab + ipilimumab group than in the PDC group (See Table 3.12). A further patient in the nivolumab + ipilimumab group died of acute heart failure; there were also more Grade 3-4 cardiovascular events in the nivolumab + ipilimumab group than in the PDC group.

3.2.4.5 Health-related quality of life

The CS noted that patients who received first-line nivolumab + ipilimumab identified their HRQoL during the treatment period as stable or improved when compared to patients who received PDC and experienced deterioration in HRQoL during the treatment and follow-up periods.[2] In the current submission, HRQoL was measured using European Quality of Life-5 Dimensions 3 levels (EQ-5D-3L) Utility Index, EQ-5D-3L visual analogue scale (VAS), and Lung Cancer Symptom Scale–Mesothelioma (LCSS-Meso) scales.[2] According to the EQ-5D-3L Utility Index, patients treated with nivolumab + ipilimumab showed improved EQ-5D scores from 0.6959 at baseline to a peak score of 0.8529 at week 84.[2] Patients treated with PDC were observed to remain stable until week 30, after which EQ-5D scores indicated a deterioration from baseline, as depicted in Figure 3.3.[2] These changes were reported to have been clinically meaningful, having exceed the MID, defined as the smallest change considered to be clinically meaningful, has been estimated to be a change from baseline of 0.08 for the EQ-5D-3L Utility Index score.[2]

Figure 3.3: EQ-5D-3L Utility Index: mean change from baseline scores by treatment group (patient-reported outcome analysis population)



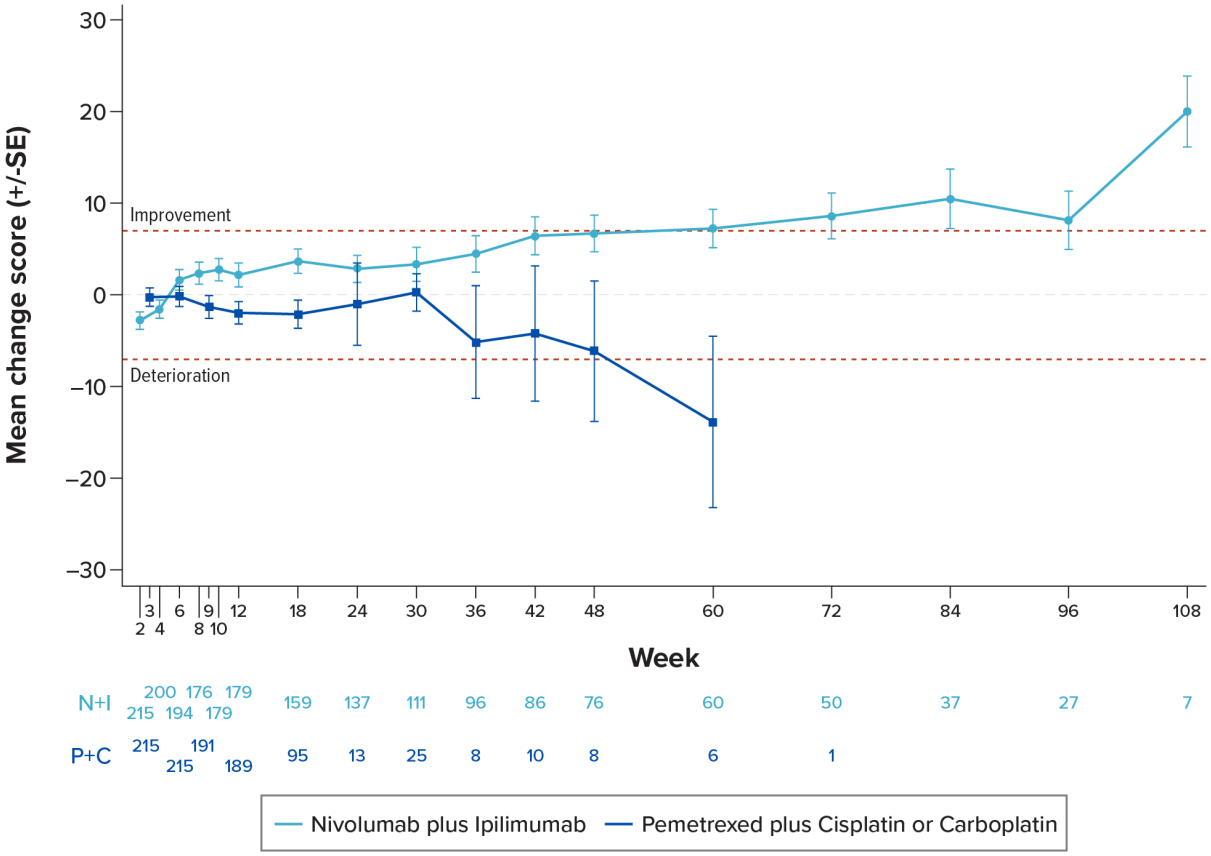
N+I = nivolumab + ipilimumab; P+C = pemetrexed + cisplatin or carboplatin; PRO = patient-reported outcome; SE = standard error.

Note: PRO analysis population includes all patients with EQ-5D baseline data and data at 1 or more postbaseline visits. The EQ-5D Utility Index score ranges from −0.594 to 1, with higher scores indicating better health state. Only time points with > 5 patients are shown.

Source: Figure 19 of the CS.[2]

The overall health of the patients was assessed using the EQ-5D VAS (see Figure 3.4).[2] There was an observed trend for improvement in the nivolumab + ipilimumab arm, which was identified as being clinically meaningful (greater than seven-point difference) from week 60.[2] Patients in the nivolumab + ipilimumab arm showed a clinically meaningful improvement in mean EQ-5D VAS scores from baseline, 69.9, to 82.7 at week 72.[2] A trend toward scores indicating deterioration was observed in the PDC arm from week 3 to week 24 and again from week 36 to week 60.[2] However, this trend was not determined to be clinically meaningful.

Figure 3.4: EQ-5D VAS: mean change from baseline scores by treatment group (patient-reported outcome analysis population)



N+I = nivolumab + ipilimumab; P+C = pemetrexed + cisplatin or carboplatin; PRO = patient-reported outcome; SE = standard error; VAS = visual analogue score.
Note: PRO analysis population includes all patients with EQ-5D baseline data and data at 1 or more postbaseline visits. The EQ-5D VAS score ranges from 0-100, with higher scores indicating better health state. Only time points with > 5 patients are shown.
Source: Figure 20 from the CS.[2]

According to the LCSS-Meso Average Symptom Burden Index (ASBI), patients treated with nivolumab + ipilimumab experienced a clinically meaningful improvement in mean score change from baseline to week 72.[2] During this time, patients in the PDC arm remained stable.[2] A similar pattern was observed for the LCSS-Meso 3IGI.[2]

ERG comment: The company described the change in EQ-5D (both 3L and VAS) as 'clinically meaningful'. It is not clear to the ERG why the decrease in EQ-5D VAS was not regarded as clinically meaningful given that it seemed to cross the seven-point threshold. Nevertheless, it does seem to be the case that the trend for nivolumab + ipilimumab indicated probable stability or improvement whereas that for PDC indicated probable deterioration.

3.2.4.6 Subsequent therapy

Subsequent systemic therapy was received by 44% and 41%; subsequent immunotherapy by 3% and 20%, and subsequent chemotherapy by 43% and 32% patients in the nivolumab + ipilimumab arm and in the PDC arm respectively, as reported in Table 6.5.3-1 in the CSR.[7]

ERG comment: The ERG asked the company to explain the differences between the two arms, with respect to the choice of subsequent therapy, and to discuss the likely implications of these differences for the relative effectiveness of nivolumab + ipilimumab vs. PDC. The ERG also requested evidence that the types of subsequent therapy in the trial are those that would also be used in England NHS practice or, if this is not the case, for the company to discuss the likely implications of any discrepancy. In response to clarification the company stated that the effect of any difference would probably be minimal given that survival on subsequent therapy is so short; they cited real-world data from the CAS registry of patients with unresectable MPM in England from January 2013-December 2017, which showed that median OS was 8.5 months from start of second-line therapy and median treatment duration of second-line therapy was 1.6 months.[4] However, the source provided and cited by the company did not report those numbers.[1] With the FAC, the company have subsequently provided the poster for that reference, which does report those figures. The company also argued that the type of subsequent therapy employed in the trial was likely to be representative of English NHS practice, again citing the same source as showing that of those who received a second-line therapy, 43.6% received second-line PDC (platinum + pemetrexed), 18.6% received second-line treatment in a clinical trial, and 24.1% received second-line vinorelbine. However, these figures could not be located by the ERG in that source.[1] Again, the poster provided with the FAC does report those figures. Nevertheless, they do appear to be quite different to those in the PDC arm of the CheckMate-743 trial of: pemetrexed (15.9%), vinorelbine (8.3%).[7] Therefore, there remains an issue as to both the effect of variation between arms and between the CheckMate-743 trial and English NHS practice.

3.2.5 Subgroup analyses

Subgroup analyses were conducted as specified in the NICE scope, i.e. according to PD-L1 status and histological subtype.

3.2.5.1 PD-L1 status

Table 3.13 shows the results of subgroup analyses by PD-L1 status. The nivolumab + ipilimumab arm among PD-L1 ≥ 1% produced a greater OS benefit than those with PD-L1 <1%.[2] The median OS among those treated with nivolumab +ipilimumab with PD-L1 <1% was 17.3 months (95% CI, 10.1to 24.3 months), whereas those treated with nivolumab + ipilimumab with PD-L1≥ 1% observed a median OS of 18.0 months (95% CI, 16.8 to 21.5 months).[2] Patients treated with PDC with PD-L1 <1% had a median OS of 16.5 months (95% CI, 13.4 to 20.5 months), whereas patients treated with PDC with PD-L1 ≥ 1% had a median OS of 13.3 months (95% CI, 11.6 to 15.4 months).[2]

When considering patients with PD-L1-positive tumours, nivolumab + ipilimumab appeared to have a beneficial effect on PFS (HR, 0.81; 95% CI, 0.64 to 1.01) when compared to patients treated with PDC. However, when considering patients with PD-L1 negative tumours, PFS favoured PDC (HR, 1.79 (95% CI, 1.21 to 2.64).[2] The company noted that the sizes of these groups were not balanced as 135 patients were included in the PD-L1 <1% group and 451 patients were in the PD-L1≥ 1% group.[2]

The ERG requested further information regarding the results of any formal statistical analyses for the comparison of all PD-L1 expression-related outcomes. However, the company reiterated all available results related to the PD-L1 subgroup had been presented.[4] The company also noted that PD-L1 was

not a stratification factor of the CheckMate-743 trial and was limited by potential imbalances in known or unknown prognostic factors.[4] Due to the small sample size and event counts in the PD-L1 negative subgroup, statistical analyses should be interpreted with caution.[4]

Table 3.13: Subgroup analyses by PD-L1 status

Outcome	PD-L1 < 1% (n=135)		PD-L1 ≥ 1% (n=451)	
	Nivolumab + ipilimumab (n=57)	PDC (n=78)	Nivolumab + ipilimumab (n=232)	PDC (n=219)
OS				
Median OS (95% CI), months[a]	17.3 (10.1-24.3)	16.5 (13.4-20.5)	18.0 (16.8-21.5)	13.3 (11.6-15.4)
HR[b] (95% CI) vs. PDC	0.94 (0.62-1.40)		0.69 (0.55-0.87)	
No. of events	40	58	150	157
PFS by BICR				
Median PFS[a] (95% CI), months	4.1 (2.7-5.6)	8.3 (7.0-11.1)	7.0 (5.8-8.5)	7.1 (6.2-7.6)
HR[b] (95% CI) vs. PDC	1.79 (1.21-2.64)		0.81 (0.64-1.01)	
No. of events	50	53	156	152
ORR per BICR				
ORR,[c] % (95% CI)	21.1 (11.4-33.9)	38.5 (27.7-50.2)	43.5 (37.1-50.2)	44.3 (37.6-51.1)
Best overall response, n (%)				
CR	0	0	3 (1.3)	0
PR	12 (21.1)	30 (38.5)	98 (42.2)	97 (44.3)
Stable disease	28 (49.1)	38 (48.7)	79 (34.1)	84 (38.4)
Progressive disease	16 (28.1)	6 (7.7)	37 (15.9)	8 (3.7)

Source: Table 15, CS.[2]
BICR = blinded independent central review; CI = confidence interval; CR = complete response; HR = hazard ratio; ORR = objective response rate; OS = overall survival; PDC = platinum-based doublet chemotherapy; PD-L1 = programmed death-ligand 1; PFS = progression-free survival; PR = partial response.
[a] Kaplan-Meier estimates.
[b] Unstratified Cox proportional hazard model.
[c] Number of (CR+PR) ÷ number of patients. CI based on the Clopper and Pearson method.

