BMJ TAG

Pembrolizumab for untreated metastatic colorectal cancer with high microsatellite instability or mismatch repair deficiency

STA Report

Source of funding

This report was commissioned by the NIHR Evidence Synthesis Programme as project number 129039T.

Title:	Pembrolizumab for untreated metastatic colorectal cancer with high microsatellite instability or mismatch repair deficiency
Produced by:	BMJ Technology Assessment Group (BMJ TAG)
Authors:	Steven J Edwards, Director of Health Technology Assessment, BMJ-TAG, London Samantha Barton, Principal Health Technology Assessment Analyst, BMJ-TAG, London Tracey Jhita, Health Economics Manager, BMJ-TAG, London Victoria Wakefield, Principal Health Technology Assessment Analyst, BMJ-TAG, London Sarah Roberts, Senior Health Economist, BMJ-TAG, London
Correspondence to:	Steve Edwards, BMJ TAG, BMJ, BMA House, Tavistock Square, London, WC1H 9JR.
Date completed:	03/12/2020
Source of funding:	This report was commissioned by the NIHR Evidence Synthesis Programme as project number 129039T.
Declared competing interests of the authors	Description of any pecuniary relationship with sponsors, both personal and of the TAR Centre. No competing interests were declared which affect the impartiality of this report. BMJ Technology Assessment Group (BMJ-TAG) and the editorial team of The BMJ work independently to one another. The views and opinions expressed in this report are those of the BMJ-TAG.
Acknowledgments:	The ERG would like to thank Professor Tim Maughan (Professor of Clinical Oncology, University of Oxford) and Dr Fiona Lofts (Consultant Medical Oncologist, St George's University Hospital) for providing clinical advice throughout the project, and for providing feedback on the clinical sections of the report.
Rider on responsibility for report:	The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.
Report reference:	Edwards SJ, Barton S, Jhita T, Wakefield V and Roberts S. Pembrolizumab for untreated metastatic colorectal cancer with high microsatellite instability or mismatch repair deficiency: A Single Technology Appraisal. BMJ Technology Assessment Group, 2020.

Contribution of authors:

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

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

Table of Contents

Table of Contents 4

List of Tables 8

List of Figures 13

List of Abbreviations 15

1 Executive summary 18

1.1 Overview of the ERG's key issues 18

1.2 Overview of key model outcomes 18

1.3 The decision problem: summary of the ERG's key issues 19

1.4 The clinical and cost effectiveness evidence: summary of the ERG's key issues 20

1.5 Summary of ERG's preferred assumptions and resulting ICER 24

2 Introduction and background 28

2.1 Introduction 28

2.2 Background 28

2.2.1 Colorectal cancer 28

2.2.2 Microsatellite instability, mismatch repair, and other genetic anomalies involved in development of colorectal cancer 29

2.3 Critique of the company's definition of the decision problem 31

2.3.1 Population 37

2.3.2 Intervention 38

2.3.3 Comparators 40

3 Clinical effectiveness 44

3.1 Critique of the methods review 44
3.2 Critique of trials of the technology of interest, the company's analysis and interpretation 47
3.2.1 Aspects of trial design 52
3.3 Clinical effectiveness results 54
3.3.1 Progression-free survival 57
3.3.2 Overall survival 61
3.3.3 Response outcomes 63
3.3.4 Health-related quality of life 65
3.3.5 Adverse effects 66
3.4 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison 69
3.4.1 CAPOX 70
3.4.2 Panitumumab in combination with FOLFOX or FOLFIRI 70
3.4.3 Cetuximab in combination with FOLFIRI or FOLFOX 72
3.5 Critique of the indirect comparison and/or multiple treatment comparison 73
3.5.1 Methods of indirect comparison 73
3.5.2 Results of indirect comparison 77
3.6 Conclusions of the clinical effectiveness section 81
4 Cost effectiveness 84
4.1 ERG comment on the company's review of cost effectiveness evidence 85
4.2 Summary and critique of company's submitted economic evaluation by the ERG 86
4.2.1 NICE reference case checklist 86

4.2.2 Population 87
4.2.3 Interventions and comparators 90
4.2.4 Modelling approach and model structure 95
4.2.5 Treatment effectiveness 98
4.2.6 Adverse events 107
4.2.7 Health-related quality of life 109
4.2.8 Resource use and costs 113
5 Cost effectiveness results 123
5.1.1 Company's cost effectiveness results 123
5.1.2 Company's sensitivity analyses 124
5.1.3 Model validation and face validity check 131
6 Additional economic analysis undertaken by the ERG 132
6.1 Model corrections 132
6.2 Exploratory and sensitivity analyses undertaken by the ERG 132
6.3 ERG scenario analysis 133
6.4 ERG preferred assumptions 137
6.5 Conclusions of the cost effectiveness sections 140
7 End of Life 145
8 References 146
9 Appendices 149
9.1 Baseline characteristics 149
9.2 Pre-specified subgroup analyses 152

9.3 Fractional polynomial NMA: ERG validation for comparators versus standard of care 154

List of Tables

Table 1. Summary of key issues 18

Table 2. Issue 1: Direct head-to-head evidence is not available for comparators listed in final scope issued by NICE 19

Table 3. Issue 2: Subgroup analyses by RAS mutation status 20

Table 4. Issue 3: Treatment regimen and resource use for pembrolizumab 22

Table 5. Issue 4: Duration of treatment with pembrolizumab 23

Table 6. Issue 5: Treatment costs for standard of care 23

Table 7. Issue 6: Time on treatment for non-KEYNOTE-177 comparators 24

Table 8. ERG's preferred assumptions - pembrolizumab vs SoC (ITT and RAS wild-type proxy analysis) 25

Table 9. ERG's preferred assumptions - pembrolizumab vs CAPOX (ITT and RAS wild-type proxy analysis) 25

Table 10. ERG's preferred assumptions - pembrolizumab vs panitumumab combination treatment (RAS wild-type proxy analysis) 26

Table 11. ERG's preferred assumptions - pembrolizumab vs cetuximab combination treatment (RAS wild-type proxy analysis) 26

Table 12. Summary of decision problem (adapted from CS, Table 1) 32

Table 13. Proportion of people receiving treatment options available in the standard of care group (reproduced from the CS, Table 16, page 43) 42

Table 14. Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem 45

Table 15. Summary of ERG's critique of the design and conduct of KEYNOTE-177,[2] the trial evaluating the technology of interest to the decision problem 48

Table 16. Summary of number of cycles of pembrolizumab received by those allocated to pembrolizumab group in KEYNOTE-177 (reproduced from company's response to clarification, question B10) 53

Table 17. Summary of duration of follow-up in KEYNOTE-177 (adapted from CS, Table 17, page 44) 53

Table 18. Summary of findings from studies evaluating CAPOX, FOLFIRI and FOLFOX in the first-line treatment of metastatic colorectal cancer 56

Table 19. Summary of event and censoring description for PFS by central imaging vendor per RECIST 1.1 (ITT population) (reproduced from CS, Tables 25–27, pages 60–62) 58

Table 20. Summary of event for OS (ITT population) (reproduced from CS, Tables 20 and 21, pages 46 and 47) 62

Table 21. Summary of best overall response by central imaging vendor per RECIST 1.1 (ITT population; database cut-off date of 19 Feb 2020) (reproduced from CS, Table 29, page 64)..... 64

Table 22. Summary of results from assessment of health-related quality of life from KEYNOTE-177 (reproduced from CS, Tables 32 and 33, page 69, and from Table 110 from Appendix L, page 537).. 65

Table 23. Summary of adverse events (ASaT population) 67

Table 24. Models assessed by company in FP NMA 74

Table 25. Time-varying hazard ratios for PFS at selected follow-up times for competing interventions versus pembrolizumab (second order FP model [p1 = 0, p2 = 0]) (adapted from CS, Table 42, page 80) 78

Table 26. Model fit estimates for fixed-effects network meta-analysis with parametric survival models for PFS (reproduced from CS, Table 41, page 79) 79

Table 27. Time-varying hazard ratios for OS at select follow-up times for competing interventions versus pembrolizumab (second order FP model [p1 = 0, p2 = 0]) (reproduced from CS, Table 36, page 75) 80

Table 28. Model fit estimates for fixed-effects network meta-analysis with parametric survival models for OS (reproduced from CS, Table 35, page 74) 80

Table 29. Company's base case results versus standard of care 84

Table 30. Company's base case results versus CAPOX 84

Table 31. Company's base case results versus panitumumab + mFOLFOX6 85

Table 32. Systematic review summary 85

Table 33. NICE reference case checklist 86

Table 34. Distribution of patients across each of the SoC regimens in KEYNOTE-177 (Table 49 of CS) 91

Table 35. Treatment regimens included in the model (adapted from Table 9 and Table 66 of the CS) 91

Table 36. Comparators by population based on the NICE final scope 94

Table 37. Overview of clinical data informing the health state transitions in the economic model 97

Table 38. Estimated life years by health state for each population 102

Table 39. ERG approach to treatment effectiveness by population 106

Table 40. Grade 3 or higher AEs with ≥ 5% incidence implemented in the model (Table 54 of the CS) 107

Table 41. Weekly adverse event rates by treatment arm (adapted from Table 54 and Table 55 of the CS) 108

Table 42. Health state utility values (adapted from Table 60 in CS) 110

Table 43. Adverse event disutility values (taken from the economic model) 110

Table 44. Alternative utility values for scenario analysis (Table 60 in the CS) 111

Table 45. Results of the utility value scenario analyses 111

Table 46. Pooled time to death utility values (adapted from Table 59 of the company submission) 112

Table 47. Drug acquisition costs (taken from the economic model) 113

Table 48. Mean time on treatment (taken from the economic model) 115

Table 49. Comparator administration costs 116

Table 50. Disease management costs (adapted from Table 68 of the company submission) 117

Table 51. Adverse event costs 117

Table 52. Proportion of patients on subsequent treatment costs (adapted from Table 57 and 58 of the company submission) 118

Table 53. Subsequent therapy costs 119

Table 54. Company's base case deterministic results: pembrolizumab versus standard of care 123

Table 55. Company's base case deterministic results: pembrolizumab versus CAPOX 123

Table 56. Company's base case deterministic results: pembrolizumab versus panitumumab + mFOLFOX6 124

Table 57. Company's base case deterministic fully incremental analysis 124

Table 58. Company's PSA results: pembrolizumab versus standard of care 124

Table 59. Company's PSA results: pembrolizumab versus vs CAPOX 126

Table 60. Company's PSA results: pembrolizumab versus panitumumab + mFOLFOX6 126

Table 61. Company's base case probabilistic fully incremental analysis 126

Table 62. Scenario analysis results 130

Table 63. Results of the ERG's scenario analyses – ITT/RAS wild-type population 134

Table 64. Scenario 9 - pembrolizumab versus cetuximab combination treatment - RAS wild-type only 137

Table 65. ERG's preferred model assumptions – ITT/ RAS wild-type population 138

Table 66. ERG's base case deterministic fully incremental analysis 139

Table 67. Summary of population baseline characteristics for KEYNOTE-177 (adapted from CS, Table 10, page 29) 149

Table 68. Time-varying hazard ratios for PFS at selected follow-up times using standard of care as the baseline treatment (second order FP model [p1 = 0, p2 = 0]) (adapted from CS, Appendix M, Table 191, page 627) 154

List of Figures

Figure 1. Kaplan–Meier estimates of PFS based on central imaging vendor per RECIST 1.1 (ITT population; database cut-off date: 19 Feb 2020) (reproduced from CS, Figure 9, page 61)..... 59

Figure 2. Kaplan–Meier estimates of OS (ITT population; database cut-off date: 19 Feb 2020) (reproduced from CS, Figure 4, page 47)..... 63

Figure 3. Network diagram for network meta-analyses of PFS and OS..... 69

Figure 4. ERG's suggested network for generation of estimates of effect for pembrolizumab versus cetuximab and panitumumab combination treatments in RAS wild-type mCRC..... 77

Figure 5. Model schematic for the state-transition and partitioned survival models (Figure 19 of the company submission)..... 96

Figure 6. Company base case extrapolation of time to progression outcomes from KEYNOTE-177 (taken from the economic model)..... 100

Figure 7. Company base case extrapolation of progression-free survival outcomes from KEYNOTE-177 (taken from the economic model)..... 101

Figure 8. Company base case extrapolation of post-progression survival outcomes from KEYNOTE-177 (taken from the economic model)..... 102

Figure 9. Comparison of time on treatment and progression-free survival for pembrolizumab in KEYNOTE-177..... 104

Figure 10. Cost-effectiveness plane: pembrolizumab versus standard of care..... 127

Figure 11. Cost-effectiveness plane: pembrolizumab versus CAPOX..... 127

Figure 12. Cost-effectiveness plane: pembrolizumab versus panitumumab + mFOLFOX6..... 128

Figure 13. OWSA tornado diagram: pembrolizumab versus standard of care..... 129

Figure 14. OWSA tornado diagram pembrolizumab vs CAPOX..... 129

Figure 15. OWSA tornado diagram pembrolizumab vs panitumumab + mFOLFOX6..... 130

Figure 16. Analysis of progression-free survival by subgroup factors by central imaging vendor per RECIST 1.1 (ITT population), based on Cox regression model with Efron's method of tie handling with treatment as a covariate (database cut-off 19 Feb 2020; reproduced from CS, Figure 10, page 63) 152

Figure 17. Analysis of overall survival by subgroup factors (ITT population), based on Cox regression model with Efron's method of tie handling with treatment as a covariate, database cut-off date 19 Feb 2020 (reproduced from CS, Figure 5, page 48) 153

Figure 18. Estimated treatment hazard ratio over time for PFS and treatments relative to SoC (second order FP model; p1=0, p2=0)) based on the ITT population (reproduced from CS, Appendix M, Figure 97, page 626) 155

Figure 19. ERG's validation of estimated treatment hazard ratio over time for PFS and treatments relative to SoC (second order FP model; p1=0, p2=0)) based on the ITT population..... 155

List of Abbreviations

AE	Adverse event
AIC	Akaike information criterion
BIC	Bayesian information criterion
BICR	Blinded independent central review
CAPOX/XELOX	Capecitabine plus oxaliplatin
CI	Confidence interval
CR	Complete response
CRD	Centre for Reviews and Dissemination
CrI	Credible interval
CS	Company submission
CSR	Clinical Study Report
DIC	Deviance information criterion
dMMR	Mismatched repair deficiency
DSU	Decision support unit
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal growth factor
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
ERG	Evidence Review Group
FOLFIRI	Folinic acid plus fluorouracil plus irinotecan
FOLFOX	Folinic acid plus fluorouracil plus oxaliplatin
FP	Fractional polynomial
HR	Hazard ratio
HRQoL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IHC	Immunohistochemistry

IPD	Individual patient data
irRECIST	Immune-related Response Evaluation Criteria in Solid Tumors
ITT	Intention to treat
KM	Kaplan–Meier
LYG	Life years gained
mCRC	Metastatic colorectal cancer
MMR	Mismatched repair
MSI-H	High microsatellite instability
NHS	National Health Service
NICE	National Institute of Health and Care Excellence
NMA	Network meta-analysis
NR	Not reported
OR	Odds ratio
ORR	Overall response rate
OS	Overall survival
OWSA	One-way sensitivity analysis
PAS	Patient access scheme
PCR	Polymerase chain reaction
PD	Progressive disease
PFS	Progression-free survival
PH	Proportional hazards
PR	Partial response
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PPS	Post-progression survival
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model

Q3W	Every 3 weeks
Q6W	Every 6 weeks
QALYs	Quality-adjusted life years
QoL	Quality of life
RAS	KRAS and NRAS (the RAS genes)
RCT(s)	Randomised controlled trial(s)
RECIST	Response Evaluation Criteria in Solid Tumours
SAP	Statistical analysis plan
SAS	Safety analysis set
SD	Standard deviation
SLR	Systematic literature review
SmPC	Summary of product characteristics
SoC	Standard of care
STA	Single Technology Appraisal
STM	State-transition model
ToT	Time on treatment
TSD	Technical support document
TTP	Time to progression

1 Executive summary

Below is a summary of aspects of the submitted evidence identified by the Evidence Review Group (ERG) as potentially important considerations in decision making. The ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs) are also reported. Supplementary details on the clinical condition, technology under assessment, evidence submitted, and discussion of non-key issues are available in the main body of the ERG report.

1.1 Overview of the ERG's key issues

Table 1 presents a summary of the ERG's key issues on the evidence submitted on the clinical and cost effectiveness of pembrolizumab.

Table 1. Summary of key issues

ID	Summary of issue	Report sections
1	KEYNOTE-177 is the only RCT reporting data for first-line treatment of MSI-H/dMMR mCRC and includes a comparator of physician's choice	1.3, 2.3.1, 2.3.3, 3.3.1, 3.5.1
2	Subgroup analyses based on RAS mutation status	1.4, 2.3.3.2, 3.3.1, 3.4, 3.5, 4.2.2, 4.2.3, 4.2.5.1
3	Treatment regimen and resource use for pembrolizumab	1.4, 4.2.3, 4.2.8.9
4	Duration of treatment with pembrolizumab	1.4, 4.2.8.9
5	Treatment costs for standard of care	1.4, 4.2.8.9
6	Time on treatment for non-KEYNOTE-177 comparators	1.4, 4.2.8.9

Abbreviations: dMMR, mismatched repair deficiency; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin mCRC, metastatic colorectal cancer; MSI-H, high microsatellite instability; RCT, randomised controlled trial.

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are around the appropriate subgroups for the analyses based on RAS mutation status and associated comparators and the cost assumptions used in the economic model.

1.2 Overview of key model outcomes

NICE technology appraisals compare how much a new technology improves length (overall survival; 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:

- Increasing progression-free survival (PFS).

Overall, the technology is modelled to affect costs by:

- Its higher unit price compared with currently available treatments in the NHS.
- Reducing the frequency of treatment cycles compared to currently available treatments in the NHS.
- Reducing the resource use associated with delivering treatment.
- Inclusion of a cap on number of treatment cycles (maximum of 35 treatment cycles).

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

- The use of post-progression survival (PPS) for all arms of the model directly impacts OS, as any gain in PFS directly equates to a gain in the estimated OS.
- Treatment regimen of 400mg once every six weeks and consultant oncologist appointments aligned to treatment cycle.
- Treatment costs for standard of care (SoC) based on mFOLFOX6 and FOLFIRI.
- Time on treatment for non-KEYNOTE-177 comparators equal to time on treatment for standard of care in KEYNOTE-177.

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

Table 2 presents the ERG's key issue with the internal and external validity of the evidence derived from KEYNOTE-177 in the context of the decision problem. Issue 1 is also applicable to the section on the ERG's issues with the evidence on clinical effectiveness.

Table 2. Issue 1: Direct head-to-head evidence is not available for comparators listed in final scope issued by NICE

Report section	2.3.3, 3.3
Description of issue and why the ERG has identified it as important	At the time of writing, KEYNOTE-177 is the only RCT enrolling specifically those with locally confirmed MSI-H/dMMR Stage IV CRC and for which data are available in the first-line setting. The comparator group in KEYNOTE-177 was SoC, which was physician's choice from one of (treatment chosen before randomisation): • mFOLFOX6, with or without bevacizumab (52.5%); • FOLFIRI, with or without bevacizumab (36.4%); • Cetuximab with either mFOLFOX6 or FOLFIRI (11.2%).

	Thus, no RCT is available in those with MSI-H/dMMR mCRC to provide head-to-head data, without breaking randomisation, for pembrolizumab versus individual comparators of interest listed in the final scope issued by NICE, including subgroup analysis based on RAS wild-type. Additionally, bevacizumab is not an available treatment option for first-line mCRC in England. Lack of direct comparative data for individual interventions means that caveats that need to be considered when interpreting available estimates of effect, which could lead to concerns in the robustness of the estimates.
What alternative approach has the ERG suggested?	For comparison with CAPOX, FOLFIRI and FOLFOX, the ERG considers the SoC group in totality from KEYNOTE-177 provides the most robust estimate of comparative treatment effectiveness for pembrolizumab, and likely underestimates the true effect of pembrolizumab, because SoC included bevacizumab and cetuximab-combination regimens. To provide more robust estimates of PFS for pembrolizumab versus cetuximab- and panitumumab- combination treatments in those with RAS wild-type mCRC, the ERG proposes that an NMA based on RAS wild-type would be more appropriate than the company's NMA in all patients.
What is the expected effect on the cost-effectiveness estimates?	Please see Issue 2 (Table 3) for a detailed description of the potential impact of carrying out the FP NMA in the subgroup of those with RAS wild-type.
What additional evidence or analyses might help to resolve this key issue?	As noted above, the ERG proposes an NMA in those with RAS wild-type and including relevant comparators listed in the NICE final scope.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; CS, company submission; dMMR, mismatched repair deficiency; ERG, Evidence Review Group; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin; FP, fractional polynomial; mCRC, metastatic colorectal cancer; MSI-H, microsatellite instability; NICE, National Institute of Health and Care Excellence; NMA, network meta-analysis; PFS, progression-free survival.

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

Issue 1 relating to the decision problem is also of relevance to this section. Table 3 to Table 7 presents the ERG's key issues with the company's cost-effectiveness analysis. All cost-effectiveness analyses presented in this report are inclusive of the company's patient access scheme (PAS) simple discount of [redacted]%.

Table 3. Issue 2: Subgroup analyses by RAS mutation status

Report section	2.3.3.2, 3.3.1, 3.4, 3.5, 4.2.2, 4.2.3, 4.2.5.1
Description of issue and why the ERG has identified it as important	The final scope issued by NICE specifies that panitumumab-combination regimens are a first-line treatment option for those with RAS wild-type mCRC, together with cetuximab plus FOLRIRI or FOLFOX for EGFR-expressing RAS wild-type. In the CS, the company provided an FP NMA that omitted cetuximab-combination treatment and was based on all patients. The company's rationale for not including cetuximab plus FOLFIRI or FOLFOX in the NMA was that cetuximab-based treatment formed part of

	SoC, and the six regimens comprising SoC were considered as a single group for the purposes of the NMA Subgroup analysis from KEYNOTE-177 indicates that the beneficial effect of pembrolizumab in PFS noted for the ITT population is maintained versus SoC for those with RAS wild-type. However, RAS status was not determined for all those enrolled in KEYNOTE-177 and was not necessarily a factor in choice of treatment in SoC, thus, some allocated to SoC might not have received the optimum intervention according to clinical practice in England. The ERG considers an FP NMA based on those with RAS wild-type in KEYNOTE-177 versus comparators relevant to clinical practice in England could substantiate the estimate of PFS observed for pembrolizumab in KEYNOTE-177 for all patients. *Post hoc* subgroup analyses of PFS based on those identified as having RAS mutations generated a markedly different result from other subgroups, with a change in direction of effect to favour SoC, albeit that the difference between pembrolizumab and SoC was not statistically significant (HR 1.19, 95% CI: 0.68 to 2.07). However, the ERG notes that the 95% CIs for estimates of PFS for RAS wild-type and non-RAS wild-type do not overlap and, as such, considers the results in the two subgroups unlikely to have arisen due to random chance. The company base case population in the economic model is the ITT population of KEYNOTE-177 with all estimates of clinical efficacy for all comparators based on this population. The ERG was concerned that subgroup analyses by RAS mutation status were not fully explored.
What alternative approach has the ERG suggested?	To carry out the FP NMA based on IPD data for those from KEYNOTE-177 identified as RAS wild-type and to produce a network including cetuximab- and panitumumab-combination regimens for PFS. The ERG appreciates that the same bias in the company's FP NMA based on "all patients" will be present in the ERG's proposed NMA, due to the different mCRC populations included in the network. However, by focusing on the RAS wild-type population, the clinical heterogeneity will be minimised as much as possible. For the RAS wild-type subgroup, the ITT cost-effectiveness analysis presented by the company may provide conservative approximations for the comparisons against mFOLFOX6/FOLFIRI, CAPOX and panitumumab combination treatment, based on a comparison of HRs from KEYNOTE-177 for pembrolizumab versus SoC for each population (ITT HR = 0.60 versus RAS wildtype HR = 0.44). However, currently there is no explicit comparison of pembrolizumab versus cetuximab combination treatment as it is blended with the SoC analysis. To provide an estimation of the cost effectiveness of pembrolizumab versus cetuximab combination treatment, the ERG considers that a simplified analysis assuming clinical equivalence between cetuximab combination treatment and panitumumab combination treatment may not be unreasonable. TA439 reports that NMA provided no statistically significant evidence to suggest that cetuximab plus FOLFOX was more effective than panitumumab plus FOLFOX at increasing PFS or OS.
What is the expected effect on the cost-effectiveness estimates?	For the illustrative analysis for pembrolizumab versus cetuximab combination treatment, the ERG estimated that pembrolizumab is the dominant treatment.

	For the non-RAS wild-type subgroup analysis, as the direction of effect favours SoC, the ERG predicts that pembrolizumab is less effective and more costly that SoC and as such would be dominated by the comparator.
What additional evidence or analyses might help to resolve this key issue?	Given the importance of RAS mutation status, not only in terms of relative treatment effectiveness but also for directing treatment choice for patients in the NHS, during the clarification stage the ERG requested the company to provide cost-effectiveness analyses for RAS wild-type and non RAS wild-type subgroups from KEYNOTE-177 and update the FP NMA to generate relative estimates of treatment effect for pembrolizumab versus the listed comparators in the NICE final scope for the relevant subgroups. The company declined to provide the requested analyses. The ERG reiterates that the preferred approach is for the company to: • Produce relative estimates of treatment effectiveness for patients with and without RAS wildtype mCRC from KEYNOTE-177 using the recommended ERG FP NMA outlined in the ERG's clarification questions; • Implement the results of the subgroup treatment effectiveness analyses in the economic model to produce cost-effectiveness estimates by relevant comparator for the RAS wild-type and non-RAS wild-type subgroups.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; CI, confidence interval; CS, company submission; EGFR, epidermal growth factor receptor; ERG, evidence review group; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; FP, fractional polynomial; HR, hazard ratio; ITT, intention-to-treat; mCRC, metastatic colorectal cancer; NMA, network meta-analysis; OS, overall survival; PFS, progression-free survival.

Table 4. Issue 3: Treatment regimen and resource use for pembrolizumab

Report section	4.2.3 and 4.2.8.9
Description of issue and why the ERG has identified it as important	In KEYNOTE-177, the treatment regimen for pembrolizumab was 200mg once every three weeks, but the company stated that it can also be administered at a 400mg dose once every six weeks. The ERG's clinical experts advised that in clinical practice, they would prefer to administer pembrolizumab using the 400mg dose once every six weeks for patient convenience and also to reduce NHS resource use. As such, a less frequent treatment regimen requires less frequent consultant oncologist appointments, which is a primary driver of the cost-effectiveness analysis.
What alternative approach has the ERG suggested?	To reflect how pembrolizumab would be used in UK clinical practice, the ERG considers the treatment regimen of 400mg once every six weeks and consultant oncologist appointments aligned to treatment cycle is appropriate.
What is the expected effect on the cost-effectiveness estimates?	The company base case ICERs after the clarification stage reduced from £7,250 to £535 for the comparison with SoC and from £27,474 to £20,736 for the comparison with CAPOX, but pembrolizumab still dominates panitumumab combination treatment under this scenario
What additional evidence or analyses might help to resolve this key issue?	The company expects that the monotherapy marketing authorisation will include an option to administer pembrolizumab at a 400mg dose once every six weeks. As such, confirmation of the treatment regimen for pembrolizumab will be available when the marketing authorisation is approved.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; mg, milligram; SoC, standard of care.

Table 5. Issue 4: Duration of treatment with pembrolizumab

Report section	4.2.8.9
Description of issue and why the ERG has identified it as important	In KEYNOTE-177, treatment with pembrolizumab was restricted to 35-cycles, which has been implemented in the cost-effectiveness analysis. However, the draft SmPC states that it can be given until disease progression or unacceptable toxicity.
What alternative approach has the ERG suggested?	The ERG performed an illustrative scenario, which assumes time on treatment is equal to PFS for pembrolizumab and removed the 35-cycle stopping rule.
What is the expected effect on the cost-effectiveness estimates?	The company base case ICERs after the clarification stage increased from £7,250 to £73,809 for the comparison with SoC and from £27,474 to £94,262 for the comparison with CAPOX, and from dominant to £44,777 for the comparison with panitumumab combination treatment.
What additional evidence or analyses might help to resolve this key issue?	As clinical efficacy for pembrolizumab in the economic model is based on a maximum of 35-cycles of treatment, it is unknown what the impact on PFS would be if treatment was given until progression.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; SmPC, summary of product characteristics; SoC, standard of care.

Table 6. Issue 5: Treatment costs for standard of care

Report section	4.2.8.9
Description of issue and why the ERG has identified it as important	In KEYNOTE-177 approximately 70% of patients received bevacizumab combination treatment, which is not recommend in the NHS. The company assumed that costs for patients who received bevacizumab combination treatment would be reflective of cetuximab combination treatment. However, cetuximab combination treatment is only recommend in the NHS for patients with RAS-wildtype mCRC. As such, the ERG considers that costs for standard of care are inflated compared to what would be incurred in the NHS.
What alternative approach has the ERG suggested?	For the ITT and RAS wildtype analysis, which focuses only on the comparison with FOLFOX and FOLFIRI (SoC), the ERG ran a scenario where treatment costs for standard of care are based only on FOLFOX and FOLFIRI treatments.
What is the expected effect on the cost-effectiveness estimates?	The company base case ICER after the clarification stage increased from £7,250 to £21,636 for the comparison with SoC.
What additional evidence or analyses might help to resolve this key issue?	As a large proportion of the direct evidence for pembrolizumab versus standard from KEYNTOTE-177 is based on patients in the comparator arm receiving treatment with bevacizumab, the uncertainty of the clinical efficacy of pembrolizumab versus FOLFOX and FOLFIRI is unresolvable. Furthermore, the ERG considers it is important to highlight that SoC treatment costs for the company base case are higher than may be incurred in NHS. However, the ERG considers that the bias in the ITT analysis and the ERG's treatment cost scenario favours SoC and as such provides a conservative estimate of the ICER.

Abbreviations: ERG, evidence review group; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; ICER, incremental cost-effectiveness ratio; ITT, intention-to treat; mCRC, metastatic colorectal cancer; SoC, standard of care.

Table 7. Issue 6: Time on treatment for non-KEYNOTE-177 comparators

Report section	4.2.8.9
Description of issue and why the ERG has identified it as important	In the base case analysis, the company assumed time on treatment to be equal to estimated progression-free survival for CAPOX and panitumumab combination treatment, in the absence of alternative data. However, mean time on treatment for CAPOX and panitumumab combination treatment is substantially longer compared with pembrolizumab and SoC, which results in inflated treatment costs. In TA439, mean treatment duration for panitumumab combination treatment was estimated to be approximately nine months, which is shorter than the company's estimated mean treatment duration of 18.5 months based on time on treatment equal to progression-free survival.
What alternative approach has the ERG suggested?	The ERG considers that a more appropriate assumption for time on treatment for non-KEYNOTE-177 comparators is to assume it is equal to KEYNOTE-177 time on treatment for SoC. Mean time on treatment for SoC was estimated to be approximately [redacted] months, which is closer to the estimates in TA439 for panitumumab combination treatment. Furthermore, given the company assume clinical outcomes for CAPOX are equal to SoC, it is not unreasonable to assume time on treatment is equal to time on treatment for SoC.
What is the expected effect on the cost-effectiveness estimates?	The scenario had a negligible impact on the ICER for CAPOX (increased from £27,474 to £29,205) but changed the ICER for panitumumab combination treatment from being dominant to £3,158.
What additional evidence or analyses might help to resolve this key issue?	The ERG considers that its scenario sufficiently explores the uncertainty around time on treatment for comparators not evaluated in KEYNOTE-177. Furthermore, mean estimates of time on treatment for panitumumab from TA439 were accepted by the committee and as such could be considered appropriate for use in the ERG's preferred analysis.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; CI, confidence interval; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; SoC, standard of care.

1.5 Summary of ERG's preferred assumptions and resulting ICER

Table 8 to Table 11 presents the results of the ERG's preferred assumptions, as well as the ERG preferred ICER for pembrolizumab compared with SoC and CAPOX for the ITT/ RAS wild-type population proxy analysis and panitumumab and cetuximab combination treatment for the RAS wild-type population analysis. The ERG could not produce probabilistic sensitivity analysis (PSA) ICERs for its base case as the PSA takes several hours to run and due to paucity of time and complexity of the model, some scenarios could not be integrated with the PSA.

Table 8. ERG's preferred assumptions - pembrolizumab vs SoC (ITT and RAS wild-type proxy analysis)

Scenario	Incremental costs	Incremental QALYs	ICER - change from company base case (£/QALY)
Company base case [after clarification]	13,497	1.86	7,250
Treatment regimen and consultant outpatient appointments for pembrolizumab – 400mg once every six weeks	996	1.86	535
Implementing the Weibull distribution after the 20-week cut-off point for the extrapolation of TTP and PFS	13,171	1.89	6,966
Removal of AE disutility	13,497	1.86	7,251
FOLFOX/FOLFIRI costs for SoC	40,278	1.86	21,636
Removal of second-line cetuximab combination treatment	13,911	1.86	7,473
ERG's preferred base case [combination of all scenarios]	**27,541**	**1.89**	**14,569**

Abbreviations: AE, adverse events; ERG, evidence review group; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; PFS, progression-free survival; QALYs, quality-adjusted life-years; TTP, time to progression; SoC, standard of care.

Table 9. ERG's preferred assumptions - pembrolizumab vs CAPOX (ITT and RAS wild-type proxy analysis)

Scenario	Incremental costs	Incremental QALYs	ICER - change from company base case (£/QALY)
Company base case [after clarification]	50,968	1.86	27,474
Treatment regimen and consultant outpatient appointments for pembrolizumab – 400mg once every six weeks	38,468	1.86	20,736
Implementing the Weibull distribution after the 20-week cut-off point for the extrapolation of TTP and PFS	50,307	1.71	29,819
Removal of AE disutility	50,968	1.86	27,383
ToT for comparator is equal to ToT for standard of care	54,180	1.86	29,205
Removal of second-line cetuximab combination treatment	51,383	1.86	27,697

ERG's preferred base case [combination of all scenarios]	**41,443**	**1.89**	**21,923**

Abbreviations: AE, adverse events; CAPOX, capecitabine plus oxaliplatin; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; ITT, intention-to-treat; PFS, progression-free survival; QALYs, quality-adjusted life-years; ToT, Time on treatment; TTP, time to progression.

