in collaboration with:

Erasmus School of Health Policy & Management

Maastricht University

Amivantamab for treating EGFR Exon 20 insertion-positive non-small-cell lung cancer after platinum-based chemotherapy (review of TA10729) [ID3836]

Produced by

Kleijnen Systematic Reviews (KSR) Ltd, in collaboration with Erasmus University Rotterdam (EUR) and Maastricht University Medical Centre+ (UMC+)

Authors

Jeremy Howick, Reviews Manager, KSR Ltd, United Kingdom (UK)
Willem Witlox, Health Economist, Maastricht UMC+, The Netherlands
Kevin McDermott, Systematic Reviewer, KSR Ltd, UK
Ben Wijnen, Health Economist, Maastricht UMC+, The Netherlands
Evangelos Danopoulos, Systematic Reviewer, KSR Ltd, UK
Thomas Otten, Health Economist, Maastricht UMC, The Netherlands
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Steven Duffy, Information Specialist, KSR Ltd, UK
Marie Westwood, Reviews Manager, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Mark Perry, Reviews Manager, KSR Ltd, UK
Robert Wolff, Managing Director, KSR Ltd, UK
Manuela Joore, Health Economist, Maastricht UMC+, The Netherlands
Jos Kleijnen, Founder and Owner, KSR Ltd

Correspondence to Jeremy Howick, Kleijnen Systematic Reviews Ltd
Unit 6, Escrick Business Park
Riccall Road, Escrick
York, YO19 6FD
United Kingdom

Date completed 21/04/2022

Source of funding: This report was commissioned by the National Institute for Health Research (NIHR) Evidence Synthesis Programme as project number STA 13/44/06.

Declared competing interests of the authors None

Acknowledgements None

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

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

Confidential comparator prices are highlighted in green throughout the report.

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

Rider on responsibility for report

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

This report should be referenced as follows:

Howick J, Witlox W, McDermott K, Wijnen B, Danopoulos E, Otten T, Ahmadu C, Otten T, Duffy S, Perry M, Westwood M, Armstrong N, Wolff R, Joore M, Kleijnen J. Amivantamab for treating EGFR Exon 20 insertion-positive non-small-cell lung cancer after platinum-based chemotherapy (review of TA10729) [ID3836]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2022.

Contributions of authors

Jeremy Howick acted as project lead and systematic reviewer on this assessment, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Willem Witlox and Ben Wijnen acted as health economic project leads, critiqued the company's economic evaluation, and contributed to the writing of the report. Thomas Otten, Charlotte Ahmadu, and Nigel Armstrong acted as health economists on this assessment, critiqued the company's economic evaluation and contributed to the writing of the report. Kevin McDermott, Evan Danopoulos, Mark Perry, and Marie Westwood acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence and contributed to the writing of the report. Steven Duffy critiqued the search methods in the submission and contributed to the writing of the report. Manuela Joore acted as health economist on this assessment, critiqued the company's economic evaluation, contributed to the writing of the report and provided general guidance. Robert Wolff and Jos Kleijnen critiqued the company's definition of the decision problem and their description of the underlying health problem and current service provision, contributed to the writing of the report and supervised the project.

Abbreviations

AACR	American Association for Cancer Research
AE	Adverse event
AEMPS	Agencia Española de Medicamentos y Productos Sanitarios
AIFA	Agenzia Italiana del Farmaco
AMI	Amivantamab
ASCO	American Society of Clinical Oncology
ATT	Average treatment effect among the treated
AWMSG	All Wales Medicines Strategy Group
BAG	Bundesamt für Gesundheit
BICR	Blinded independent committee review assessed
BMI	Body mass index
BNF	British National Formulary
BOR	Best overall response
BSC	Best supportive care
CADTH	Canadian Agency for drugs and Technologies in Health
CBR	Clinical benefit rate
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Clinical Trials
CI	Confidence interval
cm	Centimetre
CR	Complete response
CRD	Centre for Reviews and Dissemination
CS	Company submission
DARE	Database of Abstracts of Reviews of Effects
DFS	Disease free survival
DOR	Duration of response
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EGFR	Epidermal growth factor receptor
EMA	European Medicines Agency
EQ-5D-5L	EuroQoL 5-dimensions 5-levels
EQ-5D-5L VAS	EuroQol-5 dimensions 5-levels visual analogue score
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
Exon20ins	Exon 20 insertion mutations
FE	Fixing errors
FinCCHTA	Finnish Coordinating Center for Health Technology Assessment
FDA	Food and Drugs Administration
FV	Fixing violations
G-BA	Gemeinsamer Bundesausschuss
HAS	Haute Autorité de Santé
HER2	Human epidermal growth factor receptor 2
HIV	Human immunodeficiency virus
HR	Hazard ratio
HRQoL	Health-related quality of life

HTA	Health technology assessment
IASLC	International Association for the Study of Lung Cancer
ICER	Incremental cost-effectiveness ratio
iDFS	Invasive disease-free survival
ILD	Interstitial lung disease
INV	Investigator-assessed
IPD	Individual participant data
IPW	Inverse probability weighting
IRR	Infusion related reaction
IO	Immuno-oncology
ITC	Indirect treatment comparison
ITT	Intention to treat
IV	Intravenous
kg	Kilogram
KM	Kaplan-Meier
LOT	Line of therapy
LS	Least squares
LY	Life year
mg	Milligram
MET	Mesenchymal epithelial transition
MHRA	Medicines and Healthcare Products Regulatory Agency
MJ	Matters of judgement
MSCBS	Ministerio de Sanidad, Consumo y Bienestar Social
MTD	Maximum tolerated dose
NCI CTCAE	National Cancer Institute common terminology criteria for adverse events
NCPE	National Centre for Pharmacoeconomics
NCRAS	National Cancer Registration and Analysis Service
NE	Not evaluable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIPH	Norwegian Institute of Public Health
NR	Not reported
NSCLC	Non-small-cell lung cancer
NSCLC-SAQ	Non-small-cell lung cancer Symptom Assessment Questionnaire
ORR	Overall response rate
OS	Overall survival
PAS	Patient Access Scheme
pCR	Pathological complete response
PD	Progressed disease
PFS	Progression-free survival
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PHE	Public Health England
PPS	Post-progression survival
PR	Partial response
PRESS	Peer Review of Electronic Search Strategies
PRO	Patient-reported outcome

PSA	Probabilistic sensitivity analysis
PSM	Propensity score matching
PSSRU	Personal Social Services Research Unit
Pt	Platinum
QALY	Quality-adjusted life year
QoL	Quality of life
RCT	Randomised controlled trial
RECIST	Response Evaluation Criteria in Solid Tumours
RID	Residual invasive disease
RP2D	Recommended Phase 2 dose
RWD	Real world data
RWE	Real world evidence
SBU	Swedish Agency for Health Technology Assessment and Assessment of Social Services
SCLC	Small cell lung cancer
SD	Standard deviation
SLR	Systematic literature review
SMC	Scottish Medicines Consortium
SMD	Standardised mean difference
SoC	Standard of Care
SoD	Sum of diameters
STM	State transition model
STA	Single technology appraisal
TEAE	Treatment-emergent adverse event
TKI	Tyrosine kinase inhibitor
TSD	Technical Support Document
TTD	Time to treatment discontinuation
TTF	Time to treatment failure
TTNT	Time to next treatment
UK	United Kingdom
US	United States
VAS	Visual analogue scale

Table of Contents

Table of Tables

Table of Figures

1. EXECUTIVE SUMMARY

This summary provides a brief overview of the key issues identified by the Evidence Review Group (ERG) as being potentially important for decision making. If possible, it also includes the ERG's preferred assumptions and the resulting incremental cost-effectiveness ratios (ICERs).

Section 1.1 provides an overview of the key issues. Section 1.2 presents the key model outcomes. Section 1.3 discusses the decision problem, Section 1.4 issues relate to the clinical effectiveness, and Section 1.5 issues relate to the cost effectiveness. A summary is presented in Section 1.6.