3.2.5.2 Histological subtype

Table 3.14 shows the results of analyses by histological subtype; this table was populated from Appendix F.[6]

The company noted that patients treated with nivolumab + ipilimumab demonstrated an improved OS when compared to patients treated with PDC.[2] Patients treated with nivolumab + ipilimumab were noted to have a similar median OS across histologies.[2] Those with epithelioid MPM had a median OS of 18.7 months, whereas those with non-epithelioid MPM had a median OS of 18.1 months.[2] For patients treated with PDC, the median OS was observed to be lower in the non-epithelioid subgroup when compared to the epithelioid subgroup, 8.8 and 16.5 months, respectively.[2]

In patients treated with nivolumab + ipilimumab, who had non-epithelioid MPM, an improved PFS was identified when compared with treatment with PDC (HR, 0.58; 95% CI, 0.38-0.90).[2] In patients treated

with PDC, with epithelioid MPM, PFS improved over nivolumab + ipilimumab (HR, 1.14; 95% CI, 0.92 to 1.41). However, when considering patients treated with nivolumab + ipilimumab, the median PFS was noted to be of a longer duration, 8.31 months, in the non-epithelioid subgroup when compared to the epithelioid subgroup, 6.18 months.[2] In patients treated with PDC, the median PFS was shorter in the non-epithelioid subgroup, 5.59 months, when compared to the epithelioid subgroup, 7.66 months.[2]

The ERG requested further clarification regarding the results of subgroup analyses by histological subtype for response outcomes. In the response to clarification, the company stated that the assessment of outcomes by more specific non-epithelioid subgroups was limited due to the small number of patients in each non-epithelioid subtype.[4] The company also noted that formal statistical analyses were not done.[4] The company stated in their response to clarification that in real-life clinical practice in the UK, a high proportion of patients with MPM have unknown or not otherwise specified histology.[4] However, the company emphasised that the treatment effect of nivolumab + ipilimumab versus PDC was consistent across the histological subtypes.[4]

Table 3.14: Subgroup analyses by histological subtype

	Non-epithelioid (n=149)		Epithelioid (n=456)	
Outcome	**Nivolumab + ipilimumab (n=74)**	**PDC (n=75)**	**Nivolumab + ipilimumab (n=229)**	**PDC (n=227)**
OS				
Median OS (95% CI), months[a]	18.07 (12.16-22.77)	8.80 (7.43-10.15)	18.73 (16.92-21.98)	16.49 (14.88-20.47)
HR[b] (95% CI) vs. PDC	0.46 (0.31-0.68)		0.86 (0.69-1.08)	
No. of events	50	63	150	156
PFS by BICR				
Median PFS[a] (95% CI), months	8.31 (4.11-10.25)	5.45 (5.09-6.80)	5.98 (5.39-6.97)	7.75 (7.16-8.34)
HR[b] (95% CI) vs. PDC	0.56 (0.37-0.85)		1.16 (0.93-1.45)	
No. of events	51	54	167	155
ORR per BICR				
ORR,[c] % (95% CI)			38.4 (32.1-45.1)	47.6 (40.9-54.3)

Sources: Figures 4 and 5, Table 13, Appendix F.[6].
BICR = blinded independent central review; CI = confidence interval; CR = complete response; HR = hazard ratio; ORR = objective response rate; OS = overall survival; PDC = platinum-based doublet chemotherapy; PD-L1 = programmed death-ligand 1; PFS = progression-free survival; PR = partial response.
[a] Kaplan-Meier estimates.
[b] Unstratified Cox proportional hazard model.
[c] Number of (CR+PR) ÷ number of patients. CI based on the Clopper and Pearson method.

ERG comment: In terms of OS, nivolumab + ipilimumab appears to be clearly more effective than PDC in patients with MPM with PD-L1 ≥ 1% and in patients with MPM with non-epithelioid histology. Although not so clear and with a reduced difference, nivolumab + ipilimumab also appears to be more effective than PDC, for epithelioid histology. There appears to be little difference between treatments for PD-L1 < 1%.

In terms of PFS, nivolumab + ipilimumab appears to be clearly more effective than PDC for MPM with non-epithelioid histology. Although not so clear and with a reduced difference, nivolumab + ipilimumab also appears to be more effective than PDC for MPM with PD-L1 ≥ 1%. Nivolumab + ipilimumab appears to be clearly less effective than PDC for MPM with PD-L1 < 1%. Although not so clear and with a reduced difference, nivolumab + ipilimumab also appears to be less effective than PDC for MPM with epithelioid histology.

There remains a question about the potential interaction effects of these two clinically relevant subgroups; no data are available for subgroup combinations, e.g. PD-L1<1% and epithelioid histology. This factor, in combination with data immaturity, means that there remains uncertainty as to the relative effectiveness of the intervention in clinically relevant subgroups.

3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison

Because only one RCT was included, no indirect comparisons were performed.[2]

3.4 Critique of the indirect comparison and/or multiple treatment comparison

Because only one RCT was included, no indirect comparisons were performed.[2]

3.5 Additional work on clinical effectiveness undertaken by the ERG

None.

3.6 Conclusions of the clinical effectiveness section

The CS included a systematic review, which appears to have been largely well conducted.[2] This probably included more studies than are required if one considers that the most appropriate evidence is the CheckMate-743 trial, given that this can be assumed to be the only trial that compares the intervention to PDC and assuming that PDC is the only relevant comparator. The CheckMate-743 trial is an RCT that compares nivolumab + ipilimumab with PDC in MPM, the population specified in the scope, the primary outcome being OS, but also reporting all other outcomes listed in the scope including ORR, PFS, HRQoL and AEs. The population in CheckMate-743 was narrower than that of the scope, including only patients with ECOG PS 0-1, but the company confirmed that this was the population that they wanted to be considered in this appraisal. The quality of the RCT was diminished by the lack of blinding: other than that, it could be regarded as of high quality. As discussed in Section 2.2, there is also a discrepancy between the dosing of nivolumab in the trial, which was weight-based, and that of the proposed marketing authorisation, which will be fixed. The ERG did request evidence that this difference in dosing will have no effect on effectiveness, quality of life or safety.[3] However, although the company's response does provide some reassurance regarding safety, a judgment of safety/tolerance is not a substitute for actual AE rates at the given fixed dose.[4] The presentation also provided by the company stated that: *"…efficacy and safety were evaluated by characterizing the relationships between simulations of NIVO exposure and OS or grade ≥ 2 immune-mediated adverse events (grade 2+ IMAEs), respectively, using the multivariate Cox proportional-hazard model*".[5] However, it is not clear to the ERG precisely how outcomes could be estimated for a fixed dose without evidence from patients who received that dosing regimen. Subgroup analyses by weight were also provided, but again these do not show the effect of patients receiving a lower or higher dose, as would have been the case if dosing had been weight-based.[5] Therefore, an issue remains regarding the effectiveness and safety of the expected licensed dose of nivolumab.

The ability to inform a comparison with any specific form of PDC was also affected by there being IC of either carboplatin or cisplatin. Whilst such a choice is consistent with clinical practice, because only 38 patients were from the UK and the extent to which the choice of cisplatin and carboplatin would be in accordance with English NHS practice is uncertain and there remains a question both of the applicability of the comparator, as discussed in Section 2.3, and the trial generally to English NHS practice.

At the time of the interim analysis with a database lock of 3 April 2020, a statistically significant OS benefit was observed for patients who were treated with nivolumab + ipilimumab when compared to patients treated with PDC.[2] However, there was no statistically significant difference in PFS and results for ORR were similar.[2] The ERG did request results from a more recent data cut, but the company replied: "*As CheckMate-743 met its primary endpoint at the 3 April 2020 database lock, this analysis was considered the final analysis. However, follow up of CM-743 is ongoing and additional data cuts are expected, likely in Q2/Q3 2021 (TBC). As the timing of the analysis is event driven, there is uncertainty on the exact timing of future database locks*."[4] Although it is unlikely that the results will change the interpretation that nivolumab + ipilimumab is more effective in terms of OS, the precise size of the difference might be important particularly in determining whether nivolumab + ipilimumab is cost effective. The interpretation of PFS results may change, given that progression data are incomplete. This data immaturity therefore remains an issue.

The ERG asked the company to explain the differences between the two arms of the CheckMate-743 trial, in the choice of subsequent therapy, and to discuss the likely implications of these differences for the relative effectiveness of nivolumab + ipilimumab vs. PDC. The ERG also requested evidence that the types of subsequent therapy used in the trial are those that would also be used in English NHS practice or, if this is not the case, to discuss the likely implications of any discrepancy. In response to clarification the company stated that the effect of any difference would probably be minimal given that survival is so short on subsequent therapy.[4] However, the source provided by the company in response to clarification does not seem to provide those data to support this statement.[1] The company also argued that the type of subsequent therapy employed in the trial was likely to be representative of English NHS practice.[4] However, again the figures mentioned by the company could not be located by the ERG in that same source and they do appear to be quite different to those in the PDC arm of the CheckMate-743 trial in terms of pemetrexed and vinorelbine use.[7] Therefore, there remains an issue as to the effect of variation in subsequent therapy, both between arms and between the CheckMate-743 trial and English NHS practice.

The subgroup analyses specified in the scope, according to PD-L1 status and histology, were performed and did indicate some important variation in the effectiveness of nivolumab + ipilimumab versus PDC.

[redacted] There remains a question about the potential interaction effects of these two

clinically relevant subgroups; no data are available for subgroup combinations, e.g. PD-L1<1% and epithelioid histology. This factor, in combination with data immaturity, means that there remains uncertainty as to the relative effectiveness of the intervention in clinically relevant subgroups.

4. COST EFFECTIVENESS

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

This section pertains mainly to the review of cost effectiveness analysis studies. However, the search section (4.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the CS. Therefore, the following section includes searches for the cost effectiveness analysis review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

4.1.1 Searches performed for cost effectiveness section

Appendix H and I of Document C report details of an updated SLR to identify evidence about cost effectiveness, HRQoL and healthcare resource use for patients with MPM. All databases were searched from inception to 5 October 2020. Searches of NHS EED and DARE were originally undertaken in March 2018. These searches were not updated as these databases are no longer being added to. Appendix H reported searches undertaken for economic evaluations for MPM while Appendix I reported searches for HRQoL and utilities for MPM. No language limits were reported in the search strategies. A cost filter, however, was applied to 2018 searches of NHS EED which may have compromised the retrievability of potentially relevant studies. A summary of sources searched is provided in Table 4.1.

Table 4.1: Data sources for the cost effectiveness systematic review

	Resource	Host/source	Date ranges	Dates searched
Electronic databases	Embase	Embase.com	From inception	5.10.20
	MEDLINE	Embase.com		5.10.20
	NHS Economic Evaluations Database (NHS EED)	Wiley		9.5.18
	Database of Abstracts of Reviews of Effects (DARE)	Wiley		9.5.18
	MEDLINE In-Process and Ahead of print	PubMed		5.10.20
	EconLit	AEAweb.org		5.10.20
	International HTA Database			5.10.20
Conference proceedings	ASCO	https://meetinglibrary.asco.org/results/(Keywords:"Mesothelioma");page=0	2018-2020	October 2020
	ESMO	https://www.sciencedirect.com/search?qs=mesothelioma&pub=Annals%20of%20Oncology&cid=321639&years=2020&lastSelectedFacet=years	2018-2020	
	AACR	https://www.aacr.org/professionals/meetings/previous-aacr-meetings/previous-aacr-meetings-2018/	2018	

		https://www.aacr.org/professionals/meetings/previous-aacr-meetings/previous-aacr-meetings-2019/	2019	
		https://www.aacr.org/meeting/aacr-annual-meeting-2020/abstracts/ https://cancerres.aacrjournals.org/content/80/16_Supplement	2020	
	ISPOR	https://www.ispor.org/heor-resources/presentations-database/search	2018-2020	
	WCLC	https://wclc2018.iaslc.org/wp-content/uploads/2018/09/WCLC2018-Abstract-Book_vF-LR-REV-SEPT-25-2018.pdf	2018	
		https://wclc2019.iaslc.org/wp-content/uploads/2019/08/WCLC2019-Abstract-Book_web-friendly.pdf	2019	
	ELCC	https://www.jto.org/issue/S1556-0864(18)X0004-5	2018	
		https://www.sciencedirect.com/journal/annals-of-oncology/vol/30/suppl/S2?page=3#article-201	2019	
	IMIG	Not searched		

Additional resources
National Institute for Health and Care Excellence (NICE)
Scottish Medicines Consortium (SMC)
Institute for Quality and Efficiency in Health Care (IQWiG)
Autorité de Santé (HAS)
Canadian Agency for Drugs and Technologies in Health (CADTH)
Pharmaceutical Benefits Advisory Committee (PBAC)
Food and Drug Administration (FDA)
European Medicines Agency (EMA)
Gesellschaft der Epidemiologischen Krebsregister in Deutschland (GEKID)
Belgian Cancer Registry
Dutch Cancer Registry
Italian Association of Cancer Registries (ITACAN)
Red Española de Registros de Cáncer (REDECAN)
Nordic Cancer Registry (NORDCAN)
Surveillance, Epidemiology, and End Results Program (SEER)
National Lung Cancer Audit annual report

ERG comment

- The CS provided sufficient details for the ERG to appraise the literature searches. A good range of database and conference proceedings were searched, including additional grey literature resources. Reference checking was also undertaken.
- Searches overall were well-conducted and were transparent and reproducible.
- No date limits were unnecessarily applied. There was an English language restriction, but this was not applied at the searching stage.
- Cochrane Library searches conducted in March 2018 of NHS EED and DARE applied a cost filter (Appendix H). NHS EED is a database of cost evaluations and applying an additional filter will have affected the retrievability of possibly relevant records and is not recommended.[14] In response to clarification, the company confirmed that one search strategy had been used to search Cochrane Library databases and that filters had been applied to "maximise sensitivity and precision" as CENTRAL also includes cost publications. However, the ERG is concerned that the unnecessary application of a filter to a pre-filtered resource such as NHS EED compromised the sensitivity of finding potentially relevant cost studies and that this resource should have been searched separately without the application of a filter.