Table 10. ERG's preferred assumptions - pembrolizumab vs panitumumab combination treatment (RAS wild-type proxy analysis)

Scenario	Incremental costs	Incremental QALYs	ICER - change from company base case (£/QALY)
Company base case [after clarification]	-48,317	1.69	Dominant
Treatment regimen and consultant outpatient appointments for pembrolizumab – 400mg once every six weeks	-60,817	1.69	Dominant
Implementing the Weibull distribution after the 20-week cut-off point for the extrapolation of TTP and PFS	-56,167	1.68	Dominant
Removal of AE disutility	-48,317	1.65	Dominant
ToT for comparator is equal to ToT for standard of care	5,330	1.69	3,158
Removal of second-line cetuximab combination treatment	-47,946	1.69	Dominant
ERG's preferred base case [combination of all scenarios]	**-7,675**	**1.65**	**Dominant**

Abbreviations: AE, adverse events; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; QALYs, quality-adjusted life-years; ToT, Time on treatment; TTP, time to progression; SoC, standard of care.

Table 11. ERG's preferred assumptions - pembrolizumab vs cetuximab combination treatment (RAS wild-type proxy analysis)

Scenario	Incremental costs	Incremental QALYs	ICER - change from company base case (£/QALY)
Company base case [after clarification]	-	-	-
Cetuximab combination treatment is clinical equivalent to panitumumab combination treatment	-49,510	1.69	Dominant

Treatment regimen and consultant outpatient appointments for pembrolizumab – 400mg once every six weeks	-62,011	1.69	Dominant
Implementing the Weibull distribution after the 20-week cut-off point for the extrapolation of TTP and PFS	-57,446	1.68	Dominant
Removal of AE disutility	-49,510	1.65	Dominant
ToT for comparator is equal to ToT for standard of care	4,814	1.69	2,852
Removal of second-line cetuximab combination treatment	-49,140	1.69	Dominant
ERG's preferred base case [combination of all scenarios]	**-8,191**	**1.65**	**Dominant**

Abbreviations: AE, adverse events; ERG, evidence review group; ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; QALYs, quality-adjusted life-years; ToT, Time on treatment; TTP, time to progression; SoC, standard of care.

For further details of the exploratory and sensitivity analyses done by the ERG, see Section 1326.2 and 6.3.

2 Introduction and background

2.1 Introduction

Herein is a critique of the evidence submitted to the Single Technology Appraisal (STA) in support of the clinical and cost effectiveness of pembrolizumab (Keytruda®; Merck Sharp & Dohme Limited) as a regimen for adults with untreated metastatic colorectal cancer (mCRC) with high microsatellite instability (MSI-H) or mismatch repair deficiency (dMMR).[1]

2.2 Background

Within Section B.1 of the company submission (CS), the company provides an overview of:

- pembrolizumab, including its mode of action, dose and method of administration (Section B.1.2);
- mCRC with MSI-H or dMMR, including prevalence, prognosis, staging and disease management (Section B.1.3).

Based on advice from its clinical experts, the Evidence Review Group (ERG) considers the CS to present an accurate overview of the aetiology and diagnosis of mCRC that is MSI-H or dMMR, and the management of mCRC.

Direct evidence submitted in support of the clinical effectiveness of pembrolizumab in the population of interest to this STA is derived from one randomised controlled trial (RCT), KEYNOTE-177.[2] The company notes that KEYNOTE-177 is the first RCT focussing on those with untreated mCRC that is determined to be MSI-H or dMMR. To aid understanding of some points raised in the ERG's critique of the submitted evidence in the context of the decision problem, the ERG provides a summary of aspects of mCRC not covered in detail in the CS and that potentially affect response to treatment and prognosis, together with how presence of MSI-H or dMMR impacts on course of disease.

2.2.1 Colorectal cancer

Colorectal cancer (also known as bowel cancer) starts in either the colon or the rectum. The colon is about 5 feet long and is part of the beginning of the large bowel. By contrast, the rectum, which is located towards the end of the large bowel, is only 5 to 6 inches in length.[3] About two thirds of CRCs develop in the colon, with the remainder arising in the rectum.[4] Despite the difference in anatomical location of tumours, the two cancers have similar features, giving rise to cancers of the colon and rectum being grouped together.[3, 5] Although cancers of the colon and rectum have similar features,

types of symptom experienced by a person with CRC depend on whether their primary tumour is located on the left or the right side of the large intestine.[4, 5] Right-sided tumours are more frequently associated with weight loss and anaemia, whereas those located on the left side of the colon or rectum often lead to colicky pain, rectal bleeding, and bowel obstruction.

Stage of disease at presentation is the strongest prognostic factor for clinical outcomes in CRC.[6] However, tumour position, in terms of left or right side, impacts on disease progression, overall survival (OS) and response to treatment.[3] Compared with tumours on the left, tumours on the right are more likely to be at an advanced stage of disease at presentation and, therefore, have a poorer prognosis.[3] Additionally, right and left tumours differ in their molecular characteristics and histology.[3] Mutations in the MMR pathway are more commonly identified in tumours located on the right of the colon or rectum, whereas those on the left typically have more anomalies in genes involved in chromosomal instability, such as KRAS and p53. Positioning of tumour is reported to influence response to treatment,[3] with left-sided CRC tumours tending to show better response to adjuvant chemotherapy (e.g., 5-fluorouracil-based treatment) and to targeted therapy (e.g., inhibitors of epidermal growth factor receptor [EGFR]). By contrast, right-sided CRCs do not respond well to conventional chemotherapies, and are reported to have a better outcome with immunotherapies (e.g., pembrolizumab).[3]

2.2.2 Microsatellite instability, mismatch repair, and other genetic anomalies involved in development of colorectal cancer

Colorectal cancer is a heterogenous disease. Although disease stage remains the key determinant of prognosis, there is variability in clinical outcomes for the same disease stage.[7, 8] Various genetic anomalies have been identified as having a role in development of CRC, with differences in mutations across patients likely contributing to heterogeneity in clinical outcomes. Two molecular pathways involved in CRC are the MSI and the chromosomal instability (CIN) pathways.[3, 7-9] Most CRCs develop via the CIN pathway, whereas 12–15% arise from the MSI pathway that is, in turn, a consequence of dMMR.[7, 8]

Genomic stability is maintained through the DNA MMR system, which repairs errors in insertions, deletions, and base–base mismatches introduced into microsatellites during DNA replication and combination.[7, 8] Microsatellites are short, tandemly-repeated sequences (1–6 base pairs) occurring throughout the genome. As a result of their repeated structure, microsatellites are prone to mutation. Presence of microsatellites with a sequence not occurring in germline DNA indicates presence of MSI and a dMMR system, and, therefore, microsatellites are a marker of dMMR.[7, 8] The

DNA MMR system comprises four MMR genes and their encoded proteins (MLH1, MSH2, MSH6, PMS2).[7, 8] Inactivation of MLH1 and MSH2 accounts for over 90% of dMMR CRCs. Deficiency of MMR results in the production of a truncated, non-functional protein or the loss of a protein, which causes MSI. Therefore, dMMR is frequently assessed by testing for loss of an MMR protein with immunohistochemistry or for MSI using a polymerase chain reaction (PCR)-based assay. The ERG's clinical experts commented that, mostly, tests for MSI and MMR status give the same result, that is both generate a positive or negative status. However, on occasion, a person could be identified as microsatellite stable but dMMR, and, therefore, evaluating MMR proteins could be the preferred technique.

Tumours arising from a dMMR pathway have distinct features, such as origin in the right side of the colon or rectum and poorly differentiated morphology.[9] In early-stage CRC, dMMR tumours are associated with a favourable prognosis.[9] Prevalence of dMMR is low in mCRC (3.5%), which supports the proposal that dMMR tumours have a lower potential to metastasise. However, when dMMR tumours are present in mCRC, they are associated with a considerably worse outcome than those with a functional MMR system.[9] An analysis of individual patient data (IPD) from participants of four RCTs evaluating first-line treatment in mCRC found that median progression-free survival (PFS) and OS were significantly worse for those with dMMR compared with MMR tumors:[9]

- PFS:
    - 6.2 months for dMMR versus 7.6 months for MMR;
    - hazard ratio (HR) 1.33, 95% CI (confidence interval) 1.12 to 1.5;
    - p=0.001;
- OS:
    - 13.6 months for dMMR versus 16.8 months for MMR;
    - HR 1.35, 95% CI: 1.13: to 1.61;
    - p=0.001.

Other genetic anomalies involved in pathways leading to the development of CRC, and that are now targets for treatment, are those affecting proteins acting in the Ras-Raf-mitogen-activated protein kinase (MAPK) signalling pathway.[10] One such gene is BRAF, which is a modulator of the MAPK signalling pathway. BRAF mutations are present in 10% of CRC, and carriers of BRAF mutations exhibit discrete clinical characteristics and outcomes. A specific mutation, $BRAF^{v600E}$, accounts for approximately 90% of all BRAF mutations seen in CRC. Additionally, $BRAF^{v600E}$ is strongly associated with MSI. In sporadic CRCs, BRAF mutation is seen in approximately 60% of MSI-H tumours and only

5–10% of microsatellite stable tumours.[10] People with BRAF mutant CRC have low response rates to conventional therapies and poor OS. Analysis of individual patient data (IPD) from the study mentioned earlier found that median PFS and OS were significantly worse for those with BRAF mutation (includes both dMMR and MMR) compared with those with BRAF wild-type:[9]

- PFS:
    - 6.2 months for BRAF mutant versus 7.7 months for BRAF wild-type;
    - HR 1.34, 95% CI: 1.17 to 1.54;
    - p<0.001;
- OS:
    - 11.4 months for BRAF mutant versus 17.2 months for BRAF wild-type;
    - HR 1.91, 95% CI 1.66 to 2.19;
    - p<0.001.

Other genes involved in the MAPK signalling pathway are the RAS genes, KRAS and NRAS.[10] Mutations in KRAS and NRAS are present in 50% of CRCs. KRAS/NRAS mutations lead to activation of the MAPK signalling pathway downstream of EGFR, rendering these tumours resistant to anti-EGFR therapies, such as cetuximab and panitumumab.[11]

2.3 Critique of the company's definition of the decision problem

The company provided a summary of the final scope issued by the NICE, together with their rationale for any deviation from the final scope (Table 12).[1] The company highlights that the submission differs from the final scope primarily in terms of the comparators of interest to the decision problem. The key differences between the decision problem addressed in the CS and the scope are discussed in greater detail in the sections that follow.

Table 12. Summary of decision problem (adapted from CS, Table 1)

	Final scope issued by NICE	Decision problem addressed in the submission	Rationale if different from the scope	ERG comment
Population	Adults with mCRC with MSI-H or dMMR.	Adults with mCRC with MSI-H or dMMR.	N/A	Evidence derived from KEYNOTE-177[2] is aligned with the population specified in the final scope (more detailed description available in Section 2.3.1). The ERG notes that the anticipated marketing authorisation for pembrolizumab could be interpreted as encompassing first-line treatment of adults with unresectable CRC (not metastatic) that is MSI-H or dMMR, a population that is not covered in the final scope issued by NICE (more detailed discussion available in Section 2.3.1).
Intervention	Pembrolizumab	Pembrolizumab	N/A	The ERG notes two factors associated with pembrolizumab treatment that could impact on the estimates of cost effectiveness (discussed in greater detail in the Section 2.3.2 and Section 0): • choice of dosing regimen: ○ fixed dose of 200 mg over 30 minutes every 3 weeks; ○ fixed dose of 400 mg over 30 minutes every 6 weeks.

				• number of cycles of pembrolizumab administered.
Comparator(s)	**For all patients** • FOLFOX • FOLFIRI • CAPOX • Capecitabine • Tegafur with uracil (in combination with folinic acid) • Raltitrexed (only when folinic acid and fluorouracil are not tolerated or unsuitable) **For patients with RAS wild-type mCRC** • Panitumumab in combination with FOLFOX or FOLFIRI **For patients with EGFR expressing, RAS wild-type mCRC** • Cetuximab in combination with FOLFOX or FOLFIRI	**For all patients** • FOLFOX • FOLFIRI • CAPOX **For patients with RAS wild-type mCRC** • Panitumumab in combination with FOLFOX or FOLFIRI **For patients with EGFR expressing, RAS wild-type mCRC** • Cetuximab in combination with FOLFOX or FOLFIRI	• Tegafur with uracil (in combination with folinic acid) is not a relevant comparator for this appraisal as this regimen is no longer available as it was discontinued in the UK.[12, 13] • Raltitrexed (only when folinic acid and fluorouracil are not tolerated or unsuitable) is not a relevant comparator for this appraisal as this is only very rarely used in UK clinical practice. • Capecitabine is not a relevant comparator for this appraisal as it is used in elderly and frail patients who have a poor performance status (i.e. ECOG performance status score of ≥2).	The ERG's clinical experts agree with the company, and the company's rationale for each treatment, that tegafur with uracil, raltitrexed, and capecitabine are not relevant comparators for pembrolizumab in the first-line treatment of MSI-H/dMMR mCRC. No RCT is available that provides direct comparison of pembrolizumab versus any of the individual comparators specified in the final scope issued by NICE (discussed in more detail in Section 2.3.3). The comparator to pembrolizumab in KEYNOTE-177 was SoC as chosen by the investigator, with choice of treatment for each participant prespecified before randomisation. Options for SoC (total of six) were: • mFOLFOX6, with or without bevacizumab; • FOLFIRI, with or without bevacizumab; • Cetuximab with either mFOLFOX6 or FOLFIRI. The ERG notes that bevacizumab is not an available treatment option for

				first-line mCRC in England (discussed in more detail in Section 2.3.3). Data for pembrolizumab versus CAPOX in MSI-H/dMMR mCRC are derived from network meta-analysis. Comparisons of pembrolizumab versus cetuximab and panitumumab, both of which are combined with FOLFOX or FOLFIRI, were not fully explored in the CS (more detail provided in Section 2.3.3).
Outcomes	The outcome measures to be considered include: • overall survival; • progression-free survival; • response rates; • adverse effects of treatment; • health-related quality of life.	The outcome measures to be considered include: • overall survival; • progression-free survival; • response rates; • adverse effects of treatment; • health-related quality of life.	N/A	Data are available for all outcomes listed in the NICE scope; the ERG notes OS data are immature.
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.			The ERG notes that the economic model does not include costs associated with diagnostic testing for MSI status in people with mCRC as assessment of MSI or dMMR is standard clinical practice for all patients with CRC, as per recommendations in DG27.[14]

	The availability of any patient access schemes for the intervention or comparator technologies will be taken into account. The economic modelling should include the costs associated with diagnostic testing for microsatellite instability status in people with metastatic colorectal cancer who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test. See section 5.9 of the Guide to the Methods of Technology Appraisals.			
Subgroups to be considered	If the evidence allows the subgroup of people with RAS wild-type colorectal cancer will be considered.			The ERG considers that the company did not fully explore comparators of interest for the population of RAS wild-type mCRC. The company presented an NMA for all patients and that included panitumumab-based therapy and CAPOX. Given that the company has access to IPD for those with RAS wild-type from KEYNOTE-177, to align with the NICE scope, the ERG requested that the company carries out an NMA derived from IPD for those with RAS-wildtype from KEYNOTE-177 and versus the comparators specified above.

Special considerations, including issues related to equity or equality	Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.			

Abbreviations: CAPOX; capecitabine plus oxaliplatin; CS, company submission; dMMR, mismatched repair deficiency; ECOG; Eastern Cooperative Oncology Group; ERG, Evidence Review Group; FOLFIRI; folinic acid plus fluorouracil plus irinotecan; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin; mCRC, metastatic colorectal cancer; MSI-H, high microsatellite instability; MMR, mismatched repair; N/A, not applicable; SoC, standard of care.

2.3.1 Population

KEYNOTE-177 enrolled those with locally confirmed MSI-H/dMMR Stage IV CRC.[2] Participants were also required to have measurable disease at baseline as per RECIST 1.1 criteria, and as determined by the local site investigator and/or radiology assessment. Those enrolled could have recurrent or newly diagnosed mCRC. Receipt of prior systemic therapy for Stage IV CRC rendered a person ineligible for recruitment to KEYNOTE-177: previous adjuvant chemotherapy for CRC was permitted, as long as treatment had been completed at least 6 months prior to randomisation.

The ERG's clinical experts consider the characteristics of the population enrolled in KEYNOTE-177 to be representative of those in England likely to be eligible for treatment with pembrolizumab: a more detailed discussion of baseline characteristics is available in Section 3.2.

The final scope issued by NICE indicates a subgroup of interest to be those with RAS wild-type mCRC.[1] Subgroup analyses reported within the CS indicate that, in those with RAS wild-type, pembrolizumab is associated with a statistically significant improvement in PFS compared with standard of care (SoC; HR 0.44, 95% CI: 0.29 to 0.67). However, in those with mutation of KRAS or NRAS, the direction of effect favours SoC, albeit that the difference between treatments does not reach statistical significance (HR 1.19, 95% CI: 0.68 to 2.07). The ERG notes that the 95% CIs for the two subgroups do not overlap.

Estimates of PFS and OS for pembrolizumab versus the regimens specified by NICE in the final scope as being comparators of interest for RAS wild-type are not available within the CS (detailed in Section 2.3.3). The ERG appreciates that subgroup analyses are hypothesis generating. However, to substantiate the effect of pembrolizumab, the ERG considered it beneficial for the company to generate estimates of PFS and OS for RAS wild-type and relevant comparators as set out in the final scope, together with assessment of those without RAS-wild-type. As part of the clarification process, the ERG requested that the company carry out network meta-analyses (NMA) based on individual patient data (IPD) for KEYNOTE-177 in those with and without RAS wild-type (NMA discussed in more detail in Section 3.5). The ERG appreciates that, because all studies evaluating comparator regimens involve participants with mCRC and data are not available for the subgroup of MSI-H or dMMR mCRC, bias is present in the NMA. The potential direction of any bias, and its impact on the estimate of relative clinical effectiveness, are discussed in Section 3.5.

To summarise, the ERG considers three populations to be relevant to the decision problem:

- Population A: all participants (ITT population);
- Population B: subgroup of participants with RAS wild-type;
- Population C: subgroup of participants without RAS wild-type.

The ERG considers that the wording of the anticipated marketing authorisation for pembrolizumab could be interpreted as incorporating those with unresectable MSI-H or dMMR CRC that has not metastasised. The ERG's clinical experts commented that unresectable could be applied to either the primary colorectal tumour (with or without metastases) or any metastases. The ERG notes that the inclusion criteria for KEYNOTE-177 specified classification of CRC as Stage IV, which, by definition, denotes presence of metastases. Text in the Clinical Study Report for KEYNOTE-177 states that, "[redacted]".[2] As part of the clarification process, the company indicated that no person [redacted] was enrolled to KEYNOTE-177.

2.3.2 Intervention

In December 2020, the Committee for Medicinal Products for Human Use adopted a positive opinion on the use of pembrolizumab as a monotherapy for the first-line treatment of unresectable or metastatic MSI-H or dMMR CRC in adults.[15]

In KEYNOTE-177, pembrolizumab was given at a fixed dose of 200 mg infused over 30 minutes, with treatment occurring every 3 weeks (Q3W). The company anticipates that the marketing authorisation for pembrolizumab will include an option to administer pembrolizumab at a fixed dose of 400 mg infused over 30 minutes on a 6-weekly cycle (Q6W). The ERG's clinical experts commented that they would expect the two dosing schedules for pembrolizumab to be of equivalent clinical effectiveness. The ERG's experts agreed with the company's proposals that the reduction in number of hospital visits required for the Q6W schedule would lead to improvements in patients' quality of life and to increased capacity within hospitals, should this schedule become an option. The choice of Q3W or Q6W would be based on patient preference and the treating clinician's opinion.

The impact on the cost effectiveness of following the Q6W dosing schedule for pembrolizumab is assessed in Section 4.2.8.

Considering the recommended number of cycles of pembrolizumab, the ERG notes that wording of guidance on how long to continue pembrolizumab for MSI-H/dMMR mCRC differs between the CS, the CSR[2] and the draft Summary of Product Characteristics (SmPC).[16] The draft SmPC advises that, "*patients should be treated with KEYTRUDA [pembrolizumab] until disease progression or unacceptable toxicity*", with no restriction mentioned for number of treatment cycles. The SmPC also states, "*It is recommended to continue treatment for clinically stable patients with initial evidence of disease progression until disease progression is confirmed*."

In the CS, in Section B.3.2 (page 101), for the purposes of the economic model the company has applied a maximum number of 35 cycles of pembrolizumab, based on treatment administration in KEYNOTE-177, which is reported as, "*patients were to continue pembrolizumab until progressive disease, unacceptable adverse events or intercurrent illness preventing further administration of treatment, the subject has a confirmed positive serum pregnancy test or a maximum of 35 cycles of uninterrupted treatment with pembrolizumab*". Wording in Section B.2.3 of the CS (page 18) suggests that all participants receiving pembrolizumab can enter a Second Course Treatment Phase during which they could receive up to 17 additional cycles: "*pembrolizumab arm patients received up to 35 administrations of pembrolizumab (approximately 2 years) in the Initial Treatment Phase. Patients who stopped pembrolizumab with locally confirmed complete response (CR), or stable disease (SD) or better at the end of the Initial Treatment Phase may be treated in a Second Course Treatment Phase with up to 17 administrations of pembrolizumab*".

Additional detail provided in the CSR[2] for KEYNOTE-177 clarifies that participants allocated to the pembrolizumab group who stopped treatment after achieving stable disease or better and **subsequently progressed** could be eligible for an additional 17 cycles of pembrolizumab. Key criteria for re-treatment with pembrolizumab in KEYNOTE-177 are (additional criteria are available in the CSR):

- [redacted]

[redacted]

The ERG's clinical experts fed back that their preference would be to continue pembrolizumab until disease progression. The number of cycles of pembrolizumab administered has implications for cost effectiveness analyses. As part of the clarification process, the ERG requested a breakdown of the number of people receiving 35 cycles during the Initial Treatment Phase, and the proportion given the additional 17 cycles in the Second Course of Treatment Phase. The potential impact on cost effectiveness of pembrolizumab is discussed in Section 4.2.8.

2.3.3 Comparators

Based on advice from clinical experts, the ERG agrees with the company that the comparators listed below are the most relevant to the decision problem;

- **For all patients:**
    - FOLFOX;
    - FOLFIRI;
    - CAPOX.
- **For patients with RAS wild-type mCRC:**
    - Panitumumab in combination with FOLFOX or FOLFIRI.

- **For patients with EGFR expressing, RAS wild-type mCRC:**
    - Cetuximab in combination with FOLFOX or FOLFIRI.

2.3.3.1 *All patients*

No RCT is available in those with MSI-H/dMMR mCRC to provide head-to-head data, without breaking randomisation, for pembrolizumab versus individual comparators of interest listed in the final scope issued by NICE (Table 12). Data for pembrolizumab versus FOLFOX and FOLFIRI could be derived from KEYNOTE-177 through subgroup analyses of participants receiving these treatment options in the SoC group. However, estimates of clinical treatment effect would be associated with the inherent uncertainty arising from *ad hoc* subgroup analyses. Additionally, in KEYNOTE-177, FOLFOX and FOLFIRI could be given either alone or in combination with bevacizumab, which, as the company comments, is not available as a treatment option in the NHS in England. In KEYNOTE-177, 70% of those in the SoC group received FOLFOX or FOLFIRI with bevacizumab (Table 13), and, thus, the sample size to inform the subgroup analyses for FOLFOX or FOLFIRI alone is small, which would add to the uncertainty associated with the analyses. In the CS, the company provides analyses for KEYNOTE-177 that excludes participants receiving bevacizumab-containing regimens, and cautions that the results are less robust than analyses based on the full trial population. The inclusion of cetuximab-based treatment is discussed below.

CAPOX was not a treatment option in SoC for KEYNOTE-177. To generate an estimate of effect for pembrolizumab versus CAPOX, the company carried out an NMA in all patients, with the network also including panitumumab-based treatment. Studies evaluating CAPOX were in adults with mCRC, and the RCT assessing panitumumab focused on RAS wild-type mCRC. Thus, there is considerable clinical heterogeneity between the population enrolled in KEYNOTE-177 and those of other studies informing the network. The extent of bias across all studies in the network is difficult to quantify.

Taken together, and given that no RCT, other than KEYNOTE-177, reports data for solely those with untreated MSI-H/dMMR mCRC, the ERG considers the SoC group in totality from KEYNOTE-177 provides the most robust estimate of comparative treatment effectiveness for pembrolizumab versus FOLFOX, FOLFIRI and CAPOX in the specified group of all patients with MSI-H/dMMR mCRC. In reaching its conclusion, based on data available from published RCTs (details available in Section 3.3), the ERG has assumed that:

- FOLFOX, FOLFIRI and CAPOX are of comparable clinical effectiveness as first-line treatments in mCRC;[17-21]
- combining bevacizumab with FOLFOX, FOLFIRI or CAPOX improves PFS compared with the standard regimen alone in the first-line treatment of mCRC.[22, 23]

When using SoC as a proxy for FOLFOX, FOLFIRI or CAPOX alone, the ERG's assumption that bevacizumab-containing regimens are more clinically effective than FOLFOX or FOLFIRI alone introduces bias that favours SoC and therefore underestimates the likely true estimate of PFS and OS for pembrolizumab versus FOLFOX, FOLFIRI or CAPOX in those with untreated MSI-H/dMMR mCRC. Thus, the ERG considers using the ITT population to provide a conservative estimate of the clinical effectiveness of pembrolizumab. The ERG's clinical experts have fed back that the ERG's assumptions are appropriate for this STA.

Table 13. Proportion of people receiving treatment options available in the standard of care group (reproduced from the CS, Table 16, page 43)

	Standard of care group (N = 143)	
Treatment option	**Number receiving treatment**	**Percentage**
FOLFIRI	16	11.2
mFOLFOX6	11	7.7
FOLFIRI + bevacizumab	36	25.2
mFOLFOX6 + bevacizumab	64	44.8
FOLFIRI + cetuximab	11	7.7
mFOLFOX6 + cetuximab	5	3.5

Abbreviations: CS, company submission; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin.

2.3.3.2 *RAS wild-type*

The ERG considers that the company did not fully explore comparators of interest for the population of RAS wild-type mCRC. The company presented an NMA for all patients that included panitumumab-containing regimens as a comparator, but not cetuximab-based treatment with the rationale that cetuximab combination treatment was an option in the SoC group of KEYNOTE-177.

Subgroup analyses from KEYNOTE-177 in those with RAS wild-type indicated that pembrolizumab is associated with a statistically significant improvement in PFS compared with SoC (HR 0.44, 95% CI: 0.29 to 0.67). Given that RCTs are available evaluating panitumumab and cetuximab-based therapies as first-line treatments in RAS wild-type mCRC, the ERG considers it would be beneficial to carry out an NMA derived from IPD for those with RAS wild-type from KEYNOTE-177 to substantiate the estimate of effect for pembrolizumab noted in KEYNOTE-177 but versus relevant comparators set out in the final scope (Table 12).[24-27] Again, the ERG acknowledges that the NMA evaluating treatments in RAS wild-type is associated with bias that is difficult to quantify because the impact of MSI-H or dMMR is not accounted for in trials other than KEYNOTE-177. The ERG also highlights that its clinical experts commented that MSI-H/dMMR mCRC is distinct from that of RAS wild-type mCRC. Moreover, based on mode of action of pembrolizumab, pembrolizumab is likely to have benefit in MSI-H/dMMR mCRC, irrespective of RAS wild-type status.

The ERG notes the distinction between panitumumab and cetuximab in their marketing authorisations.[11] Cetuximab-based treatment is indicated for EGFR-expressing RAS wild-type mCRC and in combination with FOLFOX as a first-line treatment. By contrast, in the same setting, panitumumab is indicated for RAS wild-type mCRC, with no specification for EGFR expression, in combination with FOLFOX or FOLFIRI. Both cetuximab and panitumumab elicit their effect by binding to the EGFR, thereby preventing ligand binding and disrupting the MAPK signalling pathway (described in Section 2.2.2). Thus, the ERG considers that cetuximab and panitumumab can be included in the same NMA.

2.3.3.3 *Without RAS wild-type*

As noted in Section 2.3.1, subgroup analyses from KEYNOTE-177 indicated that, in those with mutation of KRAS or NRAS, for PFS, the direction of effect favours SoC, albeit that the difference between treatments does not reach statistical significance (HR 1.19, 95% CI: 0.68 to 2.07). Thus, for completeness, the ERG requested subgroup analysis from KEYNOTE-177 in those without RAS wild-type for pembrolizumab versus FOLFOX or FOLFIRI, with or without bevacizumab, from the SoC group, which would equate to FOLFOX, FOLFIRI, and CAPOX.

3 Clinical effectiveness

3.1 Critique of the methods review

The company undertook a broad systematic literature review (SLR) to capture studies of various designs on the efficacy and safety of pembrolizumab in metastatic colorectal cancer (mCRC) that was categorised as high microsatellite instability (MSI-H) or mismatch repair deficient (dMMR). The company carried out their SLR in accordance with guidance from the National Institute for Health and Care Excellence (NICE). Full methods and results of the SLR are reported in Appendix D of the company submission (CS). A summary of the methods, together with the Evidence Review Group's (ERG's) critique of the appropriateness of the methods adopted, is presented in Table 14.

The purpose of the SLR was to identify all relevant studies that could inform the comparison of pembrolizumab with other interventions for untreated MSI-H or dMMR mCRC. The company commented that the initial literature searches and inclusion criteria were designed for a project aimed at a global market, which therefore included interventions not relevant to UK clinical practice. To identify studies relevant to the NICE final scope, the company retrieved a subset of records from their searches focusing on interventions specified by NICE.

As noted earlier, KEYNOTE-177 is the first randomised controlled trial (RCT) evaluating clinical effectiveness of interventions in untreated mCRC that is MSI-H or dMMR, with a comparator group formed of multiple treatments considered to be standard of care (SoC) for mCRC. The company anticipated that the number of studies involving MSI-H or dMMR mCRC would be small. Thus, during the abstract-screening phase, citations meeting all other criteria, but not mentioning MSI or dMMR were included. During the first pass of full-text screening, studies fully matching the target population (or reporting outcomes for a relevant subset) were included, and publications meeting all other criteria but failing to report on MSI-H or dMMR mCRC were excluded but flagged.

As discussed in Section 2.3.3, options available in SoC in KEYNOTE-177 included treatments not available in UK clinical practice. Additionally, CAPOX and panitumumab in combination with FOLFOX or FOLFIRI were not treatment choices in the SoC group, and estimates of clinical effectiveness for pembrolizumab versus the two regimens are derived from a network meta-analysis (NMA). To identify studies to inform the NMA, the company rescreened the flagged studies excluded at an earlier stage. The ERG considers the company's methodological approach to retrieving studies to be

appropriate, but considers that cetuximab with FOLFOX or FOLFIRI should also be included in the NMA (discussed in more detail in Section 1.1.1).

Thirteen studies reported across 31 publications were identified for inclusion in the SLR (CS, Appendix D, Table 9). Of the 13 studies, three involved MSI-H/dMMR mCRC, with only one publication – KEYNOTE-177 – providing direct evidence on the comparative clinical effectiveness of pembrolizumab; the two remaining studies were retrospective analyses of response to treatment (not pembrolizumab) in a subgroup of people with MSI-H mCRC.

Through a non-systematic search of Google Scholar, the ERG identified a systematic review (preprint) of RCTs of systemic first-line treatments for mCRC, which the ERG used as a source to cross-reference the studies included by the company.[28] The ERG notes that the company excluded one RCT comparing CAPOX versus FOLFOX (Ducreux *et al.* 2011[19]; N = 316), with the reason for exclusion given as "wrong study type", that the ERG considers meets criteria for inclusion in the SLR. However, as discussed in Section 2.3.3, for the ITT population, the ERG has assumed equivalent clinical effectiveness for FOLFOX, FOLFIRI and CAPOX, a decision that was directly informed by Ducreux *et al*. 2011 (more detailed discussion available in Section 3.5).

Overall, the ERG considers the company's SLR to be of reasonable quality and likely to have retrieved all studies relevant to pembrolizumab, despite limiting inclusion to English-language publications.

Table 14. Summary of ERG's critique of the methods implemented by the company to identify evidence relevant to the decision problem

Systematic review step	Section of CS in which methods are reported	ERG assessment of robustness of methods
Data sources	Appendix D.1.1.	**The ERG considers the sources and dates searched appropriate.** Databases searched: MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials. Additional sources: Trial registry (clinicaltrials.gov), conference proceedings (ASCO and ESMO). Latest search update: 28 April 2020
Literature searches	Appendix D.1.1. Tables 1–6	**The ERG is satisfied that searches have identified all evidence relevant to the decision problem.** Search strategies combined comprehensive terms for the population (colorectal cancer) and interventions, medical subject headings, and study design filters. Terms for MSI-H or dMMR were not included, and

		studies in MSI-H or dMMR mCRC were identified through screening of abstracts and full texts. The ERG considers the company's approach to be appropriate.
Inclusion criteria	Appendix D.1.1. Tables 7 and 8	**The ERG considers it likely that no relevant evidence was excluded based on the eligibility criteria used.** Inclusion criteria for listed interventions of interest were considerably broader than the NICE final scope. The company highlighted that the literature review was carried out for a global project, which therefore included interventions not relevant to UK clinical practice. The company presented a revised set of inclusion criteria for interventions tailored to the requirements for their submission to NICE. Lists of studies excluded at full-text appraisal, together with reasons for exclusion, are provided. The ERG notes that one RCT excluded during the screening stage could have informed the company's NMA of pembrolizumab versus CAPOX (discussed in the text above). The ERG considers that the omission of the RCT was an oversight and not related to the inclusion criteria. Limited to English-language publications.
Screening and data extraction	Appendix D.1.1	**The ERG considers the methods for screening and data extraction to be robust.** Independent duplicate screening and data extraction by two reviewers against predefined criteria, with a third reviewer consulted when consensus could not be reached; screening results summarised in a PRISMA diagram.
Tool for quality assessment of included study or studies	Appendix D.1.3. Tables 23–26	**The ERG agrees with the quality assessment tool used for KEYNOTE-177 and other RCTs informing the NMA.** The ERG notes that it is unclear if quality assessment was done by one or two reviewers and, if so, whether the assessments were done independently. Detailed reasons were presented in support of the judgement of level of bias for each aspect of trial design covered in the assessment tool.

Abbreviations: ASCO, American Society of Clinical Oncology; CS, company submission; dMMR, mismatch repair deficient; ERG, Evidence Review Group; ESMO, European Society for Medical Oncology; mCRC, metastatic colorectal cancer; MSI-H, high microsatellite instability; NICE, National Institute for Health and Care Excellence; NMA, network meta-analysis; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomised controlled trial.