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

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

1.1 Overview of the ERG's key issues

Table 1.1: Summary of key issues

ID3836	Summary of issue	Report Sections
1	The narrower population considered by company may not be generalisable to the England and Wales National Health Service (NHS) population and may have led to an underestimate of adverse events (AEs).	2.1, 3.2
2	Patients in the intervention group received concomitant medications (including targeted radiotherapy) that could have exaggerated the benefits of amivantamab.	2.2, 3.2
3	Some of the comparators lack justification and could have obscured or exaggerated the benefits of amivantamab.	2.3, 3.2
4	The short follow-up time of the CHRYSALIS trial makes medium- and longer-term results uncertain.	3.2
5	The efficacy and safety populations differ in a way that is likely to exaggerate benefits and understate harms.	3.2
6	The real-world evidence (RWE) sources to identify comparators for the indirect treatment comparison were not comprehensive, leading to uncertainty in the benefits of amivantamab compared with relevant comparators.	3.3, 3.4
7	The company assumed [redacted] of the comparator basket to consist of epidermal growth factor receptor tyrosine kinase inhibitors (EGFR TKIs) which may not be consistent with UK clinical practice; the relative cost effectiveness of amivantamab is therefore unclear.	4.2.4
8	The company implemented Kaplan-Meier (KM) curves instead of parametric survival models for the survival analyses of overall survival (OS) and progression-free survival (PFS) in the standard of care (SoC) arm, leading to potential overfitting of modelled survival outcomes.	4.2.6
9	Time to treatment discontinuation (TTD) was assumed to be equal to the duration of PFS, while evidence from the CHRYSALIS trial showed that amivantamab treatment had a longer median duration	4.2.6

ID3836	Summary of issue	Report Sections
	than PFS, leading to a possible underestimate of amivantamab's relative cost.	
10	The company did not explore treatment waning in the model, whereas the Evidence Review Group (ERG) considered that the assumption of a lifelong treatment effect may not be warranted.	4.2.6
11	The company's failure to include an age-adjustment to the health state utilities in their company submission (CS) base case is not in line with good modelling practice and may have exaggerated the cost effectiveness of amivantamab.	4.2.8
12	Lack of a fully incremental analysis for all relevant comparators in the comparator basket, increasing the uncertainty of estimates of amivantamab's cost effectiveness.	5.1
13	Lack of a fixed random seed in model probabilistic sensitivity analysis (PSA) leads to fluctuations in probabilistic results and hence increased uncertainty of estimates of amivantamab's cost-effectiveness.	5.3

AEs = adverse events; CS = company submission; EGFR = epidermal growth factor receptor; ERG = Evidence Review Group; KM = Kaplan-Meier; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; OS = overall survival; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; RWE = real world evidence; SoC = standard of care; TTD = time to treatment discontinuation; TKI = tyrosine kinase inhibitor; UK = United Kingdom

The key differences between the company's preferred assumptions and the ERG's preferred assumptions include assumptions related to the population, comparators, outcomes, and sources of evidence to inform the indirect treatment comparison.

1.2 Overview of key model outcomes

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

Overall, the technology is modelled to affect QALYs by (deterministic):

- Increased post-progression survival (PPS), with an increment of 0.526 years (63% of total incremental life years (LYs)) in the amivantamab arm (1.349 years) compared with United Kingdom (UK) standard of care (SoC) (0.823 years)
- Increased progression-free survival (PFS), with an increment of 0.314 years (37% of total incremental LYs) in the amivantamab arm (0.818 years) compared with UK SoC (0.504 years).

Overall, the technology is modelled to affect costs by (deterministic):

- The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs), administration costs (additional cost of [redacted], [redacted] of total incremental costs) and post-progression disease management costs (additional cost of [redacted], [redacted] of total incremental costs).

The company performed and presented the results of probabilistic sensitivity analysis (PSA), deterministic sensitivity analysis (DSA) as well as scenario analyses. The parameters that had the greatest effect on the ICER based on the company's DSA were:

- PFS Kaplan-Meier (KM) curve for the UK SoC arm

- Drug costs in subsequent cycles for the amivantamab arm
- Health state utilities for PFS and PPS

Company submission (CS) scenarios that have the greatest impact on the ICER (not including scenarios related to discount rates and time horizon) were:

- UK SoC efficacy based on Public Health England (PHE) data (decreased ICER to £25,865)
- Using osimertinib to represent epidermal growth factor receptor tyrosine kinase inhibitors (EGFR TKIs) (decreased ICER to £31,224)
- Using investigator-assessment (INV) as a measure of progression (increased ICER to £42,249)

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

The decision problem addressed by the company in their submission is broadly in line with the final scope issued by NICE. However, there were potentially relevant differences between the populations (Table 1.2), intervention (Table 1.3), and comparator (Table 1.4)..

Table 1.2: Key issue 1. Population considered by company narrower than population in final NICE scope

Report Section	**2.1, 3.2**
Description of issue and why the ERG has identified it as important	Population considered by company is narrower than the population defined in final National Institute for Health and Care Excellence (NICE) scope; the narrower population may not be generalisable to the England and Wales National Health Service (NHS) population; and (because the company's population was "fitter"), may have led to an underestimate of adverse events (AEs).
What alternative approach has the ERG suggested?	Other than a new trial with the population specified in the final NICE scope, no alternative approach is suggested by the Evidence Review Group (ERG) who wanted to bring this to the attention of the committee.
What is the expected effect on the cost effectiveness estimates?	Could have (a) under-estimated AEs, and (b) over-estimated effectiveness and cost effectiveness.
What additional evidence or analyses might help to resolve this key issue?	N/A
AEs = adverse events; ERG = Evidence Review Group; NHS = National Health Service; N/A = not applicable; NICE = National Institute for Health and Care Excellence; UK = United Kingdom	

Table 1.3: Key issue 2. Patients in intervention group received additional medications that could have exaggerated the effects of amivantamab

Report Section	**2.2, 3.2**
Description of issue and why the ERG has identified it as important	Patients in the intervention group received a variety of concomitant medications (including targeted radiotherapy) that could have exaggerated the benefits of amivantamab.
What alternative approach has the ERG suggested?	Given that it is known which concomitant medications were received, an unbiased estimate of effectiveness of amivantamab with and without the potentially problematic concomitant medications such as targeted radiotherapy is possible.

Report Section	2.2, 3.2
What is the expected effect on the cost effectiveness estimates?	Could have over-estimated cost effectiveness.
What additional evidence or analyses might help to resolve this key issue?	Separate effectiveness and safety analyses of amivantamab with and without the problematic concomitant medications. The results of these analyses could then be input into separate cost effectiveness analyses.
ERG = Evidence Review Group	

Table 1.4: Key issue 3. Some of the comparators lack justification and could have obscured or exaggerated the benefits of amivantamab

Report Section	2.1, 3.2
Description of issue and why the ERG has identified it as important	Some of the comparators (especially tyrosine kinase inhibitors (TKIs) other than nintedanib) lack justification and could have exaggerated the benefits of amivantamab.
What alternative approach has the ERG suggested?	An exploration of the relative effects of amivantamab with the comparators suggested by the Evidence Review Group (ERG).
What is the expected effect on the cost effectiveness estimates?	Uncertain.
What additional evidence or analyses might help to resolve this key issue?	Separate effectiveness and safety analyses of the comparators with and without the problematic comparators. The results of these analyses could then be input into separate cost effectiveness analyses.
ERG = Evidence Review Group; TKI = tyrosine kinase inhibitor	

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

The ERG identified three major concerns with the evidence presented on the clinical effectiveness: the short follow-up of the included randomised controlled trials (RCTs) (see Table 1.5), the problematic choice of safety and efficacy populations (see Table 1.6), and incomplete sources of real-world evidence (see Table 1.7).