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on cost effectiveness studies, HRQoL studies and costs and resource use studies were not clearly presented in the CS, but are included in the flow charts of the three reviews in Figures 6-8 in Appendix H.[6]

ERG comment: The ERG was unsure whether the applied eligibility criteria were suitable to fulfil the company's objective to identify cost effectiveness studies in this disease area as explanations for some exclusion criteria were lacking, namely for: "Line of treatment unclear", "No SGA disease" and "No SGA LOT". The company provided justification and explanation in response to clarification question B1 and the ERG was satisfied that it was unlikely that any studies were missed.[4]

4.1.3 Conclusions of the cost effectiveness review

A total of 23 economic evaluation studies were identified, including nine with cost effectiveness analyses, which are presented in Table 24 of the CS; one more study was added in Table 21 in Appendix H. These 10 studies were summarised in Appendix H. None of these 10 economic evaluations considered nivolumab + ipilimumab for the treatment of MPM. There are only few published economic evaluations of treatments for MPM. The company also stated that *"the majority of published analyses have considered the combination treatment of pemetrexed plus cisplatin. Past analyses have been limited in scope, both in terms of time horizon and the inclusion of all relevant comparators. There is no apparent established modelling methodology at this stage, with previous analyses having adopted various approaches (from simple trial-based analyses which do not distinguish between progression-free and progressed disease, to partitioned survival modelling and Markov modelling). Preference-based quality of life data to provide utility values for cost-effectiveness analyses is a crucial data gap."*[2]

ERG comment: Eligibility criteria were suitable for the SLR performed. The CS provides an acceptable overview of the included cost effectiveness, HRQoL and resource use and costs studies.

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

4.2.1 NICE reference case checklist

Table 4.2: NICE reference case checklist

Element of health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	In line with reference case
Perspective on costs	NHS and PSS	In line with reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	In line with reference case
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	In line with reference case
Synthesis of evidence on health effects	Based on systematic review	In line with reference case
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.	In line with reference case
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	In line with reference case
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	In line with reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	In line with reference case
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	In line with reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	In line with reference case

EQ-5D = Euroqol-5D; NHS = National Health Service; PSS = personal and social services; QALY = quality-adjusted life year; UK = United Kingdom

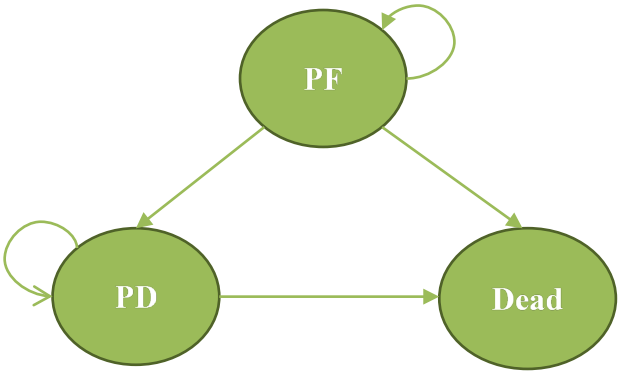
4.2.2 Model structure

The analysis was based on a three-health state partitioned survival model, using a cycle time of one week to accommodate the administration cycles for therapies considered in the model. The model was developed in Microsoft Excel and programmed using standard Excel functions, where possible.

The states in the model are progression free (PF), progressed disease (PD), and dead (Figure 4.1). The three health states represent the primary stages of disease in MPM: PF with first-line treatment, the occurrence of disease progression, and death. These health states correspond to the primary and secondary endpoints of the CheckMate-743 trial. This model structure is consistent with the approaches adopted in previous published economic evaluations within MPM and previous NICE technology appraisals of oncology products.

Patients enter the model in the PF health state. At the end of each cycle, the proportion of patients in PF, PD, and dead is calculated from parametric survival curves for PFS and OS estimated from the CheckMate-743 trial. Specifically, the number of patients occupying each state in the model is derived directly from the cumulative survival probabilities of PFS and OS (area under the curve approach), with the proportion of patients in the PD health state being calculated as the difference between OS and PFS (see CS Figure 25).[2]

Figure 4.1: Model structure



Source: Based on CS Figure 24[2]

ERG comment: The main concern of the ERG relates to the use of a partitioned survival model given the issues highlighted in NICE Decision Support Unit (DSU) Technical Support Document (TSD) 19.[15]

In clarification question B4 the company was asked to justify the use of a partitioned survival model given the issues highlighted in NICE DSU TSD 19 and to use state transition modelling to assist in verifying the plausibility of the partitioned survival model extrapolations and to address uncertainties in the extrapolation period (NICE DSU TSD 19, recommendation 11).[15] This company justified the use of the partitioned survival model based on European advisory board for the economic modelling, indicating that most advisors agreed with a partitioned survival model being used. Moreover, the company responded that it is unlikely that using a state transition model would have a large impact on outcomes as 1) post-progression treatments are administered for a short duration in this indication and 2) state transition models, in general, per se do not necessarily result in different results compared to partitioned survival models. The company did not however provide supporting evidence that the difference in this specific case would be minimal, it is unclear to the ERG why the duration of post-progression treatments is mentioned as an argument by the company. Hence the impact of the limitations related to the partitioned survival model (highlighted in NICE DSU TSD 19), such as the extrapolations of PFS and OS while assuming structural independence between these endpoints, is unclear. This is particularly relevant given the large proportion of (PF)LY that is accumulated beyond the observed data (see Section 5.1).

4.2.3 Population

The economic evaluation considers nivolumab + ipilimumab in the first-line treatment of adults with untreated unresectable MPM. The company stated that this was consistent with the study population of CheckMate-743, the decision problem and the anticipated licensed indication. No subgroup analyses were presented.

ERG comment: The main concerns of the ERG relate to: a) the population being narrower than the scope; and b) no subgroups being presented despite being listed in the scope.

a) The population in CheckMate-743 was narrower than that of the scope, limiting eligible patients to those with an ECOG status 0-1. The company clarified that no formal restriction with respect to ECOG status is made, as for many patients the ECOG status is unrecorded (see Section 2.1 of this report).
b) The company did not present subgroup cost effectiveness analyses despite relevant subgroups being listed in the scope, such as histologic subtype (epithelioid, sarcomatoid, biphasic) and level of PD-L1 expression. In response to clarification question B3, the company explained that it did *"not consider economic modelling of nivolumab + ipilimumab by histological subtype or PD-L1 expression as appropriate, given the high clinical unmet need of all patients with unresectable MPM eligible for SACT and the OS benefit seen in all subgroups in CheckMate-743."*[4] The company also considered the clinical data that was presented in the CS for the histological and PD-L1 subgroups in CheckMate-743 as descriptive in nature and that it should be interpreted with caution. Section 3.2.5 provides further detail on this issue. The ERG concludes that cost effectiveness may vary by subgroup.

4.2.4 Interventions and comparators

The intervention considered in the CS was nivolumab + ipilimumab, administered at a flat nivolumab dosage of 360 mg every three weeks, aligning with the anticipated EMA licence. This differs from the nivolumab weight-based dosage of 3 mg/kg every two weeks used in CheckMate-743. Ipilimumab is administered every six weeks at 1 mg/kg, which is in line with CheckMate-743. The CS includes a two-year stopping rule for nivolumab + ipilimumab, which is also in line with CheckMate-743.

The comparator considered was pemetrexed (500 mg/m2 every three weeks for six treatment cycles) + cisplatin (75 mg/m2 every three weeks for four treatment cycles) or carboplatin (550 mg/m2 every three weeks for four treatment cycles). According to the CS, pemetrexed + cisplatin or carboplatin is considered the standard of care therapy in the UK and is consistent with the comparator arm of the CheckMate-743 clinical trial.

The NICE scope also listed BSC and raltitrexed + cisplatin (for people for whom treatment with pemetrexed is unsuitable) as comparators, but these were not included. The company justified the selection of the comparators considering that raltitrexed was not approved for use in the UK for the first-line treatment of MPM and was not used in the NHS according to UK registry data and expert opinion. BSC was not considered an appropriate comparator because nivolumab + ipilimumab relates to a particular group of fit patients for whom BSC would not be deemed acceptable or ethical unless specifically requested by the patient.

ERG comment: The main concerns of the ERG relate to: a) the omission of potentially relevant comparators listed in the NICE scope (BSC and raltitrexed + cisplatin); b) the dosage differs between

this submission and the evidence; c) the use of a two-year stopping rule; d) pemetrexed + cisplatin and pemetrexed + carboplatin could be considered separate comparators.

a) The omission of potentially relevant comparators listed in the NICE scope (BSC and raltitrexed + cisplatin). As detailed in Section 2.3 of this report, the exclusion of BSC and raltitrexed + cisplatin was justified by the company and the ERG agrees that this is acceptable.
b) The dosage differs between this submission (which is in line with the anticipated marketing authorisation) and the evidence from CheckMate-743. As detailed in Section 2.2 of this report, uncertainty remains regarding the effectiveness and safety of the expected licensed dose of nivolumab.
c) The use of a two-year stopping rule for nivolumab + ipilimumab in the CS was in line with the evidence from CheckMate-743. It should be noted that [redacted] as shown in Figure 10 in Appendix K of the CS, despite the stopping rule stipulated in the study protocol. The impact of this on cost effectiveness results (particularly on costs) would be likely small, but it could be an important issue should this occur for further patients or should the stopping rule not be adhered to in clinical practice. The ERG therefore considers it important to explore whether more patients continued nivolumab + ipilimumab beyond 24 months and how long they continued treatment in future analyses.

4.2.5 Perspective, time horizon and discounting

The analysis was performed from the NHS and Personal Social Services (PSS) perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length is one week to accommodate the administration cycles of the included therapies, with a lifetime time horizon (20 years), and a half-cycle correction was applied.

ERG comment: In the CS, the company states a 20-year time horizon was used, and the model continues until patients reach the age of 88 years (less than 1% of patients are still alive). This was considered to represent a lifetime time horizon. The approach is in concordance with the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness used for intervention and comparators is the April 2020 database lock of CheckMate-743 (minimum follow-up for all patients was 22.1 months; 23% and 15% of patients treated with nivolumab + ipilimumab and PDC, respectively were still alive at this point). To estimate PFS and OS over the 20-year time horizon, parametric survival curves were fitted to CheckMate-743 patient-level data and used to extrapolate survival beyond the study time horizon.

4.2.6.1 Fitting and selecting procedure of the parametric survival models

Seven parametric models were considered for the extrapolation of PFS and OS (exponential, Weibull, Gompertz, log normal, log-logistic, gamma, and generalised gamma). The process for fitting and selecting parametric survival models was based on methods guidance from the Decision Support Unit at NICE and illustrated in CS Figure 26. This process included:

1. Assessing the proportional hazards assumption by examining the scaled Schoenfeld residuals (and Grambsch and Therneau's correlation test), log-cumulative hazards, log-cumulative odds, and standardised normal curve plots. In case of (non-)proportional hazards parametric survival models were (in)dependently estimated for both treatments (i.e. nivolumab + ipilimumab and pemetrexed + cisplatin or carboplatin).