3.2 Critique of trials of the technology of interest, the company's analysis and interpretation

The ERG reiterates that, for those with MSI-H/dMMR mCRC, it considers the ITT population from KEYNOTE-177 to provide the most robust estimate of comparative clinical effectiveness of pembrolizumab versus FOLFOX, FOLFIRI and CAPOX.[2]

In subsequent sections, the ERG focuses on aspects of trial design, conduct and external validity of KEYNOTE-177 that are of importance to this Single Technology Appraisal (STA). The ERG's critique of the design, conduct and internal validity of KEYNOTE-177 is summarised in Table 15. The ERG agrees with the company's assessment of KEYNOTE-177 as being at overall low risk of bias for analysis of the primary outcome, progression-free survival (PFS), based on the full trial population. The ERG notes that the results for KEYNOTE-177 have been made public through conference abstracts but, at the time of writing, areyet to be published in full in a peer-reviewed journal.

Table 15. Summary of ERG's critique of the design and conduct of KEYNOTE-177,[2] the trial evaluating the technology of interest to the decision problem

Aspect of trial design or conduct	Section of CS providing details on trial characteristic	ERG's critique
Trial conduct		
Randomisation	B.2.3 (page 18)	**Appropriate** Randomisation carried out centrally using an IVRS/IWRS. People randomised 1:1 to pembrolizumab versus SoC. No randomisation strata.
Concealment of treatment allocation	B.2.3 (page 18)	**Appropriate** Treatment allocation concealed through use of IVRS/IWRS at randomisation. For SoC group, treating clinician was required to select treatment from options available before randomisation.
Eligibility criteria	B.2.3 (page 19)	Adults aged ≥18 years with locally confirmed MSI-H or dMMR stage IV CRC. Disease must be measurable at baseline based on RECIST 1.1 as determined by the local site Investigator/radiology assessment. People must also have had a life expectancy of at least 3 months, and an ECOG performance status of 0 or 1 within 10 days prior to treatment initiation.
Baseline characteristics	B.2.3 (page 28)	Baseline characteristics were well balanced between the pembrolizumab and SoC groups in the ITT population. Baseline characteristics from KEYNOTE-177 are available in Appendix 9.1 (Table 67) of the ERG report. The ERG notes that all enrolled in KEYNOTE-177 were classified as MSI-H, which was determined by either PCR or immunohistochemistry. As outlined in Section 2.2.2, immunohistochemistry is used to analyse presence of dMMR rather than MSI-H, but MSI-H and dMMR are synonymous.

Masking appropriate	B.2.3 (page 18)	**Open label design.** However, primary analyses for disease progression-related outcomes (e.g., PFS and ORR) were based on assessment by an independent central imaging vendor who was masked to treatment allocation.
No difference between groups in treatments given, other than intervention versus control	B.2.3 (page 26)	No evidence to suggest a difference between groups in treatments given additional to allocated intervention. A proportion of people, primarily in the SoC group, went on to receive pembrolizumab on disease progression, which likely confounds analysis of long-term outcomes, such as OS (discussed in greater detail in Section 3.2.1).
Dropouts (high drop out and any unexpected imbalance between groups)	B.2.4, Table 15, page 43 Appendix D.1.2 (Figure 37, page 336)	Only two people (1.3%) were lost to follow-up, both from the pembrolizumab group. Low rate of patient withdrawal across KEYNOTE-177 (12 people [4.1%]), but a larger proportion withdrew from the SoC group (1/153 [0.7%] with pembrolizumab versus 11/154 [7.1%] with SoC). Although there is an imbalance between groups in withdrawal from the study, given the small numbers, the ERG considers that the difference does not affect the internal validity of KEYNOTE-177.
Outcomes assessed	B.2.3 (page 27)	**All clinically relevant outcomes reported.** No evidence to suggest that additional outcomes were assessed and not reported. PFS was the primary outcome for the study and was determined by an independent central imaging vendor based on the first radiologic progressive disease. RECIST 1.1 was followed for treatment decisions on trial site until verification of the site assessment of progressive disease by the blinded independent central imaging vendor. Secondary outcomes were: • OS; • ORR. Exploratory outcomes included: • PFS as assessed by irRECIST; • PFS2;[a] • Duration of response;

		• HRQoL (as assessed by EORTC QLQ-C30 and EORTC QLQ-C29).
ITT analysis carried out	B.2.4 (page 32)	**Yes.** ITT population forms the basis for analyses of efficacy.
Subgroup analyses	B.2.4 (page 39)	As outlined in Section 2.3.1, the ERG considers that the company has not fully explored the subgroup of RAS wild-type, which is specified in the final scope issued by NICE as being of interest. Prespecified subgroup analyses were based on: • Age category (≤70 vs >70 years); • Geographic region (Asia vs Western Europe/North America vs Rest of World); • Hepatic or pulmonary metastases versus other metastases; • Recurrent versus newly diagnosed stage IV CRC; • BRAF V600E (wild type vs mutant); • Surgical vs non-surgical subjects. Data are reported for additional *post hoc* subgroups based on: • KRAS/NRAS status (wild type versus mutant); • Gender; • Site of primary tumour (right versus left); • ECOG score (0 versus 1). Strata were not applied at randomisation and so all subgroup analyses break randomisation.
Statistical analysis plan		
Sample size	B.2.4 (page 38)	Based on sample size calculations to detect a difference between pembrolizumab and SoC in the primary outcome of PFS, the company reported that they planned to enrol about 300 people into KEYNOTE-177. Sample size calculations were based on the following assumptions: • PFS follows an exponential distribution with a median of 10 months in the SoC group; • enrolment period of 30 months from first patient randomised and a minimum of 13 months of follow-up after enrolment is completed; • yearly dropout rate of 5%.

Power	B.2.4 (page 38)	The final PFS analysis is planned to be performed at the time of Interim Analysis 2 when approximately 209 PFS events have occurred or 24 months after last the person has been randomised, whichever occurs first. With 209 PFS events, KEYNOTE-177 has ~98% power to detect an HR (pembrolizumab vs SoC) of 0.55 at the 1.25% (one-sided) significance level. Should fewer than 209 events be observed 24 months after randomisation of the last person enrolled, the power of the study will be reduced. For example, if 192 events were observed, then the study has 97% power for PFS.
Analysis for estimate of effect	B.2.4 (page 37)	The company reports that the overall type I error (false positive – reject null hypothesis when it is true) over the primary endpoints of PFS and OS, and the secondary endpoint of ORR is strongly controlled at 2.5% (one-sided), with initially 1.25% allocated to the PFS hypothesis, 1.25% to the OS hypothesis, and 0% to the ORR hypothesis. To account for multiple statistical tests, the company implemented an extension of the graphical method of Maurer and Bretz to strongly control the overall type I error rate for testing of multiple endpoints at the 2.5% 1-sided level (Anderson et al, unpublished data, 2018).[29] Figure 2 shows the initial 1-sided α-allocation for each hypothesis in the ellipse representing the hypotheses.

[a] PFS2 was defined as the time from randomization to disease progression on the next line of therapy, or death from any cause, whichever occurred first.

Abbreviations: CRC, colorectal cancer; CI, confidence interval; CS, company submission; dMMR, mismatch repair deficient; ECOG, Eastern Cooperative Oncology Group; EORTC, European Organisation for Research and Treatment of Cancer; ERG, Evidence Review Group; HR, hazard ratio; HRQoL, health-related quality of life; irRECIST, Immune-related Response Evaluation Criteria in Solid Tumors ; ITT, intention to treat; IVRS/IWRS, interactive voice response system/integrated web response system; MSI-H, high microsatellite instability; PFS, progression-free survival; ORR, overall response rate; OS, overall survival; PCR, polymerase chain reaction; RECIST, Response Evaluation Criteria in Solid Tumours; STA, Single Technology Appraisal.

3.2.1 Aspects of trial design

3.2.1.1 Internal validity

The ERG considers KEYNOTE-177 to be a well-designed and well-conducted RCT. As introduced in Table 15, the ERG notes that OS results from KEYNOTE-177 are likely to be severely confounded, as is widespread in studies evaluating treatments for oncological conditions where crossover is allowed. In the SoC group of KEYNOTE-177, 56 (36.4%) participants crossed over to pembrolizumab treatment on stopping initial treatment, and an additional 35 (22.7%) in the SoC group went on to receive subsequent anti-PD-1/anti-PD-L1 therapy (e.g., nivolumab and avelumab). The company presented adjusted analyses of OS to account for confounding (results available in Section 3.3).

3.2.1.2 External validity

The ERG's clinical experts consider the characteristics of the population comprising KEYNOTE-177 to be generalisable to those with MSI-H/dMMR mCRC and likely to be eligible for treatment with pembrolizumab in England. During clarification, the company informed that 20 people were recruited from the UK, with 10 people allocated to each group of pembrolizumab and SoC.

Although testing for MSI-H or dMMR is recommended at time of diagnosis of CRC,[14] at the time of writing, no treatment options are specific to those with untreated MSI-H/dMMR mCRC, with management of the condition following the same recommendations as mCRC. In England, first-line chemotherapy of mCRC is typically one of FOLFOX, FOLFIRI or CAPOX. For those identified as having RAS wild-type, additional options are cetuximab and panitumumab, both of which are administered in combination with FOLFOX or FOLFIRI. In KEYNOTE-177, panitumumab-based regimens were not available to the SoC group, and assessment of RAS status (wild-type versus mutant) was not carried out for all those enrolled. Thus, in the context of external validity to England, some of those randomised to SoC may not have been treated as they would in clinical practice. As noted in the ERG's critique of how the CS aligns with the decision problem (Section 2.3.3), the ERG considers that comparative clinical effectiveness of pembrolizumab versus comparators of interest by RAS status (wild-type versus not wild-type) warrants further analysis. Finally, as highlighted in Section 2.3.3, the availability of bevacizumab-containing regimens as treatment options for SoC does not align with clinical practice in England. To account for the effect of bevacizumab, the company carried out analyses for PFS and OS that excluded those participants receiving bevacizumab-based regimens (results presented in Section 3.3).

In Section 2.3.2, the ERG noted the potential for confusion around the maximum number of cycles of pembrolizumab. As part of the clarification process, the company confirmed that most people (150/153; 98.0%) randomised to pembrolizumab received no more than 35 cycles, with three people receiving between 36 and 52 cycles (Table 16).

Table 16. Summary of number of cycles of pembrolizumab received by those allocated to pembrolizumab group in KEYNOTE-177 (reproduced from company's response to clarification, question B10)

Measure	Cycles of pembrolizumab
Mean (SD)	19.29 (14.39)
Median (range)	16.00 (1.00 – 50.00)
Number of people receiving up to 35 cycles	150
Number of people receiving between 36 and 52 cycles	3

Abbreviations: SD, standard deviation.

Mean duration of follow-up for KEYNOTE-177 at the time of data cut-off was 25.1 months (standard deviation [SD] 13.4 months; Table 17) and 23.5 months (SD 12.5 months) in the pembrolizumab and SoC groups, respectively. At the time of analysis, 195 PFS events have occurred, which is 14 events fewer than the prespecified 209 events to achieve 98% power (Table 15). The ERG appreciates that PFS presented in the CS is not the final analysis (Table 15). Despite the small deficit in required PFS events, the ERG considers duration of follow-up of KEYNOTE-177 at the time of submission to the STA process to be sufficient for shorter-term outcomes in the ITT population, such as PFS, but not for the longer-term outcome of OS, which is immature at the specified data cut-off.

Table 17. Summary of duration of follow-up in KEYNOTE-177 (adapted from CS, Table 17, page 44)

Follow-up duration (months)[a]	Pembrolizumab (N = 153)	SoC (N= 154)
Median (range)	28.4 (0.2 to 48.3)	27.2 (0.8 to 46.6)
Mean (SD)	25.1 (13.4)	23.5 (12.5)

[a] Follow-up duration is defined as the time from randomisation to the date of death or the database cut-off date if the subject is still alive.

Database cut-off date: 19 Feb 2020.

Abbreviations: SD, standard deviation; SoC, standard of care.

3.3 Clinical effectiveness results

Results are reported, where possible, based on the three populations the ERG considers relevant to the decision problem:

- Population A: all participants (ITT population);
- Population B: subgroup of participants with RAS wild-type;
- Population C: subgroup of participants without RAS wild-type.

As described in Section 2.3.3, in KEYNOTE-177, pembrolizumab was evaluated against a SoC group, in which treatment was chosen by the treating clinician from one of:

- mFOLFOX6;
- mFOLFOX6 with bevacizumab;
- FOLFIRI;
- FOLFIRI with bevacizumab;
- Cetuximab with mFOLFOX6;
- Cetuximab with FOLFIRI.

Comparators of interest to the STA not forming part of SoC are CAPOX for all patients, and panitumumab-containing regimens for those with RAS wild-type. For the purposes of their submission, to generate an estimate of effect for pembrolizumab versus CAPOX and panitumumab-based treatment, the company carried out an NMA in all patients. Cetuximab-based therapy was omitted from the NMA, and the company indicated that cetuximab combination regimens formed part of SoC in KEYNOTE-177. The ERG's comments on the company's NMA, together with suggestions as to how to adapt the NMA to more closely align with the NICE final scope, are discussed in Section 3.5.

Here, the ERG focuses on the direct head-to-head evidence available from KEYNOTE-177. At the time of writing, KEYNOTE-177 is the only RCT reporting results for first-line treatment of MSI-H/dMMR mCRC. For reasons outlined earlier, the ERG considers results derived from the ITT population from KEYNOTE-177 to represent the most robust estimate of effectiveness for pembrolizumab in the specified population versus CAPOX, FOLFOX and FOLFIRI.

When reporting results from KEYNOTE-177, the ERG has assumed that CAPOX, FOLFIRI, and FOLFOX are of equal clinical effectiveness. RCTs in the first-line treatment of mCRC report similar median PFS

and OS for CAPOX, FOLFOX and FOLFIRI (Table 18). Additionally, it should be borne in mind that 70% of people in the SoC group of KEYNOTE-177 received a bevacizumab-based regimen, which has been shown to be more clinically effective than FOLFOX or FOLFIRI alone at improving PFS when used as a first-line treatment in mCRC: RCTs reported an improvement in median PFS of 1.2[30, 31] and 1.4[17, 23] months on addition of bevacizumab to FOLFOX, FOLFIRI or CAPOX. Thus, there is bias in the results from the ITT population that favours SoC and as a consequence provides a conservative estimate of the clinical effectiveness of pembrolizumab versus FOLFOX, FOLFIRI or CAPOX in those with untreated MSI-H/dMMR mCRC. The company reported estimates of comparative clinical effectiveness in the subgroup of those not receiving bevacizumab as part of SoC, which the ERG reports for completeness but advises that, as highlighted by the company, results be interpreted with caution as they are derived from *post hoc* subgroup analyses that are informed by small sample size and in which randomisation has been broken.

Table 18. Summary of findings from studies evaluating CAPOX, FOLFIRI and FOLFOX in the first-line treatment of metastatic colorectal cancer

Study	FOLFIRI		FOLFOX		CAPOX	
	Median PFS (months)	Median OS (months)	Median PFS (months)	Median OS (months)	Median PFS (months)	Median OS (months)
Colucci 2005[18] (N = 360)	7.0[a] (range 1 to 47)	14.0 (range 1 to 48)	7.0[a] (range 1 to 32)	15.0 (range 1 to 43)	–	–
Porschen 2007[21] (N = 476)	–	–	8.0	18.8	7.1	16.8
Hochster 2008[20] (N = 97)	–	–	8.7[a]	19.2	5.9[a]	17.2
Cassidy 2011[17] (N = 1,334)	–	–	NR	18.9	NR	19.0
Ducreux 2011[19] (N = 316)			9.3	20.5	8.8	19.9

[a] Study reported time to progression rather than PFS.

Reported studies were identified from a systematic review of RCTs evaluating first-line systemic treatments for metastatic colorectal cancer.[28]

Abbreviations: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin; OS, overall survival; PFS, progression-free survival.

3.3.1 Progression-free survival

In the ITT population, based on assessment of PFS by central imaging, pembrolizumab was associated with a 40% reduction in the risk of disease progression or death compared with SoC, with the difference between treatment groups reaching statistical significance (HR 0.60, 95% CI: 0.45 to 0.80; p=0.0002). Median PFS was reported to be 16.5 months and 8.2 months in the pembrolizumab and SoC groups, respectively (Table 19). The ERG notes that the median PFS in the SoC group is within the range of median PFS (5.9 to 9.3 months) in the RCTs evaluating CAPOX, FOLFIRI and FOLFOX in untreated mCRC (Table 18). The Kaplan–Meier (KM) plots for PFS for the ITT population from KEYNOTE-177 indicate an initial crossing of the curves at about 6 months with a subsequent separation between the curves for the pembrolizumab and SoC groups (Figure 1).

For the subgroup of people not receiving bevacizumab, estimates of effect for pembrolizumab versus SoC [redacted], pembrolizumab is [redacted] Table 18). The company selected 52 people allocated to pembrolizumab to form the pembrolizumab group not receiving bevacizumab, indicating that those selected were people who would not have received bevacizumab if they had been randomised to SoC (chemotherapy regimen for SoC was specified before randomisation)..

The company indicated that sample size required to give ~98% power to detect an HR (pembrolizumab vs SoC) of 0.55 is 209 PFS events. Reported PFS for the ITT population is based on 195 events. As a statistically significant result for PFS has been generated in the ITT population, with narrow CI and a small p value, the ERG considers the results for this population to be robust. The clinical effect of pembrolizumab on PFS compared with SoC is predominantly consistent across subgroups, although some differences are no longer statistically significant, for example, ECOG score of 1, and in those with left-sided tumours (Figure 16, Appendix 9.2).

In the CS, Figure 1 could be taken to indicate that disease progression was determined by irRECIST in the pembrolizumab arm, but centrally verified by RECIST 1.1 in the SoC group. During clarification, the company expanded on the use of irRECIST, indicating that irRECIST and RECIST 1.1 provide the same response assessment until disease progression, and that irRECIST was used to make treatment decisions beyond progression as determined by RECIST 1.1. PFS per irRECIST by central imaging

vendor is an exploratory objective of KEYNOTE-177, and was not analysed at the second interim analysis. Exploratory analysis of irRECIST will be carried out at the final analysis for KEYNOTE-177.

Data are available for PFS2 in the ITT population in Appendix L of the CS. PFS2 was defined as the time from randomisation to disease progression on the next line of therapy, or death from any cause, whichever occurs first. The ERG considers PFS2 data to be immature. PFS2 does not inform the analysis of cost effectiveness. In brief, median PFS2 was 23.5 months (95% CI: 16.6 to 32.6) for SoC but was not reached in the pembrolizumab group (HR 0.63, 95% CI: 0.45 to 0.88). Bearing in mind the immaturity of PFS2, the ERG notes that PFS and PFS2 for pembrolizumab versus SoC are similar (PFS HR 0.60, 95% CI: 0.45 to 0.80 vs PFS2 HR 0.63, 95% CI: 0.45 to 0.88).

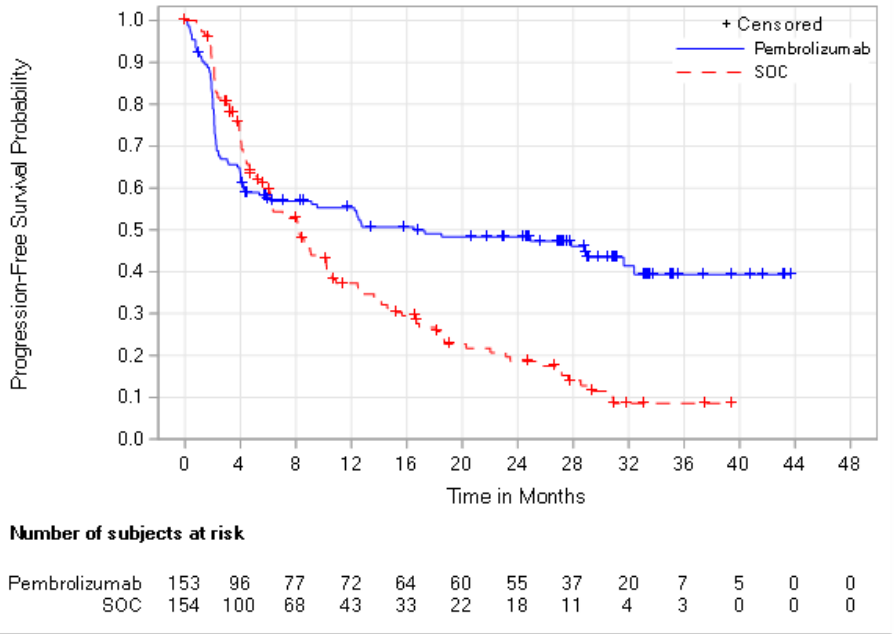
Table 19. Summary of event and censoring description for PFS by central imaging vendor per RECIST 1.1 (ITT population) (reproduced from CS, Tables 25–27, pages 60–62)

Event	ITT population		ITT population excluding those receiving bevacizumab	
	Pembrolizumab (N = 153)	SoC (N = 154)	Pembrolizumab (N =52)	SoC (N = 47)
PFS rate at time point, % (95% CI)				
6 months	57.6 (49.3 to 65.0)	59.7 (51.1 to 67.3)	[redacted]	[redacted]
9 months	56.8 (48.5 to 64.3)	45.5 (36.9 to 53.7)	[redacted]	[redacted]
12 months	55.3 (47.0 to 62.9)	37.3 (29.0 to 45.5)	[redacted]	[redacted]
18 months	49.1 (40.7 to 57.0)	26.7 (19.2 to 34.7)	[redacted]	[redacted]
24 months	48.3 (39.9 to 56.2)	18.6 (12.1 to 26.3)	[redacted]	[redacted]
Number of people with a PFS event, n (%)	82 (53.6)	113 (73.4)	[redacted]	[redacted]
Documented progression	65 (42.5)	86 (55.8)	[redacted]	[redacted]
Death	17 (11.1)	27 (17.5)	[redacted]	[redacted]
Median PFS, months (95% CI)	16.5 (5.4 to 32.4)	8.2 (6.1 to 10.2)	[redacted]	[redacted]
Subjects censored	71 (46.4)	41 (26.6)	[redacted]	[redacted]
Curative-intent surgery	12 (7.8)	12 (7.8)	[redacted]	[redacted]

New anti-cancer therapy	5 (3.3)	15 (9.7)	[redacted]	[redacted]
Last radiologic assessment showing no progression	53 (34.6)	10 (6.5)	[redacted]	[redacted]
No adequate post-baseline imaging assessment	1 (0.7)	4 (2.6)	[redacted]	[redacted]

Abbreviations: CI, confidence interval; CS, company submission; ITT, intention to treat; PFS, progression-free survival.

Figure 1. Kaplan–Meier estimates of PFS based on central imaging vendor per RECIST 1.1 (ITT population; database cut-off date: 19 Feb 2020) (reproduced from CS, Figure 9, page 61)



In terms of the two remaining populations that the ERG considers relevant based on RAS status (wild-type versus non wild-type), compared with SoC, the clinical benefit of pembrolizumab is maintained in those with RAS wild-type (HR 0.44, 95% CI: 0.29 to 0.67). CAPOX, FOLFIRI and FOLFOX are treatment options in England for RAS wild-type mCRC, and again the ERG considers that SoC in KEYNOTE-177 provides the most robust estimate of effect for pembrolizumab versus these three regimens. Also, as SoC includes a proportion of people who have received bevacizumab-based regimens, the bias in the analysis favours SoC and the estimate of effect for pembrolizumab is conservative.

Treatment comprising cetuximab, but not panitumumab, was available as an option for the treating clinician for SoC, but only a small proportion of people in the ITT population (11.2%; Table 13) received cetuximab in combination with FOLFOX or FOLFIRI. As with bevacizumab, RCTs have reported that addition of cetuximab to FOLFOX or FOLFIRI in those with RAS wild-type is associated with an improvement in PFS over FOLFOX or FOLFIRI alone, which introduces additional bias in favour of SoC. The CRYSTAL[26] and OPUS[24, 32] RCTs reported an improvement in median PFS of 1.5 months and 6.2 months with cetuximab-combination regimen compared with FOLFIRI or FOLFOX alone, respectively.

RAS status does not seem to have been a determining factor for treatment choice in KEYNOTE-177, as assessment of RAS status was not carried out for all those enrolled in the study. Thus, treatment of participants with RAS wild-type MSI-H/dMMR mCRC in the SoC group is likely not entirely representative of clinical practice in England. To substantiate the observed effect for pembrolizumab on PFS noted in the subgroup from KEYNOTE-177, for those with RAS wild-type MSI-H/dMMR mCRC, the ERG considers it important to provide estimates of effect versus those treatments that have been found to be clinically effective in this group, albeit in mCRC and not MSI-H/dMMR mCRC, that is, cetuximab and panitumumab in combination with FOLFOX or FOLFIRI. The ERG's critique of the company's NMA to generate an estimate of effect for pembrolizumab versus panitumumab combination treatment is provided in Section 3.5.

The subgroup analysis of those without RAS wild-type, that is, those with a RAS mutation, provides a markedly different result for the effect of pembrolizumab on PFS compared with SoC. The direction of effect for PFS no longer favours pembrolizumab, with pembrolizumab potentially associated with a 19% increase in risk of disease progression or death compared with SoC, albeit that the difference between the treatments does not reach statistical significance (HR 1.19, 95% CI: 0.68 to 2.07). The company comments that the results of the analyses should be interpreted with caution, given the small sample size, but does not discuss the potential implications or clinical rationale for the result. During clarification, the ERG asked that the company carry out subgroup analysis for those without RAS wild-type comparing pembrolizumab versus SoC, which the ERG defined as FOLFOX or FOLFIRI with or without bevacizumab. The company declined the ERG's request, commenting that such an analysis would be inappropriate.

Relevant comparators for those with mutations in RAS are CAPOX, FOLFOX and FOLFIRI. As with other analyses, the bias is against pembrolizumab, due to a large proportion of people having

received bevacizumab. However, as also with other analyses, the extent of bias is difficult to quantify. During clarification, the ERG requested that the company carry out an analysis of those without RAS wild-type, focusing on FOLFOX and FOLFIRI, with or without bevacizumab, as the comparator. The company commented that it would be inappropriate to carry out such an analysis as any analysis would be considerably under-powered because of the low number of events, which would likely produce false-negative results. Additionally, randomisation would be broken in the subgroups, and differences in patient baseline characteristics are noted between treatment arms that could confound the results of the analyses, rendering them unreliable. The ERG appreciates that subgroup analyses are less robust and are hypothesis generating, but considers that the difference in results for PFS in the subgroup of those with mutation of KRAS or NRAS is prominent amongst the other subgroup analyses, and the clinical effectiveness of pembrolizumab in this subgroup warrants further investigation. Given that OS data are immature, the ERG considers that additional data could be accrued on PFS in the subgroups of those with RAS mutations during longer-term follow-up.

3.3.2 Overall survival

OS data for KEYNOTE-177 are immature are the time of writing. Analysis of OS is based on 125 events and median OS has yet to be reached in the pembrolizumab group. Additionally, OS is likely to be severely confounded, with 59% of participants allocated to SoC receiving pembrolizumab (36.4%) or a PD-1/PD-L1 inhibitor (22.7%) on cessation of initial treatment, primarily because of disease progression. The ERG appreciates that, due to the potential confounding, the OS advantage of pembrolizumab, if any, will be underestimated but considers the data too immature to draw reliable conclusions.

The company attempted to mitigate against confounding through sensitivity analyses of OS, with adjustment for crossover carried out based on three models recommended by the NICE Decision Support Unit (DSU) in the Technical Support Document (TSD) 16:[33]

- simplified 2-stage model;
- rank preserving structural failure time (RPSFT) model;
- inverse probability of censoring weighting (IPCW) model.

In their analysis of cost-effectiveness, the company supplied a state transition model (STM) as their primary model, and a partitioned survival model (PSM), which was used as a validation tool. Adjusted OS data were implemented only in the PSM. Within the STM, the company assumes no

additional benefit in OS for patients in the post-progression state, which the ERG considers a reasonable assumption that removes the uncertainty associated with the estimation of OS from the trial (assumption is discussed in more detail in Section 4.2.4.1). As adjusted OS data are only applicable to the PSM, the ERG does not discuss further the estimates for adjusted OS.

In the ITT population, there were 56 (36.6%) events in the pembrolizumab group compared with 69 (44.8%) events in the SoC group (Table 20). The direction of effect favours pembrolizumab but the difference between treatments does not reach statistical significance (HR 0.77, 95% CI: 0.54 to 1.09; p = 0.0694). The KM curves for OS for the ITT population depict a crossing of the curves at about 6 months with a subsequent separation between the curves for the pembrolizumab and SoC groups (Figure 2). At the time of analysis, both the pembrolizumab and SoC curve appear to reach a plateau at about 35 months. The ERG considers the plateaus observed in the OS curves for pembrolizumab and SoC to be a common artefact observed in the "tails" of KM curves, and is likely to be due to small patient numbers, few events, and heavy censoring.

Estimates of effect for pembrolizumab were consistent across pre-specified subgroups, [redacted] (Figure 17, Appendix 9.2). The ERG [redacted].

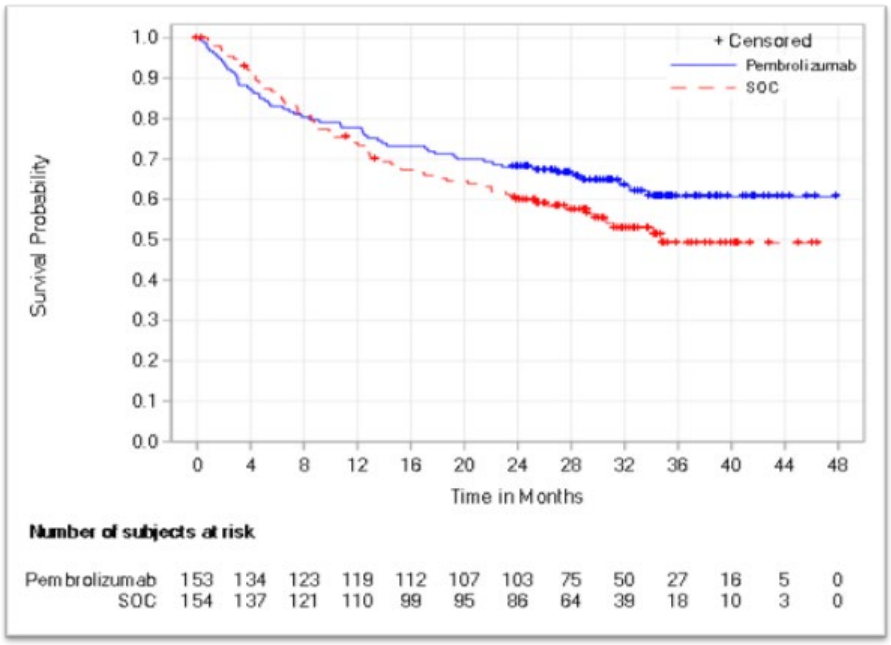
Table 20. Summary of event for OS (ITT population) (reproduced from CS, Tables 20 and 21, pages 46 and 47)

Event	ITT population	
	Pembrolizumab (N = 153)	SoC (N = 154)
OS rate at time point, % (95% CI)		
6 months	83.0 (76.1 to 88.1)	86.0 (79.4 to 90.7)
9 months	79.7 (72.5 to 85.3)	78.7 (71.3 to 84.4)
12 months	77.8 (70.3 to 83.6)	74.0 (66.2 to 80.3)
18 months	71.2 (63.4 to 77.7)	65.9 (57.7 to 72.9)
24 months	68.0 (59.9 to 74.7)	59.8 (51.5 to 67.2)

Number of people with a OS event, n (%)	56 (36.6)	69 (44.8)
Median OS, months (95% CI)	NR (NR to NR)	34.8 (26.3 to NR)

Abbreviations: CI, confidence interval; CS, company submission; ITT, intention to treat; NR, not reached; OS, overall survival.

Figure 2. Kaplan–Meier estimates of OS (ITT population; database cut-off date: 19 Feb 2020) (reproduced from CS, Figure 4, page 47)



3.3.3 Response outcomes

Within the ITT population of KEYNOTE-177, a larger proportion of people in the pembrolizumab group achieved a complete response (CR) to treatment as assessed by central imaging vendor per RECIST 1.1 (11.1% with pembrolizumab versus 3.9% with SoC; Table 21). Best overall response rate (ORR), which combines CR and partial response (PR), was also higher in the pembrolizumab group (43.8% with pembrolizumab versus 33.1% with SoC). The absolute difference in ORR between pembrolizumab and SoC was 10.7% higher with pembrolizumab, a difference that approaches statistical significance (95% CI: –0.2% to +21.3%; one sided p value of 0.0275).

Mean and median time to response were similar for pembrolizumab and SoC, but duration of response was considerably longer with pembrolizumab:

- Median time to response: 2.2 (range 1.8 to 18.8) months with pembrolizumab versus 2.1 (range 1.7 to 24.9) months with SoC;
- Mean time to response: 4.0 (SD 3.7) months with pembrolizumab versus 3.6 (SD 4.1) months with SoC;
- Proportion of participants with duration of response at ≥12 months: 85.1% with pembrolizumab versus 43.8% with SoC;
- Proportion of participants with duration of response at ≥24 months: 82.6% with pembrolizumab versus 35.3% with SoC.

Table 21. Summary of best overall response by central imaging vendor per RECIST 1.1 (ITT population; database cut-off date of 19 Feb 2020) (reproduced from CS, Table 29, page 64)

Response evaluation	Pembrolizumab (N = 153)			SoC (N = 154)		
	n	%	95% CI[a]	n	%	95% CI[a]
CR	17	11.1	(6.6 to 17.2)	6	3.9	(1.4 to 8.3)
PR	50	32.7	(25.3 to 40.7)	45	29.2	(22.2 to 37.1)
Stable disease	32	20.9	(14.8 to 28.2)	65	42.2	(34.3 to 50.4)
PD	45	29.4	(22.3 to 37.3)	19	12.3	(7.6 to 18.6)
Objective response (CR+PR)	67	43.8	(35.8 to 52.0)	51	33.1	(25.8 to 41.1)
Disease control (CR+PR+stable disease)	99	64.7	(56.6 to 72.3)	116	75.3	(67.7 to 81.9)
Not evaluable	3	2.0	(0.4 to 5.6)	2	1.3	(0.2 to 4.6)
No assessment[b]	6	3.9	(1.5 to 8.3)	17	11.0	(6.6 to 17.1)

Only confirmed responses are included.