Table 1.5: Key issue 4. Short follow-up time of included randomised trials

Report Section	3.2
Description of issue and why the ERG has identified it as important	The short follow-up time of the CHRYSALIS trial makes medium- and longer-term results uncertain.
What alternative approach has the ERG suggested?	None suggested.
What is the expected effect on the cost effectiveness estimates?	Uncertainty of the medium and long-term effects of amivantamab.
What additional evidence or analyses might help to resolve this key issue?	Updated data with longer follow-up times.
ERG = Evidence Review Group	

Table 1.6: Key issue 5: Problematic choice of populations

Report Section	3.2
Description of issue and why the ERG has identified it as important	The efficacy and safety populations seem to differ substantially.
What alternative approach has the ERG suggested?	Safety and efficacy analyses with populations that are the same (or at the very least not so different).
What is the expected effect on the cost effectiveness estimates?	Exaggeration of cost effectiveness estimates.
What additional evidence or analyses might help to resolve this key issue?	At least in an exploratory basis, use the intention to treat (ITT) population for both safety and efficacy.
ERG = Evidence Review Group; ITT = intention to treat	

Table 1.7: Key issue 6. Incomplete real world evidence sources for the indirect treatment comparison

Report Section	3.3, 3.4
Description of issue and why the ERG has identified it as important	The real world evidence (RWE) sources to identify comparators for the indirect treatment comparison were not comprehensive, leading to uncertainty in the benefits of amivantamab compared with relevant comparators.
What alternative approach has the ERG suggested?	A full search for and incorporation of all relevant studies.
What is the expected effect on the cost effectiveness estimates?	Increased uncertainty regarding the cost effectiveness of amivantamab relative to relevant comparators.
What additional evidence or analyses might help to resolve this key issue?	An updated intention to treat (ITT), conducted after a full search for and incorporation of all relevant studies.
ERG = Evidence Review Group; ITC = intention to treat; RWE = real world evidence	

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

A full summary of the cost effectiveness evidence review conclusions can be found in Section 6.4 of this report. The company's cost effectiveness results are presented in Section 5, the ERG's summary and detailed critique in Section 4, and the ERG's amendments to the company's model and results are presented in Section 6. The ERG identified seven major issues with the cost effectiveness evidence are discussed in the Tables 1.8 to 1.14 below.

Table 1.8: Key issue 7: Representativeness of the comparator basket effectiveness to UK clinical practice is unclear, leading to uncertainty in relative cost effectiveness of amivantamab

Report Section	4.2.4
Description of issue and why the ERG has identified it as important	The company assumed 19% of the comparator basket to consist of epidermal growth factor receptor tyrosine kinase inhibitors (EGFR TKIs) which may not be consistent with United Kingdom (UK) clinical practice; the relative cost effectiveness of amivantamab is therefore unclear.

Report Section	4.2.4
What alternative approach has the ERG suggested?	An analysis where EGFR TKI therapies are excluded from the United States (US) real world data (RWD) informing the comparator basket.
What is the expected effect on the cost effectiveness estimates?	Increased uncertainty regarding amivantamab cost effectiveness.
What additional evidence or analyses might help to resolve this key issue?	1. Updated economic model excluding the costs and effects of EGFR TKIs. 2. Updated assessment of the NICE DSU TSD 14 criteria for survival analyses without EGFR TKIs in the standard of care (SoC) basket to support curve selection.

ERG = Evidence Review Group; EGFR = epidermal growth factor receptor; NICE DSU TSD 14 = National Institute for Health and Care Excellence Decision Support Unit Technical Support Document 14; RWD = real world data; SoC = standard of care; TKI = tyrosine kinase inhibitors; UK = United Kingdom; US = United States

Table 1.9: Key issue 8: Implementation of parametric survival curves instead of KM curves for SoC

Report Section	4.2.6
Description of issue and why the ERG has identified it as important	The company implemented Kaplan-Meier (KM) curves instead of parametric survival models for the survival analyses of overall survival (OS) and progression-free survival (PFS) in the standard of care (SoC) arm, leading to potential overfitting of modelled survival outcomes.
What alternative approach has the ERG suggested?	Implement parametric models based on NICE DSU TSD 14 for survival analyses of OS and PFS in the SoC arm.
What is the expected effect on the cost effectiveness estimates?	Depends on selected curves. Using a Weibull model for OS and a log-logistic model for PFS, the incremental cost effectiveness ratio (ICER) slightly increased.
What additional evidence or analyses might help to resolve this key issue?	N/A

ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; KM = Kaplan Meier; N/A = not applicable; NICE DSU TSD 14 = National Institute for Health and Care Excellence Decision Support Unit Technical Support Document 14; OS = overall survival; PFS = progression-free survival; SoC = standard of care

Table 1.10: Key issue 9: Time to treatment discontinuation

Report Section	4.2.6
Description of issue and why the ERG has identified it as important	Time to treatment discontinuation (TTD) was assumed to be equal to the duration of progression-free survival (PFS), while evidence from the CHRYSALIS trial showed that amivantamab treatment had a longer median duration than PFS, leading to a possible underestimate of amivantamab's relative cost.
What alternative approach has the ERG suggested?	The Evidence Review Group (ERG) suggested applying a parametric survival model to TTD based on CHRYSALIS evidence.

What is the expected effect on the cost effectiveness estimates?	Applying an exponential model to TTD based on CHRYSALIS evidence increased the ICER.
What additional evidence or analyses might help to resolve this key issue?	Details of NICE DSU TSD 14 criteria assessment to support TTD curve selection.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; NICE DSU TSD 14 = National Institute for Health and Care Excellence Decision Support Unit Technical Support Document 14; PFS = progression-free survival; TTD = time to treatment discontinuation	

Table 1.11; Key issue 10: Treatment waning

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The company did not explore treatment waning in the model, whereas the Evidence Review Group (ERG) considered that the assumption of a lifelong treatment effect may not be warranted.
What alternative approach has the ERG suggested?	An updated economic model including treatment waning scenarios.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	An updated economic model including treatment waning scenarios. Additional evidence to support the company's statement that treatment waning would be implicitly captured in the selected curves.
ERG = Evidence Review Group	

Table 1.12: Key issue 11: Exclusion of age-adjusted health state utilities in the CS base case

Report Section	**4.2.8**
Description of issue and why the ERG has identified it as important	In the company submission (CS) base case, the company did not include an age-adjustment to the health state utilities given the relatively short time horizon of the model, which is not in line with, good modelling practice, and exaggerated the cost effectiveness of amivantamab. This was subsequently provided as a scenario analysis by the company at clarification question stage.
What alternative approach has the ERG suggested?	Include age-adjusted health state utilities in the CS base case.
What is the expected effect on the cost effectiveness estimates?	Minor exaggeration of the incremental cost effectiveness ratio (ICER).
What additional evidence or analyses might help to resolve this key issue?	N/A
CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; N/A = not applicable	

Table 1.13: Key issue 12: Lack of a fully incremental analysis for all relevant comparators in the comparator basket.

Report Section	5.1
Description of issue and why the ERG has identified it as important	Amivantamab was compared to a basket of treatments. The comparator effectiveness and costs are therefore based on the average clinical effectiveness and costs across all the treatments included in the comparator basket, rather than a fully incremental analysis of all relevant comparators in the comparator basket. This increased uncertainty of estimates of amivantamab's cost effectiveness.
What alternative approach has the ERG suggested?	A fully incremental analysis of all relevant comparators in the comparator basket.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	A fully incremental analysis of all relevant comparators in the comparator basket.
ERG = Evidence Review Group	

Table 1.14: Key issue 13: lack of a fixed random seed in model PSA

Report Section	5.3
Description of issue and why the ERG has identified it as important	The differences between the probabilistic results when running the same model multiple times (i.e., without changing model settings, likely due to the lack of a fixed random seed in the model, adds to slightly different random draws each time the model runs, and consequent added uncertainty of the cost effectiveness estimates of amivantamab.
What alternative approach has the ERG suggested?	Implement fixed random seed to model PSA.
What is the expected effect on the cost effectiveness estimates?	The implementation of a fixed random seed will make the results of the model PSA reproducible.
What additional evidence or analyses might help to resolve this key issue?	Implement a fixed random seed to the model PSA.
ERG = Evidence Review Group; PSA = probabilistic sensitivity analysis	

1.6 Summary of the ERG's view

The CS base case probabilistic and deterministic ICERs were £40,246 and £39,764 per QALY gained, respectively. According to the company's model amivantamab is set to influence cost effectiveness by: 1) increased PPS, with an increment of 0.526 years (63% of total incremental LYs) in the amivantamab arm (1.349 years) compared with UK SoC (0.823 years); 2) increased PFS, with an increment of 0.314 years (37% of total incremental LYs) in the amivantamab arm (0.818 years) compared with UK SoC (0.504 years); and 3) the higher drug costs, administration costs and post-progression disease management costs. The two (probabilistic) ERG base case analyses resulted in ICERs of £55,043 per

QALY gained (when assuming all ERG changes and the inverse probability weighting (IPW) approach to determine comparative effectiveness) and £49,273 per QALY gained (when assuming all ERG changes and the propensity score matching (PSM) approach to determine comparative effectiveness). Time to treatment discontinuation (TTD) informed by parametric curves based on the CHRYSALIS trial protocol had the biggest impact in the ICER compared to the CS base case. The ICER increased most in the scenario analysis in which health state utilities were based on CHRYSALIS health-related quality of life (HRQoL) data. The ICER decreased most when assuming time to next treatment (TTNT) as a proxy for treatment discontinuation in SoC. It should be noted that the latter scenario assumes that TTNT is a good approximation to TTD, which is questionable according to the ERG (as discussed in Section 4.2.6. of this report).