2. Assessing fit to the observed data by examining goodness-of-fit statistics (AIC/BIC). This includes using rules of thumb that indicate that models that have an AIC/BIC of <4/<2 higher than the lowest AIC/BIC are considered the best fitting models based on the Burnham and Anderson rule of thumb (AIC) and Raftery rule of thumb (BIC)[16, 17].
3. Assessing clinical plausibility and external validation of the extrapolated survival estimates by considering data from the systemic anticancer therapy (SACT) population with newly diagnosed MPM from the Cancer Analysis System (CAS)[1] and the MAPS trial investigating bevacizumab + pemetrexed + cisplatin compared with pemetrexed + cisplatin for the treatment of patients with newly diagnosed unresectable MPM [18]. Particularly, the survival function as well as the shape of the hazard function were considered. Additionally, UK clinical experts were consulted on the expected survival with current treatments. The clinical input received indicated that five-year survival would be expected at 5%, 7.5-year survival at 2%, and 10-year survival at 0-2%.

4.2.6.2 Overall survival

The fitting and selecting procedure for OS is described considering the above-mentioned three criteria (see also CS Table 32).

1. **Proportional hazards assumption.** Based on CS Figure 30, non-proportional hazards were assumed, and the parametric survival models were fitted separately for both nivolumab + ipilimumab and PDC.
2. **Fit to the observed data.** For nivolumab + ipilimumab based on statistical goodness-of-fit and the abovementioned rules of thumb, the parametric survival models with the *Weibull, gamma and Gompertz distributions* might be considered the best fitting models (i.e. difference in AIC/BIC of <4/<2 compared with the lowest AIC/BIC; the *generalised gamma distribution* as well when only considering the AIC), see CS Table 28. For PDC these were the parametric survival models with the *gamma and log-logistic distributions* (the *generalised gamma and Weibull distributions* as well when only considering the AIC), see CS Table 30.
3. **Clinical plausibility and external validation of the extrapolated survival.** It was considered that, for both treatments, the modelled **hazard function** of the selected distribution should have an initial increase in hazards followed by long-term decreasing hazards (based on CS Figure 28, derived from MAPS data; according to clarification response B6 smoothed hazard plots based on CheckMate-743 and SACT data provided similar shapes for the hazard function). This was only observed for the parametric survival models using the *log-logistic and log-normal distributions* and for PDC using the *generalised gamma distribution* as well (though the decline of the hazard over time was smaller than for the log-logistic and log-normal distributions).
   Parametric survival models with distributions with **predicted survival** for PDC slightly below the survival observed in MAPS are appropriate; predictions aligned with the MAPS data were considered neutral, and predictions above or significantly below survival in MAPS were considered inappropriate. For nivolumab + ipilimumab, predicted survival that is lower than that observed for PDC in MAPS was considered inappropriate. Based on this criterion, *the log-logistic and log-normal distributions* were appropriate for nivolumab + ipilimumab while these were *the exponential, log-logistic and generalised gamma distributions* for PDC.

Based on these findings, the company selected a piecewise approach combining Kaplan-Meier (KM) data (up to the 22 months break point) with independently estimated parametric survival models for extrapolation (i.e. assuming non-proportional hazards). The selected parametric survival models were based on the log-logistic and exponential distributions for nivolumab + ipilimumab and PDC

respectively. The 22 months break point was selected as it was the approximate minimum patient follow-up at the database lock of CheckMate-743, and most censoring in the OS data in both treatment arms occurred after this point (see CS Figure 12).

4.2.6.3 Progression-free survival

The fitting and selecting procedure for PFS is described considering the above-mentioned three criteria.

1. **Proportional hazards assumption.** Based on CS Figure 30, non-proportional hazards are assumed, and the parametric survival models are fitted separately for both nivolumab + ipilimumab and PDC.
2. **Fit to the observed data.** For nivolumab + ipilimumab based on statistical goodness-of-fit and the abovementioned rules of thumb, the parametric survival model with the *generalised gamma distribution* might be considered the best fitting models (i.e. difference in AIC/BIC of <4/<2 compared with the lowest AIC/BIC), see CS Table 33. For PDC this was the parametric survival model with the *log-logistic distribution*, see CS Table 35.
3. **Clinical plausibility and external validation of the extrapolated survival.** The validation of this criterion for PFS is not explicitly described in the CS. Notably, the company stated that for PFS the selection of the parametric survival models was primarily guided by statistical and visual fit to the CheckMate-743 data for both treatment arms. As it has been shown previously that PFS for immunotherapies does not follow the same pattern as for other oncology treatments, the MAPS data were not considered appropriate for validating PFS for nivolumab + ipilimumab.

Based on these findings, the company selected parametric survival models with the generalised gamma and log-logistic distributions for nivolumab + ipilimumab and PDC respectively. If PFS is greater than OS at any time, the PFS is assumed to be equivalent to OS to avoid inconsistencies between OS and PFS.

4.2.6.4 Potential waning of treatment effect

In the CS base-case no treatment waning was assumed, i.e. the PFS and OS were assumed to be different for nivolumab + ipilimumab and PDC for the whole duration of the time horizon.

ERG comment: The main concerns of the ERG relate to: a) the approach to estimate OS; b) plausibility of long-term extrapolation of PFS; c) assuming no treatment waning in the CS base-case.

a) The selection of a piecewise approach to estimate OS for both nivolumab + ipilimumab and PDC was clarified in response to clarification question B5. Here the company stated: "*The decision to utilise a piecewise model was primarily guided by the PDC arm. As presented in the CS, distributions with the best statistical and visual fit to the KM data for the PDC arm did not provide plausible long-term extrapolations. The chosen base-case distribution for PDC (exponential) provided the most plausible long-term extrapolation that was aligned with clinical expert input but had a relatively poor fit to the within-trial data (underestimating within-trial survival). Thus, to overcome this limitation for the within-trial period the piecewise approach was selected. The same issue of fit to the within-trial data was not seen to the same extent in the nivolumab + ipilimumab arm. However, for consistency the approach was applied to both arms in the model.*"[4] Although this clarifies the company's preference for the piecewise approach, using KM data up to 22 months to overcome poor fit to the observed data, the combination of the specific distributions (i.e. exponential and log-logistic) with the KM data is not clearly justified. These combinations might be evaluated differently than reported in CS

Table 32 for the different distributions (without using KM data up to 22 months). The ERG generally does not prefer using KM curves for economic models as it might overfit the trial data which seems suboptimal for decision-making in UK clinical practice. Moreover, NICE DSU TSD 21 on flexible methods for survival analysis highlights that the selected 22 months break point may be arbitrary and potentially importantly influence the results of an analysis. Finally, deviation from standard parametric survival models (opting for a piecewise approach) because of suboptimal fit to the observed data might not be warranted given the large majority of LY gains are accumulated beyond the observed data period (See Table 5.2)

In addition to the above, based on the company's response to clarification question B5c it became clear that the estimation and implementation of the piecewise models incorporated in the economic model deviate from common practice and the piecewise models described in NICE DSU TSD 21. The implemented piecewise models are using parametric survival models estimated from baseline (time = 0; using the full dataset) instead of being estimated specifically from the break point (of 22 months). This approach is flawed according to the ERG as these parametric survival models, estimated from baseline, are not intended to be used after the break point only as the proportion of patients surviving up to this break point (i.e. conditional survival) using these parametric survival models might differ from the conditional survival based on the KM curve.

Given the abovementioned limitations of the company's piecewise approach and the lack of justifications for the selecting the distributions for the piecewise approach, the ERG prefers to use a standard parametric approach to estimate OS in its base-case. Specifically, the log-logistic distribution for both treatment arms is considered a plausible alternative, as illustrated in CS Table 32 considering the goodness of fit (AIC and BIC), the appropriateness of the hazard function as well as survival extrapolations (i.e. aligned with the MAPS data). Moreover, the CS section "Heuristics for selection of survival extrapolation for OS based on external validation" describes identical hazard functions for both PDC and nivolumab + ipilimumab (i.e. the hazard function of the selected distribution should have an initial increase in hazards followed by long-term decreasing hazards) that is consistent with the log-logistic distribution. Therefore, the log-logistic distribution is used for both treatment arms in the ERG base-case.

b) Based on fit to the observed data, the company's selected approach to estimate PFS seems appropriate (using the generalised gamma distribution for nivolumab + ipilimumab and the log-logistic for PDC), these were also confirmed by clinical experts as described in response to clarification question B8. Moreover, the company provided justification for using different distributions for nivolumab + ipilimumab and PDC, highlighting the different mechanism of action that nivolumab + ipilimumab has compared to PDC. Notably, given the large majority of PFLY gains are accumulated beyond the observed data period (See Table 5.2), the plausibility of long-term extrapolation and PFS gains is arguably the most important criterion to consider. In response to clarification question B8, the company indicated that the estimated PFS for PDC was in line with MAPS trial data up to five years. The MAPS data were not considered appropriate for validating PFS for nivolumab + ipilimumab. Given the substantial uncertainty related to the plausibility of the extrapolated PFS, the ERG performed two scenario analyses to examine the impact of alternative assumptions, selected based on statistical goodness-of-fit, related to estimated PFS: 1) use log-logistic distributions for both treatment arms and 2) use generalised gamma distributions for both treatment arms.

c) In the CS base-case no treatment waning was assumed, i.e. the PFS and OS were assumed to be different for PDC and nivolumab + ipilimumab for the whole duration of the time horizon. The company justified this by stating *"there is long-term evidence of a robust and durable treatment effect lasting beyond discontinuation for immunotherapies"* (response to clarification

question B10)[4] and referring to a publication by Antonia et al.[19] considering four-year survival with nivolumab in patients with previously treated advanced non-small-cell lung cancer. Additionally, the company provided a scenario analysis where the treatment effect is assumed to start deteriorating at year five and then decrease linearly to no treatment effect at year 10. This scenario resulted in a substantial increase in the ICER which would most likely increase further when assuming no treatment effect at year five as for instance preferred by the committee in ID1585 considering nivolumab for treating squamous cell carcinoma of the head and neck; appraisal consultation document section 3.15[20]. Given that it is unclear whether assuming a continued treatment effect over the lifetime horizon of the model is plausible and the uncertainty related to the long-term extrapolations (only three patients were at risk at 36 months according to Table 13 in the clarification letter), treatment waning was assumed after five years in the ERG base-case. Although there is precedence to use the five-year treatment waning time point (as highlighted above), the ERG acknowledges that the selected time point is arbitrary.

4.2.7 Adverse events

The only source of evidence on treatment adverse events used for intervention and comparators was CheckMate-743. In the model, only treatment-related adverse events ≥ grade 3 adverse events with an incidence ≥ 2% were included (Table 39 of the CS).

ERG comment: The ERG was concerned about the exclusion of many adverse events from the model based on the company's inclusion criteria. In particular, the ERG noted that AE rates used in the model (Table 39 of the CS and later updated in response to clarification question B11[4]) were smaller in the nivolumab + ipilimumab arm compared with the PDC arm, whilst the company's Table 17 of the CS suggests that more AEs occurred in the nivolumab + ipilimumab arm compared with the PDC arm (whether treatment-related or all-cause AEs, any grade or only grade 3-4). In response to clarification question B11[4], the company clarified the source of AEs reported in Table 39 of the CS as Tables 8.5-2 and S.6.2.2 in the CheckMate-743 CSR. The latter Table S.6.2.2 was not made available to the ERG and the AE rates included with ≥ 2% and < 5% incidence used in the model could therefore not be verified.

4.2.8 Health-related quality of life

The utility values were estimated, using EQ-5D-3L (UK scoring algorithm) data obtained in CheckMate-743, for the following health states: PF and progressed disease. These health state utility values were assumed to be treatment dependent.

4.2.8.1 Health-related quality of life data identified in the review

According to the CS, the SLR identified a total of 13 studies that met the eligibility criteria for the review; however, none of the studies evaluated nivolumab + ipilimumab or used the EQ-5D in an appropriate population. Therefore, HRQOL data from CheckMate-743 were used in this submission.

4.2.8.2 Health state utility values

Patient-level utility data from CheckMate-743 were used to derive progression-based utility values for the model. Model fit based on regression models with or without treatment specific utility values were assessed. The analysis showed that treatment had a statistically significant impact on the utility values (P = 0.000). Therefore, treatment dependent health state utilities were selected for the CS base-case. Alternative treatment independent utilities were tested in scenario analyses. A summary of all these utility values is provided in Table 4.3.

Table 4.3: Health state utility values

Health state utility (standard error)	Nivolumab + ipilimumab	Pemetrexed + cisplatin or carboplatin	Difference
CS base-case: treatment dependent			
Progression free	0.737 (0.012)	0.733 (0.012)	0.004
Progressed disease	0.652 (0.014)	0.580 (0.015)	0.072
CS scenario: treatment independent			
Progression free	0.734 (0.008)	0.734 (0.008)	0.000
Progressed disease	0.620 (0.010)	0.620 (0.010)	0.000

4.2.8.3 Disutility values

Specific AE–related disutilities (retrieved from the literature, see CS Table 40) were not incorporated in the CS base-case as it was assumed that the estimated health state utilities already accounted for the AE–related disutilities. In CS scenario 5 AE–related disutilities were of [redacted] and [redacted] were implemented for nivolumab + ipilimumab and PDC respectively.