[a] Based on binomial exact confidence interval method.

[b] No Assessment: subject had no post-baseline imaging.

Abbreviations: CI, confidence interval; CR, complete response; CS, company submission; PD, progressive disease; PR, partial response; SoC, standard of care.

3.3.4 Health-related quality of life

Health-related quality of life (HRQoL) was assessed using three tools, EQ-5D (3L), EORTC QLQ-C30 and EORTC QLQ-C29. The EQ-5D tool is used to evaluate general health status, whereas EORTC questionnaires are the instruments most frequently used to measure QoL in people with cancer. Assessments were based on responses from the full analysis set rather than the ITT population and captured changes in QoL at 18 weeks of follow-up.

Results from EQ-5D indicate that, compared with SoC, pembrolizumab is associated with improvements in utility score and on the visual analogue scale (Table 22), although the improvement over SoC in utility score is reported by the company not to be clinically meaningful. Responses to EORTC QLQ-C30 identified a small, but reported to be a clinically important, improvement in QoL with pembrolizumab over SoC (Table 22).

Table 22. Summary of results from assessment of health-related quality of life from KEYNOTE-177 (reproduced from CS, Tables 32 and 33, page 69, and from Table 110 from Appendix L, page 537)

Treatment	Baseline		Week 18		Change from baseline at week 18	
	N	Mean (SD)	N	Mean (SD)	N	LS mean (95% CI)[a]
EQ-5D utility score						
Pembrolizumab	142	0.77 (0.195)	102	0.84 (0.175)	151	0.04 (0.00 to 0.08)
SoC	133	0.75 (0.197)	82	0.77 (0.199)	141	–0.01 (–0.05 to 0.02)
Pembrolizumab versus SoC					Difference in LS means (95% CI): 0.05 (0.00 to 0.10) p = 0.0311	
EQ-5D VAS						
Pembrolizumab	142	70.12 (18.862)	102	76.86 (17.916)	151	4.50 (1.16 to 7.83)
SoC	133	70.83 (19.788)	82	70.76 (18.198)	141	–2.88 (–6.46 to 0.69)
Pembrolizumab versus SoC					Difference in LS means (95% CI): 7.38 (2.82 to 11.93) p = 0.0016	
EORTC QLQ-C30 Global Health Status/QoL scales						
Pembrolizumab	141	66.19 (21.029)	102	72.14 (20.530)	151	3.33 (–0.05 to 6.72)

SoC	131	66.60 (20.737)	82	62.60 (17.677)	141	–5.63 (–9.32 to –1.94)
Pembrolizumab versus SoC					Difference in LS means (95% CI): 8.96 (4.24 to 13.69) p = 0.0002	

[a] Based on cLDA model with the PRO scores as the response variable, and treatment by study visit interaction as covariates.

Abbreviations: CI, confidence interval; CS, company submission; EORTC, European Organisation for Research and Treatment of Cancer; LS, least squares; QoL, quality of life; SD, standard deviation; SoC, standard of care; VAS, visual analogue scale.

3.3.5 Adverse effects

Adverse events (AEs) in KEYNOTE-177 were coded according to the Medical Dictionary for Regulatory Activities (MedDRA) and grouped system organ class, and were graded by investigators using the Common Terminology Criteria for Adverse Events (CTCAE), version 4.0. The All Subjects as Treated (ASaT) population was used for the analyses of safety data in KEYNOTE-177. The ASaT population comprised all randomised patients who received at least one dose of study medication, with patients included in the treatment group corresponding to the study medication they received rather than the randomised treatment group, unless they only received the incorrect study medication for one cycle but received the correct treatment for all other cycles. The company provided a summary of the AEs data from KEYNOTE-177 in the CS with further details in Appendix F of the CS.

The median duration of study drug exposure was longer in the pembrolizumab group (11.1 months) compared with the SoC group (5.7 months), but the proportion of participants with at least one AE in those who received pembrolizumab was slightly smaller (97.4%) than in those allocated to SoC (99.3%). In addition, compared with SoC, pembrolizumab was associated with a lower frequency of Grade 3 to 5 AEs (56.2% vs 77.6%), serious adverse events (SAEs; 40.5% vs 52.4%), and investigator assessed drug-related AEs (79.7% vs 98.6%; Table 23) compared with SoC. Only one person from KEYNOTE-177, who was allocated to the SoC group, was categorised as having died as a result of an AE. The proportion of participants who experienced an AE resulting in treatment discontinuation was slightly larger in the pembrolizumab group compared with the SoC group (13.7% vs 11.9%).

The eleven most frequently experienced treatment-related AEs in the pembrolizumab group were: diarrhoea (24.8%); fatigue (20.9%); pruritis (13.7%); nausea (12.4%); rash (11.1%); increased aspartate aminotransferase (11.1%); arthralgia (10.5%); hypothyroidism (10.5%), increased alanine aminotransferase (9.8%); increased blood alkaline phosphatase (7.8%); and decreased appetite

(7.8%). As already noted, the frequency of drug-related AEs was higher with SoC, and the AE profile in the SoC group differed from that of pembrolizumab, with commonly occurring AEs being: diarrhoea (52.4%); nausea (55.2%); fatigue (44.1%); decreased appetite (34.3%); stomatitis (30.1%); vomiting (28.0%); decreased neutrophil count (23.1%); neutropenia (21.0%); and peripheral sensory neuropathy (20.3%). The ERG's clinical experts reported that the AEs observed in KEYNOTE-177 were as expected, and did not have any concerns relating to the AEs experienced by those randomised to pembrolizumab.

The company also conducted an exposure-adjusted analysis of AEs that showed similar results to the data for unadjusted AEs. The ERG notes that the company used the AE data from KEYNOTE-177 for Grade 3+ AEs that occurred in at least 5% of patients to inform the AEs in the economic model. The AE data and their implementation in the economic model are discussed in Section 4.2.6.

Table 23. Summary of adverse events (ASaT population)

Adverse effect	Pembrolizumab (N = 153)		SoC (N = 143)	
	n	(%)	n	(%)
One or more AE	149	(97.4)	142	(99.3)
No AE	4	(2.6)	1	(0.7)
Drug-related[a] AEs	122	(79.7)	141	(98.6)
Toxicity Grade 3–5 AEs	86	(56.2)	111	(77.6)
Toxicity Grade 3-5 drug-related AEs	33	(21.6)	94	(65.7)
Serious AEs	62	(40.5)	75	(52.4)
Serious drug-related AEs	25	(16.3)	41	(28.7)
Death	6	(3.9)	7	(4.9)
Death due to a drug-related AE	0	(0.0)	1	(0.7)
Discontinued[b] drug due to an AE	21	(13.7)	17	(11.9)

Discontinued[b] drug due to a drug-related AE	15	(9.8)	8	(5.6)
Discontinued[b] drug due to a serious AE	12	(7.8)	13	(9.1)
Discontinued[‡] drug due to a serious drug-related AE	7	(4.6)	5	(3.5)

[a] Determined by the investigator to be related to the drug.

[b] All study medications withdrawn.

Notes: Grades are based on NCI CTCAE version 4.03. MedDRA preferred terms "Neoplasm progression", "Malignant neoplasm progression" and "Disease progression" not related to the drug are excluded. Non-serious adverse events up to 30 days of last dose and serious adverse events up to 90 days of last dose are included. Database Cutoff Date: 19 Feb 2020.

Abbreviations: AE, adverse event; ASaT, All Subjects as Treated; CTCAE, Common Terminology Criteria for Adverse Events; MedDRA, Medical Dictionary for Regulatory Activities; SoC, standard of care.

3.3.5.1 *Summary of adverse events of special interest*

A higher incidence of AEs of special interest (AEOSIs) was recorded with pembrolizumab than with SoC (47 [30.7%] vs 18 [12.6%]). Serious AEOSIs were experienced by 10.5% of participants allocated to pembrolizumab compared with 0.7% of participants randomised to SoC. However, the number of AEOSIs that led to discontinuation of pembrolizumab treatment was low (6.5 %). The company reported that most AEOSIs were manageable with systemic corticosteroids, supportive care, and dose interruption. No fatal AEOSIs was reported.

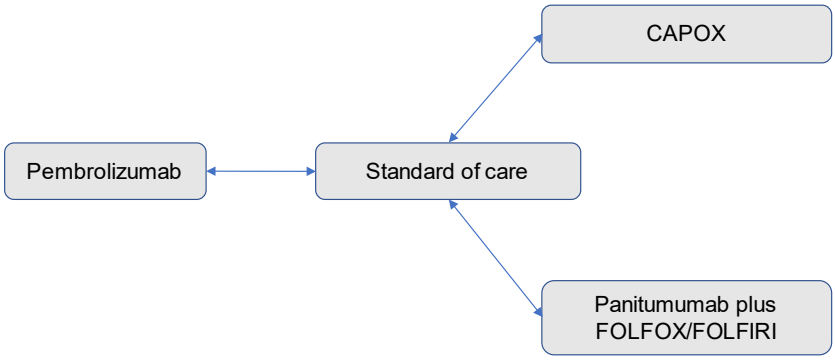
The most common AEOSIs categories, with ≥2% incidence, in the pembrolizumab group were: hypothyroidism; hyperthyroidism; colitis; pneumonitis; adrenal insufficiency; hepatitis; and infusion reactions. Of particular note, the incidence of hypothyroidism was much higher in the pembrolizumab group (19 participants [12.4%]) than in the SoC group (3 participants [2.1%]). However, all hypothyroidism AEOSIs occurring with pembrolizumab group were of Grade 1 or 2 in severity. Colitis AEOSIs were more frequent in the pembrolizumab group than the SoC group (6.5% vs 0%). Colitis AEOSIs experienced with pembrolizumab comprised five participants with Grade 1 or 2 events, three with a Grade 3 event, and two with a Grade 4 event. Five participants had colitis events deemed to be drug-related SAEs, and a colitis AEOSI led to four participants discontinuing treatment, with two events categorised as serious AEs. Of the 65 AEOSIs episodes experienced during treatment with pembrolizumab, 23.1% were initially treated with high-dose corticosteroids. Further details on the breakdown of the AEOSIs are provided in the CS (Table 46 and Appendix F).

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

NMAs were used to generate estimates of comparative clinical effectiveness for PFS and OS for pembrolizumab versus comparators not included in SoC of KEYNOTE-177 but listed in the final scope issued by NICE, that is, CAPOX and panitumumab in combination with FOLFIRI or FOLFOX. The company decided against including cetuximab in combination with FOLFIRI or FOLFOX as a separate comparator as cetuximab-based therapy was an option in SoC of KEYNTOE-177, a decision with which the ERG does not agree (more detail available in Section 3.5).

Using the same SLR carried out to identify studies providing head-to-head evidence on pembrolizumab in MSI-H/dMMR mCRC, five studies, including KEYNOTE-177, were identified as relevant to the NMA.[2, 17, 20, 21, 25] The company's NMA was informed by a population of "all patients", which, based on RCTs informing the network, encompasses those with mCRC, with RAS wild-type mCRC, and with MSI-H/dMMR mCRC, and, as such with overlap across populations because MSI-H/dMMR status is unknown for studies in mCRC and RAS wild-type. As the population is based on all patients, an approach with the ERG disagrees (more detail in Section 3.5.1.2), only one network of interventions is required (Figure 3). Not all identified studies provided data on PFS or OS in a format that met the requirements of the company and, as a result, not all RCTs inform both networks.

Figure 3. Network diagram for network meta-analyses of PFS and OS



Within the Appendices forming part of the CS, for the five studies included in the NMA, the company provides:

- detailed descriptions of study characteristics (Appendix D1.1, Tables 13 to 19);
- tabulated data (Appendix D1.1, Table 20);
- quality assessments (Appendix D1.3, Tables 23 to 26).

As highlighted by the company, differences are noted across FOLFOX regimens, with RCTs evaluating FOLFOX4, FOLFOX6 and mFOLFOX6. The ERG's clinical experts advised that the various FOLFOX regimens are of similar clinical effectiveness and it is appropriate to combine data from all studies evaluating FOLFOX.

The five RCTs included in the company's networks are:

- Pembrolizumab versus SoC:
    - KEYNOTE-177, contributes to PFS and OS.[2]
- CAPOX versus SoC:
    - NO16966, contributes to OS;[17]
    - Porschen 2007, contributes to PFS and OS;[21]
    - TREE-1, contributes to OS.[20]
- Panitumumab plus FOLFOX/FOLFIRI versus SoC:
    - PRIME, contributes to PFS and OS.[25]

3.4.1 CAPOX

The ERG considers the three identified RCTs evaluating CAPOX to be relevant to the decision problem as set out by the company. The ERG identified an additional RCT comparing CAPOX versus FOLFOX6 that met the company's SLR inclusion criteria and provided data to inform both the PFS and OS networks.[19] However, as discussed in Section 3.3, the ERG has assumed CAPOX to be of similar clinical effectiveness to FOLFIRI and FOLFOX in the first-line treatment of mCRC, and so considers it appropriate to exclude CAPOX from the NMA: in the company's response to clarification, base case results for CAPOX are based on equivalency with SoC, from which the ERG has inferred that the company considers the ERG's assumption of similar clinical effectiveness for CAPOX, FOLFIRI and FOLFOX to be appropriate. Here, the ERG does not describe in full or critique further the four RCTs evaluating CAPOX versus SoC. Should the ERG have considered it necessary to generate estimates of effect for pembrolizumab versus CAPOX, to minimise clinical heterogeneity within the networks, the ERG would have requested that the networks be separated by RAS wild-type (more detail available in Section 3.5).

3.4.2 Panitumumab in combination with FOLFOX or FOLFIRI

For panitumumab-combination treatment, the company identified one RCT — PRIME — that compared panitumumab plus FOLFOX4 versus FOLFOX4 alone.[25] PRIME is an open-label Phase III

study that enrolled adults aged ≥18 years with an ECOG score of ≤2 and first occurrence of mCRC, and had a primary end point of PFS.[25] As noted by the ERG undertaking the assessment that informed TA439, estimates of effectiveness of panitumumab plus FOLFOX4 versus FOLFOX4 alone in RAS wild-type tumours were derived from a subgroup of the full trial population.[13]

During the design and conduct of PRIME, research emerged on the impact of KRAS and NRAS mutations in reducing the clinical effectiveness of EGFR inhibitors, such as panitumumab. In the case of PRIME, extended RAS subgroup analysis was noted alongside a protocol amendment to restrict analysis of the ITT population to compare PFS and OS according to KRAS status.[13] In the first-line setting, panitumumab is indicated for the treatment of adults with RAS wild-type mCRC in combination with FOLFOX or FOLFIRI.[34]

Baseline characteristics of the RAS wild-type subgroup from PRIME are comparable with those of the ITT population of KEYNOTE-177, in terms of age and ECOG score. Median age of participants was 61 years in PRIME compared with 63 years in KEYNOTE-177. A small proportion of participants enrolled in PRIME had ECOG score of 2 at baseline (6%), whereas KEYNOTE-177 specified an ECOG score of 0 or 1 as an inclusion criteria: an ECOG score of 2 indicates a poorer health status than score of 0 or 1. Data on site of primary tumour, in terms of left or right location, are not available for PRIME.

In its critique of the quality assessment of PRIME, the ERG on project TA439 considered that randomisation allocation was adequate for the full population of PRIME but that relevant data based on RAS wild-type were derived from a subgroup. However, the ERG went on to comment that they considered the biological rationale for the re-evaluation by RAS status to support the validity of the estimates of effect. As an open-label study, there is risk of bias in assessment of subjective outcomes, such as PFS. In PRIME, an independent monitoring committee reviewed interim analyses of safety and one descriptive interim analysis of PFS.[13] Uncertainties associated with effect estimates from subgroup analyses potentially arise from ascertainment bias and selection bias, but, in TA439, the ERG's assessment was that the potential for the two types of bias was minimal. Lack of statistical power to identify a true difference in effect between two treatments is also associated with use of subgroup data. In TA439, the ERG considered that the underlying rationale of tumour biology and consistency of effect estimates for panitumumab supported the validity of estimates of effect.[13] The ERG for this STA agrees with the points raised in TA439 and considers estimates of effect for

panitumumab plus FOLFOX4 derived from the subgroup analysis based on RAS wild-type from PRIME to provide the best estimate for panitumumab-based treatment in the population in which panitumumab has been shown to have clinical benefit, and are relevant to the decision problem.

3.4.3 Cetuximab in combination with FOLFIRI or FOLFOX

TA439 described two RCTs evaluating cetuximab in combination with FOLFIRI or FOLFOX – CRYSTAL[26] and OPUS in first-line treatment of mCRC.[24, 32] In contrast to PRIME, research indicating the negative impact of RAS mutations on the clinical effectiveness of EGFR inhibitors was not available at the protocol development stage of CRYSTAL and OPUS. Thus, the re-analyses of the ITT populations of the two studies based on RAS status are retrospective in nature. Nevertheless, the analyses formed the basis for revision of the licensed population for cetuximab.[35] Cetuximab is indicated for the treatment of EGFR-expressing, RAS wild-type mCRC in combination with irinotecan-based chemotherapy (FOLFIRI) and, in first-line, in combination with FOLFOX.

CRYSTAL and OPUS are Phase III randomised, open-label studies.[13] Primary outcome assessed in OPUS was the proportion of participants who had an objective response rate, with tumour response assessed by an independent review committee according to modified World Health Organisation criteria. By contrast, primary endpoint in CRYSTAL was PFS. Baseline characteristics in the ITT populations of the two studies were balanced between treatment groups and were similar. The ERG for project TA439 noted that, because relevant data are derived from retrospective subgroup analyses, and randomisation was not stratified by RAS status, randomisation has been broken in all comparisons, increasing the risk of selection bias.[13] However, the ERG stated that, from the evidence provided during the project (published and unpublished), they were able to verify that the treatment groups in CRYSTAL and OPUS were adequately similar at baseline on a range of prognostic factors for the RAS wild-type population. Characteristics were similar across the ITT and RAS wild-type populations, suggesting a low risk of selection bias in the RAS tested trial population.[13] The internal validity of OPUS and CRYSTAL is affected by the same issues associated with PRIME, in terms of data are derived from *post hoc* subgroup analyses, with the resulting lack of statistical power. However, as with PRIME, the biological rationale for carrying out the retrospective analyses, together with consistency of estimates of effect for cetuximab plus FOLFIRI or FOLFOX, support the validity of the analyses.

Subsequent to the publication of TA439, results from the TAILOR RCT became available.[27] TAILOR is an open-label Phase III randomised study set in China and designed to evaluate prospectively the

efficacy and tolerability cetuximab in combination with FOLFOX4 versus FOLFOX4 alone in the first-line treatment of RAS wild-type mCRC. No strata was applied at randomisation. Baseline characteristics were balanced between treatment groups. However, the study did not stipulate detectable tumour EGFR expression as an inclusion criteria, which is not in line with the marketing authorisation for cetuximab. The primary endpoint in TAILOR is PFS as assessed by RECIST 1.0, with a blinded review of imaging and clinical data carried out by an independent review committee.[27] Analysis was based on a modified ITT population, comprising 393 out of 397 randomised. As a prospective study, the risk of selection bias is minimal compared with PRIME, CRYSTAL, and OPUS.

As with PRIME, baseline characteristics for CRYSTAL, OPUS and TAILOR are not available by MSI-H/dMMR status. The ERG acknowledges that there is likely clinical heterogeneity between the population enrolled in KEYNOTE-177 and those of the three studies evaluating cetuximab-combination treatments, but considers that data from CRYSTAL, OPUS and TAILOR provide the best estimate of effect for cetuximab-based regimens in the population in which cetuximab has been shown to have clinical benefit.

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

3.5.1 Methods of indirect comparison

3.5.1.1 *Methods and assumptions*

Section D1.1 of the Appendices component of the CS provides an overview of the methods followed by the company to produce estimates of effect for pembrolizumab versus CAPOX and panitumumab in combination with FOLIRI or FOFOX. The ERG considers the company's approach and rationales for decisions made to be mostly appropriate. Here, the ERG provides a brief summary of the company's methods, together with comments where the ERG has a suggestion for an alternative approach.

Initially, the company assessed assumption of proportional hazards (PHs), which is required for NMA based on constant HRs. KM curves for PFS and OS for included studies, when available, were visually inspected for crossing survival curves, together with more formal assessments of proportionality. Through their analyses, the company identified that the assumption of PHs was violated for some studies for PFS and OS.

Fractional polynomials (FPs) offer an alternative method to generate estimates of effect via NMA when the assumption of PHs is violated.[36] For the FP NMA, to generate estimates for OS and PFS, the

company used the Weibull and Gompertz distributions, which are both first order FPs, and also second order FPs with varying values of the two powers required for second order (Table 24). The deviance information criterion (DIC) was used to compare the goodness-of-fit of competing survival models, with a difference in DIC of about 5 points considered meaningful. Plausibility of the curves produced was also considered. The ERG considers that, based on methods reported in Jansen,[36] that evaluation of a wider range of powers for the first and second order FPs would have been helpful to underscore the company's choice of FP.

Table 24. Models assessed by company in FP NMA

Model
Weibull: First order FP (p = 0); treatment effects on 1 scale (d0) and 1 shape parameter (d1)
Gompertz: First order FP (p = 1); treatment effects on 1 scale (d0) and 1 shape parameter (d1)
Second order FP: p1 = 0, p2 = 0; treatment effects on 1 scale (d0) and 1 shape parameter (d1)
Second order FP: p1 = 0, p2 = 1; treatment effects on 1 scale (d0) and 1 shape parameter (d1)
Second order FP: p1 = 1, p2 = 0; treatment effects on 1 scale (d0) and 1 shape parameter (d1)
Second order FP: p1 = 1, p2 = 1; treatment effects on 1 scale (d0) and 1 shape parameter (d1)

Abbreviations: FP, fractional polynomial; NMA, network meta-analysis.

Within Appendix M, the company presents results for analyses based on constant HR and FPs. Given that assumption of PHs does not hold for PFS for KEYNOTE-177, the ERG considers constant HR NMA to be inappropriate and has a preference for the results derived from the FP NMA. Moreover, results from the FP analysis are implemented in the economic model. Given the ERG's preference for the FP analysis, the ERG does not discuss the results from the constant HR NMA further.

During the clarification process, the company provided the code and digitised data used in the FP NMA. Using the company's data set, and following the company's methodology, the ERG was able to validate the company's choice of FP model for PFS in all patients. The ERG focused on PFS as OS data are immature. Using the preferred model, the ERG was able to replicate the company's estimates of effect of PFS at the time points reported for CAPOX and pembrolizumab versus SoC, but notes deviations from the company's results for panitumumab plus FOLFOX versus SoC (provided in Section 3.5.2.1).

Given that the RCTs included in the FP NMA involve different patient populations in mCRC, with KEYNOTE-177 enrolling only those with MSI-H/dMMR mCRC and PRIME focusing on RAS wild-type mCRC, clinical heterogeneity is present and introduces bias into the analyses, the impact of which on estimates of effect is difficult to quantify. For the NMA of all patients, given that MSI-H/dMMR mCRC is associated with a poor prognosis and SoC in KEYNOTE-177 included bevacizumab treatment, it is likely that bias in the analysis is against pembrolizumab and generated estimates of effect might underestimate the true effect of pembrolizumab.

Differences are noted across studies in the FOLFOX and FOLFIRI regimens administered. However, the ERG's clinical experts have advised that the variations in schedule are unlikely to have an impact on clinical effectiveness.

In summary, the ERG considers the company's use of FP NMA to derive estimates of effect to be appropriate, but notes that evaluation of further powers would be valuable.

3.5.1.2 *Network structure*

As highlighted in Section 2.3.3, the final scope issued by NICE specifies that panitumumab-combination regimens are a first-line treatment option for those with RAS wild-type mCRC, together with cetuximab plus FOLRIRI or FOLFOX for EGFR-expressing RAS wild-type, and CAPOX is an option for all mCRC. The network of treatments created by the company (Figure 3) omits cetuximab-combination treatment and is based on all patients. The rationale for not including cetuximab plus FOLFIRI or FOLFOX in the NMA was that cetuximab-based treatment formed part of SoC, and the six regimens comprising SoC were considered as a single group for the purposes of the NMA. The ERG notes that only 11.2% (16/143) of people allocated to SoC received a cetuximab-containing treatment, and thus removing those who received cetuximab from the SoC group for purposes of NMA would leave a relatively large sample size.

Given that cetuximab-combination treatment is considered in the cost effectiveness analysis, the ERG considers it appropriate to generate an estimate of effect for pembrolizumab versus cetuximab-based treatment via FP NMA that is based on available RCTs, as was done for panitumumab plus FOLFOX. To align more closely with the final scope issued by NICE, during clarification, the ERG requested that the company carry out FP NMA for the subgroup of people from KEYNOTE-177 with RAS wild-type tumours as depicted in Figure 4. The RCTs included in the network are based on studies identified by two systematic reviews.[13, 28] For the FP NMA, the ERG defined SoC as FOLFOX or

FOLFIRI, with or without bevacizumab, and assumed equivalence of FOLFOX and FOLFIRI regimens. The company declined to carry out the requested FP NMA, citing that It would be inappropriate to perform the analyses as:

- the analyses would be considerably under-powered and likely to produce false-negative results.
- randomisation would be broken for treatment comparisons in the subgroups.

The company also commented that, "*subgroup analyses for the subgroup of patients with RAS-wildtype colorectal cancer are not necessary for the purposes of this appraisal, as the results from the analyses in the overall population would be reasonable and appropriate proxies for what would be observed in RAS-wildtype patients*". The ERG agrees with the company's points around the limitations associated with *post hoc* subgroup analyses. However, the ERG counters that the network suggested by the ERG is more relevant to the decision problem than that carried out by the company, minimising clinical heterogeneity through exclusion of studies evaluating CAPOX in mCRC with no analysis of biomarker status and by focusing on RAS wild-type. Subgroup analysis by RAS wild-type in KEYNOTE-177 shows a clear benefit for pembrolizumab versus SoC (HR 0.44, 95% CI: 0.29 to 0.67; Figure 16), a benefit that is also likely underestimated because of the inclusion of cetuximab- and bevacizumab-based treatment. To provide the best available estimate of effect to inform the cost effectiveness analysis that is derived from available evidence, the ERG considers it important to evaluate the effectiveness of pembrolizumab in RAS wild-type versus all relevant comparators, albeit that MSI-H/dMMR status is not known for comparator studies.

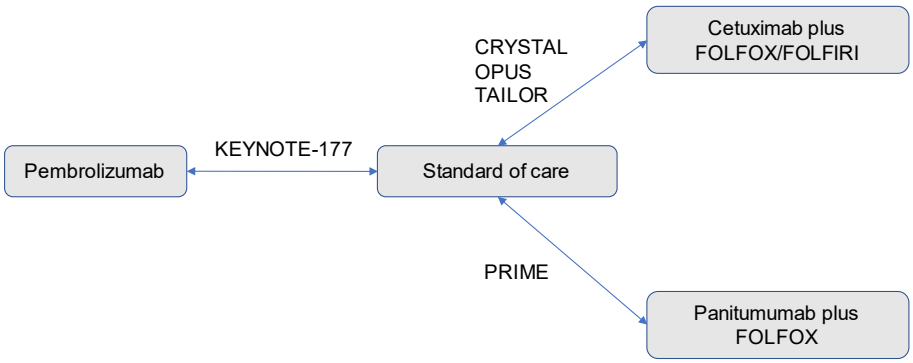
Given that estimates of effect are not available for the ERG's preferred network, and the company's network is informing estimates of effect for CAPOX, the ERG reiterates that the network for all patients omits an RCT that could inform the comparison of CAPOX versus SoC (discussed in Section 3.1).[19] As the ERG is assuming equivalent clinical effectiveness for CAPOX, FOLFIRI and FOLFOX in the first-line treatment of mCRC, the ERG did not request that the company reanalyse the data to include the RCT in their network as part of the clarification process. Additionally, within KEYNOTE-177, because RAS status was not necessarily a factor in choice of treatment in SoC, some allocated to SoC might not have received the optimum intervention according to clinical practice in England.

No estimate of effect for pembrolizumab versus cetuximab-combination treatment in the first-line setting is available from the FP NMA, thus, based on results from NMA presented in TA439,[13] for

economic analyses the ERG has assumed that panitumumab and cetuximab-combination regimens have equivalent clinical effectiveness in RAS wild-type mCRC. TA439 presented results from NMA based on separate networks for FOLFIRI and FOLFOX-based treatments. [13] As noted by the company, and TA439, no study is available that evaluates panitumumab plus FOLFIRI, and so indirect evidence on cetuximab and panitumumab was reported for FOLFOX combinations. TA439 reports that NMA provided no statistically significant evidence to suggest that cetuximab plus FOLFOX was more effective than panitumumab plus FOLFOX at increasing PFS or OS, with HRs reported as:

- PFS: cetuximab plus FOLFOX versus panitumumab plus FOLFOX: HR 0.74, 95% credible interval (CrI): 0.36 to 1.49;
- OS: cetuximab plus FOLFOX versus panitumumab plus FOLFOX: HR 1.22, 95% CrI: 0.71 to 2.11.

Figure 4. ERG's suggested network for generation of estimates of effect for pembrolizumab versus cetuximab and panitumumab combination treatments in RAS wild-type mCRC



3.5.2 Results of indirect comparison

3.5.2.1 Progression-free survival

A second order FP was deemed to be the model with the best fit to the PFS data available from the RCTs informing the network for the ITT population (Table 26). By the company's FP NMA, compared with SoC, CAPOX and panitumumab plus FOLFOX, the clinical benefit for pembrolizumab noted in KEYNOTE-177 for PFS [redacted] (Table 25; [redacted]). At follow-up of 8 months and beyond, [redacted].

Using the data set and code supplied by the company, and applying the same number of burn-in, chains, iterations as the company, the ERG generated the [redacted]

[redacted] (Table 25). [redacted]. The ERG also validated the company's FP NMA for comparisons versus SoC. [redacted]: results for company's and ERG's FP NMA using SoC as the baseline treatment are presented in Appendix 1.1 (Table 68, Figure 18, Figure 19).

Table 25. Time-varying hazard ratios for PFS at selected follow-up times for competing interventions versus pembrolizumab (second order FP model [p1 = 0, p2 = 0]) (adapted from CS, Table 42, page 80)

Month	HR versus pembrolizumab (95% CrI)[a]					
	CAPOX Company	CAPOX ERG	Panitumumab + FOLFOX Company	Panitumumab + FOLFOX ERG	SoC Company	SoC ERG
4	[redacted]					
8	[redacted]					
12	[redacted]					
16	[redacted]					
20	[redacted]					
24	[redacted]					
28	[redacted]					
32	[redacted]					

36	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
40	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] HR >1 favours pembrolizumab, HR <1 favours comparator.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; CS, company submission; CrI, credible interval; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin; HR, hazard ratio; PFS, progression-free survival; SoC, standard of care.

Table 26. Model fit estimates for fixed-effects network meta-analysis with parametric survival models for PFS (reproduced from CS, Table 41, page 79)

Model	Dbar	pD	DIC
Weibull: First order FP (p = 0); treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Gompertz: First order FP (p = 1); treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Second order FP: p1 = 0, p2 = 0; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Second order FP: p1 = 0, p2 = 1; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Second order FP: p1 = 1, p2 = 0; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Second order FP: p1 = 1, p2 = 1; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]

Abbreviations: CS, company submission; DIC, deviance information criterion; FP, fractional polynomial; NMA, network meta-analysis; PFS, progression-free survival.

3.5.2.2 *Overall survival*

For analysis of OS, the ERG cautions that OS data from KEYNOTE-177 are immature and OS data are likely to be confounded due to crossover from SoC to pembrolizumab on progression of disease. A second order FP was deemed to be the model with the best fit to the OS data available from the RCTs informing the network for the ITT population (Table 27).

By FP NMA, pembrolizumab [redacted] (Table 28; [redacted]).

Table 27. Time-varying hazard ratios for OS at select follow-up times for competing interventions versus pembrolizumab (second order FP model [p1 = 0, p2 = 0]) (reproduced from CS, Table 36, page 75)

Month	HR versus pembrolizumab (95% CrI)[a]		
	CAPOX	Panitumumab + FOLFOX	SoC
4	[redacted]	[redacted]	[redacted]
8	[redacted]	[redacted]	[redacted]
12	[redacted]	[redacted]	[redacted]
16	[redacted]	[redacted]	[redacted]
20	[redacted]	[redacted]	[redacted]
24	[redacted]	[redacted]	[redacted]
28	[redacted]	[redacted]	[redacted]
32	[redacted]	[redacted]	[redacted]
36	[redacted]	[redacted]	[redacted]
40	[redacted]	[redacted]	[redacted]

[a] HR >1 favours pembrolizumab, HR <1 favours comparator.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; CS, company submission; CrI, credible interval; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin; HR, hazard ratio; OS, overall survival; SoC, standard of care.

Table 28. Model fit estimates for fixed-effects network meta-analysis with parametric survival models for OS (reproduced from CS, Table 35, page 74)

Model	Dbar	pD	DIC
Weibull: First order FP (p = 0); treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Gompertz: First order FP (p = 1); treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Second order FP: p1 = 0, p2 = 0; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Second order FP: p1 = 0, p2 = 1; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]
Second order FP: p1 = 1, p2 = 0; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]

Second order FP: p1 = 1, p2 = 1; treatment effects on 1 scale (d0) and 1 shape parameter (d1)	[redacted]	[redacted]	[redacted]

Abbreviations: CS, company submission; DIC, deviance information criterion; FP, fractional polynomial; NMA, network meta-analysis; OS, overall survival.

3.6 Conclusions of the clinical effectiveness section

The ERG considers the company's SLR to be of reasonable quality and likely to have retrieved all studies relevant to pembrolizumab, despite limiting inclusion to English-language publications. In the context of the FP NMA for all patients, the ERG notes that there is an RCT, which was identified but excluded by the company, that meets the SLR inclusion criteria and provides data relevant to the company's NMA.

The ERG considers KEYNOTE-177 to be a well-designed and well-conducted RCT, with an overall low risk of bias and high internal validity. KEYNOTE-177 is the only RCT, at the time of writing, reporting clinical effectiveness of first-line treatments of those specifically with MSI-H/dMMR mCRC. For the primary outcome of PFS, the ERG considers that the evidence derived from the ITT population of KEYNOTE-177 supports the company's proposal that pembrolizumab markedly improves PFS compared with SoC. Pembrolizumab was associated with a median PFS of 16.5 months compared with a median of 8.2 months for SoC, a difference that was statistically significant (HR 0.60, 95% CI: 0.45 to 0.80; p=0.0002).