In conclusion, there remains uncertainty about the effectiveness and relative effectiveness of amivantamab, which can be at least partly resolved by the company by conducting further analyses (e.g., incorporate the results of the indirect treatment comparison excluding TKIs in the model, perform a fully incremental analysis, and explore treatment waning). Moreover, the current assessment does not provide an appropriate estimation of the comparators listed in the scope. Therefore, the ERG believes that the CS nor the ERG report contains an unbiased ICER of amivantamab compared with relevant comparators (see Table 1.15).

Table 1.15: Summary of ERG's preferred assumptions and ICER

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS deterministic base case					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	39,764
UK SoC	[redacted]	[redacted]			
Fixing violation (1-Exclusion of age-adjustment to the health state utilities)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	40,293
UK SoC	[redacted]	[redacted]			
Matter of judgement (2-Use of PSM approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	45,790
UK SoC	[redacted]	[redacted]			
Matter of judgement (3-Implementation of parametric survival curves in SoC arm)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	41,401
UK SoC	[redacted]	[redacted]			
Matter of judgement (4-TTD informed by the CHRYSALIS trial protocol)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	55,695
UK SoC	[redacted]	[redacted]			
Deterministic ERG base case 1 (IPW approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	56,799
UK SoC	[redacted]	[redacted]			
Probabilistic ERG base case 1 (IPW approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	54,418
UK SoC	[redacted]	[redacted]			
Deterministic ERG base case 2 (PSM approach)					

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	52,185
UK SoC	[redacted]	[redacted]			
Probabilistic ERG base case 2 (PSM approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	49,880
UK SoC	[redacted]	[redacted]			

CS = company submission; ERG = Evidence Review Group; ICER = incremental cost-effectiveness ratio; IPW = inverse probability weighting; PSM = propensity score matching; QALY = quality-adjusted life year; SoC = standard of care; UK = United Kingdom

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM

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

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
Population	Adults with EGFR Exon 20 insertion-positive NSCLC after previous platinum-based chemotherapy.	Adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon20ins, whose disease has progressed on or after platinum-based chemotherapy.	Aligned with the licensed indication for amivantamab.	The population considered by the company is different than the population defined in the final NICE scope in a way that leads to potentially biased estimates of amivantamab efficacy, safety, and cost effectiveness.
Intervention	Amivantamab	Amivantamab monotherapy, administered via IV infusion 1,050 mg for patients with body weight <80 kg 1,400 mg for patients with body weight ≥80 kg	In line with the intervention received by patients falling within the licensed indication in the registrational CHRYSALIS trial.	No comment.
Comparator(s)	Established clinical management without amivantamab including: • Atezolizumab • Nivolumab (subject to an ongoing NICE appraisal)	UK SoC consisting of TKIs, IO agents, platinum-based chemotherapy and non-platinum-based chemotherapy.	Aligned with the final NICE scope. Further details can be found in Section B.1.3.2	There are differences between the comparators considered by the company and those listed in the final NICE scope (including the inclusion of TKIs other than nintedanib). These differences could have led to an exaggeration of the relative benefits of amivantamab.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	• Pembrolizumab (for disease with PD-L1 >1%) • Chemotherapy such as docetaxel alone or with nintedanib, pemetrexed and carboplatin			Because the company used the term TKIs without qualification, the ERG had assumed that this included nintedanib. However, in the FAC the company stated: "Beginning at submission and at any timepoint afterwards, nintedanib was not treated as a TKI in the Company's classification of treatments." Therefore, it appears that when the company stated TKIs what was intended was EGFR TKIs.
Outcomes	The outcome measures to be considered include: • overall survival • progression-free or disease-free survival • response rate • TTD • adverse effects of treatment • HRQoL	Key outcomes from the CHRYSALIS trial include: • ORR • CBR • DOR • PFS • TTF • OS • AEs • HRQoL	Additional outcomes (CBR) and DOR were included to capture the most important health benefits for amivantamab.	The ERG is satisfied with this justification.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be	The cost effectiveness of the treatments evaluated in this appraisal is expressed in	The genetic test for the EGFR Exon20ins mutation, with a scope covering small variant detection, is included in the National Genomic Test Directory. The directory specifies which	The ERG is satisfied with this justification.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	expressed in terms of incremental cost per QALY. 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 availability of any commercial arrangements for the intervention, comparator and subsequent treatment technologies will be taken into account. The use of amivantamab is conditional on the presence of an EGFR mutation. The economic modelling should include the costs associated with diagnostic testing for	terms of incremental cost per QALY. A lifetime time horizon was adopted to capture all relevant costs and health-related utilities. All costs and utilities were discounted at a rate of 3.5% per year in alignment with the NICE guide to the methods of technology appraisal. Costs were considered from an NHS and PSS perspective. The cost of diagnostic testing for EGFR Exon20ins mutations has not been included within the economic analysis.	genomic tests are commissioned by the NHS in England and is available at: https://www.england.nhs.uk/publication/national-genomic-test-directories/ EGFR Exon20ins mutations can be tested as part of the EGFR test conducted at diagnosis for all NSCLC patients. As such, Janssen, considers there are no additional costs likely to be incurred by the NHS over and above the current standard of care EGFR testing requirements for all NSCLC patients. Thus, the economic modelling excludes the costs associated with diagnostic testing for EGFR in people with NSCLC. This approach is aligned with that taken in previous appraisals in which testing for a specific mutation would be required (such as TA595, TA643 and TA670).[1-3] Some treatments comprising UK SoC (such as atezolizumab, pembrolizumab, nivolumab, afatinib and nintedanib) are subject to PASs. Due to their confidential nature, these discounts are not taken into account in the base case cost effectiveness analysis.	

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	EGFR in people with NSCLC 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				N/A
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.	None identified.	N/A – in line with the NICE final scope.	The sources cited by the company to support claims that there are special considerations related to equity or equality do not provide the support claimed by the company.

Based on Table 1 and pages 10 to 12 of the CS[4]

AE = adverse event; CBR = clinical benefit rate; CS = company submission; DCIS = ductal carcinoma in situ; DOR = duration of response; eBC = early breast cancer; EGFR = epidermal growth factor receptor; ERG = Evidence Review Group; Exon20ins = Exon 20 insertion mutations; HER2 = human epidermal growth factor receptor 2; IDFS = invasive disease-free survival; HRQoL = health-related quality of life; IO = immuno-oncology; IV = intravenous; mg = milligram; kg = kilogram; N/A = not applicable; NHS = National Health Service; NICE = National Institute of Health and Care Excellence; NSCLC = non-small-cell lung cancer; ORR = overall response rate; OS = overall survival

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
PAS = Patient Access Scheme; PD-L1 = progressed disease (level 1); PFS = progression free survival; pCR = pathological complete response; QALY = quality-adjusted life year; RID = residual invasive disease; SoC = standard of care; TKI = tyrosine kinase inhibitor; TTF = Time to treatment failure; UK = United Kingdom				

2.1 Population

The population defined in the scope is adults with EGFR Exon 20 insertion-positive non-small-cell lung cancer (NSCLC) after previous platinum-based chemotherapy.[5] The population in the CS is limited to *"Adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon20ins, whose disease has progressed on or after platinum-based chemotherapy."*[4]

According to the company the decision problem addressed in the CS is slightly different from the population specified in the final NICE scope. The main difference between the population defined in the NICE scope listed below (CS, Table 1, page 10),[4] is that, whereas the final NICE scope includes all those with EGFR Exon 20 insertion-positive NSCLC, the company limits the population to those with locally advanced or metastatic NSCLC with activating EGFR Exon20ins.