ERG comment: The main concerns of the ERG relate to: a) the data and methods used to estimate health state utilities and b) the duration of the utility benefits.

a) Details regarding the data and methods used to estimate health state utilities were lacking in the CS. In response to clarification question B12, these details were provided. Mixed models were fitted to the data (using SAS PROC MIXED), to account for repeated EQ5D assessments per subject. According to the company, no strong patterns in the missing data were indicated, and 96.2% (582/605) of all randomised patients had at least one EQ5D utility value. Moreover, the EQ5D utility data was found to be 89% complete (4,899/5,488 EQ5D assessments) with completion rates of above 80% at all on-treatment visits except week 8 and week 108 at 78% (similar in both treatment arms). Given the clarifications provided by the company, the approach used for the CS base-case seems reasonable.
b) The treatment dependent utilities, used in the CS base-case, result in utility benefits for nivolumab + ipilimumab compared to PDC. This is 0.004 and 0.072 for the PF and PD health states. The face validity of the PD utility gain for nivolumab + ipilimumab compared to PDC, as well as its representativeness for UK clinical practice might be an important consideration. Additionally, in the CS base-case, these utility benefits are maintained for the whole duration of the time horizon. The plausibility of this assumption can be debated. Although the company's responses to clarification question B12 were informative and seemed to indicate that there might be a utility benefit even when patients are off treatment (clarification response Tables 26 and 27), the duration/extrapolation of the utility benefit is unclear. Therefore, the ERG base-case adopted the treatment dependent utilities (with the nivolumab + ipilimumab utility benefit) up to three years and treatment independent utilities afterwards (three years was selected given the limited data, only three patients were at risk, at this point).

4.2.9 Resources and costs

The cost categories included in the model were treatment acquisition and administration costs, monitoring and management of the disease, end-of-life costs, costs of managing AEs, and costs associated with subsequent therapy.

Unit prices were mostly based on the NHS reference prices[21], British National Formulary (BNF)[22], the Department of Health Drugs and pharmaceutical electronic market information tool (eMIT)[23], and Personal Social Services Research Unit (PSSRU)[24].

4.2.9.1 Resource use and costs data identified in the review

An SLR was conducted to identify costs and resource use in the first-line treatment and ongoing management of patients with MPM as described in Appendix J.[6] The literature search identified no relevant studies reporting the cost and resource use burden associated with MPM's first-line treatment. Due to the limited availability of cost and resource use data in the first-line setting, data irrespective of the line of treatment can also be considered. Three cost analyses were identified, conducted in: Italy, the UK and France[25-27]. These were, however, not used by the company.

4.2.9.2 Treatment costs (with PAS)

A flat nivolumab dosage of 360 mg every three weeks, aligning with the anticipated EMA licence, was used in the base-case analysis. The model includes the option to use the weight-based dose of 3 mg/kg every two weeks that was used in CheckMate-743. The weight-based dose is used in a scenario analysis. Ipilimumab is administered every six weeks at 1 mg/kg, which was in line with CheckMate-743. The CS includes a stopping rule of two years for nivolumab + ipilimumab, which was in line with CheckMate-743 Costs per dose are reported in Table 41 of the CS. There are simple PASs for nivolumab ([redacted]) and ipilimumab ([redacted]) approved by the Department of Health.

The comparator considered was pemetrexed (500 mg/m2 every three weeks for six treatment cycles) + cisplatin (75 mg/m2 every three weeks for four treatment cycles) or carboplatin (550 mg/m2 every three weeks for four treatment cycles). Costs per dose are reported in Table 41 of the CS. In the pemetrexed combination, 33% of patients were assumed to use cisplatin and 67% to use carboplatin, based on CheckMate-743.

The duration of treatment in the model was based on the duration of treatment recorded in CheckMate-743. Given the minimum follow-up was 22.1 months in CheckMate-743 and that the maximum duration of treatment for the nivolumab + ipilimumab arm is 24 months, complete duration of treatment data were available for the pemetrexed + cisplatin or carboplatin arm and data for 98.3% of patients are available for the nivolumab + ipilimumab arm. Thus, use of KM data for duration of treatment would be a viable option instead of parametric survival analyses (described in Appendix K). Both use of KM data and parametric survival analyses were explored, but the former was only used in a scenario and the latter not incorporated in the model. Instead, the company used the mean number of doses reported in CheckMate-743, which were: [redacted] (adjusted from [redacted] to reflect three-weekly doses instead of two-weekly doses) for nivolumab, and [redacted] for ipilimumab. For the PDC arm, the mean number of doses received for pemetrexed, cisplatin, and carboplatin was [redacted], [redacted], and [redacted], respectively. Missed or delayed doses were not corrected for in addition when using this approach, as these were already captured by the approach of using mean doses. Treatment costs were calculated using the mean number of doses and applied in the first model cycle. This approach was chosen over the use of KM data or parametric survival analysis as, according to the company, it *"most accurately captures treatment costs because it accounts for delayed or missed doses and provides values for each treatment within the regimens"*.[4]

Administration costs associated with all treatments are shown in Table 42 of the CS. Nivolumab is administered every three weeks and ipilimumab every six weeks. The cost for delivering complex parenteral chemotherapy is applied when both treatments are administered; the cost for delivering simple parenteral chemotherapy is applied when only nivolumab is administered. Total administration

costs are calculated using the mean number of doses from CheckMate-743 and are also applied in the first model cycle with the company's mean doses approach.

Monitoring costs reflect treatment-specific resource use such as laboratory tests and scans that are required to ensure patients are tolerating the treatment well (Table 43 of the CS). Monitoring costs were modelled for as long as patients stay on treatment, based on the KM data for duration of treatment from CheckMate-743. Monitoring costs were applied to the proportion of patients on treatment in each model cycle using separate KM curves for nivolumab + ipilimumab and for pemetrexed + cisplatin or carboplatin.

4.2.9.3 Subsequent treatment costs

According to the company, on failure with first-line treatment of nivolumab + ipilimumab or pemetrexed + cisplatin or carboplatin (i.e. on entry to the PD health state), proportions of 44.22% of patients on nivolumab + ipilimumab and 40.73% of patients in the PDC arm were modelled to go on to a subsequent treatment. The distribution of subsequent therapies received by initial treatment was based on CheckMate-743. Four subsequent treatment strategies were omitted because of low usage (< 1%). The median duration of 1.7 months assumed for all subsequent therapies (irrespective of therapy or treatment arm) was based on the publication by Waterhouse et al [28] (both distribution and duration of subsequent treatments are presented in Table 44 of the CS). Dosing details of all subsequent treatments are presented in Table 45 and administration costs in Table 46 of the CS.

4.2.9.4 Health state costs

Health state costs related to the PF health state (Table 47 of the CS), the progressed disease health state (Table 48 of the CS), and the end of life/terminal care cost health state (Table 49 of the CS). The weekly disease management costs for the PF state includes as outpatient visits chest radiography, CT scans (chest), CT scans (other), and electrocardiograms and the frequencies for these were obtained from TA531[29], amounting to a total cost per week of £42.60. For the PD state, weekly costs in addition include GP home visits, and therapist visits, with frequencies (every other week for both GP and therapist visits) also based on TA531[29], amounting to a total cost per week of £107.85. End of life/terminal care costs of £5,018.27 were applied as a one-off cost upon entering the death state. These costs included community nurse visits, GP home visits, Macmillan nurse, drugs and equipment, terminal care in hospital and terminal care in hospice, which frequencies obtained from TA531.

4.2.9.5 Event costs

Cost of treatment-related AEs (grade ≥ 3 AEs with an incidence rate of ≥ 2%) are shown in Table 4.4. Combined with the incidence of AEs in both treatment arms shown in Table 39 of the CS, this resulted in AE costs of £106.13 for nivolumab + ipilimumab, and £726.23 for the PDC arm, which are applied as a one-off in the first model cycle.

Table 4.4: Costs per weekly cycle

	Nivolumab + ipilimumab arm		PDC arm	
Treatment cost (£)	Nivolumab 360 mg Q3W, up to 2 years (company base-case)	[redacted]	Pemetrexed 500 mg/m^2 Q3W for 6 treatment cycles§	300.00
	Nivolumab 3 mg/kg Q2W, up to 2 years (company scenario)§	[redacted]	Cisplatin 75 mg/m^2 Q3W for 4 treatment cycles§	1.89

	Nivolumab + ipilimumab arm		PDC arm	
	Ipilimumab 1 mg/kg Q6W, up to 2 years§	[redacted]	Carboplatin 550 mg/m^2 Q3W for 4 treatment cycles§	7.91
Treatment administration cost (£)	Nivolumab + ipilimumab*	101.12	Pemetrexed + cisplatin or carboplatin**	88.09
Monitoring cost (£)	CS Table 43	50.33	CS Table 43	50.02
Subsequent treatment cost (£)	CS Tables 44 and 45	[redacted]	CS Tables 44 and 45	[redacted]
Health state cost (£)	PF state	42.60	PF state	42.60
	PD state	107.85	PD state	107.85
	End of life / terminal care	5,018.27	End of life / terminal care	5,018.27
Adverse event cost (£)	CS Table 50	106.13	CS Table 50	726.28

CS = company submission; PD = progressed disease; PDC = platinum-based doublet chemotherapy; PF = progression-free
§Mean patient characteristics used for calculations of weekly drug costs
*Based on company's mean doses approach, calculated over median TTD of approximately 24 weeks
**Based on company's mean doses approach, calculated over median TTD of approximately 15 weeks

ERG comment: The main concerns of the ERG relate to: a) proportional use of cisplatin versus carboplatin in the comparator arm; b) the approach to estimating treatment duration; c) the approach to including subsequent treatments in the analysis; d) costs related to AEs.

a) The ERG was concerned that the proportions to which the comparator included carboplatin versus cisplatin were unclear. The company clarified this and also performed a minor correction to the model: carboplatin was used by 66% of patients and cisplatin by 34% of patients in CheckMate-743. Regarding the generalisability of these proportions to UK clinical practice, the company clarified in response to clarification question B14[3] that "*Data from the EU cross-sectional study for the cohort of 248 UK patients suggest a similar proportion of carboplatin and cisplatin use. In the UK,* [redacted] *[10]. The proportions used in the model are more similar to estimates from The UK National Mesothelioma Audit 2020 in which pemetrexed with carboplatin was the most common regimen used (48%), followed by pemetrexed with cisplatin (20%), in patients who received chemotherapy."*[9] However, as noted in the ERG critique in Section 2 of this report, these numbers suggest some variation in the proportion of use of carboplatin versus cisplatin. Due to the low weekly cost of carboplatin and cisplatin, the magnitude of proportional use has a minor impact on cost effectiveness outcomes, as demonstrated by the company's scenario assuming an equal split between carboplatin and cisplatin use, which increased the ICER by £23 per QALY gained.
b) The use of mean number of doses to estimate treatment duration may introduce bias because this method does not take account of right-censoring[30]. According to Appendix N, the health economic experts agreed that *"that the CM-743 time-to-treatment discontinuation K-M curves were the best available evidence to inform treatment duration"*[6]. The ERG agrees with the experts, but also considers that parametric survival analysis on this evidence may potentially

be preferred, using dose intensity to reflect missed and delayed doses, and reflecting the stopping rule for nivolumab + ipilimumab by discontinuing all patients still on treatment at 24 months. The ERG requested this analysis from the company at clarification stage, but the company did not provide this analysis. In response to question B13 the company claimed that UK clinical experts considered the mean number of doses approach the most appropriate[3]. However, this was not supported by a reference (the company points to page 100 of their report, however this statement is not made there). The company's rationale for maintaining the mean number of doses approach is that the KM data, or parametric distributions do not account for missed or delayed doses. However, this can be addressed by using dose intensity as observed in CheckMate-743, as was proposed by the ERG in the clarification letter and is not considered by the ERG as a strong argument for not performing this analysis. In terms of the bias that might be introduced by using the mean number of doses approach, the company states that because the data are mature (minimum follow-up time is 22.1 months, the median is 29.7 months and the stopping rule for treatment is at 24 months), right-censoring would have minimal impact on the final estimates of doses received. The company also committed to providing updated duration of treatment data once these are available. Whilst the ERG considers the company's argument plausible, it would prefer to see the impact explored in scenario analysis using parametric survival models fitted to the time-to-treatment discontinuation KM data and using dose intensity. The mean dose intensity for each treatment was not made available by the company, but Table 6.1-1 of the CheckMate-743 CSR indicated dose intensity of around [redacted] for all treatments[7]. Because the company did not provide parametric survival analysis for TTD and mean dose intensity estimates were not available, the ERG used TTD KM estimates and 100% dose intensity for all treatments in a scenario. As far as generalisability is concerned, TTD KM estimates for the PDC arm were compared with available data for SACT from the CAS registry of patients with unresectable MPM in Figure 13 in the CQ response and showed that median treatment duration was [redacted] in the CAS registry compared with CheckMate-743[3]. No potential reasons for this discrepancy were provided.

c) There is remaining uncertainty about the modelling of subsequent treatments. These are only used by a proportion of patients in the PD state: 44.22% in the nivolumab + ipilimumab arm and 40.73% in the PDC arm as per CheckMate-743. The company confirmed in response to clarification question B15 that these proportions were aligned with clinical expectations according to clinical experts consulted during the development of the economic model, but the company also acknowledged that these proportions could be higher or lower in clinical practice. Subsequent treatments used are in line with CheckMate-743 but their use may not be in line with UK clinical practice. According to the company's response to question B15, there is no standard second-line therapy in MPM used in NHS clinical practice and this was also confirmed by UK clinical experts. For example, nivolumab + ipilimumab will only be used in the first-line setting, not in second-line. Re-treatment or re-challenge with nivolumab + ipilimumab is also not supported by any data currently, according to the company. Re-treatment with pemetrexed + cisplatin/carboplatin was shown by the company to be in line with UK clinical practice.