SoC in KEYNOTE-177 comprised FOLFOX and FOLFIRI given either alone or in combination with bevacizumab or cetuximab. Evidence from RCTs indicates that bevacizumab- and cetuximab-based regimens are more clinically effective at improving PFS and OS than FOLFIRI or FOLFOX alone. In KEYNOTE-177, 70% of those in the SoC group received FOLFOX or FOLFIRI with bevacizumab, and 10% cetuximab with FOLFOX or FOLFIRI. Additionally, bevacizumab is not available as a treatment option for clinicians in England. Given that bevacizumab-based treatment is associated with improved PFS and OS, the ERG considers that bias in the comparison of pembrolizumab versus SoC is likely to be against pembrolizumab, resulting in a conservative estimate of treatment effectiveness. For comparison with CAPOX, FOLFIRI and FOLFOX, the ERG considers the SoC group in totality provides the most robust estimate of comparative treatment effectiveness for pembrolizumab versus CAPOX, FOLFIRI and FOLFOX in the specified group of all patients with MSI-H/dMMR mCRC.

In the first-line setting for treatment of mCRC in England, cetuximab and panitumumab in combination with either FOLFIRI or FOLFOX are additional options if a person is identified as having

RAS wild-type mCRC. *Post hoc* subgroup analyses from KEYNOTE-177 indicate that, in those with RAS wild-type, pembrolizumab is associated with a statistically significant improvement in PFS compared with SoC (HR 0.44, 95% CI: 0.29 to 0.67). In KEYNOTE-177, panitumumab-based regimens were not available to the SoC group, and assessment of RAS status (wild-type versus mutant) was not carried out for all those enrolled. Thus, in the context of external validity to England, some of those randomised to SoC might not have received treatment as they would in clinical practice.

Direct head-to head data for pembrolizumab versus cetuximab- and panitumumab-combination regimens in RAS wild-type mCRC are not available, and to generate estimates of effect required that the company carry out an NMA. In the CS, the company presented an NMA for all patients that included panitumumab-containing regimens as a comparator, but not cetuximab-based treatment, with the rationale that cetuximab combination treatment was an option in the SoC group of KEYNOTE-177. The ERG considers the subgroup analysis based on RAS wild-type from KEYNOTE-177 provides the best estimate for PFS for pembrolizumab versus CAPOX, FOLFOX and FOLFIRI in that subgroup. However, given that RCTs are available evaluating panitumumab and cetuximab-based therapies as first-line treatments in RAS wild-type mCRC, and clinical effectiveness informs the economic analyses, the ERG considers more robust estimates for pembrolizumab versus comparators relevant to the decision problem could be derived from an NMA based on IPD for those with RAS wild-type from KEYNOTE-177. The ERG appreciates that the evidence from KEYNOTE-177 on the benefit of pembrolizumab in improving PFS is strong, but considers that the requested NMA could substantiate the estimate of effect for pembrolizumab noted in KEYNOTE-177. The ERG acknowledges that, as with the company's original NMA in all patients, because the impact of MSI-H or dMMR is not accounted for in trials other than KEYNOTE-177, there will be bias in an NMA evaluating treatments in RAS wild-type subgroup.

Considering the company's NMA, the ERG agrees with the company's rationale for choice of FP NMA, but considers evaluation of a wider range of powers than those reported by the company would be valuable. Using the data and code supplied by the company during the clarification process, the ERG found similar results to the company for pembrolizumab versus interventions of interest to the decision problem, and also for CAPOX and pembrolizumab versus SoC. Considering panitumumab + FOLFOX versus SoC, the ERG's results were similar up to 16 months, but diverged thereafter. The ERG has no explanation for the discrepancy in the results.

Post hoc subgroup analyses of PFS based on those identified as having RAS mutations generated a markedly different result from other subgroups, with a change in direction of effect to favour SoC, albeit that the difference between pembrolizumab and SoC was not statistically significant (HR 1.19, 95% CI: 0.68 to 2.07). However, the ERG notes that the 95% CIs for estimates of PFS for RAS wild-type and non-RAS wild-type do not overlap and, as such, considers the results in the two subgroups unlikely to have arisen due to random chance.

OS data for KEYNOTE-177 are immature are the time of writing. Analysis of OS is based on 125 events and median OS has yet to be reached in the pembrolizumab group. Additionally, OS is likely to be severely confounded as crossover from SoC to pembrolizumab was allowed: 59% of participants allocated to SoC received pembrolizumab (36.4%) or a PD-1/PD-L1 inhibitor (22.7%) on cessation of initial treatment, primarily because of disease progression. The ERG appreciates that, due to the potential confounding, the OS advantage of pembrolizumab, if any, will be underestimated but considers the data too immature to draw reliable conclusions. Additionally, OS data are not implemented in the cost effectiveness analysis, with the company assuming that post progression survival is the same for pembrolizumab as for SoC.

4 Cost effectiveness

A summary of the company's deterministic and probabilistic cost-effectiveness results for the pembrolizumab versus standard of care (SoC), CAPOX and panitumumab + mFOLFOX6 are presented in Table 29 to Table 31.

Table 29. Company's base case results versus standard of care

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
Standard of Care	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	13,497	3.145	1.862	7,250
Probabilistic results							
Standard of Care	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	18,199	3.133	1.858	9,795

Abbreviations: LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 30. Company's base case results versus CAPOX

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
CAPOX	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	50,968	3.145	1.855	27,474
Probabilistic results							
CAPOX	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	55,820	3.133	1.852	30,143

Abbreviations: CAPOX, capecitabine plus oxaliplatin; LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 31. Company's base case results versus panitumumab + mFOLFOX6

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Deterministic results							
mFOLFOX6 + panitumumab	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	-48,317	3.145	1.688	Dominant
Probabilistic results							
Panitumumab + mFOLFOX6	167,456	4.212	2.590	-	-	-	-
Pembrolizumab	[redacted]	7.022	4.274	-43,910	2.81	1.684	Dominant

Abbreviations: mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

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

The company performed two systematic literature reviews (SLRs) to identify published studies that could inform the cost-effectiveness evaluation of pembrolizumab for adult patients with untreated metastatic colorectal cancer (mCRC) with high microsatellite instability (MSI-H) or mismatch repair deficiency (dMMR). The first search attempted to identify resource use, costs and cost-effectiveness studies, while the second search was designed to identify health-related quality of life (HRQoL) estimates.

The searches focused on identifying cost-effectiveness evidence for the MSI-H and dMMR mCRC patient cohort. A summary of the Evidence Review Group's (ERG) critique of the company's SLR is given in Table 32.

Table 32. Systematic review summary

Systematic review step	Section of CS in which methods are reported			ERG assessment of robustness of methods
	Cost effectiveness evidence	HRQoL evidence	Resource use and costs evidence	
Search strategy	Appendix G	Appendix H	Appendix I	Appropriate
Inclusion/exclusion criteria	Appendix G	Appendix H	Appendix I	Appropriate

Screening	Appendix G	Appendix H	Appendix I	Appropriate
Data extraction	Appendix G	Appendix H	Appendix I	Appropriate
Quality assessment of included studies	Appendix G	Appendix H	Appendix I	Appropriate

Abbreviations: CS, company submission; ERG, evidence review group; HRQoL, health-related quality of life.

The cost-effectiveness SLR identified one study, Chu 2019[37], that fulfilled the inclusion criteria. A further nine studies were listed in the PRISMA chart that covered first-line mCRC cancer treatment, however these were not included in the final evidence table. The HRQoL literature review included 38 studies. The company states that there were 21 included studies, but only included details for the 2 in the MSI-H/dMMR subpopulation.

None of the studies identified in in the SLR were used for parameterising the model. Instead, the data sources used in the model came from the KEYNOTE-177 trial, a multiple technology assessment of cetuximab and panitumumab (TA439), the company's clinical experts, the NHS Reference Cost Schedule and the PSSRU Unit Costs of Health and Social Care. The ERG considers the data sources used by the company to be reasonable.

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

4.2.1 NICE reference case checklist

Table 33 summarises the ERG's appraisal of the company's economic evaluation against the requirements set out in the National Institute for Health and Care Excellence (NICE) reference case checklist for the base-case analysis, with reference to the NICE final scope outlined in Section 2.

Table 33. 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	All relevant health effects for adult patients with untreated unresectable or mCRC with MSI-H or dMMR have been included.
Perspective on costs	NHS and PSS	All relevant costs have been included and are based on the NHS and PSS perspective.

Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Cost-utility analysis with fully incremental analysis has been provided by the company.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Lifetime horizon (40 years).
Synthesis of evidence on health effects	Based on systematic review	The company performed an appropriate systematic review.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	QALYs based on EQ-5D-3L data from KEYNOTE-177 used in the base case analysis.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	EQ-5D-3L data obtained directly from patients in KEYNOTE-177.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Patients in KEYNOTE-177 are representative of the UK population.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	The economic evaluation matches the 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	Costs included in the analysis have been sourced using NHS reference costs,[38] MIMS,[39] eMIT[40] and published literature and are reported in pounds sterling for the price year 2019.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Discount rate of 3.5% has been used for both costs and health effects.

Abbreviations: dMMR, mismatched repair deficiency; ERG, evidence review group; mCRC, metastatic colorectal cancer; MSI-H, high microsatellite instability; NHS, national health service; PSS, personal social services; QALY, quality adjusted life year, UK, United Kingdom.

4.2.2 Population

The population considered by the company for this Single Technology Appraisal (STA) is based on the proposed marketing authorisation, which includes adult patients with untreated unresectable or metastatic colorectal with MSI-H or dMMR. The population can be further divided by RAS mutation status (RAS wild-type or non-RAS wild-type). The population included in the company's base case

analysis is the intention-to-treat (ITT) population of KEYNOTE-177, which includes all MSI-H or dMMR patients irrespective of RAS mutation status. In addition, in their response to clarification questions, the company confirmed that no patients in KEYNOTE-177 had unresectable disease at baseline.

In KEYNOTE-177, approximately 70% of standard of care (SoC) patients received bevacizumab combination treatments, but NICE guidance does not recommend its use for treating mCRC.[41, 42] The company presented subgroup analyses for the ITT population excluding patients treated with bevacizumab, but note that the analyses are not robust due to small sample size and imbalances in baseline characteristics. The ERG agrees that the analyses for the "ITT population minus bevacizumab" patients are likely to be subject to substantial uncertainty and that including patients receiving treatment with bevacizumab is likely to be biased against pembrolizumab, resulting conservative estimates of treatment effectiveness. Please see Section 2.3 for further detail on the population of KEYNOTE-177.

The ERG's primary concern with the population of the model is that in addition to the analysis using the ITT population, the company should have presented subgroup analyses by RAS mutation status. The NICE final scope specifies that, "If the evidence allows the subgroup of people with RAS wild-type colorectal cancer will be considered".[1] Treatment options in the UK differ if a patient has RAS wild-type mCRC (see Section 4.2.3 for further details). According to the NICE pathway for managing mCRC, first-line biological therapy (cetuximab or panitumumab combination treatments) are only recommended for patients with RAS wild-type mCRC.[42] Furthermore and as mentioned in Section 2.3.1, subgroup analyses reported within the company submission (CS) indicate that, in those with RAS wild-type, pembrolizumab is associated with a statistically significant improvement in progression-free survival (PFS) compared with SoC (hazard ratio [HR] 0.44, 95% confidence interval [CI]: 0.29 to 0.67). However, in the non-RAS wild-type subgroup, the direction of effect favours SoC, albeit that the benefit with SoC would not be considered statistically significant (HR 1.19, 95% CI: 0.68 to 2.07). The ERG notes that the 95% CIs for the two subgroups do not overlap and, as such, considers the difference in the two subgroups unlikely to be due to random chance.

Given the importance of RAS mutation status, not only in terms of relative treatment effectiveness but also for directing treatment choice for patients in the NHS, during the clarification stage the ERG requested the company to provide cost-effectiveness analyses for RAS wild-type and non RAS wild-type subgroups from KEYNOTE-177 and update the fractional polynomial (FP) network meta-analysis

(NMA) to generate relative estimates of treatment effect for pembrolizumab versus the listed comparators in the NICE final scope for the relevant subgroups.

The company declined to provide subgroup analyses based on RAS mutation status, stating that the requested analyses were inappropriate because the results would be underpowered and likely to produce false-negative results due to small numbers of events, randomisation would be broken and that RAS mutation status was unknown for 27% of the trial population. However, the ERG notes that the issues highlighted by the company for the subgroup analyses apply to the ITT minus bevacizumab subgroup analyses the company provided within their submission.

In addition, the company stated that analyses for the RAS wild-type subgroup are unnecessary as the results for the ITT population (used for the company base case analyses) would be appropriate proxies for what would be observed in this patient group. For the ITT population, pembrolizumab is associated with a statistically significant improvement in PFS compared with SoC (HR 0.60, 95% CI: 0.45 to 0.80), which the ERG considers is more conservative compared with the results for the RAS wild-type subgroup. However, there are several issues with the company's assertation that the ITT analysis can be used as a proxy for the RAS wild-type subgroup. Cetuximab combination treatment is included within the SoC comparison, but only approximately 10% of patients in KEYNOTE-177 received this treatment. However, it is not clear whether all RAS wild-type patients in KEYNOTE received cetuximab combination treatment, as this was not routinely tested for, as would happen in the NHS. Furthermore, cetuximab combination treatment costs are used as a proxy for the patients who received bevacizumab combination treatment and as such the analysis implicitly assumes that approximately 80% of the model population in the SoC arm have RAS wild-type mCRC.

The ERG is also concerned that the direction of effect for the non-RAS wild-type subgroup in KEYNOTE-177 favours SoC. If cost-effectiveness estimates were provided for the non-RAS wild-type subgroup by the company, the ERG predicts it would have demonstrated that pembrolizumab is less effective and more costly that SoC and as such would be dominated by the comparator. However, it should be noted that as mentioned earlier in this section, SoC in KEYNOTE-177 is made up of approximately 70% of patients who have received bevacizumab combination treatment and thus may not be reflective of what would be seen in UK clinical practice. While the use of bevacizumab containing treatments may be inflating the effectiveness of SoC in the non-RAS wild-type subgroup to infer that SoC is potentially more effective than pembrolizumab, the ERG still considers that it is likely that if this bias was corrected for that SoC and pembrolizumab may well have similar

effectiveness in this subgroup. This is illustrated by the HR of 0.90 (95% CI: 0.24 to 3.39) for the non-RAS wild-type subgroup analyses for the ITT minus bevacizumab subgroup, though caution should be applied when interpreting these results due to small sample sizes.

Section 4.2.3 discusses the issue of comparators and Section 4.2.5.1 discusses the issues around treatment effectiveness for the subgroup analyses based on RAS mutation status and within each section outlines the ERG's preferred approach to the subgroup analyses. The ERG highlights that clinical data provided within the CS and used in the model from KEYNOTE-177 are based on the second interim analysis (database cut-off date: 19 February 2020). Thus, the ERG considers that longer term data from KEYNOTE-177 will be crucial to determine whether the direction of effect for the non-RAS wild-type subgroup will be substantiated in the longer term and also to reduce the overall uncertainty in the cost-effectiveness analysis.

4.2.3 Interventions and comparators

The intervention considered for the economic analysis is pembrolizumab. The primary comparator considered by the company is SoC and is based on the comparators included in KEYNOTE-177, which are mFOLFOX6 (folinic acid plus fluorouracil plus oxaliplatin), FOLFIRI (folinic acid plus fluorouracil plus irinotecan), cetuximab in combination with mFOLFOX6 or FOLFIRI, and bevacizumab in combination with mFOLFOX6 or FOLFIRI. Over 70% of SoC patients in KEYNOTE-177 received bevacizumab combination treatments, but NICE guidance does not recommend its use for treating mCRC.[41, 42] Thus, the company adjusted the drug acquisition costs in the model so that the proportion of patients who received bevacizumab combination treatments in KEYNOTE-177 were costed as receiving cetuximab combination treatments. Table 34 outlines the distribution of patients across each of the SoC regimens in KEYNOTE-177.

Table 34. Distribution of patients across each of the SoC regimens in KEYNOTE-177 (Table 49 of CS)

Treatment regimen	Percentage of patients (n=154)*
mFOLFOX6 – 9.1%	9.1%
FOLFIRI – 11.0%	11.0%
Cetuximab + mFOLFOX6 or FOLFIRI	10.4%
Bevacizumab + mFOLFOX6 or FOLFIRI[a]	69.5%

[a] Assumed in the model to be cetuximab + mFOLFOX6 or FOLFIRI.
Abbreviations: FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin.
* based on all randomized subjects in treatment arm whether or not treatment was administered

In line with the NICE final scope, the company has also presented cost-effectiveness analyses for CAPOX (capecitabine plus oxaliplatin) and panitumumab in combination with mFOLFOX6. Table 35 presents the treatment regimens included in the economic analysis. All treatments are delivered by intravenous (IV) infusion, except capecitabine, which is given orally.

The company expects that the pembrolizumab monotherapy licence will include an option to administer pembrolizumab at a 400mg dose once every six weeks. The ERG clinical experts advised that in clinical practice, they would prefer to administer pembrolizumab using the 400mg dose once every six weeks for patient convenience and also to reduce NHS resource use and as such, the ERG considers this regimen to be appropriate to include in its preferred analysis, presented in Section 6.4.

Table 35. Treatment regimens included in the model (adapted from Table 9 and Table 66 of the CS)

Treatment	Dose	Treatment cycle
Pembrolizumab	200 mg IV over 30 minutes	Once every three weeks
mFOLFOX6	Oxaliplatin 85 mg/m^2 IV over 2 hours, day 1 Leucovorin 400 mg/m^2 IV over 2 hours, day 1 5-FU 400 mg/ m^2 IV bolus, day 1, then 5-FU 1200 mg/m^2/day x 2 days (2400 mg/m^2 over 46-48 hours) IV continuous infusion	Once every two weeks

FOLFIRI	Irinotecan 180 mg/m^2 IV over 30-90 minutes, day 1 Leucovorin 400 mg/m2 IV infusion to match duration of irinotecan infusion, day 1 5-FU 400 mg/m^2 IV bolus day 1, then 1200 mg/m^2/day x 2 days (total 2400 mg/m^2 over 46-48 hours) IV continuous infusion	Once every two weeks
Cetuximab + FOLFIRI	Cetuximab: 400 mg/m^2 IV over 2 hours first infusion, then 250 mg/m^2 IV over 1 hour weekly FOLFIRI: Irinotecan 180 mg/m^2 IV over 30-90 minutes, day 1 Leucovorin 400 mg/m^2 IV infusion to match duration of irinotecan infusion, day 1 5-FU 400 mg/m^2 IV bolus day 1, then 1200 mg/m^2/day x 2 days (total 2400 mg/m^2 over 46-48 hours) IV continuous infusion	Once every two weeks
Cetuximab + mFOLFOX6	Cetuximab: 400 mg/m^2 IV over 2 hours first infusion, then 250 mg/m^2 IV over 1 hour, **weekly** mFOLFOX6: Oxaliplatin 85 mg/m^2 IV over 2 hours, day 1 Leucovorin* 400 mg/m^2 IV over 2 hours, day 1 5-FU 400 mg/ m^2 IV bolus, day 1, then 5-FU 1200 mg/m^2/day x 2 days (2400 mg/m^2 over 46-48 hours) IV continuous infusion	Once every two weeks
Panitumumab + mFOLFOX6	Panitumumab: 6mg/kg IV via an infusion pump over 30 – 90 minutes, day 1 mFOLFOX6: Leucovorin 300 mg/m^2 IV over 2 hours, day 1 5-FU 400 mg/ m^2 IV bolus, day 1, then 5-FU 1200 mg/m^2/day x 2 days (2400 mg/m^2 over 46-48 hours) IV continuous infusion	Once every two weeks
CAPOX	Oxaliplatin 130mg/m^2 over 2 hours, day 1 Capecitabine 1000mg/m^2 orally twice daily for first two weeks and then one week off	Once every three weeks

Abbreviations: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; IV, intravenous; m2, metre squared; mg, milligram;

In the NICE final scope, panitumumab in combination with FOLFIRI was listed as a relevant comparator, but the company stated that no studies were found for this combination that could be included in the NMA and based on clinical expert opinion in TA439, the company considered FOLFOX and FOLFIRI to be clinically equivalent.[13] Furthermore, the NICE final scope also listed tegafur with uracil (in combination with folinic acid), capecitabine and raltitrexed (only when folinic acid and fluorouracil are not tolerated or unsuitable) as comparators, but the company excluded these from the analysis.[1] The company stated that tegafur with uracil (in combination with folinic acid) is no longer available in the UK, capecitabine is only used in the elderly and frail (most often with an ECOG status of >1), and that raltitrexed is rarely used in UK clinical practice. Based on advice from the ERG's clinical experts, the company's rationale to exclude tegafur with uracil (in combination with folinic acid), capecitabine, and raltitrexed is considered reasonable by the ERG. With regards to assuming clinical equivalence for FOLFOX and FOLFIRI for the panitumumab combination treatment comparison, the ERG agrees with the assumption of clinical equivalence based on a study evaluating FOLFOX versus FOLFIRI which was supported by the ERG's clinical experts.[18] Please refer to Section 3.3 for further detail on the FOLFOX and FOLFIRI trials.

As mentioned in Section 4.2.2, a key issue with the company's approach to the cost-effectiveness analysis is the lack of analyses of relevant comparators based on the different indicated populations, as outlined by the NICE final scope. The company used the ITT population of KEYNOTE-177 for all comparators even though, in the UK, treatment regimens that are combined with panitumumab are indicated for the treatment of patients with RAS wild-type mCRC and cetuximab combination treatments are further specified for patients with epidermal growth factor receptor (EGFR) expressing RAS wild-type mCRC. The ERG notes the company's rationale for their approach, as RAS wild-type is not a consideration in the proposed marketing authorisation for pembrolizumab. Nonetheless, to enable appropriate cost-effectiveness estimates of pembrolizumab compared with cetuximab and panitumumab combination regimens, the appropriate population for the analysis should be the RAS wild-type subgroup of the ITT population in KEYNOTE-177. Table 36 presents the relevant comparators by population, which the ERG considers more appropriately reflects the NICE final scope compared with the approach adopted by the company for its base case analysis.

Table 36. Comparators by population based on the NICE final scope

Population from KEYNOTE-177	Comparator
A – ITT	• mFOLFOX6/ FOLFIRI • CAPOX
B – RAS wild-type	• Cetuximab in combination with FOLFOX or FOLFIRI • Panitumumab in combination with FOLFOX or FOLFIRI • mFOLFOX6/ FOLFIRI • CAPOX
C – non-RAS wild-type	• mFOLFOX6/ FOLFIRI • CAPOX

Abbreviations: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; ITT, intention-to-treat.

As mentioned in Section 4.2.2, at the clarification stage the ERG requested the company to produce cost-effectiveness analyses based on the populations and associated comparators listed in Table 36. The company declined to provide the subgroup analyses by RAS-mutation status as they considered them to be inappropriate because the results would be underpowered and likely to produce false-negative results due to small numbers of events, randomisation would be broken and that RAS mutation status was unknown for 27% of the trial population. Furthermore, the company considered that the ITT analysis serves as an appropriate proxy for the RAS wild-type subgroup. However, as mentioned previously, the ERG considers the company's reasoning to exclude subgroup analysis by RAS mutation status to be inconsistent as the company provided subgroup cost-effectiveness analyses for the ITT minus bevacizumab subgroup analyses within their submission.

For population B (RAS wild-type), the ITT cost-effectiveness analysis presented by the company may provide conservative approximations for the comparisons against mFOLFOX6/FOLFIRI, CAPOX and panitumumab combination treatment, based on a comparison of HRs from KEYNOTE-177 for pembrolizumab versus SoC for each population (ITT HR = 0.60 versus RAS wildtype HR = 0.44). However, currently there is no explicit comparison of pembrolizumab versus cetuximab combination treatment as it is blended with the SoC analysis. As presented in Table 34, approximately 10% of patients were randomised to receive cetuximab combination treatment, whereas in UK clinical practice this figure could be higher as patients are routinely tested for gene mutations when

diagnosed with mCRC.[1] However, as the company has assumed the costs of cetuximab combination treatment for the proportion of patients who received bevacizumab combination treatment, SoC is assumed to be made up of 80% of patients receiving cetuximab combination treatment and implicitly assumes that these patients have RAS wild-type mCRC (based on the NICE pathway[1]).

As mentioned in Section 3.5.1.2, TA439 reports that NMA provided no statistically significant evidence to suggest that cetuximab plus FOLFOX was more effective than panitumumab plus FOLFOX at increasing PFS or OS.[13] To provide an estimation of the cost effectiveness of pembrolizumab versus cetuximab combination treatment in lieu of the ERG analysis requested at the clarification stage, the ERG considers that a simplified analysis assuming clinical equivalence between cetuximab combination treatment and panitumumab combination treatment may not be unreasonable. Please refer to Section 4.2.5.1 for the ERG's critique on treatment effectiveness and Section 6.3 for the results of the scenario assuming clinical equivalence between panitumumab and cetuximab combination treatment. However, the ERG reiterates that its preferred approached is that outlined in Table 17 and detailed in the ERG's clarification questions, as that would provide a more robust estimation of the relative treatment effects of pembrolizumab compared with the listed comparators.

As a secondary issue, the ERG's clinical experts advised that those patients in KEYNOTE-177 who received bevacizumab combination treatments would have likely been given FOLFOX or FOLFIRI in UK clinical practice as cetuximab is only indicated for EGFR expressing RAS wild-type mCRC patients, whereas bevacizumab is indicated for all mCRC patients. Please refer to Section 4.2.8 for more details on the issue of comparator costs used in the economic model.

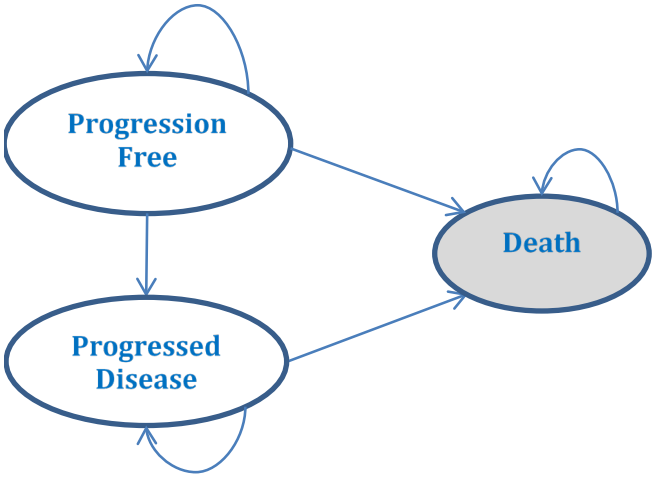
4.2.4 Modelling approach and model structure

Two *de novo* economic models were developed by the company to assess the cost-effectiveness of pembrolizumab for adult patients with untreated unresectable or mCRC with MSI-H or dMMR using Microsoft© Excel. The company base-case model was a five health state-transition model (STM) and the second model was a three health state partitioned survival model (PSM), which the company used as a validation tool.

The STM included two health states (progression-free and progressive disease) for patients who underwent surgery with curative intent (Figure 20 of the company submission). The proportion of surgery patients in each arm of the model was below 10% and the company assumed that clinical

outcomes for these patients (PFS and overall survival [OS]) would be equal. However, the difference in surgery rates between pembrolizumab and the comparators was less than 1%. As surgery rates are similar between treatment arms in KEYNOTE-177 and the clinical outcomes of surgery are the same for all patients, the ERG requested the company during the clarification stage to amend the STM to exclude surgery health states as it is unlikely to be a primary driver of cost-effectiveness and it can be considered an appropriate simplification of the model. In addition, by excluding the surgery health states from the STM, the structure of the model aligns with the structure of the PSM, allowing the PSM to better validate the STM. Figure 5 presents the company's revised base-case structure for the STM and the original structure of the PSM.

Figure 5. Model schematic for the state-transition and partitioned survival models (Figure 19 of the company submission)



As mentioned in Section 3.3.2, OS data from KEYNOTE-177 are immature and are likely to be confounded for SoC due to crossover to pembrolizumab. The PSM relies on immature OS data (adjusted for crossover using the two-stage method) from KEYNOTE-177. In KEYNOTE-177, 36.4% of SoC patients switched over to pembrolizumab monotherapy after discontinuation of the protocol treatment and a further 22.7% switched over to another anti-PD1 or PD-L1 treatment. Given the large amount of crossover in the SoC arm and the immaturity of the OS data, PSM is subject to a substantial amount of uncertainty. As such, for the remainder of the report, the ERG focuses its description and critique around the company's revised three health state STM.

All patients enter the model in the progression-free health state. During each model cycle, patients in the progression-free health state can be either on-treatment or off-treatment, with different treatment discontinuation rules applied depending on treatment regimen (see Section 4.2.8 for

more details). Furthermore, from the progression-free health state, patients can transition to either the progressed disease health state when they experience disease progression or die (thus transitioning to the death health state). When patients transition to the progressed disease health state, they remain there until death. Table 37 presents an overview of clinical data informing the health state transition probabilities. Please refer to Section 4.2.5 for more detail on the clinical data informing the economic model.

Table 37. Overview of clinical data informing the health state transitions in the economic model

Health state transition	Clinical data informing the transition probability
Progression-free to progression-free	PFS data from KEYNOTE-177. For non-trial comparators, time-varying hazard ratios obtained from the fractional polynomial network meta-analysis applied to SoC PFS data from KEYNOTE-177.
Progression-free to progressed disease	Time to progression (TTP) data from KEYNOTE-177. For non-trial comparators, time-varying hazard ratios obtained from the fractional polynomial network meta-analysis applied to SoC TTP data from KEYNOTE-177.
Progression-free to death	PFS minus TTP.
Progressed disease to death	Post-progression survival (PPS) data from KEYNOTE-177. For SoC, PPS is assumed to be equal to PPS for pembrolizumab to mitigate the issue of crossover. For non-trial comparators, PPS is also equal to PPS for pembrolizumab.

Abbreviations: PFS, progression-free survival; PPS. Post-progression survival SoC, standard of care; TTP, time to progression;

A model cycle length of one week with half-cycle correction applied was implemented in the model and the model time horizon was set to 40 years (lifetime), as the mean age in KEYNOTE-177 at baseline was 61 years. The perspective of the analysis was based on the UK NHS, with costs and benefits discounted using a rate of 3.5%, as per the NICE reference case.[43]

4.2.4.1 *ERG critique*

The ERG considers the revised structure of the company's STM to be appropriate, capturing all relevant health states and clinically plausible transitions between health states that are largely similar to other oncology models submitted for NICE appraisal. The one-week cycle length used in the model is suitable to capture important changes in the health state of patients, allowing for robust estimates of costs and benefits to be calculated for each treatment. Half-cycle correction has

been appropriately applied in the model to prevent over or under-estimation of costs and quality-adjusted life-years (QALYs).

With regards to the clinical transitions between the health states in the model, the ERG notes that as post-progression survival (PPS) is used in lieu of mature OS data from KEYNOTE-177, any gain in PFS directly equates to a gain in the estimated OS, which becomes more significant as the company has assumed that PPS for all comparators is equal to PPS for pembrolizumab. Please refer to Section 4.2.5 for further detail on how PFS and PPS are used in the model.

4.2.5 Treatment effectiveness

Clinical data included in the economic model for pembrolizumab and SoC are based on individual patient level data from KEYNOTE-177 and include time to progression (TTP), PFS, PPS, adverse events (AEs) and time on treatment (ToT) outcomes. Please refer to Sections 4.2.6 and 4.2.8 for further details of AEs and ToT included in the model. For panitumumab combination treatment, "per cycle" HRs derived from the FP NMA described in Section 3.5 are used and the application of the estimates is described later in this section. During the clarification stage, the ERG outlined its preferred assumption that, based on studies evaluating FOLFOX versus FOLFIRI[18] and CAPOX versus FOLFIRI,[17, 19-21] the treatments can be considered clinically equivalent, which the company accepted and implemented in its revised base case analysis.

Overview of the company's approach to survival analysis

To extrapolate the KEYNOTE-177 Kaplan-Meier (KM) data, the company followed the guidelines for survival model selection outlined in the NICE Decision Support Unit (DSU) Technical Support Document (TSD) 14.[44] The company first tested whether the assumption of proportional hazards (PH) held for TTP, PFS and PPS outcomes for the ITT population (excluding surgery patients) by producing log-cumulative hazard and Schoenfeld residual plots. The company used the outcomes of the PH assessment to decide to either jointly or independently fit survival distributions. Extrapolations of the KM data were then explored using standard parametric survival distributions (exponential, Weibull, Gompertz, log-normal, log-logistic and generalised gamma) and piecewise models with different time-point cut-offs. To select an appropriate distribution for the extrapolation of each outcome, the company assessed the fit of each modelled curve against the KM data using statistical goodness of fit statistics, including Akaike information criterion (AIC) and Bayesian

information criterion (BIC) statistics, visual inspection of the curves and clinical plausibility of the extrapolation over the time horizon of the model.

During the clarification stage, the ERG requested the company exclude post-surgery health states from the economic model and as such, the company updated the data used for TTP, PFS and PPS to include outcomes for patients who received surgery. However, as the original curve selection process indicated proportional hazards didn't hold, the company did not present updated tests for proportional hazards in their clarification response and modelled curves independently as per the original base case approach.

Time to progression

The company fit standard parametric distributions as well as two-piece models (KM data and standard parametric distribution after cut-off point) with 10- and 20-week cut-off points to the TTP KM data from KEYNOTE-177 for pembrolizumab and SoC. Based on statistical fit (Table 11 of the company's clarification response) and visual inspection of the curves, the company selected the two-piece exponential model with a cut-off point of 20 weeks to model TTP. Based on the ERG's preferred approach, the company assumed that TTP for SoC and CAPOX are clinically equivalent. For panitumumab combination treatment, the company applied time-varying "per cycle" hazard ratios estimated from the FP NMA for PFS to the SoC TTP extrapolation. General population mortality was applied to TTP estimates as a minimum, such that the probability of a patient being alive and progression-free could never be below background mortality.

Figure 6 presents the company's base case TTP extrapolations for pembrolizumab, SoC (CAPOX) and panitumumab combination treatment.

Figure 6. Company base case extrapolation of time to progression outcomes from KEYNOTE-177 (taken from the economic model)

Abbreviations: FP, fractional polynomial; KM, Kaplan-Meier; KN177, KEYNOTE-177; NMA, network meta-analysis; SoC, standard of care; TTP, time to progression; XELOX, CAPOX.