The population in the clinical trial for amivantamab in this indication, the CHRYSALIS trial, is: "*Adult patients (aged ≥18 years) with confirmed metastatic or unresectable NSCLC who failed or were ineligible for SoC therapy. Patients in part two of the study had measurable disease, with qualifying EGFR mutations or MET mutations or amplifications. Previous treatment with investigational EGFR Exon 20 ins-targeted TKIs was prohibited in the EGFR Exon20ins expansion cohort*."[4] The company also notes that they present data for a subset of the population in the CHRYSALIS trial related to: "*patients with EGFR Exon20ins mutations who had received previous treatment with platinum-based chemotherapy.*"[4]

On May 2021, the US Food and Drug Administration (FDA) approved amivantamab for "*adult patients with non-small cell lung cancer whose tumors have specific types of genetic mutations: epidermal growth factor receptor (EGFR) exon 20 insertion mutations*."[6] The European Medicines Agency (EMA) granted conditional approval for amivantamab for: "*adult patients with advanced non-small cell lung cancer (NSCLC) with activating epidermal growth factor receptor (EGFR) Exon 20 insertion mutations, after failure of platinum-based therapy.*"[7] Amivantamab was granted an innovation passport by the Medicines & Healthcare Products Regulatory Agency (MHRA) on 8th April 2021. On 15th November 2021, the MHRA granted a marketing authorisation for amivantamab for adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon 20 insertion mutations, whose disease has progressed on or after platinum-based chemotherapy.[8]

In their response to clarification, the company confirmed that "*the population in the submission is narrower than the NICE final scope population and is aligned with the licensed indication: adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon20ins mutations, whose disease has progressed on or after platinum-based chemotherapy.*"[9]

In addition, the CHRYSALIS trial had several inclusion criteria that made the studied population narrower than the one in the final NICE scope. These include: (i) histologically- or cytologically-confirmed NSCLC that was metastatic or unresectable; and (ii) an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1. In their response to clarification questions about this, the company stated: "*A situation in which the licensed indication is broader than the inclusion criteria of the pivotal clinical trial is not unusual as it permits equitable access to new therapies for patients who are not able to enrol in clinical trials. NICE appraise and make recommendations based on the licensed indication population. [10-13] The differences between the licensed indication and the CHRYSALIS trial population are common for oncology treatments (for example restricting to patients with ECOG status of 0 or 1), and mean that trial populations are generally, slightly fitter than the population in UK clinical practice for the reasons outlined in the bullets above*."[9]

The company consulted a clinical expert to inform their responses to clarification questions. The expert stated: "*Clinicians would consider amivantamab as a treatment option in some patients who are ECOG >1 if it was commissioned in such patients.*"[14]

ERG comment:

- The population specified in the decision problem appears to be different from the population defined in the final NICE scope. Although the specification of the mutation uses different wording, 'activating' can be regarded as implied because the insertion that is being referred to is one that is only relevant because it is activating i.e., causes activation of the EGFR pathway. However, only the experience of platinum-based chemotherapy is specified in the scope as opposed to having progressed to advanced, either locally advanced or metastatic disease, as expressed in the decision problem. Precisely which patients would be excluded is unclear, but presumably those who had not progressed. However, this should not be regarded as a key issue if NICE can only make a recommendation for the licensed population.
- With respect to the CHRYSALIS trial having a narrower population than the one defined in the final NICE scope (in the ways described above), the ERG notes that:
    - the results in the narrower trial population may not apply to patients in routine practice who may eventually receive amivantamab in the NHS setting; and
    - the company acknowledge that the patients in the trial might be "*fitter*"[9]
    - the clinical expert commissioned by the company stated that some patients in NHS clinical practice with an ECOG >1 (see above) would be considered for amivantamab. Therefore, the exclusion of patients with higher ECOGs may have led to an understatement of AEs, as these might have been more likely to arise in patients with worse performance statuses. It may also have impacted upon the effectiveness, as patients with the ECOG status required for admission to the CHRYSALIS trial may have been more likely to respond.

2.2 Intervention

The intervention (amivantamab) is broadly in line with the scope. In their submission, the company specifies that amivantamab is administered via IV (intravenous) infusion, and that the dose is 1,050 mg for patients with body weight <80 kg, and 1,400 mg for patients with body weight ≥80 kg.

The company also noted that patients receiving amivantamab also received *"any concomitant medications or treatments deemed necessary to provide adequate supportive care except for those listed as prohibited therapies.*"[4] The allowed concomitant medications included: symptomatic treatment, prophylactic medications. localised limited radiotherapy of short duration (e.g., 5 days) for palliative purposes only after discussion with approval by the sponsor's medical monitor.[4] The company provided a full list of concomitant medications in Table 43 of their responses to our clarification questions.[9]

The ERG asked the company whether the targeted radiotherapy could have been a confounder.[9] The company replied that "*the administration of these concomitant therapies would not have had an impact on ORR or DOR.*"[9] The company also consulted a clinical expert to inform their responses to clarification questions, and the clinical expert appeared to confirm that targeted radiotherapy could lead to a clinical benefit in a subset of patients: "*Palliative radiotherapy is part of supportive care and does not tend to cause any improvement in efficacy, except in patients who develop brain metastases treated by SRS. The latter population could derive clinical benefit from targeted radiotherapy (SRS).*"[14]

ERG comment: The ERG notes uncertainty regarding whether targeted radiotherapy confounded the results of the study. To confirm whether targeted radiotherapy confounded the study, the ERG would

need to see additional analyses that explored the effect of targeted radiotherapy on all outcomes (not just ORR or DOR).

2.3 Comparators

The description of the comparators in the NICE scope is as follows: *"Established clinical management without amivantamab including: atezolizumab, nivolumab (subject to an ongoing NICE appraisal), pembrolizumab (for disease with PD-L1>1%), or chemotherapy such as docetaxel alone or with nintedanib, pemetrexed and carboplatin."*[5]

The comparator chosen by the company is a pooled treatment basket in the form of real-world data to estimate clinical effectiveness and SoC in the cost effectiveness analysis: "*UK standard of care (SoC) consisting of TKIs, IO agents, platinum-based chemotherapy and non-platinum-based chemotherapy.*"[4] Table 52 (page 118) of the CS notes that immuno-oncology (IO) agents included atezolizumab, pembrolizumab and nivolumab, and the TKIs included afatinib and osimertinib.[4] Because the company used the term TKIs without qualification, the ERG had assumed that this included nintedanib. However, in the FAC the company stated: "Beginning at submission and at any timepoint afterwards, nintedanib was not treated as a TKI in the Company's classification of treatments." Therefore, it appears that when the company stated TKIs what was intended was EGFR TKIs.

The ERG asked the company to provide analyses in which TKIs other than nintedanib were excluded. In their response, the company provided hazard ratios (HRs) for PFS, OS, and TTNT for the base case, and a scenario excluding all TKIs (apart from nintedanib). The HRs were slightly higher in the scenario analysis with all EGFR TKIs excluded.

The company also claims that there is no SoC (CS, page 28), and notes a variety of treatments offered to this population (CS page 29, 30). The company therefore determined SoC with an advisory board with UK clinical experts who "*confirmed that patients with EGFR Exon20ins mutated NSCLC are treated in a manner broadly similar to patients without EGFR or anaplastic lymphoma kinase (ALK) mutations (i.e. no gene mutation or fusion protein), per NICE Guideline 122. Therefore, treatment options for patients in the UK may include the three pathways outlined in Table 4 below*."