   Subsequent treatment duration of 1.7 months regardless of the subsequent treatment received and prior treatment allocation is considered by the ERG unlikely to be a good reflection of clinical practice. This was based on a poster by Waterhouse et al[28], in which the mix of second-line treatments differed from that in CheckMate-743 and the model. There was large variation in subsequent treatment duration (interquartile range of 1 – 11.90 in Waterhouse et al), and the differences may partly be driven by the type of subsequent treatment or prior treatment. First,

treatment duration may be longer with immunotherapies than with chemotherapies. Second, treatment duration may differ by initial treatment allocation, as post-progression survival appears to be longer in the modelled nivolumab + ipilimumab arm compared with the PDC arm. The company stated that the assumption of equal subsequent treatment duration would be conservative, given that immunotherapies would be expected to have a longer duration of treatment compared with chemotherapies and there was a higher proportion of immunotherapies in the PDC arm. However, the ERG considers that uncertainty remains about subsequent treatment duration and that it would ideally be able to implement differential subsequent treatment durations for each model arm (currently not enabled in the model), possibly based on data from CheckMate-743 once these are available, or expert opinion. The company provided scenario analyses to explore the impact of different assumptions surrounding subsequent treatments in Table 30 of the clarification response[3]. These scenarios, for example increasing subsequent treatment duration in both arms, resulted in only relatively small changes to the base-case ICER. In addition, the ERG performed a scenario setting subsequent treatment costs in the nivolumab + ipilimumab arm equal to the subsequent costs in the PDC arm and increasing treatment duration in both arms to three months. The impact of this was still minor.

d) As pointed out in Section 4.2.7, there may be selection bias in the included AEs which may result in an over-estimation of the incremental AE costs for the PDC arm versus the nivolumab arm. The company did not provide cost effectiveness analyses with all-causality (treatment-emergent) AEs instead of only treatment-related AEs, and the restriction on the incidence changed to 1% as requested by the ERG in the clarification letter. Equal AE rates for both treatment arms (using currently included AEs) would result in an increase in the ICER of slightly less than £1,000 per QALY gained and it should be noted that it only affects costs. Despite this not being a very impactful issue the ERG considers that the impact of AEs on the two treatment arms is currently likely mis-represented in the model.

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

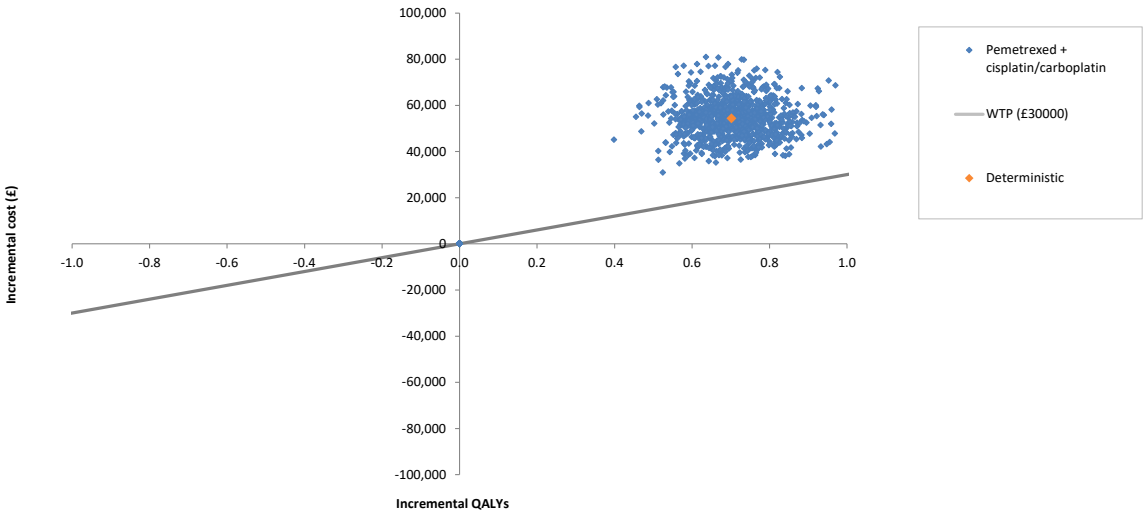
The CS base-case cost effectiveness results (probabilistic) indicated that nivolumab + ipilimumab is both more effective (incremental QALYs of 0.706) and more costly (additional costs of £55,423) than PDC amounting to an ICER of £77,127 per QALY gained (Table 5.1). Moreover, the 95% percentiles for the probabilistic incremental costs and QALYs were (£39,156 - £72,154) and (0.543 - 0.882) respectively (Figure 5.1). The probabilities of nivolumab + ipilimumab being cost effective, at thresholds of £20,000, £30,000 and £50,000 per QALY gained, compared to PDC are 0%, 0% and 1% respectively.

Overall, nivolumab + ipilimumab is modelled to affect QALYs in the company base case by:

- Increased mean PFS (undiscounted time in the PF health state: 18.5 vs 10.5 months) and mean OS (undiscounted survival: 34.4 vs 20.6 months) compared with PDC.
- Increased health state utility values for the PF (0.74 vs 0.73) and PD (0.65 vs 0.58) health states compared with PDC.
- The PFS, OS and health state utility benefits are maintained for the whole duration of the time horizon (i.e. no waning of these treatment benefits).

These effects combined result in the majority (55%) of the QALY gains (58% of the undiscounted LYs) being accumulated in the pre-progression state (CS Appendix L Tables 49 and 50). The majority (92%) of the additional costs are also accumulated due to increased drug acquisition costs followed by increased PD (4%) and PF (2%) health state costs (CS Appendix L Table 51).

Figure 5.1: CS base-case cost effectiveness plane



Source: Economic model

Table 5.1: CS base-case results

	Total costs (£)	Total LY	Total QALY	Incremental costs (£)	Incremental LY	Incremental QALYs	ICER (£)
Deterministic							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	54,397	0.916	0.702	77,502
Probabilistic (1,000 iterations)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	55,423	0.921	0.706	77,127
Source: CS Table 55[2] and economic model							

5.1.1 Company's subgroup analyses

No subgroup analyses were performed.

ERG comment: The main concerns of the ERG relate to: a) minor errors in the original CS base-case and b) extent and plausibility of the observed gains accumulated beyond the observed data period; c) not exploring cost effectiveness for subgroups listed in the scope.

a) In the clarification responses, the company started section B with the highlighting of two errors identified in the economic model used to calculate the CS base-case. These errors related to:
    a. One drug-related AE occurring in ≥ 2% of patients was omitted (see response to question B11, part a)
    b. The proportions of cisplatin and carboplatin use in combination with pemetrexed were incorrect (see response to question B14, part a)

   The company corrected these errors in their revised base-case. Compared with the original CS base-case, these corrections did not impact the estimated effectiveness (LY/QALYs), the company's revised base-case (deterministic) only slightly increased the estimated (incremental) costs as well as the ICER (increased from £77,502 to £77,531). Probabilistic results for the revised company base-case were not provided.
b) In clarification question B17, the ERG requested the company to provide a comparison of the observed survival as well as progression free survival for instance using restricted mean survival time (RMST) and the undiscounted LY as well as undiscounted progression free LY (PFLY) and elaborate on the plausibility of the differences. Unfortunately, the company stated that RMST was not reported in CheckMate-743. The RMST can be easily calculated from the KM data provided in the economic model. Therefore, the ERG calculated the RMST for LY as well as PFLY using different truncation points (Table 5.2). Based on these calculations it can be derived that the proportion of (PF)LY accumulated beyond the observed data is substantially larger for nivolumab + ipilimumab than for PDC. Moreover, considering the increments, approximately [redacted] of the LYs are gained beyond the observed data period for nivolumab + ipilimumab compared with PDC while this is even larger (approximately [redacted] for PFLY. While the company's response to clarification questions B5 and B8 give some indication about the plausibility of the long-term extrapolations, the findings presented in Table 5.2 indicate that the large majority of gains are accumulated beyond the observed data period and hence additional explanation of the mechanism by which the model generated these differences as well as a justification for why they are plausible based upon available evidence is warranted (as requested but not provided in the company's response to clarification question B17). This includes verifying the plausibility of the partitioned survival model extrapolations (see Section 4.2.2). Additionally, this highlights that the generated differences beyond the observed data period are a key issue while the estimated (PF)LY for the observed data period (i.e. whether to use KM data due to suboptimal fit in the observed data period) might have less priority.
c) The NICE scope mentioned subgroups based on histologic subtype (epithelioid, sarcomatoid, biphasic) and level of programmed death-ligand 1 (PD-L1) expression. These subgroups were not considered in the cost effectiveness sections of the CS despite these were prespecified subgroup analysis in CheckMate-743 (CS Table 7) and the relative effectiveness might differ between these subgroups (CS Figure 23 regarding OS hazard ratios per subgroup; tests for interactions were unfortunately not provided by the company despite requested in clarification question A13, while for PFS Section 3.2.5 suggests qualitative interactions regarding relative treatment effectiveness for these subgroups). Therefore, it might be informative to consider

subgroups specific cost effectiveness analyses. See also section 3.2.5 and 4.2.3 for further details.

Table 5.2: CS base-case comparing observed and estimated undiscounted (PF)LYs

	Observed	Modelled	
	Restricted mean survival time (RMST)[a]	**Estimated (lifetime time horizon)**	**Proportion beyond observed data[a]**
OS - RMST period / truncation point: 30 months (selected based on patients at risk Table)			
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]
Increment[a]	[redacted]	[redacted]	[redacted]
OS - RMST period / truncation point: 22 months (break point for piecewise approach[b])			
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]
Increment[a]	[redacted]	[redacted]	[redacted]
OS - RMST period / truncation point: [redacted] months (latest KM data point: [redacted] months)			
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]
Increment[a]	[redacted]	[redacted]	[redacted]
PFS - RMST period / truncation point: 30 months (selected consistently with OS)			
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]
Increment[a]	[redacted]	[redacted]	[redacted]
PFS - RMST period / truncation point: 22 months (break point for piecewise approach[b])			
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]
Increment[a]	[redacted]	[redacted]	[redacted]
PFS - RMST period / truncation point: [redacted] months (latest KM data point: [redacted] months)			
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]
Increment[a]	[redacted]	[redacted]	[redacted]

Source: economic model

[a]Calculated by the ERG (based on information on the "KM Data Store" worksheet), the estimated numbers might be subject to rounding errors

[b]The company justified the 22 months break point by stating that it was the approximate minimum patient follow-up at the database lock of CheckMate-743, and most censoring in the OS data in both treatment arms occurred after this point

5.2 Company's sensitivity and scenario analyses

The company performed and presented the results of probabilistic sensitivity analyses (PSA), deterministic sensitivity analyses (DSA) as well as scenario analyses.