Progression-free survival

As with TTP, the company fit standard parametric distributions as well as two-piece models with 10- and 20-week cut-off points to the PFS KM data from KEYNOTE-177 for pembrolizumab and SoC. Based on statistical fit (Table 12 of the company's clarification response) and visual inspection of the curves, the company selected the two-piece exponential model with a cut-off point of 20 weeks to model PFS. Based on the ERG's preferred approach, the company assumed that PFS for SoC and CAPOX are clinically equivalent. For panitumumab combination treatment, the company applied time-varying "per cycle" hazard ratios estimated from the FP NMA for PFS to the SoC PFS extrapolation. General population mortality was applied to PFS estimates as a minimum, such that the probability of a patient being alive and progression-free could never be below background mortality.

Figure 7 presents the company's base case PFS extrapolations for pembrolizumab, SoC (CAPOX) and panitumumab combination treatment.

Figure 7. Company base case extrapolation of progression-free survival outcomes from KEYNOTE-177 (taken from the economic model)

Abbreviations: FP, fractional polynomial; KM, Kaplan-Meier; KN177, KEYNOTE-177; NMA, network meta-analysis; PFS, progression-free survival; SoC, standard of care; XELOX, CAPOX.

Post-progression survival

As mentioned in Section 3.3.2, in KEYNOTE-177 patients in the SoC arm were allowed to crossover to pembrolizumab upon disease progression. To mitigate the issue of crossover, the company assumed that PPS for all comparators was the same as PPS for pembrolizumab. As such, the company fit standard parametric distributions to PPS KM data from KEYNOTE-177 for pembrolizumab. Based on statistical fit (Table 13 of the company's clarification response), visual inspection of the curves and external validation against real world data, the company selected the Weibull distribution to extrapolate PPS.

Figure 8. Company base case extrapolation of post-progression survival outcomes from KEYNOTE-177 (taken from the economic model)

Abbreviations: FP, fractional polynomial; KM, Kaplan-Meier; KN177, KEYNOTE-177; NMA, network meta-analysis; PPS, post-progression survival; SoC, standard of care; XELOX, CAPOX.

As mentioned in Section 4.2.4, as PPS is used in lieu of mature OS data from KEYNOTE-177, any gain in PFS directly equates to a gain in the estimated OS. Table 38 presents the mean life years estimated for each treatment by population based on the company's extrapolation of the KEYNOTE-177 data, application of the time-varying "per cycle" HRs and subsequent calculation of transition probabilities for the health states in the model.

Table 38. Estimated life years by health state for each population

Health state	Pembrolizumab	SoC	CAPOX	Panitumumab + mFOLFOX6
PFS	4.56	1.21	1.21	1.55
PPS	2.37	2.57	2.57	2.55
Total	**6.93**	**3.78**	**3.78**	**4.10**

Abbreviations: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; PFS, progression-free survival; PPS, post-progression survival; SoC, standard of care.

4.2.5.1 *ERG critique*

Overall, the ERG considers the company's general approach to extrapolating outcomes for TTP, PFS and PPS to be appropriate. The use of piecewise models, with a cut-off point of 20 weeks for TTP and PFS appropriately captures the change in the hazard observed in log-cumulative hazard plots presented in Figure 22 and Figure 27 of the CS. Use of the exponential model after the cut-off point of 20 weeks can be considered appropriate for the pembrolizumab arm based on the log-cumulative hazards plots, but for the SoC arm hazards are increasing. Therefore, use of the Weibull model for both pembrolizumab and SoC may be more appropriate as the hazard function can increase monotonically, but also reduces down to the exponential function if the hazard rate is constant over time (i.e. $\gamma = 1$).[44] However, the ERG notes that diagnostic plots were only provided for the company's original analysis which excluded surgery patients from the clinical outcomes data, whereas the revised base case analysis includes these patients and as such there may be a some change to the hazards over time for pembrolizumab and SoC, but the ERG anticipates it is not likely to be significantly different based on a comparison of the KM curves for TTP and PFS with and without surgery patients included.

The ERG ran a scenario using a piecewise model with a cut-off point of 20 weeks and the Weibull model implemented as the second piece and found this did not have a substantial impact on the incremental cost-effectiveness ratio (ICER). Please refer to Section 6.3 for results of the Weibull scenario.

In KEYNOTE-177, 36.4% of SoC patients switched over to pembrolizumab monotherapy after discontinuation of the protocol treatment and a further 22.7% switched over to another anti-PD1 or PD-L1 treatment. To mitigate the impact of crossover on OS outcomes (and as such PPS outcomes) for the SoC arm, the company assumed that PPS outcomes for SoC would be equal to PPS for the pembrolizumab arm. The company also extended the PPS assumption for the comparison with CAPOX and panitumumab. The ERG notes that the company did explore crossover adjustment methods using OS data from KEYNOTE-177 and this analysis was implemented in the company's PSM. However, the company state that due to the substantial amount of crossover in the SoC arm and the immaturity of OS data in KEYNOTE-177, the OS analyses are subject to a substantial amount of uncertainty. The ERG considers that adjustment of OS using crossover adjustment methods is more robust, but as OS data are immature, the company's simplified assumption of equal PPS for all arms of the model may be acceptable even though it bypasses the issue altogether.

The company's PPS assumption implies that the treatment effect with pembrolizumab is not extended beyond the progression-free health state. The ERG considers that it is not unreasonable that the treatment effect of pembrolizumab is assumed to only be in the progression-free health state as long-term overall survival data are immature, therefore estimating the duration of treatment effect beyond progression is currently problematic. Furthermore, in KEYNOTE-177 pembrolizumab is given for 35 cycles, with a mean treatment duration of 13.3 months, median 16.5 months and the median duration of follow-up was 28.4 months, but beyond discontinuation of treatment there continues to be an ongoing separation of the PFS curves (See Figure 9), though the data for PFS are still immature. Thus, longer term data from KEYNOTE-177 will determine the duration of the treatment effect for pembrolizumab.

Figure 9. Comparison of time on treatment and progression-free survival for pembrolizumab in KEYNOTE-177

Abbreviations: KM, Kaplan-Meier; PFS, progression-free survival; ToT, time on treatment; yrs, years.

The ERG notes that the use of PPS for all arms of the model directly impacts OS, as any gain in PFS directly equates to a gain in the estimated OS. There is uncertainty around whether the PFS to OS relationship holds or whether patients on pembrolizumab are likely to have an accelerated mortality rate upon progression compared to patients on other treatments, particularly for RAS wild-type patients treated with biological therapies (panitumumab and cetuximab). However, in KEYNOTE-177 median PFS2 was [redacted] months (95% CI: [redacted]) for SoC [redacted] pembrolizumab group (HR [redacted], 95% CI: [redacted]). Bearing in mind the immaturity of PFS2,

the ERG notes that [redacted] for pembrolizumab versus SoC [redacted] (PFS HR 0.60, 95% CI: 0.45 to 0.80 vs PFS2 [redacted]). As such, the ERG considers it may not be unreasonable that [redacted], thus providing initial support for the assumption that there isn't an accelerated mortality rate for pembrolizumab upon progression compared with SoC. However, mature OS data from KEYNOTE-177 is required to mitigate the uncertainty around long term survival outcomes.

As mentioned throughout this report, the ERG's primary issue with the cost-effectiveness analysis is that subgroup analysis by RAS mutation status has not been explored by the company. All the clinical data in the model relates to the ITT population of KEYNOTE-177 and thus, estimates of relative treatment effect for non-trial comparators are also for all mCRC patients, irrespective of RAS mutation status. Moreover, for non-trial comparators, there is no evidence in the MSI-H/dMMR population, however this is an unresolvable issue. The company's analysis of PFS by subgroup factors (Figure 10 of the company submission) indicates that, in those with RAS wild-type, pembrolizumab is associated with a statistically significant improvement in PFS compared with SoC (HR 0.44, 95% CI: 0.29 to 0.67). However, in those with a RAS mutation (assumed to be reflective of the non-RAS wild-type subgroup), the direction of effect favours SoC, albeit that the benefit for SoC would not be considered statistically significant (HR 1.19, 95% CI: 0.68 to 2.07). Furthermore, the ERG notes that the 95% CIs for the two subgroups do not overlap and, as such, considers the difference in the two subgroups unlikely to be due to random chance. For the ITT population, the estimate PFS HR is 0.60 (95% CI: 0.45 to 0.80).

At the clarification stage, the ERG requested subgroup analyses by RAS mutation status, but the company stated that the ITT analysis is preferred as it reflects the marketing authorisation for pembrolizumab, and it can also be considered a reasonable proxy for the RAS wild-type subgroup as it is more conservative. However, the company made no mention of estimates of treatment effect for the non-RAS wild-type subgroup. As such, to align the analysis with NICE final scope and provide the committee with estimates of cost-effectiveness for subgroups by RAS mutation status, the ERG outlines its assumptions of treatment effectiveness for the subgroups by RAS mutation status in Table 20. In particular, as the company has not provided separate analysis comparing pembrolizumab versus cetuximab combination treatment, the ERG has made a simplified assumption that cetuximab combination treatment is clinically equivalent to panitumumab combination treatment. As mentioned in Section 3.5.1.2, TA439 reports that NMA provided no

statistically significant evidence to suggest that cetuximab plus FOLFOX was more effective than panitumumab plus FOLFOX at increasing PFS or OS.[13]

The ERG considers the approach outlined in Table 39 only provides illustrative estimates of cost-effectiveness for populations B and C and reiterates that the preferred approach is for the company to:

- Produce relative estimates of treatment effectiveness for patients with and without RAS wildtype mCRC from KEYNOTE-177 using the recommended ERG FP NMA outlined in the ERG's clarification questions;
- Implement the results of the subgroup treatment effectiveness analyses in the economic model to produce cost-effectiveness estimates by relevant comparator for population B and C, as presented in Table 39.

New estimates of cost-effectiveness relate to the cetuximab combination treatment comparison for the RAS wild-type subgroup and the results of this scenario can be found in Section 6.3.

Table 39. ERG approach to treatment effectiveness by population

Population from KEYNOTE	Comparator	Treatment effectiveness approach/ assumptions
A – ITT	mFOLFOX6/ FOLFIRI	Company base case - based on ITT SoC clinical data from KEYNOTE-177.
	CAPOX	Company base case - assumed to be clinically equivalent to mFOLFOX6/ FOLFIRI.
B – RAS wild-type	mFOLFOX6/ FOLFIRI	Company suggested proxy - ITT analysis assumed as proxy for subgroup.
	Panitumumab in combination with FOLFOX or FOLFIRI	Company suggested proxy - ITT analysis assumed as proxy for subgroup.
	Cetuximab in combination with FOLFOX or FOLFIRI	Assumed to be clinically equivalent to panitumumab in combination with FOLFOX or FOLFIRI.
	CAPOX	ITT analysis assumed as proxy for subgroup.

C – non-RAS wild-type	mFOLFOX6/ FOLFIRI	Subgroup analysis based on data from KEYNOTE-177 suggest that relative treatment effect favours SoC (HR 1.19, 95% CI: 0.68 to 2.07). It can be reasonably assumed that because pembrolizumab is more expensive and less effective than SoC, it is dominated by the comparator.
	CAPOX	Assumed to be clinically equivalent to mFOLFOX6/ FOLFIRI. As such, it can be reasonably assumed that because pembrolizumab is more expensive and less effective than SoC, it is dominated by the comparator.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; CI, confidence interval; ERG, evidence review group; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; HR, hazard ratio; ITT, intention to treat; SoC, standard of care.

4.2.6 Adverse events

For the base case analysis, the company included grade 3 or higher treatment-emergent adverse events (TEAEs) that were reported by at least 5% of patients in the safety population in either treatment arm of KEYNOTE-177, presented in Table 40.

Table 40. Grade 3 or higher AEs with ≥ 5% incidence implemented in the model (Table 54 of the CS)

Adverse events	Pembrolizumab (%)	SoC (%)
Anaemia	5.2	10.5
Neutropenia	0.0	15.4
Diarrhoea	5.9	11.2
Abdominal pain	5.2	5.6
Fatigue	3.9	9.1
Neutrophil count decreased	0.0	16.8
Hyponatraemia	5.2	2.8
Hypokalaemia	1.3	6.3
Hypertension	7.2	4.9

Abbreviations: SoC, standard of care.

The company estimated weekly adverse event rates for each adverse event by treatment arm by multiplying the number of patients experiencing a particular adverse event by mean number of episodes of the adverse event per patient and the averaging over the total time on treatment for all patients in a treatment arm. For CAPOX and panitumumab in combination with mFOLFOX6, the company performed an NMA of grade 3 or higher TEAEs in the ITT population, to generate odds ratios (ORs). The estimated ORs are presented in Table 55 of the CS. The ORs were applied to the SoC weekly AE rates for SoC. Table 41 presents the weekly AE rate by treatment.

Table 41. Weekly adverse event rates by treatment arm (adapted from Table 54 and Table 55 of the CS)

Adverse event	Pembrolizumab (%)	SoC (%)	CAPOX (%)	Panitumumab + mFOLFOX6 (%)
Anaemia	0.0009	0.0033	0.0026	0.0075
Neutropenia	0.0000	0.0058	0.0045	0.0131
Diarrhoea	0.0010	0.0039	0.0031	0.0089
Abdominal pain	0.0012	0.0019	0.0015	0.0043
Fatigue	0.0007	0.0029	0.0023	0.0066
Neutrophil count decreased	0.0000	0.0070	0.0055	0.0159
Hyponatraemia	0.0009	0.0010	0.0008	0.0024
Hypokalaemia	0.0002	0.0025	0.0019	0.0056
Hypertension	0.0014	0.0027	0.0021	0.0061

Abbreviations: CAPOX, capecitabine plus oxaliplatin; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; SoC, standard of care.

The impact of AEs on patients' quality of life is considered in the model and is described further in Section 4.2.7, while the costs of managing AEs are discussed in Section 4.2.8.

4.2.6.1 *ERG critique*

The ERG considers that there are no major issues with the company's approach to estimating AEs for inclusion in the economic model.

4.2.7 Health-related quality of life

4.2.7.1 Health state utility values

In KEYNOTE-177, HRQoL data for the ITT population were collected at treatment cycles one, two, three, four, five and seven and at end of treatment, or at one year, depending on which occurred first, and at the 30-day post-treatment follow up. Measurement of HRQoL was based on the EQ-5D-3L instrument and UK preference scores.

The company explored two approaches to estimate utility values from the EQ-5D-3L data: time-to-death utilities and utilities based on progression status. The time-to-death approach captures the time-based decline in HRQoL cancer patients experience as they progress through the disease and as a result it provides more health states than progression-based approaches.

For the time to death approach, the company split the utility data from KEYNOTE-177 into four time point categories: >360 days; 180-360 days; 30-180 days and <30 days. Utility values were estimated by treatment (pembrolizumab and SoC) and the company also provided pooled utility values.

The company states that while the preference is to use time-to-death utilities in the economic model, due to the low number of patients in each group, even after pooling the results, the sample sizes were small, particularly in the <30 days to death group, with wide confidence intervals. (See Table 59 of the CS).

The company instead preferred the standard approach used in NICE appraisals, using a binary pre-progression/progression disease approach to the utility valuation for the model. The progression status approach allowed for more robust estimates with narrower confidence intervals (see Table 60 of the CS). Again, the utility data were split by treatment arm, with a pooled estimate also provided. Patient data from KEYNOTE-177 was used for all patients, including those who received surgery with curative intent.

The company used treatment specific utility values in the progression-free health state for the base case, with the justification that the confidence intervals for the utility estimates between the intervention and SoC arms did not overlap. For the progressed disease health state, the company used a pooled utility value based on all patients, regardless of treatment arm. An overview of the utility values used in the model is presented in Table 42.

Table 42. Health state utility values (adapted from Table 60 in CS)

Health state	Pembrolizumab	SoC
Progression free	0.843	0.787
Progressed disease	0.730	0.730

Abbreviations: SoC, standard of care.

The company also included age-related utility decrements in the economic model using a published algorithm by Ara and Brazier 2010[45].

4.2.7.2 *Adverse event utility values*

For adverse events, rather than an event-specific rate, utility decrements were estimated based on the difference between progression-free utility values with and without status any serious adverse events (grade 3+) present. The adverse event utility data came from the KEYNOTE-177 trial. Adverse event disutility values for each treatment included in the model are presented in Table 43.

Table 43. Adverse event disutility values (taken from the economic model)

Treatment	Disutility value
Pembrolizumab	0.031
SoC	0.031
CAPOX	0.025
Panitumumab + mFOLFOX6	0.065

Abbreviations: AE, Adverse event; QALY, Quality Adjusted Life Year; SoC, Standard of Care.

4.2.7.3 *ERG Critique*

The company generated the utility values from the KEYNOTE-177 trial, which the ERG agrees is preferable to estimates derived from the literature, even with the uncertainty around the estimates used in the model.

The use of progression-based utility values is appropriate for the base case. However, the ERG considers that the company did not sufficiently explore plausible scenarios for alternative utility values in the CS. Furthermore, the ERG queried the use of treatment specific utilities for the progression-free health state, as the difference between pembrolizumab and SoC was substantial

(0.056). In the CS, the company stated that use of treatment specific utility values was based on statistical tests and during the clarification stage, explained that the confidence intervals for the utility estimates between the intervention and SoC arms did not overlap rather than statistical tests directly comparing the differences in the utility estimates.

The ERG confirmed the use of treatment specific utilities with its clinical experts, who supported the use of treatment specific values as pembrolizumab might result in improved quality of life compared with SoC in the progression free state, as it is a monotherapy and as such would require shorter duration and less frequent hospital visits. Nonetheless, at the clarification stage, the ERG requested that the company explore scenario analyses using pooled utility estimates rather than the treatment specific estimates and also where the pooled progression-free utility value was applied only in the first cycle, after which the progression-free no AE utility value was implemented. The utility values for the requested scenarios are presented in Table 44.

Table 44. Alternative utility values for scenario analysis (Table 60 in the CS)

Health state	Utility value
Progression-free - pooled	0.819
Progressed disease - pooled	0.730
Progression-free – no AE	0.833

Abbreviations: SoC, standard of care.

The results of the ERG requested scenarios are presented in Table 45 and show that changes had minimal impact on the ICERs.

Table 45. Results of the utility value scenario analyses

Pembrolizumab vs.	Company base case	Pooled utility value scenario	Pooled progression-free AE utility value in first cycle
Standard of Care	7,250	7,762	7,615
CAPOX	27,474	29,432	28,869
Panitumumab + mFOLFOX6	Panitumumab + mFOLFOX6 is dominated	Panitumumab + mFOLFOX6 is dominated	Panitumumab + mFOLFOX6 is dominated

Abbreviations: AE, adverse event; CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin

The ERG considers that because the company's used utility values that do not distinguish between patients with and without an AE, that the impact of AEs is inherently included in the overall utility value. As such, it is the ERG's preference to exclude disutility associated with AEs and has used this assumption for its base case analysis presented in Section 6.4.

The ERG also requested a scenario where the time to death estimates reported in the company submission were used for the QALY calculations in the model; however, the company did not provide the scenario. Thus, the ERG implemented the time-to-death scenario using the pooled utility estimates from the company's submission for each arm of the model, presented in Table 46. The ERG's scenario is based on estimating the proportion of patients in each cycle who die and utilities for the model cycle are applied based on which time-to-death category it falls within. The results of the scenario using time-to-death utility values can be found in Section 6.3 and show that the scenario had minimal impact on the ICER. However, as mentioned earlier in this section, the time-to-death utility estimates are immature, with small sample sizes and large confidence intervals, particularly for the categories closest to death.

Table 46. Pooled time to death utility values (adapted from Table 59 of the company submission)

Time to death	Pooled estimate
<30 days to death	0.781
30-180 days to death	0.762
180-360 days to death	0.701
>360 days to death	0.499

As mentioned throughout this report, the ERG requested subgroup analysis for the RAS wild-type population and at the clarification stage requested that utilities values for patients with RAS wild-type mCRC were explored in a scenario. However, the company declined to provide the analyses by RAS mutation status. The company's SLR identified studies which contained utility values for patients with RAS wild-type mCRC as well as estimates for total metastatic disease utilities. The ERG examined the SLR results and found that most of the studies that included utility values for RAS wild-type mCRC relied on estimates from four papers, of which only one provided utility values based on RAS wild-type mutation status (Bennett *et al*. 2011[46]). The baseline utility score in the study by Bennett *et al*. 2011 for KRAS wildtype colorectal cancer was 0.778, which is similar to the SoC PFS

utility value from KEYNOTE-177.[46] As such, the ERG considers the available evidence doesn't support a difference in HRQoL if a patient has RAS wild-type mCRC.

4.2.8 Resource use and costs

The company included the following costs in the economic model: drug acquisition costs, administration costs, disease management costs, adverse reaction costs, subsequent treatment costs and terminal care costs. The details for each of these are given in the following subsections.

Unit costs used in the model were inflated to 2018/2019 prices using the PSSRU hospital and community health services pay and prices index.[47]

4.2.8.1 *Drug acquisition costs*

Pembrolizumab is a fixed dose drug given as a 200mg infusion in two 100ml vials every three weeks. The list price per vial is £2,630, and the total cost of a dose is £5,260. There is currently a patient access scheme (PAS) discount in place for pembrolizumab of [redacted], therefore the price per dose of the intervention drug is [redacted]. Pembrolizumab can also be given as a 400mg infusion every 6 weeks and assuming the same discount and usage of 100mg vials, the drug cost of the 400mg dosage is [redacted].

The comparator drugs are dosed according to the patients' body surface area or body mass, depending on the drug. The company details eight standard of care drug dosing schedules as options (please refer to Section 4.2.3 for the dosing regimens). Details of the costs and frequencies are presented in Table 47.

Table 47. Drug acquisition costs (taken from the economic model)

Drug treatment	Total cost	Frequency
mFOLFOX6	£35.51	Cost per 2 weeks
FOLFIRI	£39.85	Cost per 2 weeks
CAPOX	£15.15	Cost per week (applied ever 3 weeks)
	£0.66	Cost per week (applied 2 weeks on 1 week off)
Cetuximab + mFOLFOX6	£35.51	Cost per 2 weeks

Cetuximab + FOLFIRI	£39.85	Cost per 2 weeks
	£1289.44	One-off cost
	£805.90	Cost per 1 week
Panitumumab + FOLFOX6	£1643.31	Cost per week, applied every 2 weeks

Abbreviations: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin

Drug acquisition costs are taken from eMIT (electronic market information tool)[40] and MIMS (Monthly Index of Medical Specialities)[39] databases.

4.2.8.2 *Relative dose intensity*

The company accounted for the relative dose intensity (RDI) of each of the treatment regimens to factor in doses that were not received and therefore did not incur costs. The RDIs were calculated as the percentage of planned doses that were administered and these were then multiplied by the drug acquisition costs per cycle. Details of the RDI are shown in Table 64 of the CS. The RDI for pembrolizumab was higher than for SoC (96.5% vs 88.6%) and reflects the more intensive and toxic nature of SoC treatments. For instance, in KEYNOTE-177 the adverse event profile of SoC was higher than pembrolizumab both in percentage of patients affected (56.2% for pembrolizumab vs 77.6% for SoC) and by mean number of episodes (2.64 for pembrolizumab vs 3.42 for SoC). Vial sharing was assumed in the SoC arm. It was not considered for pembrolizumab as it is a flat dose for all patients.

4.2.8.3 *Time on treatment*

Time on treatment (ToT) was modelled for pembrolizumab and SoC separately using data from KEYNOTE-177. For the pembrolizumab arm, the company used ToT KM data from the trial and included a 35-cycle stopping rule (reflective of the stopping rule in KEYNOTE-177), while for the SoC arm, KM ToT data were extrapolated using an exponential distribution (Figure 39 of the CS). Details of the company's methods for survival analysis and model selection can be found in Section 4.2.5. For non-trial comparators (CAPOX and panitumumab combination treatment), the company assumed that ToT was equal to PFS. Mean time on treatment for each arm of the model is shown in Table 48.

Table 48. Mean time on treatment (taken from the economic model)

Arm	Mean time on treatment (weeks)
Pembrolizumab	[redacted]
SoC	[redacted]
CAPOX	63.1
Panitumumab + mFOLFOX6	80.3

Abbreviation: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin SoC, Standard of Care

In KEYNOTE-177, >90% of pembrolizumab patients discontinued treatment at around 2 years (see Figure 38 in CS and in Figure 6). Only around 10% of patients were still on pembrolizumab after 2 years, and all had discontinued by 3 years. Based on the extrapolation, most patients on SoC are estimated to discontinue treatment after 4 years.

A scenario exploring the cost of retreatment with pembrolizumab was provided by the company at the clarification stage to capture the small proportion of patients (2.6%) who had treatment for longer than 35 cycles (presented in Section 5.1.2.2). The cost of treatment for pembrolizumab increased by [redacted] under this scenario.

4.2.8.4 *Administration costs*

Administration costs for the infusion of pembrolizumab were based on the simple parenteral chemotherapy at first attendance cost code (SB12Z) from NHS reference costs 2018-19, as has been used for previous submissions to NICE.[13, 38] The unit cost of SB12Z is given as £254.14.

The company assumed the number of treatment days for SoC and panitumumab combination treatment as three days and for CAPOX was one day. Table 49 presents a summary of the comparator administration costs included in the economic model. The ERG notes that administration costs for oral capecitabine were only applied in the first model cycle.

Table 49. Comparator administration costs

Treatment	Number of days of treatment per cycle	Unit cost	Cost description
Pembrolizumab	1	£254.14	NHS reference cost code SB12Z: Deliver Simple Parenteral Chemotherapy at First Attendance, Daycase and Reg Day/Night.[38]
SoC	3	£385.25	Day 1. NHS reference cost code SB14Z - Deliver complex chemotherapy, including prolonged infusion treatment at first attendance, Daycase and Reg Day/Night.[38]
		£362.35	Day 2 & 3. NHS reference cost code SB15Z - Deliver Subsequent Elements of a Chemotherapy Cycle, Daycase and Reg Day/Night.[38]
CAPOX	1	£385.25	Oxaliplatin IV infusion. NHS reference cost code SB14Z - Deliver complex chemotherapy, including prolonged infusion treatment at first attendance, Daycase and Reg Day/Night.[38]
		£216.08	Oral capecitabine. NHS reference cost code SB11Z - Deliver Exclusively Oral Chemotherapy, Daycase and Reg Day/Night.[38]
Panitumumab +mFOLFOX6	3	£385.25	Day 1. NHS reference cost code SB14Z - Deliver complex chemotherapy, including prolonged infusion treatment at first attendance, Daycase and Reg Day/Night.[38]
		£362.35	Day 2 & 3. NHS reference cost code SB15Z - Deliver Subsequent Elements of a Chemotherapy Cycle, Daycase and Reg Day/Night.[38]

Abbreviations: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; IV, intravenous; NHS, National Health Service; SoC, standard of care

4.2.8.5 *Disease management costs*

A SLR was undertaken to identify resource use evidence for the dMMR/MSI-H population (see Section 4.1); however, most of the resource use assumptions were obtained from TA439.[13] The unit costs are summarised in Table 50.

Table 50. Disease management costs (adapted from Table 68 of the company submission)

Resource	Unit Cost	Pre-progression monthly units	Post-progression monthly units	Total monthly cost
Consultant outpatient cost	£187	2.17	0	£405.79
Tumour marker test	£14	0.25	0	£3.50
Liver function test	£29	1.25	0	£36.25
CT scan	£116	0.33	0	£38.28
MRI scan	£206	0.25	0	£51.50
Best supportive care	£1600	0	1	£1600

Abbreviations: CT, Computer Tomography; MRI, Magnetic Resonance Imaging.

The company submission does not include the cost of testing for dMMR and MSI-H status as these are routinely performed in UK clinical practice, as per NICE guidance DG27.[14]

4.2.8.6 *Adverse events costs*

The unit costs of treating AEs are given in Table 70 of the company submission. Unit costs were derived from the NHS Reference Costs schedule 2018-19.[38] The company stated that guidance on appropriate AE unit costs was informed by previous submissions for pembrolizumab. In the model these costs are calculated first as a weekly cost for each of the 22 subtypes of adverse event, multiplied by the normalised weekly risk of each adverse event (please refer to Section 4.6.2), and then multiplied by the average length of treatment to get a total cost of adverse events that is then applied as a one-off cost to the first cycle of the model. Table 51 presents the total costs of AEs for each comparator considered in the economic model.

Table 51. Adverse event costs

Treatment	Weekly adverse event cost	Mean treatment duration in months	Total adverse event cost (as one-off cost)
Pembrolizumab	£5.85	13.3	£337.74
SoC	£21.38	8	£743.77
CAPOX	£16.69	8	£580.55

Panitumumab + mFOLFOX6	£48.76	8	£1696.30

Abbreviations: CAPOX, capecitabine plus oxaliplatin; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin SoC, Standard of Care

4.2.8.7 *Subsequent treatment costs*

Upon progression, patients in the model are assigned second-line treatment costs. The base case model assumes that there is no difference in subsequent treatment distribution between the arms of the model after patients experience disease progression. Based on data from KEYNOTE-177, 53.7% of pembrolizumab patients received second-line therapy compared to 83.2% of SoC. The treatment distribution used in the model is based on either estimate from the company's experts of clinical practice in the UK. The proportion from the KEYNOTE-177 trial was explored in scenario analysis requested by the ERG during the clarification stage (presented in Section 5.1.2.2). Details of the second-line treatment scenarios are shown in Table 52.

Table 52. Proportion of patients on subsequent treatment costs (adapted from Table 57 and 58 of the company submission)

Second line treatment	Company's clinical experts estimate (both arms)	KEYNOTE-177	
		Pembrolizumab	SoC
No second line treatment	46.3%	46.3%	16.8%
mFOLFOX6	0.0%	13.2%	6.9%
FOLFIRI	37.6%	10.0%	8.2%
Cetuximab + mFOLFOX6	0.0%	1.6%	0.0%
Cetuximab + FOLFIRI	16.1%	1.3%	0.0%
Bevacizumab + FOLFOX6	0.0%	15.6%	30.9%
Bevacizumab + FOLFIRI	0.0%	11.9%	37.1%
Pembrolizumab	0.0%	0.0%	0.0%

Abbreviations: CAPOX, capecitabine plus oxaliplatin; FOLFIRI, folinic acid plus fluorouracil plus irinotecan; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin

The unit cost of second-line treatment is assumed to be the same for patients in both the pembrolizumab and the standard of care arms at £8,449. This is split between drug acquisition costs of £2,577 and administration costs of £5,872 across 20 weeks.

The total subsequent treatment costs for pembrolizumab and standard of care in the two different approaches is given in Table 53.

Table 53. Subsequent therapy costs

Subsequent therapy scenario	KEYNOTE-177 estimate	Clinical expert estimate
Pembrolizumab costs	£7,913	£8,305
SoC costs	£12,941	£8,086

Abbreviations: SoC, standard of care.

4.2.8.8 *Terminal care costs*

The company did not assume a separate cost for terminal care in their base case, instead choosing to base post-progression resource on 'Best Supportive Care' costs from estimated from Färkkilä *et al.* 2015[48]. However, the company did include an option in the model to include a terminal care cost based on Round *et al.* 2015.[49] The terminal care cost option is £5,156.50, a proportion of which is allocated to each cycle the patient is in the progressed health state. This is compared to the £1,600 monthly cost for best supportive care, which is the direct health costs of being in a 'palliative' health state from Färkkilä *et al.* 2015 made up of outpatient, inpatient and travel costs.[48] Best Supportive Care costs are applied for the whole time the patient is in the progressed health state within the model.

4.2.8.9 *ERG Critique*

The ERG considers the company's methods regarding the estimation of unit costs and resource use to be generally reasonable. The ERG considers the administration costs presented in the model to be acceptable. In the CS, the administration cost for the intervention reflects an outpatient simple cost and the comparator a complex day case, based on the treatment days in a cycle of between 1 and 3 days for the comparator and half a day for pembrolizumab. The company's approach reflects the time and complexity required to administer each treatment and the difference in intensity between the intervention and compactor regimens. Furthermore, the ERG considers the decision to not include testing for dMMR or MSI-H status to be correct, as these tests are routinely performed in the NHS.

In KEYNOTE-177, the treatment regimen for pembrolizumab was 200mg once every three weeks, but the company stated that it can also be administered at a 400mg dose once every six weeks. The ERG

clinical experts advised that in clinical practice, they would prefer to administer pembrolizumab using the 400mg dose once every six weeks for patient convenience and also to reduce NHS resource use. The ERG requested that the company to provide a scenario exploring the less frequent dosing regimen for pembrolizumab at the clarification stage. The company provided the results of the ERG requested scenario (presented in Section 5.1.2.2), which improved the ICERs, but not substantially. The company expects that the monotherapy licence will include an option to administer pembrolizumab at a 400mg dose once every six weeks. As such, the ERG considers includes the use of the less frequent treatment regimen for pembrolizumab in its preferred analysis, as that is likely to be how it is used in UK clinical practice.

Related to the less frequent treatment regimen for pembrolizumab is the frequency of consultant oncologist appointments, which is a primary driver of cost-effectiveness. In the company's base case, consultant oncologist appointments occurred according to the SoC treatment cycle (every 2 weeks), even when the treatment was given less frequently (once every three weeks), as was the case for pembrolizumab and CAPOX. The ERG asked at the clarification stage for a scenario analysis where consultant oncologist appointments and liver function tests were included once per cycle of treatment and this was implemented by the company. The ERG made a minor correction to the company's resource use scenario as its implementation affected costs for comparators with regimens that were once every two weeks. For the scenario a formula was used to derive the number of appointments per month, resulting in a slightly higher estimation (~2.174). However, in the company's base case, consultant oncologist appointment per month were hardcoded in the economic model as 2.17, which is just a rounded figure of the calculation, which the ERG used to correct the company's scenario.

The results of the ERG corrected one-way sensitivity analysis show that the model is sensitive to the assumption around the frequency of consultant oncologist appointments in particular, with the ICER changing from £7,250 to £4,043 for the comparison with SoC and from £27,474 to £25,275 for the comparison with CAPOX, but pembrolizumab still dominates panitumumab combination treatment under this scenario. Further details of the resource use scenario are presented in Section 5.1.2.2. The ERG also ran another scenario to align consultant oncologist appointments to be used in conjunction with the scenario exploring the pembrolizumab treatment regimen of 400mg once every six weeks, which reduce the ICER further to £535 for the comparison with SoC, £20,736 for CAPOX, with pembrolizumab still dominant against panitumumab combination treatment. The ERG believes that oncologist appointments and liver function tests aligned to treatment cycle is more reflective of

clinical practice in the UK, but only consultant oncologist appointments had a substantial impact on the ICER and as such has included in its base case, further details of which can be found in Section 6.4.