The comparator chosen by the company is a pooled treatment basket in the form of real-world data. However, as specified in the scope, established clinical management depends upon line of therapy (first or later) and PD-L1 status. The ERG requested data from the company on the appropriateness of the comparator chosen by the company. More specifically, the company was asked to provide separate clinical effectiveness analyses (indirect treatment comparisons) by line of therapy and PD-L1 subgroup using only the comparators that would be standard care for the specific subgroup e.g., only pembrolizumab or nivolumab for PD-L1 positive patients. The response[9] from the company was as follows:

"Overall, Janssen maintain that a basket of comparators is the most appropriate comparator to amivantamab given expert feedback and the real-world evidence (RWE) indicating the heterogenous mix of treatments that patients receive in practice. Further, it is not considered appropriate to split the RWE data for SoC into subgroups given that this introduces additional uncertainty given the smaller sample sizes involved in such analyses, thus limiting their robustness. However, in order to provide some of the information requested in the ERGs question, subgroup analyses by line of therapy have been provided below. HRs are consistent with results from the base case;[9] however, these relative treatment effects are estimated for a restricted population and are therefore associated with greater uncertainty." In short, the company provided some, but not all, of the evidence requested by the ERG

The ERG also asked the company to further justify their definition of SoC.[9] The company responded by providing real world evidence (RWE) to "show that there is heterogeneity in the treatments used for this patient population with no definitive standard of care."[9]

ERG comment:

- There is evidence that EGFR TKIs (i.e., excluding nintedanib are unlikely to be effective against this EGFR Exon 20 insertion-positive NSCLC. In general *EGFR* Exon20ins mutations are known to be resistant to EGFR TKIs;[15-17] EGFR TKIs have shown limited to no activity in patients with Exon20ins mutations. Given the limited activity in this population, existing EGFR TKIs are rarely used in patients with *EGFR* Exon20ins mutation-positive NSCLC following platinum-based chemotherapy[18] (i.e., the position of the anticipated mobocertinib licence).
- EGFR TKIs are not included in the final scope for the ongoing appraisal of mobocertinib for treating *EGFR* Exon20ins mutation-positive NSCLC following platinum-based chemotherapy (ID3836). Given the amivantamab and mobocertinib appraisals target the same patient population, the comparators should be identical.
- Regarding the company's refusal to fully respond to the ERG's request to provide separate clinical effectiveness analyses by line of therapy and PD-L1 subgroup (to align with the final NICE scope):
    - The ERG understands that the smaller sample sizes in subgroups will lead to lower power. However, the extent to which the smaller subgroups would not be informative can only be verified after doing them. Comparison with the correct comparator in each subgroup is intended to address any potential bias notwithstanding the uncertainty.
    - Even if the estimates based on smaller subgroups are uncertain, they are appropriate to the decision problem defined in the final NICE scope.
    - Therefore, the ERG believes that the data should be presented according to different lines of therapy and PD-L1 status.
- With respect to the heterogeneity of SoC in standard practice, the ERG notes that heterogeneity in clinical practice does not imply that the company's determination is the correct one. The company might have explored a range of scenarios to explore whether a different choice of treatment basket would have made a difference to the main outcomes.

2.4 Outcomes

The NICE final scope lists the following outcome measures:[5]

- OS
- PFS or DFS
- Response rate
- TTD
- AEs of treatment
- HRQoL

The outcomes considered by the company were broadly in line with those listed in the final NICE scope, with two differences. Firstly, the company added two additional outcomes: clinical benefit rate, and duration of response. In addition, whereas the final NICE scope listed TTD as an outcome, the company listed time to treatment failure (TTF). In their response to clarification questions, the company states: "*TTF is identical to time to treatment discontinuation as it encompasses treatment discontinuation due to "any reason.*"

These were all assessed in the CHRYSALIS trial.

ERG comment:

- The ERG agrees that the company's definition of TTF (encompassing discontinuation for any reason) is the same as time to treatment discontinuation mentioned in the final NICE scope.
- With respect to the additional outcomes used by the company that go over and above those listed in the final NICE scope, the ERG recommends focusing on the outcomes listed in the final NICE scope.

2.5 Other relevant factors

2.5.1 Innovation

According to the company, amivantamab is innovative because it is the first targeted treatment for adult patients with EGFR Exon20ins mutated NSCLC.[4] The company claim that this is a population with a high unmet need and a poor prognosis.

An innovation passport was granted to amivantamab by the Medicines and Healthcare Products Regulatory Agency (MHRA) as part of the Innovative and Licensing and Access Pathway and enabled Janssen to apply for marketing authorisation under the MHRA accelerated regulatory pathway.[19]

ERG comment: The ERG agrees that amivantamab is innovative in that it meets the needs of an underserved population.

2.5.2 Equity and equality

The company states that the introduction of amivantamab to UK clinical practice has the potential to improve health inequity related to the stigma that can be associated with a lung cancer diagnosis, the relevance of cultural differences on treatment-seeking behaviours, and the impact of the COVID-19 pandemic on time to diagnosis.[4] The company emphasises that people of Asian heritage are more likely to receive a positive diagnosis for EGFR Exon20ins is also supported by the references they cite (including reference 3).[20] In their submission (Section B.1.4) The company claims that people who are diagnosed with lung cancer can be stigmatised due to its association with smoking.

The company also notes in their submission (Section B.1.4) that the stigma associated with a lung cancer diagnosis may be disproportionately high in Asian populations, and that this could be exacerbated by the COVID-19 pandemic. They state: "*Since many symptoms of lung cancer mimic those of COVID-19 (especially the persistent cough), people of Asian heritage who display lung cancer symptoms in public may face race-based prejudice and even outright racism as a result of public misunderstanding about the origins of the virus*."[4] On the basis of this background, in their submission on page 16, the company states that *"[t]his raises the prospect of patients being disproportionately disadvantaged on the basis of race*."[4]

The company concludes that these equity considerations are not inherently captured within the cost per QALY or budget impact frameworks.

ERG comment:

- The company presents strong evidence that there is stigma associated with lung cancer diagnoses.
- The references they use to support the claim that people of Asian origin may experience additional prejudice since some lung cancer symptoms overlap with those of COVID-19 (17, 18)[REFS 17, 18] were published in 2017 and 2016 respectively (before the pandemic) so do not support the company's claim. Therefore, the ERG notes that the claim about disproportionate prejudice or stigma towards people of Asian origin is highly speculative.

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The CS states that a de novo clinical systematic literature review (SLR) was conducted in January 2021, and updated September 2021, to identify relevant evidence on clinical efficacy and safety outcomes in patients with EGFR Exon20ins mutated NSCLC. The SLR was designed to capture data specifically in EGFR Exon20ins mutated NSCLC that was reported in both interventional and observational studies. Full details of the SLR search strategy, study selection process and results were reported in Appendix D.[21]

3.1.1 Searches

The following Section contains a summary and critique of all searches related to clinical effectiveness presented in the CS.[4, 21] The Canadian Agency for Drugs and Technologies in Health (CADTH) evidence-based checklist for the Peer Review of Electronic Search Strategies (PRESS), was used to inform this critique.[22, 23] The CS was checked against the single technology appraisal (STA) specification for company/sponsor submission of evidence.[24]

Appendix D of the CS provided details of the literature searches conducted for the SLR of clinical efficacy and safety outcomes.[21] Database searches were conducted in January 2021, then updated in September 2021. A summary of the resources searched is provided in Table 3.1.

Table 3.1: Resources searched for the clinical effectiveness systematic review (as reported in the CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print	Ovid	1946 to 19 January 2021 1946 to 24 September 2021	20/01/21 27/09/21
Embase	Ovid	1974 to 2021 January 19 1974 to 2021 September 24	20/01/21 27/09/21
CDSR	Cochrane Library, Wiley	Issue 1 of 12, January 2021 Issue 9 of 12, September 2021	20/01/21 27/09/21
CENTRAL	Cochrane Library, Wiley	Issue 1 of 12, January 2021 Issue 9 of 12, September 2021	20/01/21 27/09/21
DARE	University of York CRD platform	Issue 2 of 4, April 2015	20/01/21

Resource	Host/Source	Date Ranges	Dates searched
Additional resources			
ClinicalTrials.gov	Internet	06/05/20 05/10/21	06/05/20 05/10/21
AACR	Internet	2018 to 2021	January 2021 September 2021
ASCO			January 2021 September 2021
ESMO			January 2021 September 2021
ESMO ELCC			January 2021 September 2021
ESMO Asia			January 2021 September 2021
IASLC World Conference on Lung Cancer			January 2021 September 2021
IASLC European Conference on Lung Cancer			January 2021 September 2021
AACR = American Association for Cancer Research; ASCO = American Society of Clinical Oncology; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Clinical Trials; CRD = Centre for Reviews and Dissemination; DARE = Database of Abstracts of Reviews of Effects; ELCC = European Lung Cancer Congress; ESMO = European Society for Medical Oncology; IASLC = International Association for the Study of Lung Cancer			

ERG comment:

- The CS provided full details of the literature searches for the ERG to appraise.[4, 21]
- A good range of databases, clinical trials registry, grey literature resources, and a comprehensive range of relevant conference proceedings were searched.
- Full details of the database searches, including the database name, host platform, date range and date searched, were provided.
- Full details of the conference proceeding searches were provided. The search terms used, URL links, date range, date of searches, and results, were reported.
- Full details of the ClinicalTrials.com search was provided, including the search terms used (with an explanation that automatic synonym searching occurs in ClinicalTrials.com), all fields selected, date searched, and results.
- The database search strategies were well structured, transparent, and reproducible. They included truncation, proximity operators, synonyms, and subject headings (MeSH in MEDLINE and the Cochrane Library, and EMTREE in Embase). There were no language or date limits.
- It would have been preferable for the database search strategies to be presented exactly as run, rather than copied into a tabular format, as item 8 of the PRISMA-S checklist recommends.[25] The Cochrane Handbook also recommends that *"…bibliographic database search strategies should be copied and pasted into an appendix exactly as run and in full, together with the search set numbers and the total number of records retrieved by each search strategy. The search strategies should not be re-typed, because this can introduce errors"*.[26]