The parameters that have the greatest effect on the ICER (based on the company's sensitivity analyses illustrated in CS Figure 43) are:

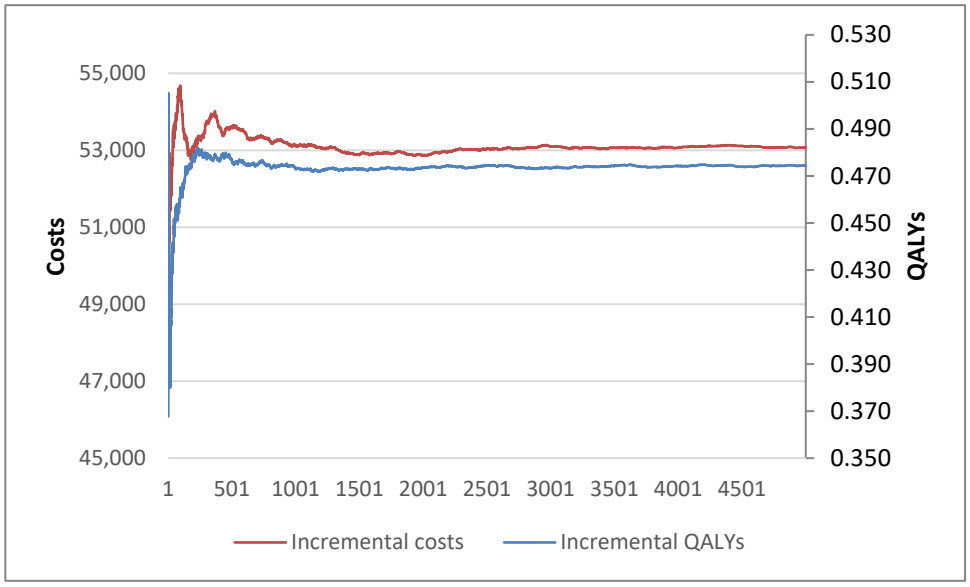
- PF and PD health state utility values for nivolumab + ipilimumab
- PF and PD health state utility values for PDC
- Discount rates for outcomes and costs
- Nivolumab and ipilimumab dosing
- Pemetrexed dosing
- Cohort starting age

Consistently, modelling assumptions that relate to these parameters likely have the greatest effect on the ICER. This is illustrated by the following CS scenarios that have a substantial impact on the ICER:

- CS scenarios 1-3: estimating OS using an alternative approach
- CS scenario 5: using treatment independent utility values
- CS scenario 4: estimating PFS using an alternative approach
- CS scenario 6: using nivolumab weight-based dosing

ERG comment: The main concern of the ERG relates to the number of PSA iterations being insufficient. The convergence plots in Figure 5.2 show that incremental costs and QALYs were not yet completely stable at 1,000 iterations in the ERG base-case. This might particularly hamper the comparison/interpretation of scenarios with very similar (incremental) results. It should also be noted that in scenarios using TTD estimates or KM estimates (in the piecewise approach), these should be included in the PSA, but it appears as if they are not (given these are not included in the parameter sheet).

Figure 5.2: PSA convergence plot for ERG base-case



5.3 Model validation and face validity check

5.3.1 Face validity assessment

During the development of the economic model, external clinical and health economic experts were consulted to ensure an appropriate approach was taken and that the model had clinical validity. Three advisory boards including UK clinical and HTA experts were held for this purpose.

5.3.2 Technical verification

The company did not provide detail on the technical verification of their model.

5.3.3 Comparisons with other technology appraisals

No comparisons with other technology appraisals were provided.

5.3.4 Comparison with external data

The company undertook comparisons between their modelled OS extrapolations, CheckMate-743 OS data (used to develop this model) and OS data from the MAPS dataset (not used to develop the model).

ERG comment: The main concerns of the ERG relate to: a) internal validity efforts, and b) lack of cross-validation.

a) The internal validity or technical verification was not detailed by the company, but in response to clarification question B18, the company clarified that *"the model was quality controlled and all calculations and data were checked by an independent researcher"*[3]. The ERG also requested that a checklist be filled in, such as the TECH-VER checklist[31], but the company did not provide this. Although the internal validity was not fully demonstrated, the ERG was able to reproduce life year gains, QALY gains and costs of the company's base-case.
b) No cross-validation with other technology appraisals was provided. In response to clarification question B19, the company stated that TA135 was not suitable for cross-validation but did not explore potential cross-validation with other appraisals. The ERG acknowledges that cross-validation with other appraisals would be limited since there are no published appraisals in MPM to date (apart from TA135).

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Table 6.1 summarises the key issues related to the cost effectiveness categorised according to the sources of uncertainty as defined by Grimm et al. 2020[32]:

- Transparency (e.g. lack of clarity in presentation, description, or justification)
- Methods (e.g. violation of best research practices, existing guidelines, or the reference case)
- Imprecision (e.g. particularly wide confidence intervals, small sample sizes, or immaturity of data)
- Bias and indirectness (e.g. there is a mismatch between the decision problem and evidence used to inform it in terms of population, intervention/comparator and/or outcomes considered)
- Unavailability (e.g. lack of data or insight)

Identifying the source of uncertainty can help determine what course of action can be taken (i.e. whether additional clarifications, evidence and/ or analyses might help to resolve the key issue). Moreover, Table 6.1 lists suggested alternative approaches, expected effects on the cost effectiveness, whether it is reflected in the ERG base-case as well as additional evidence or analyses that might help to resolve the key issues.

Based on all considerations in the preceding Sections of this ERG report, the ERG defined a new base-case. This base-case included multiple adjustments to the original base-case presented in the previous sections. These adjustments made by the ERG form the ERG base-case and were subdivided into three categories (derived from Kaltenthaler 2016)[33]:

- Fixing errors (FE) (correcting the model where the company's submitted model was unequivocally wrong)
- Fixing violations (FV) (correcting the model where the ERG considered that the NICE reference case, scope or best practice had not been adhered to)
- Matters of judgement (MJ) (amending the model where the ERG considers that reasonable alternative assumptions are preferred)

6.1.1 ERG base-case

Adjustments made by the ERG, to derive the ERG base-case (using the CS base-case as starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the ERG base-case. The ERG did not identify any errors or violations and all adjustments pertained to matters of judgement.

6.1.1.1 Matters of judgement

1. The use of piecewise KM estimates for OS extrapolation (Section 4.2.6)
   ERG adjustment: do not use the piecewise approach
2. The use of log-logistic and exponential distributions for OS in nivolumab + ipilimumab and PDC arms respectively (Section 4.2.6)
   ERG adjustment: use log-logistic distributions for OS in both treatment arms
3. Assumption that treatment effect will persist through lifetime (Section 4.2.6)
   ERG adjustment: implement treatment waning from five years onwards by adjusting the nivolumab + ipilimumab OS and PFS hazards to align with those of the PDC arm after this time point
4. Assumption that treatment effect on utilities will persist throughout lifetime (Section 4.2.8)

ERG adjustment: change to treatment independent utilities at three years

6.1.2 ERG exploratory scenario analyses

The ERG performed the following exploratory scenario analyses to explore the impact of alternative assumptions conditional on the ERG base-case.

6.1.2.1 Exploratory scenario analyses

1. Uncertainty about PFS (Section 4.2.6)
   a) ERG adjustment: Use log-logistic distributions for both arms
   b) ERG adjustment: Use generalised gamma distributions for both arms
2. Likely selection bias in AEs (Section 4.2.7)
   ERG adjustment: set AE rates equal in both treatment arms
3. Potentially biased approach to time-on-treatment estimation (Section 4.2.9)
   ERG adjustment: use TTD KM estimates with 100% dose intensity
4. Likely bias in subsequent treatment duration estimate (Section 4.2.9)
   ERG adjustment: set equal nivolumab + ipilimumab arm subsequent treatment costs to PDC arm subsequent treatment costs and increase subsequent treatment duration for both to 3 months

6.1.3 ERG subgroup analyses

No subgroup analyses were performed by the ERG.

Table 6.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 5.1)

Key issue pertaining to cost effectiveness (See Section 1)	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case
6) No state transition model provided to validate the partitioned survival analysis model	4.2.2	Methods	State transition model	+/-	No
7) No subgroup analysis provided	4.2.3	Methods	Subgroup analysis	+/-	No
8) Two-year stopping rule may not be observed in CheckMate-743 (although included in study protocol)	4.2.4	Indirectness	Correct for this if necessary (proportion of patients not adhering to stopping rule in cost estimation)	Could be +, if applicable	No
9) Treatment effectiveness and extrapolation of OS and PFS for nivolumab + ipilimumab highly uncertain due to immature data, limited long-term validation	4.2.6	Methods, unavailability	Alternative approaches for estimating PFS and OS as well as assumptions related to treatment waning	+/-	Partly, data immaturity cannot be currently resolved
10) Duration of treatment effect on HRQoL uncertain	4.2.8	Methods, unavailability	Treatment independent utilities from certain time point	+	Partly, explore appropriate time point
11) Estimation of time on treatment potentially biased	4.2.9	Methods	Use TTD KM estimates and parametric survival analysis and dose intensity	+	Partly, dose intensity adjustment needed
12) Duration of subsequent treatments potentially biased, remaining uncertainty about subsequent treatment use	4.2.9	Imprecision, indirectness	Longer subsequent treatment duration in both arms and set costs equal	+	Partly, differential implementation of subsequent treatment duration per arm needed
13) Selection bias in AE rates and therefore likely bias in AE associated costs	4.2.9	Indirectness	Set AE rates equal, or preferable incorporate all cause AEs	+	Partly, enable all cause AEs

Key issue pertaining to cost effectiveness (See Section 1)	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base-case
14) Large proportion of (PF)LY accumulated beyond the observed data	5.1	Unavailability	Using CheckMate-743 data with additional follow-up data.	+/-	No

[a] Likely conservative assumptions (of the intervention versus all comparators) are indicated by '-'; while '+/-' indicates that the bias introduced by the issue is unclear to the ERG and '+' indicates that the ERG believes this issue likely induces bias in favour of the intervention versus at least one comparator

AE = adverse events; ERG = Evidence Review Group; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; KM = Kaplan-Meier; LY = life years; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation

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

In Section 6.1 the ERG base-case was presented, which was based on various changes compared to the company base-case. Table 6.2 shows how individual changes impact the results plus the combined effect of all changes simultaneously. The exploratory scenario analyses are presented in Table 6.3. These are all conditional on the ERG base-case. The analyses numbers in Tables 6.2 and 6.3 correspond to the numbers reported in Section 6.1. The submitted model file contains technical details on the analyses performed by the ERG (e.g. the "ERG" sheet provides an overview of the cells that were altered for each adjustment).

Table 6.2: ERG base-case (deterministic unless indicated)

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Company's corrected base-case							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	54,417	0.916	0.702	77,531
Matter of judgement 1: do not use piecewise approach (key issue 9)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	54,579	0.943	0.719	75,867
Matter of judgement 2: use log-logistic distributions for OS in both treatment arms (using piecewise) (key issue 9)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,269	0.700	0.576	92,413
Matter of judgement 3: implement treatment waning from 5 years onwards (key issue 9)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	52,988	0.540	0.443	119,543
Matter of judgement 4: change to treatment-independent utilities from 3 years onwards (key issue 10)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	54,417	0.916	0.678	80,206
ERG base-case (Changes 1-4)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,327	0.617	0.476	112,005

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
ERG base-case probabilistic (5,000 runs)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,076	0.612	0.474	111,898

Table 6.3: Deterministic scenario analyses (conditional on ERG base-case)

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
ERG base-case							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,327	0.617	0.476	112,005
Scenario 1a: PFS log-logistic distribution for both arms (key issue 9)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,861	0.617	0.460	117,179
Scenario 1b: PFS generalised gamma distribution for both arms (key issue 9)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,602	0.617	0.467	114,786
Scenario 2: set AE rates equal in both treatment arms (key issue 13)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,927	0.617	0.476	113,267
Scenario 3: use TTD KM estimates with 100% dose intensity (key issue 11)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	59,726	0.617	0.476	125,446
Scenario 4: set equal nivol + ipi arm subsequent treatment costs to PDC & increase treatment duration to 3 months (key issue 12)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,812	0.617	0.476	113,024

Table 6.4: Probabilistic scenario analyses (conditional on ERG base-case, 1,000 iterations unless stated otherwise)

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Company base-case							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	54,396	0.926	0.710	76,633
ERG base-case (5,000 iterations)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,076	0.612	0.474	111,898
Scenario 1a: PFS log-logistic distribution for both arms (key issue 9)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,981	0.618	0.460	117,281
Scenario 1b: PFS generalised gamma distribution for both arms (key issue 9)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,147	0.611	0.464	114,466

Technologies	Total costs (£)	Total LYs	Total QALYs	Incremental costs (£)	Incremental LYs	Incremental QALYs	ICER (£/QALY)
Scenario 2: set AE rates equal in both treatment arms (key issue 13)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,505	0.611	0.475	112,539
Scenario 3: use TTD KM estimates with 100% dose intensity (key issue 11)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	59,693	0.617	0.477	125,139
Scenario 4: set equal nivol + ipi arm subsequent treatment costs to PDC & increase treatment duration to 3 months (key issue 12)							
Nivolumab + ipilimumab	[redacted]	[redacted]	[redacted]				
Pemetrexed + cisplatin or carboplatin	[redacted]	[redacted]	[redacted]	53,981	0.614	0.475	113,612

6.3 ERG's preferred assumptions

The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was £111,898 per QALY gained. The probabilistic ERG base-case analyses indicated cost effectiveness probabilities of 0%, 0% and 0% at willingness to pay thresholds of £20,000, £30,000 and £50,000 per QALY gained. The most influential adjustments were implementing treatment waning from five years onwards and using the log-logistic distribution for estimating OS in the PDC arm. The ICER increased most in the scenario analysis using TTD estimates with 100% dose intensity instead of the number of mean doses approach. Since dose intensity was likely lower in the trial, this may be regarded as a the upper bound of the ICER using alternative scenarios on time on treatment.