As mentioned previously, in KEYNOTE-177 approximately 70% of patients received bevacizumab combination treatment, which is not recommend in the NHS.[41] The company assumed that costs for patients who received bevacizumab combination treatment would be reflective of cetuximab combination treatment. However, cetuximab combination treatment is only recommend in the NHS for patients with RAS-wildtype mCRC. As such, for the ERG's preferred ITT and RAS wildtype analysis (population A and B),which focuses only on the comparison with FOLFOX and FOLFIRI (SoC), the ERG ran a scenario where treatment costs for standard of care are based only on FOLFOX and FOLFIRI treatments – this increased the ICER from £7,250 to £21,636 (See Section 6.3 for further details). The ERG notes that the results of this scenario should be interpreted with caution as the clinical efficacy of SoC includes the efficacy of bevacizumab and as such breaks the alignment of costs with observed efficacy in KEYNOTE-177. The ERG considers it is important to highlight that SoC treatment costs for the company base case are higher than may be incurred in UK clinical practice and as such, the scenario can be considered conservative.

Another key issue with the costs used in the model, is the assumption of ToT used for the non-trial comparators, CAPOX and panitumumab combination treatment. In the base case analysis, the company assumed ToT to be equal to estimated PFS for CAPOX and panitumumab combination treatment, in the absence of alternative data. As presented in Table 48, mean ToT for CAPOX and panitumumab is substantially longer compared with pembrolizumab and SoC, which results in inflated treatment costs. In TA439, mean treatment duration for panitumumab combination treatment was estimated to be approximately nine months, which is shorter than the company's estimated mean treatment duration of 18.5 months based on ToT equal to PFS.[13] As such, the ERG considers that a more appropriate assumption for ToT for non-trial comparators is to assume it is equal to KEYNOTE-177 ToT for SoC. Mean ToT for SoC was estimated to be approximately [redacted] months, which is closer to the estimates in TA439 for panitumumab combination treatment. Furthermore, given the company are assuming clinical outcomes for CAPOX are equal to SoC, it is not unreasonable to assume ToT is equal to ToT for SoC. The ERG performed a scenario for the non-trial comparators, where ToT is equal to ToT for SoC and this had a negligible impact on the ICER for CAPOX, but changed the ICER for panitumumab combination treatment from being dominant to

£3,158 and as such is included in the ERG's preferred analysis. Further details of this scenario can be found in Section 6.3 and 6.4.

The ERG notes that the estimation of treatment costs for pembrolizumab is based on KM data from KEYNOTE-177 and the company has also included a 35-cycle stopping rule to reflect the maximum number of cycles of treatment patients could receive in the trial. The ERG is satisfied with the approach taken in the pembrolizumab arm of the model, but highlights that as per the draft SmPC, [redacted]. As such, the ERG has run an illustrative scenario, which assumes ToT is equal to PFS for pembrolizumab and removed the 35-cycle stopping rule, presented in Section 6.3.

Subsequent treatment costs for patients with progressed disease were calculated as a single 20-week post-progression treatment cost applied as a one-off cost upon progression. The ERG considers this assumption is appropriate as the ERG's clinical experts advised that post-progression survival for patients with mCRC is limited. However, the distribution of subsequent treatment costs was split between FOLFIRI and cetuximab + FOLFIRI treatments. NICE guidance does not recommend cetuximab for use as a monotherapy or in combination therapies as second line treatment for mCRC, thus the ERG has run a scenario where cetuximab + FOLFIRI has been removed from the subsequent treatment costs.[50] The ERG removed the proportion of patients receiving second-line cetuximab + FOLFIRI (16.1%) and added this to the proportion of patients assumed to receive FOLFIRI (37.6%+16.1%). The ERG's subsequent treatment scenario reduces the total cost of the subsequent treatment cost to £5,509 from £8,449, but has limited impact on all ICERs, as presented in Section 6.3.

The ERG is satisfied with the company's approach for relative dose intensity, adverse event costs and terminal care costs. Best supportive care costs were tested as part of the model validation process and the model was not found to be sensitive to extremes in these values.

5 Cost effectiveness results

5.1.1 Company's cost effectiveness results

The results of the company's base-case analysis are given in Table 54, showing an incremental cost-effectiveness ratio (ICER) of £7,250 per QALY gained for pembrolizumab versus standard of care (SoC). For the comparison with CAPOX, the cost per QALY was £27,474 (Table 55), and pembrolizumab dominates panitumumab + mFOLFOX6 (Table 56). The results include the company's agreed patient access scheme (PAS) discount for pembrolizumab of [redacted] on the list price. The company's fully incremental cost-effectiveness analysis is presented in Table 57.

The ERG notes that the ICERs for CAPOX presented in the company's clarification responses were incorrect when compared with the economic model submitted with the response. The company confirmed that on the day of submission of the clarification response, an amendment was made to the economic model. Therefore, the ICERs for CAPOX presented in the ERG report are obtained from the company's economic model submitted with the clarification response.

Table 54. Company's base case deterministic results: pembrolizumab versus standard of care

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Standard of Care	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	13,497	3.145	1.862	7,250

Abbreviations: LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 55. Company's base case deterministic results: pembrolizumab versus CAPOX

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
CAPOX	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	50,968	3.145	1.855	27,474

Abbreviations: CAPOX, capecitabine plus oxaliplatin; LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 56. Company's base case deterministic results: pembrolizumab versus panitumumab + mFOLFOX6

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Panitumumab + mFOLFOX6	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	-48,317	2.825	1.688	Dominant

Abbreviations: mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 57. Company's base case deterministic fully incremental analysis

Technologies	Total costs (£)	Total QALYs	Total LYs	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)
CAPOX	[redacted]	[redacted]	[redacted]	-	-	-
SoC	[redacted]	[redacted]	[redacted]	37,472	-0.01	Dominated
Pembrolizumab	[redacted]	[redacted]	[redacted]	50,968	1.86	27,379
Panitumumab + mFOLFOX6	[redacted]	[redacted]	[redacted]	48,317	-1.68	Dominated

Abbreviations: CAPOX, capecitabine plus oxaliplatin; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

5.1.2 Company's sensitivity analyses

The company conducted a probabilistic sensitivity analysis (PSA) based on 5,000 samples. The results for pembrolizumab against SoC, CAPOX and mFOLFOX6 + panitumumab are shown in Table 58 to Table 60. Table 62 presents the company's probabilistic fully incremental analysis. The cost-effectiveness planes for these results are given in Figure 10 to Figure 12. It should be noted that the company only explored variation around clinical efficacy for the estimates derived from the proportional hazards network meta-analysis, which is not used for the base case, resulting in similar estimates of life years between the deterministic and PSA results. As such, the ERG considers that the uncertainty around clinical efficacy has not been fully explored in the PSA.

Table 58. Company's PSA results: pembrolizumab versus standard of care

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Standard of Care	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	18,199	3.133	1.858	9,795

Abbreviations: LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 59. Company's PSA results: pembrolizumab versus vs CAPOX

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
CAPOX	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	55,820	3.133	1.852	30,143

Abbreviations: CAPOX, capecitabine plus oxaliplatin; LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 60. Company's PSA results: pembrolizumab versus panitumumab + mFOLFOX6

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Panitumumab + mFOLFOX6	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	-43,910	2.81	1.684	Dominant

Abbreviations: mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin; LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

Table 61. Company's base case probabilistic fully incremental analysis

Technologies	Total costs (£)	Total QALYs	Total LYs	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)
CAPOX	[redacted]	[redacted]	[redacted]	-	-	-
SoC	[redacted]	[redacted]	[redacted]	37,620	-0.006	Dominated
Pembrolizumab	[redacted]	[redacted]	[redacted]	18,199	1.858	9,795
Panitumumab + mFOLFOX6	[redacted]	[redacted]	[redacted]	43,910	-1.684	Dominated

Abbreviations: CAPOX, capecitabine plus oxaliplatin; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

Figure 10. Cost-effectiveness plane: pembrolizumab versus standard of care

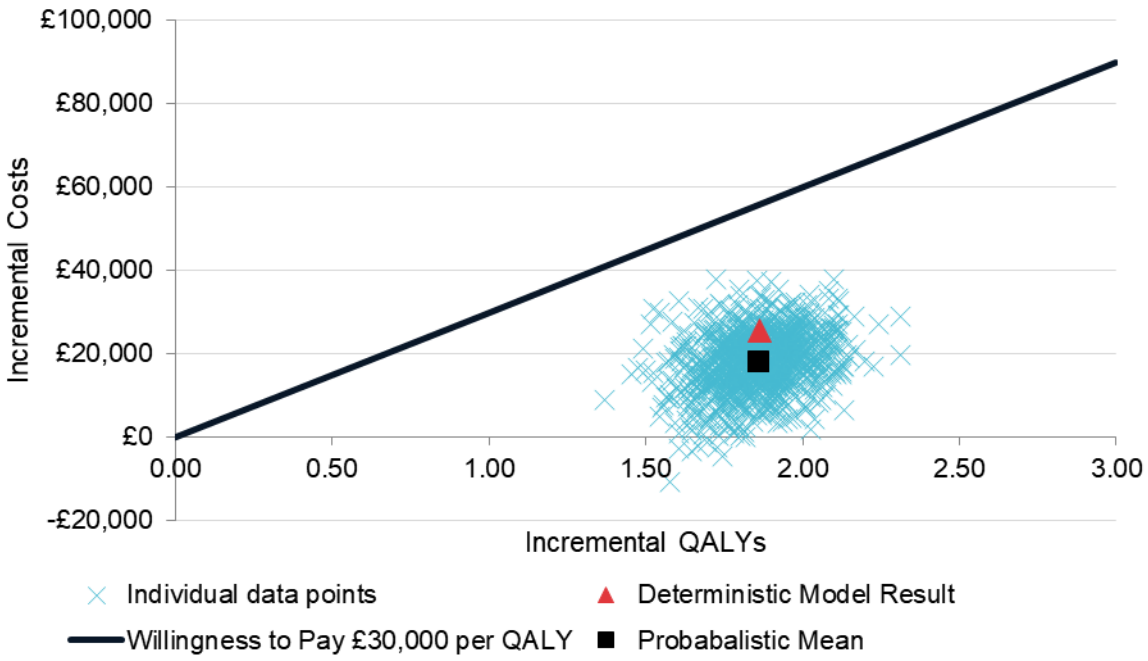


Figure 11. Cost-effectiveness plane: pembrolizumab versus CAPOX

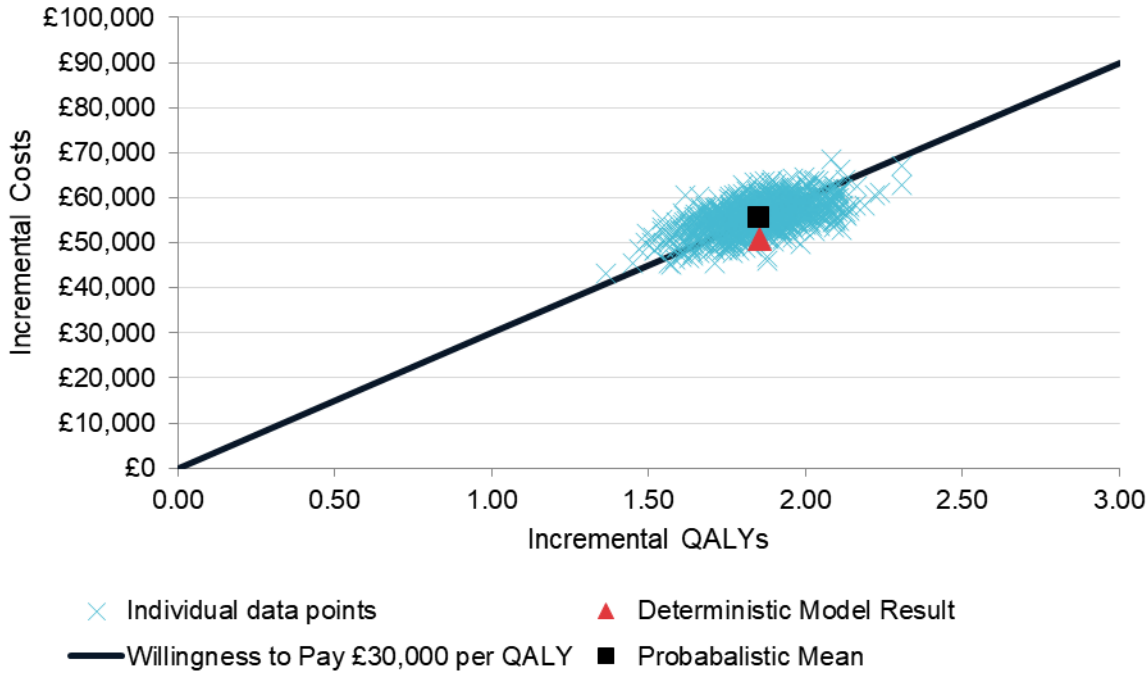
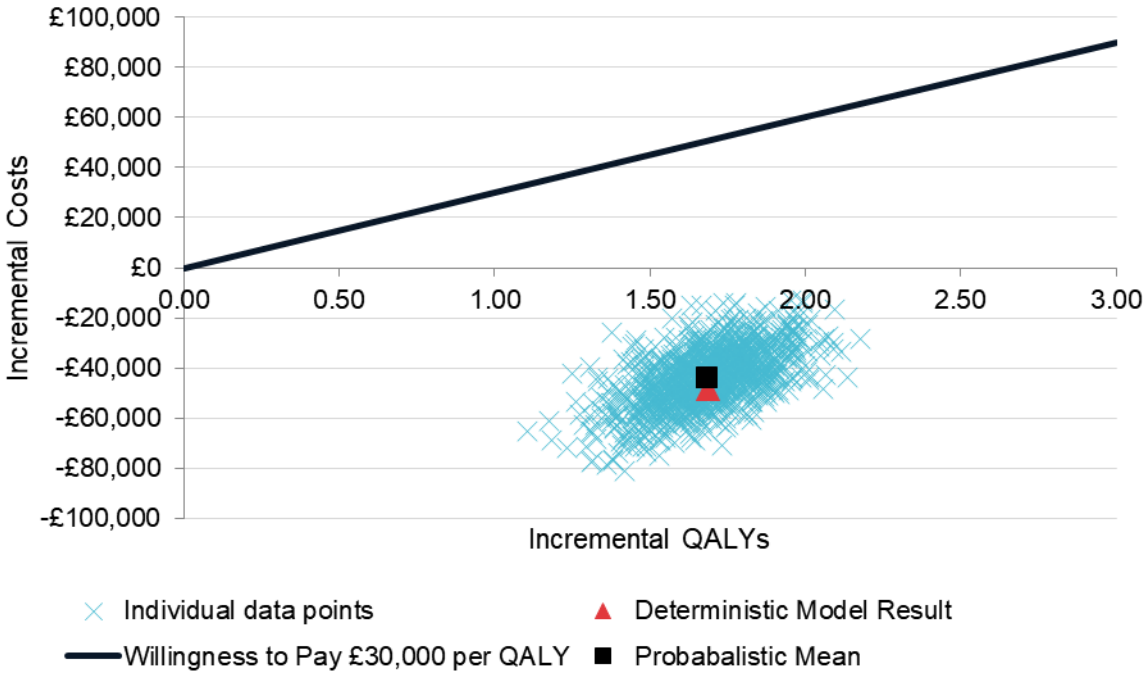


Figure 12. Cost-effectiveness plane: pembrolizumab versus panitumumab + mFOLFOX6



5.1.2.1 *One-way sensitivity analysis*

The company conducted a range of one-way sensitivity analyses (OWSAs) to test the impact that plausible changes on parameters have on the overall results. The tornado plot in Figure 13 shows the parameters that had the greatest impact, with the Best Supportive Care (BSC) costs having the largest impact with ICERs ranging between £3,575 and £10,925.

Consultant treatment had the biggest impact on the ICERs for the CAPOX comparator, with values between £26,176 and £28,772 (Figure 14). For the mFOLFOX6 + panitumumab comparator the main driver of the difference in costs is best supportive care, with ICERs remaining dominated (Figure 15).

Figure 13. OWSA tornado diagram: pembrolizumab versus standard of care

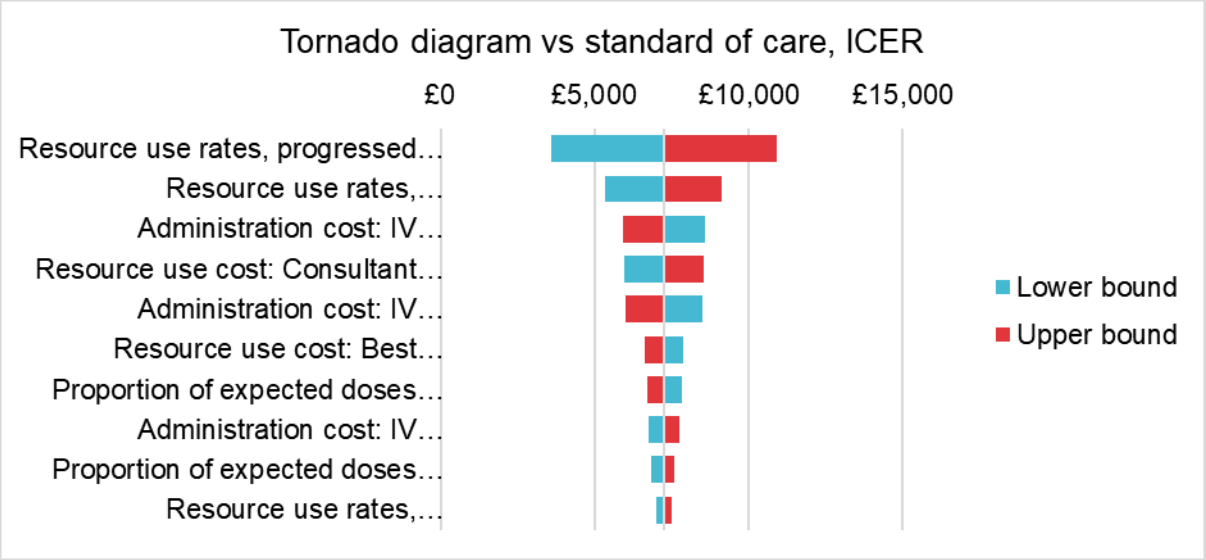


Figure 14. OWSA tornado diagram pembrolizumab vs CAPOX

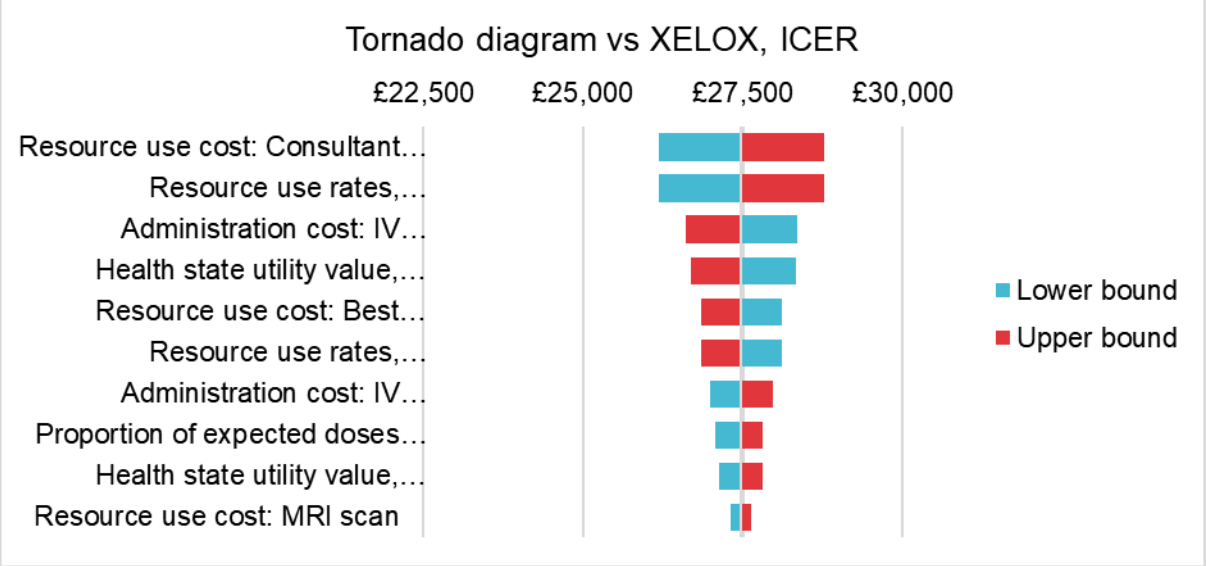
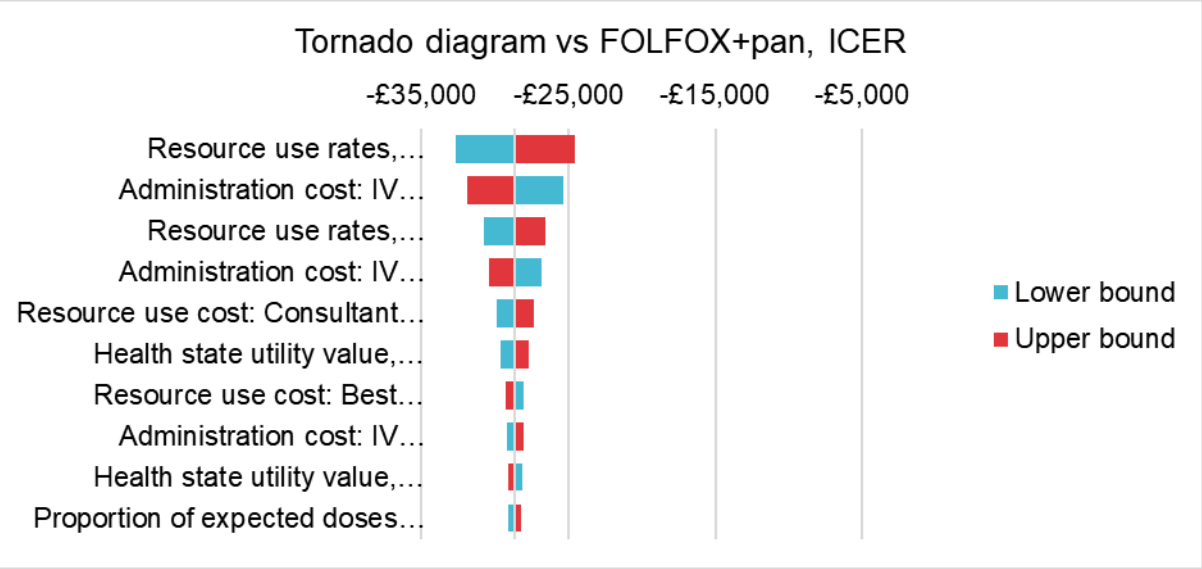


Figure 15. OWSA tornado diagram pembrolizumab vs panitumumab + mFOLFOX6



5.1.2.2 *Scenario analysis*

The company provided a range of scenario analyses around their base case, which are detailed in full in Table 79 of the CS. The impact of the scenario analyses on the ICER for pembrolizumab versus SoC, CAPOX and panitumumab combination treatment is shown in Table 62.

Table 62. Scenario analysis results

Scenario	ICER (£) – pembrolizumab versus.		
	SoC	CAPOX	Panitumumab + mFOLFOX6
Base case	**7,250**	**27,474**	**Dominant**
Company scenarios provided in the company submission			
Time horizon- 30 years	7,115	27,484	Dominant
PH NMA results used	7,250	27,474	Dominant
Time to progression and Progression free survival (pembrolizumab & SoC) -Two-piece Weibull with 10- week cut-off point	6,548	20,971	Dominant
Post progression survival (pembrolizumab & SoC) - Lognormal	6,046	27,419	Dominant
Vial sharing assumed	5,961	27,455	Dominant

Scenarios requested by the ERG at the clarification stage			
Pooled utility values for the progression-free health state from KEYNOTE-177	7,762	29,422	Dominant
Pooled progression-free AE utility value is applied only in the first cycle, after which the progression-free no AE utility value is implemented	7,615	28,860	Dominant
Second phase retreatment with pembrolizumab (as per KEYNOTE-177)	7,976	28,203	Dominant
Subsequent treatment proportions based on KEYNOTE-177	4,467	24,681	Dominant
Pembrolizumab treatment regimen – 400mg once every 6 weeks	6,967	27,190	Dominant
Oncologist outpatient appointment and liver function test aligned to treatment cycle (**ERG corrected scenario**)	4,043	25,275	Dominant

Abbreviations: NMA, network meta-analysis; PH, proportional hazard; SoC, standard of care.

5.1.3 Model validation and face validity check

For the model validation, the company stated that quality control checks were performed by model developers to ensure calculations were correct and consistent with the model specification. In addition, health economists not involved with model development performed quality assurance of the model for coding errors, inconsistencies and plausibility of model parameters and results. The company performed external validation of the model by comparing estimated clinical outcomes from the model against the observed data in KEYNOTE-177 as well as using real world data to validate extrapolations of post-progression survival used in the model.[51] The ERG considers the company's model validation and face validity check to be robust and as such has not identified any errors in the model.

6 Additional economic analysis undertaken by the ERG

6.1 Model corrections

The Evidence Review Group (ERG) did not identify any model errors.

6.2 Exploratory and sensitivity analyses undertaken by the ERG

In Section 4 of this report, the ERG has described several scenarios that warrant further exploration in addition to the company's own sensitivity and scenario analyses to ascertain the impact of these changes on the incremental cost-effectiveness ratio (ICER). The deterministic scenarios the ERG has produced are applied to the company's revised base case analysis for the ITT population, with the underlying assumption that the ITT analyses are a proxy for the RAS wild-type subgroup for FOLFOX, FOLFIRI, CAPOX and panitumumab combination treatment. A separate scenario, estimating the cost-effectiveness of pembrolizumab versus cetuximab combination treatment is provided in Table 64. The scenarios that the ERG has performed are as follows:

1. Implementing the Weibull distribution after the 20-week cut-off point for the extrapolation of time to progression (TTP) and progression-free survival (PFS) – Section 4.2.5.1

2. Time to death utility values – Section 4.2.7.3

3. Removal of AE disutility – Section 4.2.7.3

4. Using only mFOLFOX6 (50%) and FOLFIRI (50%) costs for SoC treatment costs – Section 4.2.8.9

5. Assuming time on treatment (ToT) for panitumumab and CAPOX is equal to ToT for standard of care (SoC) – Section 4.2.8.9

6. Exploring ToT for pembrolizumab equal to PFS and removing the 35-cycle stopping rule. The draft SmPC states pembrolizumab [redacted], but was restricted to 35 cycles in KEYNOTE-177 – Section 4.2.8.9

7. Removal of subsequent cetuximab combination treatment – Section 4.2.8.9

8. Less frequent treatment regimen and consultant oncologist appointments for pembrolizumab (400mg once every six weeks) based on advice from ERG's clinical experts – Section 4.2.8.9

9. RAS wild-type subgroup analysis – cetuximab combination treatment is clinical equivalent to panitumumab combination treatment– Sections 4.2.2, 4.2.3 and 4.2.5.1.

6.3 ERG scenario analysis

Table 63 presents the results of the ERG exploratory analyses described in Section 6.2. Results reported include the company's proposed patient access scheme (PAS) of [redacted].

Table 63. Results of the ERG's scenario analyses – ITT/RAS wild-type population

	Results per patient	Pembrolizumab (1)	SoC (2)	CAPOX (3)	Panitumumab +mFOLFOX6 (RAS wild-type only) (4)	Incremental value (1-2)	Incremental value (1-3)	Incremental value (1-4)
0	**Company base case**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	13,497	50,968	-48,317
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.86	1.86	1.69
	ICER (£/QALY)					7,250	27,474	Dominant
1	**Weibull distribution applied after 20-week cut-off point for TTP and PFS**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	13,171	50,307	-56,167
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.89	1.88	1.68
	ICER (£/QALY)					6,966	26,698	Dominant
2	**Implementation of time-to-death utility values from KEYNOTE-177**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	13,497	50,968	-48,317
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.71	1.71	1.50
	ICER (£/QALY)					7,896	29,819	Dominant

3	**Removal of AE disutility**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	13,497	50,968	-48,317
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.86	1.86	1.65
	ICER (£/QALY)					7,251	27,383	Dominant
4	**FOLFOX and FOLFIRI costs used for SoC**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	40,278	50,968	-48,317
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.86	1.86	1.69
	ICER (£/QALY)					21,636	27,474	Dominant
5	**ToT for CAPOX and panitumumab + mFOLFOX6 equal to ToT for SoC**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	13,497	54,180	5,330
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.86	1.86	1.69
	ICER (£/QALY)					7,250	29,205	3,158
6	**ToT for pembrolizumab equal to PFS and removal of 35-cycle stopping rule**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	137,400	17,4871	75,586
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.86	1.86	1.69
	ICER (£/QALY)					73,809	94,262	44,777

7	**Removal of second-line cetuximab combination treatment**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	13,911	51,383	-47,946
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.86	1.86	1.69
	ICER (£/QALY)					7,473	27,697	Dominant
8	**Treatment regimen and consultant outpatient appointments for pembrolizumab – 400mg once every six weeks**							
	Total costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	996	38,468	-60,817
	QALYs	[redacted]	[redacted]	[redacted]	[redacted]	1.86	1.86	1.69
	ICER (£/QALY)					535	20,736	Dominant

Abbreviations: ICER, incremental cost-effectiveness ratio; PFS, progression-free survival; QALY, quality adjusted life year; TTP, time to progression

Table 64. Scenario 9 - pembrolizumab versus cetuximab combination treatment - RAS wild-type only.

Interventions	Total Costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Cetuximab + mFOLFOX6/FOLFIRI	[redacted]	[redacted]	[redacted]	-	-	-	-
Pembrolizumab	[redacted]	[redacted]	[redacted]	-49,510	2.83	1.69	Dominant

Abbreviations: LYG, Life Years Gained; QALYs, Quality Adjusted Life Years; ICER, incremental cost effectiveness ratio

6.4 ERG preferred assumptions

In this section, the ERG presents its base-case ICER for ITT/ RAS wild-type proxy analyses. For the ERG base case, the assumption of cetuximab combination treatment is clinically equivalent to panitumumab combination treatment in the RAS wild-type subgroup has been included and all relevant assumptions applied. Deterministic results and fully incremental analyses are presented in Table 65 and Table 66 and incorporates the company's patient access scheme (PAS) simple discount of [redacted]. The ERG could not produce PSA ICERs for its base case as the PSA takes several hours to run and due to paucity of time and complexity of the model, some scenarios could not be integrated with the PSA.

Table 65. ERG's preferred model assumptions – ITT/ RAS wild-type population

Preferred assumption	Section in ERG report	Pembrolizumab vs Standard of Care		Pembrolizumab vs CAPOX		Pembrolizumab vs Panitumumab + mFOLFOX6 (RAS wild-type only)		Pembrolizumab vs Cetuximab + FOLFIRI/ mFOLFOX6 (RAS wild-type only)	
		ICER (£/QALY)	Cumulative ICER (£/QALY)	ICER (£/QALY)	Cumulative ICER (£/QALY)	ICER (£/QALY)	Cumulative ICER (£/QALY)	ICER (£/QALY)	Cumulative ICER (£/QALY)
Company base case	-	7,250	-	27,474	-	Dominant	-	-	-
Cetuximab combination treatment is clinical equivalent to panitumumab combination treatment	*4.2.2, 4.2.3 and 4.2.5.1*	-	-	-	-	-	-	Dominant	-
Treatment regimen and consultant outpatient appointments for pembrolizumab – 400mg once every six weeks	*4.2.3 and 4.2.8.9*	535	535	20,736	20,736	Dominant	Dominant	Dominant	Dominant
Implementing the Weibull distribution after the 20-week cut-off point for the extrapolation of TTP and PFS	*4.2.5.1*	6,966	164	26,698	19,872	Dominant	Dominant	Dominant	Dominant
Removal of AE disutility	*4.2.7.3*	7,251	164	27,383	19,808	Dominant	Dominant	Dominant	Dominant

ToT for comparator is equal to ToT for standard of care	*4.2.8.9*	-	164	29,205	21,684	3,158	Dominant	2,852	Dominant
FOLFOX/FOLFIRI costs for SoC	*4.2.8.9*	21,636	14,330	-	21,684	-	Dominant	-	Dominant
Removal of second-line cetuximab combination treatment	*4.2.8.9*	7,473	14,569	27,697	21,923	Dominant	Dominant	Dominant	Dominant
ERG preferred ICER	**-**	**-**	**14,569**	**-**	**21,923**	**-**	**Dominant**	**-**	**Dominant**

Abbreviations: ERG, evidence review group; ICER, incremental cost-effectiveness ratio; QALY, quality adjusted life year

Table 66. ERG's base case deterministic fully incremental analysis

Technologies	Total costs (£)	Total QALYs	Total LYs	Inc. costs (£)	Inc. QALYs	ICER (£/QALY)
CAPOX	[redacted]	[redacted]	[redacted]	-	-	-
Standard of care	[redacted]	[redacted]	[redacted]	13,902	0.00	Dominated
Pembrolizumab	[redacted]	[redacted]	[redacted]	41,443	1.89	21,923
Panitumumab + mFOLFOX6	[redacted]	[redacted]	[redacted]	7,675	-1.65	Dominated
Cetuximab + mFOLFOX6/FOLFIRI	[redacted]	[redacted]	[redacted]	8,191	-1.65	Dominated

Abbreviations: CAPOX, capecitabine plus oxaliplatin; mFOLFOX6, folinic acid plus fluorouracil plus oxaliplatin ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years

6.5 Conclusions of the cost effectiveness sections

Overall, the case made by the company to demonstrate the cost-effectiveness of pembrolizumab compared with relevant comparators based on the intention-to-treat (ITT) population of the primary trial, KEYNOTE-177, is considered by the ERG to be generally robust. In KEYNOTE-177, progression-free survival (PFS) outcomes for pembrolizumab demonstrate a statistically significant improvement compared with standard of care (SoC) in the trial. The ERG notes that standard of care in KEYNOTE-177 comprised mFOLFOX6 (folinic acid plus fluorouracil plus oxaliplatin), FOLFIRI (folinic acid plus fluorouracil plus irinotecan), cetuximab in combination with mFOLFOX6 or FOLFIRI, and bevacizumab in combination with mFOLFOX6 or FOLFIRI. Over 70% of SoC patients in KEYNOTE-177 received bevacizumab combination treatments, but NICE guidance does not recommend its use for treating mCRC.[41, 42] However, bevacizumab combination regimens are likely to be more effective than FOLFOX and FOLFIRI alone. As such, the ERG considers that receiving treatment with bevacizumab is likely to be biased against pembrolizumab, resulting conservative estimates of treatment effectiveness.