- The CS reported in Appendix D that the searches "aimed to capture a broader evidence base of EGFR + NSCLC", when the search strategy population facet actually combined 'NSCLC + advanced/metastatic + EGFR + mutations/TKI resistant': a much more focused approach.[21] It is unclear what effect this may have had on recall. The suggested broader approach would have been better and might have identified additional useful references. The searches would certainly have been improved if the set of search terms for 'mutations' had been omitted or had at least included search terms for 'exon 20 insertions'. However, the search strategy did not include an intervention/comparator facet of search terms, which probably compensated for the narrow focus of the population facet.
- Study design search filters for RCTs and observational studies were included in the search strategies. The search filters used were not cited, as current practice recommends.[25]
- Separate searches for safety outcomes were not conducted. It is unlikely that efficacy searches that include study design filters for RCTs and observational studies will be sensitive enough to identify safety data. Ideally, searches for AEs should be carried out alongside the searches for efficacy.[27]
- Targeted searches were conducted, as described in D.1.1.6: "*Ovid (MEDLINE and Embase), Google Scholar, and Google were additionally searched using terms for "exon 20 insertions" and "non-small cell lung cancer" to identify any additional, relevant studies for inclusion not identified via the database searches or other supplementary sources*".[21] It was not clear why search terms for 'exon 20 insertions' were not included in the main search strategies in the first instance, negating the need to conduct targeted searches. Full details of the search strategies or search terms used, date range, date searched, and results were not provided. In response to the ERG clarification letter the company explained that "*as no search terms specific to Exon 20 insertions (Exon20ins) were included in the database search strategies, additional targeted searches were conducted to increase the comprehensiveness of the review*" and full details of the search strategies were provided.[9]
- As the SLR for clinical efficacy and safety did not identify relevant evidence, the company conducted an additional SLR of prognostic patient and disease characteristics to identify potential confounders for the adjusted treatment comparison.[28] Searches were conducted for clinical guidelines, SLRs, and real-world observational studies in Embase and MEDLINE. The searches were limited to English language studies published between 2018 and 2020 and were conducted on 31st August 2020. Full details of the search strategies were not reported. Details of the search terms used were provided in response to the ERG clarification letter.[9]

3.1.2 Inclusion criteria

The eligibility criteria used in the search strategy for RCTs and non-RCTs is presented in Table 3.2.

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

	Inclusion criteria	Exclusion criteria
Population	Patients with metastatic or surgically unresectable EGFR mutation positive NSCLC harbouring exon 20 insertion mutations, specifically: • Patients with stage IIIB, IIIC or IV disease • Studies with patients only specified as "stage 3" will be eligible only if stage 4 patients	• Patients with lung cancer not otherwise specified • Patients with NSCLC not otherwise defined • Patients with locally advanced disease not otherwise specified • Patients with stage 3 disease not otherwise specified, with no stage 4 patients included in the same study

	are also included within the study population • Studies wherein staging was unclear, but patients received targeted therapy (e.g., TKIs), and were confirmed to harbour EGFR Exon20ins, were considered relevant unless their use was clearly out-of-scope (e.g., adjuvant use, or a trial specifically investigating interventions in early-stage patients)	• Patients without an EGFR Exon20ins • Patients that include eligible and ineligible populations but where results for the eligible population are not presented separately
Interventions	Any therapeutic or palliative intervention administered within the healthcare system	In addition to the comparator stated in the scope (BSC), other interventions (both first and second-line) were searched in the systematic review. Studies where patients received a therapy as first-line treatment were later excluded for the purpose of this submission.
Comparators	Any comparator (or none)	N/A
Outcomes	• Baseline characteristics of eligible patients, including: Demographics Disease characteristics • Clinical efficacy outcomes, including: OS PFS DFS ORR CBR/DCR Relapse/recurrent free survival DOR TTTD TTNT • Safety outcomes, including but not limited to: AEs SAEs • QoL outcomes • Patient-reported outcomes (RCTs only)	• Economic outcomes • Epidemiological outcomes • Patient-reported outcomes (non-RCTs and observational studies)
Study design	• RCTs • Interventional non-RCTs (i.e., non-randomised and uncontrolled clinical studies), including compassionate use programmes	• Editorials, commentaries, narrative reviews, guidelines, letters (unless they contain novel data) • Case reports

	• Observational studies (e.g., prospective/retrospective cohorts, case-control, chart reviews) • Case series • Conference abstracts published in the last 3 years • Post-hoc/pooled analyses • SLRs and (network) meta-analyses These will be considered relevant at the title/abstract review stage and hand searched for relevant primary studies, but will be excluded during the full-text review stage unless they themselves present original research	
Language and other restrictions	• Human subjects • English language abstract/full text	N/A
Based on table 7, appendix D, CS[21] AE = adverse event; BSC = best supportive care; CBR = clinical benefit rate; CS = company submission; DCR = disease control rate; DFS = disease-free survival; DOR = duration of response; EGFR = epidermal growth factor receptor; N/A = not applicable; NSCLC = non-small-cell lung cancer; ORR = overall response rate; OS = overall survival; PFS = progression free survival; QoL = quality of life; RCT = randomised controlled trial; RFS = relapse-free survival; SAE = serious adverse event; SLR = systematic literature review; TKI = tyrosine kinase inhibitor, TTNT = time to next treatment; TTTD = Time to treatment discontinuation.		

ERG comment:

Inclusion criteria - The ERG in its clarification letter asked the company to discuss how the SLR eligibility criteria on population is relevant to the NICE final scope for this submission. In discussing the submission population in their response, the company stated that, *"This submission focused on adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon20ins mutations, whose disease has progressed on or after platinum-based chemotherapy. This is in line with the UK marketing authorisation for amivantamab, but is narrower than the population defined in the final scope from NICE as locally advanced or metastatic disease is specified."*[9] The implications of this narrower population have already been discussed in Section 2.1 of this report. Concerning the SLR population, they explained that *"Whilst disease staging eligibility criteria for the SLR were narrower than that of the final scope, the SLR included a slightly broader population than the NICE scope in terms of treatment experience. Specifically, treatment naïve and chemotherapy naïve patients were included in the SLR; however, studies conducted in patients progressing on or after platinum-based chemotherapy were reported separately in the SLR write-up as these data were considered most relevant to the submission."*[9] As the company considered only one of the 88 studies identified during the SLR as being eligible for data extraction, to provide relevant evidence for the efficacy and safety of amivantamab in the submission population, it is unlikely that relevant studies may have been omitted due to a narrower disease staging criteria in the SLR.

Language restrictions - The ERG notes that an English language only restriction was applied to the SLR search. The ERG considers excluding non-English language studies to be inappropriate for obtaining robust evidence on the treatment of adults with advanced NSCLC.

3.1.3 Critique of data extraction

Appendix D of the CS provides clarity on the process of data extraction. Studies that were deemed to meet the inclusion criteria after screening were split into two categories, whereby they underwent either abbreviated or full data extraction.

Studies that reported only qualitative data on patients harbouring EGFR Exon20ins, containing individual participant data (IPD) only, or indicating that patients with Exon 20 insertions had been enrolled but no further details have been provided had an abbreviated data extraction. This involved the collecting of qualitative study characteristics and trial details.

Studies where quantitative data on patients harbouring EGFR Exon20ins were reported, either comprising the entire population studied or a separately reported exon 20 insertion subgroup underwent full data extraction. This involved obtaining detailed characteristics of the study and the participants, along with extracted numerical data on various efficacy, safety, and quality of life outcomes.

The CS reports that data extraction was performed by a single reviewer and then a second reviewer independently checked and verified the extracted data. Any disagreements or discrepancies were discussed between the two reviewers until a consensus could be reached. A third reviewer provided arbitration where consensus between the first two reviewers could not be achieved.