6.4 Conclusions of the cost effectiveness section

The company's cost effectiveness model was well built and complied with the NICE reference case. The main critique points are modelling choices and assumptions. The overarching challenge was the immaturity of the data from CheckMate-743, which results in the ICER being very uncertain. The company's approach of using a PSM was questioned, especially given that a large proportion of life years and QALYs gains could be attributed to the time period beyond available trial data. The most influential issue was the extrapolation of OS. The ERG considered the company's piecewise approach not to offer any improvements over conventional survival analysis and replaced it by conventional survival analysis in the ERG base-case. Given current evidence, the ERG also questioned the company's choice of distributions (log-logistic and exponential) and preferred the log-logistic distribution in both arms. The ERG furthermore questioned the company's implicit assumption of a lifelong treatment effect (OS and PFS) and relaxed this by implementing treatment waning from five years onwards in the ERG base-case, which had a significant impact on the ICER. The ERG also explored the impact of different PFS distributions in scenarios, which was smaller compared with OS modifications. AEs may be mis-represented in the cost effectiveness analysis model because of the company's applied selection criteria, which could result in underestimation of AE-related costs in the model. The impact of this on cost effectiveness results is likely small. In terms of HRQoL, the main uncertainty related to whether the treatment effect on HRQoL was lifelong and the ERG relaxed this assumption in the ERG base-case. The ERG questioned the method of using number of mean doses for estimating treatment duration, which may be biased due to right-censoring. Even though the company highlighted that the data were mature and right-censoring therefore unlikely to be a significant problem, the ERG considered that since treatment duration was a key driver of the model, the impact of using parametric survival analysis using TTD data should be explored. Subsequent treatments and their treatment duration were also subject to uncertainty and may warrant further investigation, even though the impact on cost effectiveness may be relatively small. No subgroup analyses were provided, but the ERG considered that cost effectiveness may vary by subgroup.

The company's corrected deterministic ICER was £77,531 per QALY gained and no corrected probabilistic ICER was presented. The ERG's replication of the company base-case probabilistic analysis resulted in an ICER of £76,633 per QALY gained. The estimated ERG base-case ICER (probabilistic), based on the ERG preferred assumptions highlighted in Section 6.1, was £111,898 per QALY gained. The probabilistic ERG base-case analyses indicated cost effectiveness probabilities of 0% and 0% at willingness to pay thresholds of £30,000 and £50,000 per QALY gained. The most influential adjustments were implementing treatment waning from five years onwards and using the log-logistic distribution for estimating OS in the PDC arm. The ICER increased most in the scenario analysis using TTD estimated with 100% dose intensity instead of the mean doses approach. Since dose

intensity was likely lower in the trial, this may be regarded as a the upper bound of the ICER using alternative scenarios on time on treatment.

There is large remaining uncertainty about the effectiveness and relative effectiveness of nivolumab + ipilimumab versus PDC, which can be at least partly resolved with future analyses of CheckMate-743 data. In view of the immaturity of the CheckMate-743 study it was not possible for the ERG to quantify all uncertainty now. Further data cuts could potentially result in additional survival gains for the nivolumab + ipilimumab arm. However, it is currently questionable whether nivolumab + ipilimumab can be cost effective compared to PDC.

7. END OF LIFE

The company claim that the end of life criteria are fulfilled:[2]

- Most patients die less than two years after diagnosis, with a median survival of 13 months in unresectable patients with MPM treated with SACT.[34]
- Interim results from CheckMate-743 show a median 4-month survival benefit with nivolumab + ipilimumab versus PDC, with a median OS follow-up of 29.7 months.

ERG comment: As reported in Section 3.2.4.1, the ERG notes also that PDC had a median OS of 14.1 months (95% CI: 12.4 to 16.2 months), which would be consistent with a survival that was lower than two years. In additions, the company's base-case model supports this as it results in an undiscounted mean OS of 1.7 years (Table 5.2). However, the ERG base-case indicates possible undiscounted mean OS of exactly two years. The ERG can also verify the increase in survival of four months given that those treated with nivolumab + ipilimumab were noted to have a median OS of 18.1 months (95% CI: 16.8 to 21.4 months). The company's and ERG's base-case analyses support the survival gain of > 3 months.

8. REFERENCES

[1] Baas P, Daumont MJ, Lacoin L, Penrod J, Carroll R, Tanna N. Treatment patterns and outcomes in malignant pleural mesothelioma in England: a nationwide CAS registry analysis from the I-O Optimise initiative. Poster presented at European Society for Medical Oncology (ESMO) Virtual Congress; 19-21 September 2020.

[2] Bristol-Myers Squibb. *Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma [ID1609]: submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*. Uxbridge: Bristol-Myers Squibb, 2020. 129p.

[3] National Institute for Health and Care Excellence. *Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma [ID1609]: clarification letter*. London: NICE, 2021. 21p.

[4] Bristol-Myers Squibb. *Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma [ID1609]: response to request for clarification from the ERG*. Uxbridge: Bristol-Myers Squibb, 2021. 85p.

[5] Tsao AS, Baas P, Nowak AK, Zalcman G, Fujimoto N, Peters S, et al. Evaluation of flat dosing for nivolumab + ipilimumab in first-line unresectable malignant pleural mesothelioma: CheckMate 743. Presented at ESMO Immuno-Oncology Virtual Congress 2020; 9-12 December 2020.

[6] Bristol-Myers Squibb. *Nivolumab with ipilimumab for untreated unresectable malignant pleural mesothelioma [ID1609]: document C [Appendices]. Submission to National Institute for Health and Care Excellence. Single technology appraisal (STA)*. Uxbridge: Bristol-Myers Squibb, 2020. 129p.

[7] Bristol-Myers Squibb. *Nivolumab + ipilimumab: final clinical study report for study CA209743. A phase III, randomized, open label trial of nivolumab in combination with ipilimumab versus pemetrexed with cisplatin or carboplatin as first line therapy in unresectable pleural mesothelioma*. Lawrenceville (NJ): Bristol-Myers Squibb, 3 August 2020. 152p.

[8] Woolhouse I, Bishop L, Darlison L, De Fonseka D, Edey A, Edwards J, et al. British Thoracic Society Guideline for the investigation and management of malignant pleural mesothelioma. *Thorax* 2018;73(Suppl 1):i1-i30.

[9] Royal College of Physicians. *National mesothelioma audit report 2020 (for the audit period 2016-18) [Internet]*. London: RCP, May, 2020 [accessed 2.2.21]. 27p. Available from: https://www.rcplondon.ac.uk/file/21631/download

[10] Moore AB, Hart K, McDonald L, McKenna M, Daumont MJ. *Malignant pleural mesothelioma: treatment patterns and humanist burden of disease in Europe [manuscript in prepration]*, 2021

[11] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.0 (updated July 2019) [Internet]*: Cochrane, 2019 [accessed 6.7.20] Available from: https://training.cochrane.org/handbook

[12] National Institute for Health and Care Excellence. *Single Technology Appraisal: specification for manufacturer/sponsor submission of evidence [Internet]*. London: NICE, 2012 [accessed ??]. 76p. Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/Specification-for-manufacturer-sponsor-submission-of-evidence-June-2012.doc

[13] Baas P. First-line nivolumab + ipilimumab vs chemotherapy in unresectable malignant pleural mesothelioma: CheckMate 743. Presented at International Association for the Study of Lung Cancer (IASLC) World Conference on Lung Cancer (WCLC); 8 August 2020.

[14] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 23.3.11] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[15] Woods B, Sideris E, Palmer S, Latimer N, M. S. *NICE DSU technical support document 19. Partitioned survival analysis for decision modelling in health care: a critical review [Internet]*. Sheffield: University of Sheffield, 2017 [accessed 2.2.21]. 72p. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf

[16] Burnham K, Anderson D. *Model selection and multimodel inference: a practical information-theoretic approach*. 2nd ed. New York: Springer, 2004.

[17] Raftery AE. Bayesian model selection in social research. *Sociol Methodol* 1995;25:111-63.

[18] Zalcman G, Mazieres J, Margery J, Greillier L, Audigier-Valette C, Moro-Sibilot D, et al. Bevacizumab for newly diagnosed pleural mesothelioma in the Mesothelioma Avastin Cisplatin Pemetrexed Study (MAPS): a randomised, controlled, open-label, phase 3 trial. *Lancet* 2016;387(10026):1405-1414.

[19] Antonia SJ, Borghaei H, Ramalingam SS, Horn L, De Castro Carpeno J, Pluzanski A, et al. Four-year survival with nivolumab in patients with previously treated advanced non-small-cell lung cancer: a pooled analysis. *Lancet Oncol* 2019;20(10):1395-408.

[20] National Institute for Health and Care Excellence. Nivolumab for treated squamous cell carcinoma for the head and nedk after platinum-based chemotherapy (CDF Review TA490) [ID1585] [Internet]. London: NICE, 2021 [accessed 9.3.21]. Available from: https://www.nice.org.uk/guidance/indevelopment/gid-ta10538/documents

[21] NHS Improvement. 2018/2019 National cost collection data [Internet]. Leeds: NHS Digital, 2020 [accessed 2.2.21]. Available from: https://improvement.nhs.uk/resources/national-cost-collection/

[22] National Institute for Health and Care Excellence. BNF [Internet]. London: NICE, 2021 [accessed 9.2.21]. Available from: https://bnf.nice.org.uk/

[23] Department of Health and Social Care. *Drugs and pharmaceutical electronic market information tool (eMIT) [Internet: Excel spreadsheet]*. Leeds: Department of Health and Social Care, 4 March, 2020 [accessed 2.2.21] Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[24] Curtis LAB, A. *Unit Costs of Health and Social Care 2019 [Internet]*. Canterbury: Personal Social Services Research Unit, University of Kent, 2019 [accessed 2.2.21]. 184p. Available from: https://kar.kent.ac.uk/79286/11/UCFinalFeb20.pdf

[25] Ippoliti R, Falavigna G, Grosso F, Maconi A, Randi L, Numico G. The economic impact of clinical research in an Italian public hospital: the malignant pleural mesothelioma case study. *International journal of health policy and management* 2018;7(8):728-737.

[26] Verma V, Ahern CA, Berlind CG, Lindsay WD, Grover S, Culligan MJ, et al. Facility volume and postoperative outcomes for malignant pleural mesothelioma: a National Cancer Data Base analysis. *Lung Cancer* 2018;120:7-13.

[27] Chouaid C, Assié JB, Andujar P, Blein C, Tournier C, Vainchtock A, et al. Determinants of malignant pleural mesothelioma survival and burden of disease in France: a national cohort analysis. *Cancer Med* 2018;7(4):1102-1109.

[28] Waterhouse D, Nwokeji E, Boyd M, Penrod JR, Espirito J, Robert NJ, et al. Treatment patterns and outcomes of advanced malignant pleural mesothelioma (MPM) patients in a community practice setting. Poster presented at the International Association for the Study of Lung Cancer (IASLC) World Conference on Lung Cancer; 7-10 September 2019; Barcelona, Spain

[29] National Institute for Health and Care Excellence. *Pembrolizumab for untreated PD-L1 positive metastatic non-small-cell lung cancer [TA531]. Committee papers [Internet]*. London: NICE, 2018 [accessed 2.2.21]. 305p. Available from: https://www.nice.org.uk/guidance/ta531/documents/committee-papers

[30] Wijeysundera HC, Wang X, Tomlinson G, Ko DT, Krahn MD. Techniques for estimating health care costs with censored data: an overview for the health services researcher. *ClinicoEconomics and outcomes research: CEOR* 2012;4:145.

[31] Büyükkaramikli NC, Rutten-van Mölken M, Severens JL, Al M. TECH-VER: A verification checklist to reduce errors in models and improve their credibility. *Pharmacoeconomics* 2019;37(11):1391-408.

[32] Grimm SE, Pouwels X, Ramaekers BLT, Wijnen B, Knies S, Grutters J, et al. Development and Validation of the TRansparent Uncertainty ASsessmenT (TRUST) Tool for assessing uncertainties in health economic decision models. *Pharmacoeconomics* 2020;38(2):205-16.

[33] Kaltenthaler E, Carroll C, Hill-McManus D, Scope A, Holmes M, Rice S, et al. The use of exploratory analyses within the National Institute for Health and Care Excellence single technology appraisal process: an evaluation and qualitative analysis. *Health Technol Assess* 2016;20(26):1-48.

[34] Health and Safety Executive. *Mesothelioma statistics for Great Britain 2020 [Internet]*. London: HSE, 2020 [accessed 2.2.21]. 12p. Available from: https://www.hse.gov.uk/statistics/causdis/mesothelioma/mesothelioma.pdf