One of the ERG's primary concerns with the company's approach to the cost-effectiveness analysis is with the population of the economic model. The ERG considers that in addition to the analysis using the ITT population, the company should have presented subgroup analyses by RAS mutation status. The NICE final scope specifies that, "If the evidence allows the subgroup of people with RAS wild-type colorectal cancer will be considered".[1] Treatment options in the UK differ if a patient has RAS wild-type mCRC. According to the NICE pathway for managing mCRC, first-line biological therapy (cetuximab or panitumumab combination treatments) are only recommended for patients with RAS wild-type mCRC.[42] Furthermore, subgroup analyses reported within the company submission (CS) indicate that, in those with RAS wild-type, pembrolizumab is associated with a statistically significant improvement in PFS compared with SoC (hazard ratio [HR] 0.44, 95% confidence interval [CI]: 0.29 to 0.67). However, in the non-RAS wild-type subgroup, the direction of effect favours SoC, albeit that the benefit with SoC would not be considered statistically significant (HR 1.19, 95% CI: 0.68 to 2.07). The ERG notes that the 95% CIs for the two subgroups do not overlap and, as such, considers the difference in the two subgroups unlikely to be due to random chance. The ERG concludes that the company should have presented cost-effectiveness analyses by RAS mutation status.

In lieu of the RAS mutation subgroup analyses, the ERG considers that using the ITT analyses as a proxy for the RAS wild-type population may provide a conservative estimate of the cost-

effectiveness of pembrolizumab, but for the non-RAS wild-type subgroup, the ERG predicts that pembrolizumab is less effective and more costly that SoC and as such would be dominated by the comparator. With regards to cetuximab combination treatment, in the company's current analysis, it is blended with SoC, but approximately 10% of SoC patients in KEYNOTE-177 received this treatment. As such, the ERG considers that a separate analysis for cetuximab combination treatment is appropriate, as it is only recommended for patients with RAS wild-type mCRC. TA439 reports an NMA that provided no statistically significant evidence to suggest that cetuximab plus FOLFOX was more effective than panitumumab plus FOLFOX at increasing PFS or OS and as such, the ERG implemented a scenario assuming cetuximab and panitumumab combination treatment are clinically equivalent to provide an illustrative estimate of cost-effectiveness. However, the ERG reiterates that its preferred approach for the subgroup analyses by RAS mutation status is for the company to:

- Produce relative estimates of treatment effectiveness for patients with and without RAS wildtype mCRC from KEYNOTE-177 using the recommended ERG FP NMA outlined in the ERG's clarification questions;
- Implement the results of the subgroup treatment effectiveness analyses in the economic model to produce cost-effectiveness estimates by relevant comparator for RAS wild-type and non-RAS wild-type subgroups.

Setting aside the issue of the subgroup analyses by RAS-mutation status and appropriate comparators, the ERG interrogated the company's ITT state-transition model (STM). The structure of model was considered appropriate, capturing all relevant health states and clinically plausible transitions between health states that are largely similar to other oncology models submitted for NICE appraisal. With regards to the clinical transitions between the health states in the model, the ERG notes that as post-progression survival (PPS) is used in lieu of mature OS data from KEYNOTE-177, any gain in PFS directly equates to a gain in the estimated OS, which becomes more significant as the company has assumed that PPS for all comparators is equal to PPS for pembrolizumab.

The ERG considers the company's general approach to extrapolating outcomes for TTP, PFS and PPS to be appropriate. The use of piecewise models, with a cut-off point of 20 weeks for TTP and PFS appropriately captures the change in the hazard observed in log-cumulative hazard plots presented in the CS. However, the ERG questioned the use of an exponential model after the 20-week cut-off point for the TTP and PFS extrapolations for SoC, as the log-cumulative hazard plots indicate

increasing hazards. As such, use of the Weibull model after the 20-week cut-off point was considered more appropriate.

For their base case, the company assumed that PPS for all comparators was equal to PPS for pembrolizumab. The company's PPS assumption implies that the treatment effect with pembrolizumab is not extended beyond the progression-free health state. The ERG considers that it is not unreasonable that the treatment effect of pembrolizumab is assumed to only be in the progression-free health state as long-term overall survival data are immature, therefore estimating the duration of treatment effect beyond progression is currently problematic.

The ERG notes that the use of PPS for all arms of the model directly impacts OS, as any gain in PFS directly equates to a gain in the estimated OS. There is uncertainty around whether the PFS to OS relationship holds or whether patients on pembrolizumab are likely to have an accelerated mortality rate upon progression compared to patients on other treatments, particularly for RAS wild-type patients treated with biological therapies (panitumumab and cetuximab). However, the ERG notes that PFS and PFS2 for pembrolizumab versus SoC are similar (PFS HR 0.60, 95% CI: 0.45 to 0.80 vs PFS2 HR 0.63, 95% CI: 0.45 to 0.88), though this data are not mature. As such, the ERG considers it may not be unreasonable that gains in PFS are maintained in PFS2, thus providing initial support for the assumption that there isn't an accelerated mortality rate for pembrolizumab upon progression compared with SoC. However, mature OS data from KEYNOTE-177 is required to mitigate the uncertainty around long term survival outcomes.

Health-related quality of life (HRQoL) in the model was based on utility values obtained from the KEYNOTE-177 trial, which the ERG agrees is preferable to estimates derived from the literature. The use of progression-based utility values is appropriate for the base case. However, the ERG considers that the company did not sufficiently explore plausible scenarios for alternative utility values in the CS. Furthermore, the ERG queried the use of treatment specific utilities for the progression-free health state, as the difference between pembrolizumab and SoC was substantial (0.056). However, the ERG's clinical experts supported the use of treatment specific values as pembrolizumab might result in improved quality of life compared with SoC in the progression free state, as it is a monotherapy and as such would require shorter duration and less frequent hospital visits.

The ERG considers the company's methods regarding the estimation of unit costs and resource use to be generally reasonable. However, a key driver of the cost-effectiveness analysis is the frequency

of pembrolizumab treatment and in turn the resource use associated with different regimens. In KEYNOTE-177, the treatment regimen for pembrolizumab was 200mg once every three weeks, but the company stated that it can also be administered at a 400mg dose once every six weeks. The ERG's clinical experts advised that in clinical practice, they would prefer to administer pembrolizumab using the 400mg dose once every six weeks for patient convenience and also to reduce NHS resource use. The company expects that the monotherapy marketing authorisation will include an option to administer pembrolizumab at a 400mg dose once every six weeks. As such, the ERG considers includes the use of the less frequent treatment regimen for pembrolizumab in its preferred analysis, as that is likely to be how it is used in UK clinical practice. Furthermore, the ERG's clinical experts also advised that consultant oncologist appointments would be aligned to treatment cycle. The ERG ran a scenario which combined the pembrolizumab treatment regimen of 400mg once every six weeks with consultant oncologist appointments once every six weeks, which resulted in a substantial reduction in the ICER. The ERG considers that a key benefit of pembrolizumab is around the reduction in the need for frequent treatment resulting in decreased patient burden and NHS resource use savings. However, the ERG considers that the strength of the evidence for pembrolizumab remains around the comparison with SoC/CAPOX, as uncertainties still remain for the comparison with panitumumab and cetuximab combination treatments.

As mentioned previously, in KEYNOTE-177 approximately 70% of patients received bevacizumab combination treatment, which is not recommend in the NHS.[41] The company assumed that costs for patients who received bevacizumab combination treatment would be reflective of cetuximab combination treatment. However, cetuximab combination treatment is only recommend in the NHS for patients with RAS-wildtype mCRC. As such, for the ERG's preferred ITT and RAS wildtype analysis (population A and B),which focuses only on the comparison with FOLFOX and FOLFIRI (SoC), considers that a scenario exploring only including FOLFOX and FOLFIRI costs is important to highlight that SoC treatment costs for the company base case are higher than may be incurred in UK clinical practice. Using only FOLFOX and FOLFIRI costs in a scenario analysis substantially reduced the costs of SoC and as such increased the ICER, resulting in a conservative estimate of the cost effectiveness of pembrolizumab.

Another key issue with the costs used in the model, is the assumption of time on treatment (ToT) equal to PFS for the non-trial comparators, CAPOX and panitumumab combination treatment. Mean ToT for CAPOX and panitumumab is substantially longer compared with pembrolizumab and SoC, which results in inflated treatment costs. In TA439, mean treatment duration for panitumumab

combination treatment was estimated to be approximately nine months, which is shorter than the company's estimated mean treatment duration of 18.5 months based on ToT equal to PFS.[13] As such, the ERG considers that a more appropriate assumption for ToT for non-trial comparators is to assume it is equal to KEYNOTE-177 ToT for SoC (estimated to be approximately [redacted] months), which is closer to the estimates in TA439 for panitumumab combination treatment.

Subsequent treatment costs for patients with progressed disease in the company's base case was split between FOLFIRI and cetuximab + FOLFIRI treatments. NICE guidance does not recommend cetuximab for use as a monotherapy or in combination therapies as second line treatment for mCRC, thus the ERG considers that second line treatment should only reflect FOLFOX and FOLFIRI, but changing this assumption has minimal impact on the ICERs.

7 End of Life

The company has not made a case for pembrolizumab to be considered as an end-of-life treatment, which the ERG considers is appropriate.

8 References

1. NICE. Pembrolizumab for untreated metastatic colorectal cancer with high microsatellite instability or mismatch repair deficiency - Final scope. 2020. Available from: https://www.nice.org.uk/guidance/gid-ta10420/documents/final-scope. Date accessed: 20 Oct 2020.
2. Merck Sharp & Dohme Limited. KEYNOTE-177 Clinical Study Report. 2020.
3. Baran B, Mert Ozupek N, Yerli Tetik N, Acar E, Bekcioglu O, Baskin Y. Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. *Gastroenterology Res* 2018; **11**: 264-73.
4. Tidy C. Colorectal cancer. 2020. Available from: https://patient.info/doctor/colorectal-cancer. Date accessed: 21 Oct 2020.
5. Cancer.org. What is colorectal cancer? 2020. Available from: https://www.cancer.org/cancer/colon-rectal-cancer/about/what-is-colorectal-cancer.html. Date accessed: 5 Nov 2020.
6. Zlobec I, Lugli A. Prognostic and predictive factors in colorectal cancer. *Postgrad Med J* 2008; **84**: 403-11.
7. Kawakami H, Zaanan A, Sinicrope FA. Microsatellite instability testing and its role in the management of colorectal cancer. *Curr Treat Options Oncol* 2015; **16**: 30.
8. Nojadeh JN, Behrouz Sharif S, Sakhinia E. Microsatellite instability in colorectal cancer. *EXCLI J* 2018; **17**: 159-68.
9. Venderbosch S, Nagtegaal ID, Maughan TS, Smith CG, Cheadle JP, Fisher D, et al. Mismatch repair status and BRAF mutation status in metastatic colorectal cancer patients: a pooled analysis of the CAIRO, CAIRO2, COIN, and FOCUS studies. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2014; **20**: 5322-30.
10. Clarke CN, Kopetz ES. BRAF mutant colorectal cancer as a distinct subset of colorectal cancer: clinical characteristics, clinical behavior, and response to targeted therapies. *J Gastrointest Oncol* 2015; **6**: 660-7.
11. Garcia-Foncillas J, Sunakawa Y, Aderka D, Wainberg Z, Ronga P, Witzler P, et al. Distinguishing Features of Cetuximab and Panitumumab in Colorectal Cancer and Other Solid Tumors. *Front Oncol* 2019; **9**: 849.
12. NICE. Guidance on the use of capecitabine and tegafur with uracil for metastatic colorectal cancer [TA61]. 2003. Available from: https://www.nice.org.uk/Guidance/TA61. Date accessed: 19 Oct 2020.
13. NICE. Cetuximab (review of TA176) and panitumumab (partial review of TA240) for the first line treatment of metastatic colorectal cancer ID794 - Committee papers. 2015. Available from: https://www.nice.org.uk/guidance/ta439/documents/committee-papers. Date accessed: 19 Oct 2020.
14. NICE. Molecular testing strategies for Lynch syndrome in people with colorectal cancer [DG27]. 2017. Available from: https://www.nice.org.uk/guidance/dg27. Date accessed: 30 Oct 2020.
15. European Medicine Agency (EMA). Keytruda. 2020. Available from: https://www.ema.europa.eu/en/medicines/human/summaries-opinion/keytruda-1#:~:text=Opinion-,Opinion,is%20Merck%20Sharp%20%26%20Dohme%20B.V. Date accessed: 4 January 2021.
16. Merck Sharp & Dohme Limited. KEYTRUDA 25 mg/mL concentrate for solution for infusion. Draft Summary of Product Characteristics. 2020.
17. Cassidy J, Clarke S, Diaz-Rubio E, Scheithauer W, Figer A, Wong R, et al. XELOX vs FOLFOX-4 as first-line therapy for metastatic colorectal cancer: NO16966 updated results. *Br J Cancer* 2011; **105**: 58-64.

18. Colucci G, Gebbia V, Paoletti G, Giuliani F, Caruso M, Gebbia N, et al. Phase III randomized trial of FOLFIRI versus FOLFOX4 in the treatment of advanced colorectal cancer: a multicenter study of the Gruppo Oncologico Dell'Italia Meridionale. *J Clin Oncol* 2005; **23**: 4866-75.
19. Ducreux M, Bennouna J, Hebbar M, Ychou M, Lledo G, Conroy T, et al. Capecitabine plus oxaliplatin (XELOX) versus 5-fluorouracil/leucovorin plus oxaliplatin (FOLFOX-6) as first-line treatment for metastatic colorectal cancer. *Int J Cancer* 2011; **128**: 682-90.
20. Hochster HS, Hart LL, Ramanathan RK, Childs BH, Hainsworth JD, Cohn AL, et al. Safety and efficacy of oxaliplatin and fluoropyrimidine regimens with or without bevacizumab as first-line treatment of metastatic colorectal cancer: results of the TREE Study. *J Clin Oncol* 2008; **26**: 3523-9.
21. Porschen R, Arkenau HT, Kubicka S, Greil R, Seufferlein T, Freier W, et al. Phase III study of capecitabine plus oxaliplatin compared with fluorouracil and leucovorin plus oxaliplatin in metastatic colorectal cancer: a final report of the AIO Colorectal Study Group. *J Clin Oncol* 2007; **25**: 4217-23.
22. Hurwitz H, Fehrenbacher L, Novotny W, Cartwright T, Hainsworth J, Heim W, et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 2004; **350**: 2335-42.
23. Saltz LB, Clarke S, Diaz-Rubio E, Scheithauer W, Figer A, Wong R, et al. Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: a randomized phase III study. *J Clin Oncol* 2008; **26**: 2013-9.
24. Bokemeyer C, Bondarenko I, Makhson A, Hartmann JT, Aparicio J, de Braud F, et al. Fluorouracil, leucovorin, and oxaliplatin with and without cetuximab in the first-line treatment of metastatic colorectal cancer. *J Clin Oncol* 2008; **27**: 663-71.
25. Douillard JY, Siena S, Cassidy J, Tabernero J, Burkes R, Barugel M, et al. Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. *J Clin Oncol* 2010; **28**: 4697-705.
26. Van Cutsem E, Kohne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009; **360**: 1408-17.
27. Qin S, Li J, Wang L, Xu J, Cheng Y, Bai Y, et al. Efficacy and Tolerability of First-Line Cetuximab Plus Leucovorin, Fluorouracil, and Oxaliplatin (FOLFOX-4) Versus FOLFOX-4 in Patients With RAS Wild-Type Metastatic Colorectal Cancer: The Open-Label, Randomized, Phase III TAILOR Trial. *J Clin Oncol* 2018; **36**: 3031-9.
28. Xu S, Sak A, Erol YB. Systemic therapy as first-Line treatment for patients with metastatic colorectal cancer: a systematic review and network meta-analysis of randomized clinical trials. 2020. Available from: https://www.researchsquare.com/article/rs-12790/v1. Date accessed: 3 Nov 2020.
29. Maurer W, Bretz F. Multiple testing in group sequential trials using graphical approaches. *Stat Biopharm Res* 2013; **5**: 311-20.
30. Passardi A, Nanni O, Tassinari D, Turci D, Cavanna L, Fontana A, et al. Effectiveness of bevacizumab added to standard chemotherapy in metastatic colorectal cancer: final results for first-line treatment from the ITACa randomized clinical trial. *Ann Oncol* 2015; **26**: 1201-7.
31. Passardi A, Scarpi E, Cavanna L, Fontana A, Vertogen B, Ruscelli S, et al. Effectiveness of bevacizumab added to gold standard chemotherapy in metastatic colorectal cancer (mCRC): Final results from the Itaca randomized clinical trial. *J Clin Oncol.* 2013; **31**: 3517.
32. Bokemeyer C, Kohne CH, Ciardiello F, Lenz HJ, Heinemann V, Klinkhardt U, et al. FOLFOX4 plus cetuximab treatment and RAS mutations in colorectal cancer. *European journal of cancer* 2015; **51**: 1243-52.
33. Latimer NR, Abrams KR. NICE DSU Technical Support Document 16: Adjusting Survival Time Estimates in the Presence of Treatment Switching. 2014. Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD16_Treatment_Switching.pdf. Date accessed: 13 Nov 2020.

34. Electronic Medicines Compendium (eMC). Vectibix 20 mg/mL concentrate for solution for infusion. 2020. Available from: https://www.medicines.org.uk/emc/product/6178. Date accessed: 21 Nov 2020.
35. Electronic Medicines Compendium (eMC). Erbitux 5 mg/mL solution for infusion. 2020. Available from: https://www.medicines.org.uk/emc/product/317. Date accessed: 21 Nov 2020.
36. Jansen JP. Network meta-analysis of survival data with fractional polynomials. *BMC Med Res Methodol* 2011; **11**: 61.
37. J.N C, J C, S O, J.A T, K.L R, A T, et al. Cost-effectiveness of immune checkpoint inhibitors for microsatellite instability–high/mismatch repair–deficient metastatic colorectal cancer. *Cancer* 2019; **15**: 278-89.
38. NHS. 2018/19 National Cost Collection data. 2019. Available from: https://www.england.nhs.uk/national-cost-collection/#ncc1819. Date accessed.
39. MIMS. Monthly Index of Medical Specialties (MIMS). Available at: https://www.mims.co.uk/.
40. Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT). Available at: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit. 2018.
41. NICE. Bevacizumab in combination with oxaliplatin and either fluorouracil plus folinic acid or capecitabine for the treatment of metastatic colorectal cancer [TA212]. 2010. Available from: https://www.nice.org.uk/guidance/TA212. Date accessed.
42. NICE. NICE Pathways - Managing metastatic colorectal cancer - First-line biological therapy. 2020. Available from: https://pathways.nice.org.uk/pathways/colorectal-cancer/managing-metastatic-colorectal-cancer#content=view-node%3Anodes-first-line-biological-therapy. Date accessed.
43. NICE. Guide to the methods of technology appraisal [PMG9]. 2013. Available from: https://www.nice.org.uk/process/pmg9/chapter/foreword. Date accessed.
44. (DSU) DSU. NICE DSU Technical Support Support Document 14: Survival analysis for economic evaluations alongside clinical trials – extrapolation with patient-level data. 2013. Available from: http://nicedsu.org.uk/technical-support-documents/survival-analysis-tsd/. Date accessed.
45. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010; **13**: 509-18.
46. Bennett L, Zhao Z, Barber B, Zhou X, Peeters M, Zhang J, et al. Health-related quality of life in patients with metastatic colorectal cancer treated with panitumumab in first- or second-line treatment. *Br J Cancer* 2011; **105**: 1495-502.
47. Curtis LAaB, Amanda. Unit Costs of Health and Social Care 2019. 2019.
48. Farkkila N, Torvinen S, Sintonen H, Saarto T, Jarvinen H, Hanninen J, et al. Costs of colorectal cancer in different states of the disease. *Acta Oncol* 2015; **54**: 454-62.
49. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. *Palliat Med* 2015; **29**: 899-907.
50. NICE. Cetuximab, bevacizumab and panitumumab for the treatment of metastatic colorectal cancer after first-line chemotherapy: Cetuximab (monotherapy or combination chemotherapy), bevacizumab (in combination with non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-line chemotherapy[TA242]. 2012. Available from: https://www.nice.org.uk/guidance/ta242. Date accessed.
51. Tougeron D, Sueur B, Zaanan A, de la Fouchardiere C, Sefrioui D, Lecomte T, et al. Prognosis and chemosensitivity of deficient MMR phenotype in patients with metastatic colorectal cancer: An AGEO retrospective multicenter study. *Int J Cancer* 2020; **147**: 285-96.

9 Appendices

9.1 Baseline characteristics

Table 67. Summary of population baseline characteristics for KEYNOTE-177 (adapted from CS, Table 10, page 29)

Characteristic	Pembrolizumab (N = 153) n (%)	SoC (N = 154) n (%)	Total (N = 307) n (%)
Gender			
Male	71 (46.4)	82 (53.2)	153 (49.8)
Female	82 (53.6)	72 (46.8)	154 (50.2)
Age			
<65	80 (52.3)	83 (53.9)	163 (53.1)
≥65	73 (47.7)	71 (46.1)	144 (46.9)
<70	105 (68.6)	112 (72.7)	217 (70.7)
≥70	48 (31.4)	42 (27.3)	90 (29.3)
Mean (SD)	61.9 (14.9)	60.6 (14.8)	61.2 (14.8)
Median (range)	63.0 (24 to 93)	62.5 (26 to 90)	63.0 (24 to 93)
Race			
Asian	24 (15.7)	26 (16.9)	50 (16.3)
Black or African American	9 (5.9)	5 (3.2)	14 (4.6)
White	113 (73.9)	116 (75.3)	229 (74.6)
Missing	7 (4.6)	7 (4.5)	14 (4.6)
Ethnicity			
Hispanic or Latino	11 (7.2)	10 (6.5)	21 (6.8)
Not Hispanic or Latino	128 (83.7)	131 (85.1)	259 (84.4)
Not reported	10 (6.5)	10 (6.5)	20 (6.5)
Unknown	2 (1.3)	2 (1.3)	4 (1.3)

Missing	2 (1.3)	1 (0.6)	3 (1.0)
Geographic region			
Asia	22 (14.4)	26 (16.9)	48 (15.6)
Western Europe/North America	109 (71.2)	113 (73.4)	222 (72.3)
Rest of World	22 (14.4)	15 (9.7)	37 (12.1)
ECOG			
0	75 (49.0)	84 (54.5)	159 (51.8)
1	78 (51.0)	70 (45.5)	148 (48.2)
Site of primary tumour[a]			
Right	102 (66.7)	107 (69.5)	209 (68.1)
Left	46 (30.1)	42 (27.3)	88 (28.7)
Other	4 (2.6)	5 (3.2)	9 (2.9)
Missing	1 (0.7)	0 (0.0)	1 (0.3)
Metastases location			
Hepatic or pulmonary	86 (56.2)	73 (47.4)	159 (51.8)
Other metastases	67 (43.8)	81 (52.6)	148 (48.2)
Diagnosed stage			
Recurrent	80 (52.3)	74 (48.1)	154 (50.2)
Newly diagnosed stage	73 (47.7)	80 (51.9)	153 (49.8)
Prior systemic therapy			
Adjuvant only	33 (21.6)	37 (24.0)	70 (22.8)
Neoadjuvant only	2 (1.3)	3 (1.9)	5 (1.6)
Neoadjuvant and adjuvant	3 (2.0)	5 (3.2)	8 (2.6)
None	115 (75.2)	109 (70.8)	224 (73.0)
Mutation status[b]			
BRAF/KRAS/NRAS all wild type	34 (22.2)	35 (22.7)	69 (22.5)

KRAS/NRAS mutant and BRAF V600E not mutant	33 (21.6)	38 (24.7)	71 (23.1)
BRAF V600E mutant and KRAS/NRAS not mutant	34 (22.2)	40 (26.0)	74 (24.1)
BRAF V600E and KRAS/NRAS mutant	0 (0.0)	3 (1.9)	3 (1.0)
Other	52 (34.0)	38 (24.7)	90 (29.3)
MSI-High status[c]			
Positive	153 (100.0)	153 (99.4)	306 (99.7)
Negative	0 (0.0)	1 (0.6)	1 (0.
Oncologic surgery with curative intent[d]			
Received surgery	14 (9.2)	13 (8.4)	27 (8.8)
Did not receive surgery	139 (90.8)	141 (91.6)	280 (91.2)

[a] If there were primary tumours in both left side and right side, the subject would be categorized into Other.

[b] When none of BRAF V600E, KRAS and NRAS was mutant, if at least one of the mutation statuses was undetermined or missing, or the type of BRAF mutation was not V600E, the subject was categorized into Other.

[c] MSI status by PCR test or IHC test at local site laboratory.

[d] Oncologic surgery that was with curative intent and occurred after subject randomisation and before initiation of new anti-cancer therapy, crossover treatment and second course treatment.

Database Cutoff Date: 19FEB2020.

Abbreviations: ECOG, Eastern Cooperative Oncology Group; IHC, immunohistochemistry; MSI, microsatellite instability; PCR, polymerase chain reaction; SD, standard deviation; SoC, standard of care.

9.2 Pre-specified subgroup analyses

Figure 16. Analysis of progression-free survival by subgroup factors by central imaging vendor per RECIST 1.1 (ITT population), based on Cox regression model with Efron's method of tie handling with treatment as a covariate (database cut-off 19 Feb 2020; reproduced from CS, Figure 10, page 63)

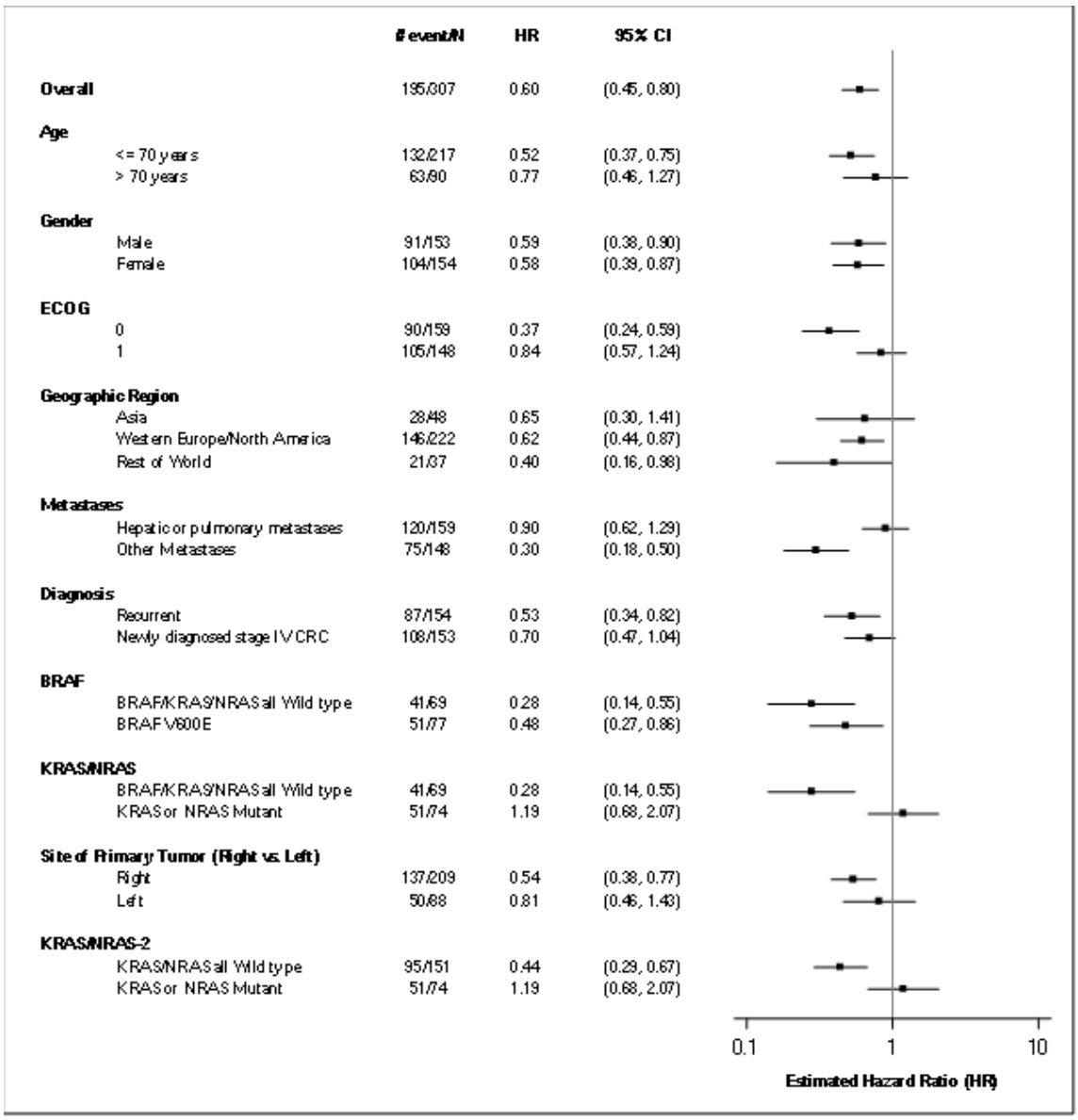
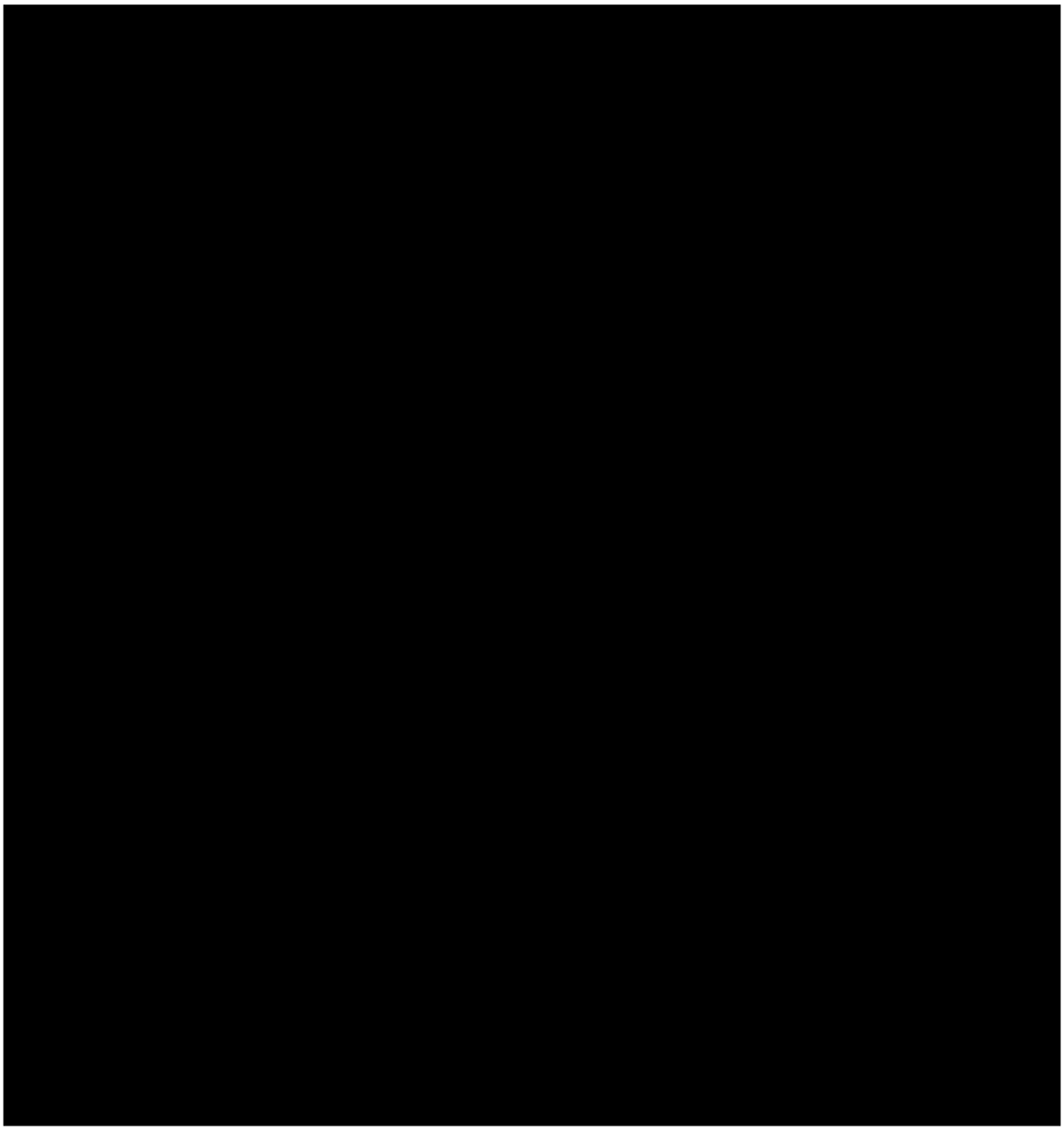


Figure 17. Analysis of overall survival by subgroup factors (ITT population), based on Cox regression model with Efron's method of tie handling with treatment as a covariate, database cut-off date 19 Feb 2020 (reproduced from CS, Figure 5, page 48)



9.3 Fractional polynomial NMA: ERG validation for comparators versus standard of care

Table 68. Time-varying hazard ratios for PFS at selected follow-up times using standard of care as the baseline treatment (second order FP model [p1 = 0, p2 = 0]) (adapted from CS, Appendix M, Table 191, page 627)

Month	HR versus SoC (95% CrI)[a]					
	CAPOX Company	CAPOX ERG	Panitumumab + FOLFOX Company	Panitumumab + FOLFOX ERG	Pembrolizumab Company	Pembrolizumab ERG
4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
8	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
12	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
16	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
20	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
24	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
28	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
32	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
36	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
40	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

[a] HR >1 favours SoC, HR <1 favours comparator.

Abbreviations: CAPOX, capecitabine plus oxaliplatin; CS, company submission; CrI, credible interval; FOLFOX, folinic acid plus fluorouracil plus oxaliplatin; HR, hazard ratio; PFS, progression-free survival; SoC, standard of care.

Figure 18. Estimated treatment hazard ratio over time for PFS and treatments relative to SoC (second order FP model; p1=0, p2=0)) based on the ITT population (reproduced from CS, Appendix M, Figure 97, page 626)

Figure 19. ERG's validation of estimated treatment hazard ratio over time for PFS and treatments relative to SoC (second order FP model; p1=0, p2=0)) based on the ITT population