ERG comment: The methodology represents an accepted and viable process although the optimal process would be to ensure two independent data extraction processes. The CS does not clarify whether the third reviewer independently examined and extracted the data and then compared this data to the first extraction and check, or whether a further verification was provided of the initial extraction, before deciding.

3.1.4 Quality assessment

The CS reports that quality assessment of all included RCTs that underwent full data extraction was completed by one reviewer and then verified by a second independent reviewer. RCTs were appraised using the quality assessment tool developed by the York University Centre for Reviews and Dissemination (CRD), while interventional non-RCTs and observational studies that underwent full data extraction was assessed using the ROBINS-1 checklist.

ERG comment: No information is provided to determine how disagreements were resolved. As for data extraction, the optimal process would be to ensure two completely independent quality evaluation processes.

3.1.5 Evidence synthesis

A clinical SLR was conducted in January 2021, and updated September 2021, to identify relevant evidence on clinical efficacy and safety outcomes in patients with EGFR Exon20ins mutated NSCLC. The SLR was designed to capture data specifically in EGFR Exon20ins mutated NSCLC that was reported in both interventional and observational studies. Because the company only used one trial (CHRYSALIS) for their analysis, they did not conduct a meta-analysis.

The CHRYSALIS trial provided data for primary outcome, namely ORR, and secondary outcomes, DOR, PFS, TTF, OS and HRQoL. A subgroup analysis was also conducted on ORR by age (four categories), sex, race (Asian versus non-Asian), ECOG status (0 versus ≥1), history of smoking and prior immunotherapy.

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

This Section of the report details the sources of evidence in the CS for the clinical effectiveness of amivantamab. According to Section B.2.1 of the CS[4] only one study identified CHRYSALIS (NCT02609776). CHRYSALIS is a Phase 1b, single arm, first-in-human, open-label, multicentre, 2-part trial. The study provided data on efficacy and safety on patients with EGFR-mutated NSCLC with Exon20ins receiving amivantamab treatment. Further details of this study are outlined in this Section.

Comparative evidence for the study was provided via two RWE sources. Data from PHE and a US pooled cohort (pooled data from Flatiron Health Spotlight, ConcertAI and COTA data sources). These datasets are discussed in Section 3.3.

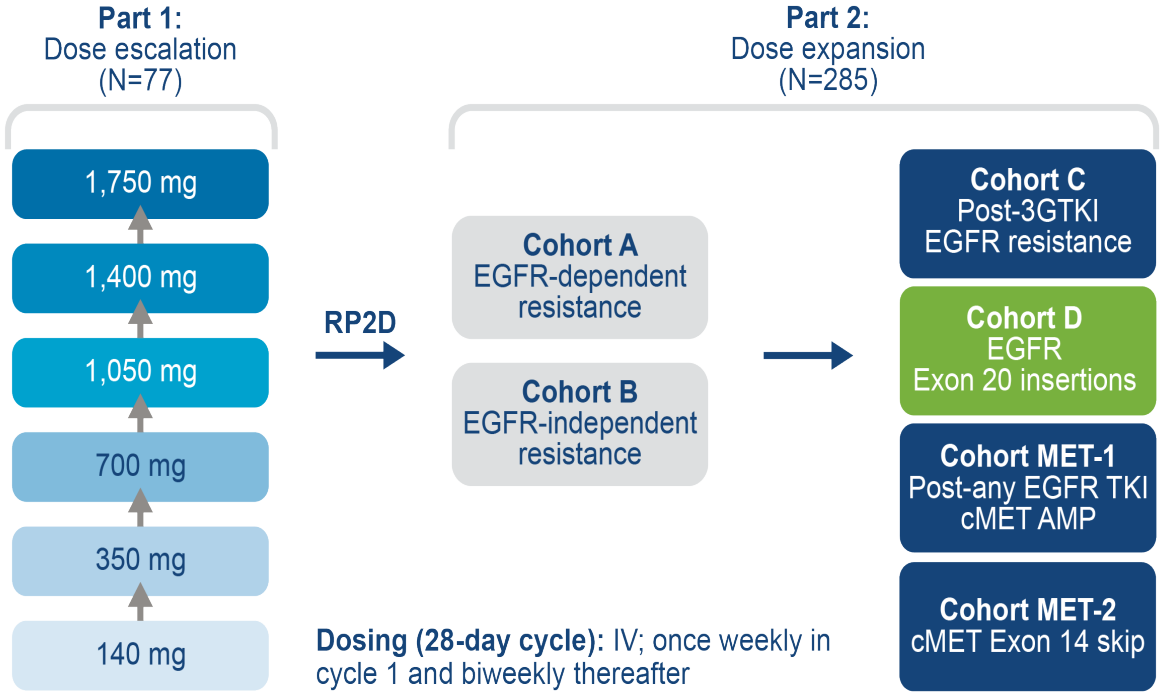
3.2.1 Details of the included trial

The CHRYSALIS trial is an ongoing Phase 1b trial in patients at least 18 years of age, with advanced NSCLC. The study was also used to support the conditional marketing authorisation[8]. The study tested both amivantamab as monotherapy and in combination with lazertinib. In the CS[4] only the monotherapy results are presented and discussed.

CHRYSALIS is a two-part trial consisting of a dose escalation phase (Part 1) and a dose expansion phase (Part 2) (see Figure 3.1 for the study design's overview). The aim of Part 1 (N=77) was to determine the recommended phase 2 dose (RP2D) of amivantamab monotherapy based on safety, pharmacokinetic, pharmacodynamic, and anti-tumour activity data. Six doses were tested from 140 mg to 1,750 mg. It concluded on a weight-based determination of the RP2D at 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.

The aim of Part 2 (N=285) was to better define the safety and pharmacokinetics at the RP2D, and to examine clinical activity within subgroups defined by tumour molecules. This part of the study consisted of six molecularly defined Cohorts: A, B, C, D, mesenchymal epithelial transition (MET)-1 and MET-2 (Figure 3.1), including patients with locally advanced or metastatic NSCLC patients with activating EGFR and/or MET mutation. Further patient eligibility criteria are provided in Table 3.3 as detailed in Table 8 of the CS[4]. The patients used for the CS are a subset of Cohort A and Cohort D, from now on termed D+ (N=114). It includes patients treated at the RP2D, aligned to the decision problem criteria defined in Table 1 of the CS[4] as *"adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon20ins, whose disease has progressed on or after platinum-based chemotherapy"*.[4]

Figure 3.1: Study design, CHRYSALIS (NCT02609776) study



Source: Figure B.2.1 of the CS [4]
CS = company submission; EGFR = epidermal growth factor receptor; IV = intravenous; MET = mesenchymal epithelial transition; RP2D = recommended Phase 2 dose; TKI = tyrosine kinase inhibitor

Table 3.3: Summary of study methodology, CHRYSALIS

Study design	International, multicentre, Phase 1b, single arm, first-in-human, open-label, 2-part trial
Study objective[29]	Primary objectives Part 1 Dose Escalations • Determine the maximum tolerated dose (MTD), if one exists (Part 1 monotherapy dose escalation only), and the recommended Phase 2 dose (RP2D) for subjects with NSCLC Part 2 (Expansion) • Determine the safety, tolerability, and anti-tumour activity at the RP2D • Estimate the anti-tumour activity at the RP2D, and in selected populations of subjects with documented EGFR or MET mutation(s) who have progressed after treatment with SoC
	Secondary objectives: • Assess additional measures of clinical benefit • Assess the PK and immunogenicity in subjects with NSCLC
Locations	90 sites in 11 countries, including the UK (three sites)
Eligibility criteria for participants	Key inclusion criteria: • Adult patients (≥18 years of age) • Histologically- or cytologically-confirmed NSCLC that was metastatic or unresectable • Progressed on or after prior therapy or were not candidates for currently available approved therapeutic options

	• Must have measurable disease according to RECIST v1.1 • An ECOG performance status of 0 or 1 • Qualifying EGFR mutations or MET mutations or amplifications • Previously diagnosed activating EGFR Exon20ins not previously treated with a TKI having known activity in Exon20ins disease (e.g., poziotinib) but previously treated with a platinum-based chemotherapy regimen • Adequate organ and bone marrow function, as assessed by laboratory measurements of haemoglobin, absolute neutrophil count, platelets, alanine aminotransferase, aspartate aminotransferase, total bilirubin and serum creatine Key exclusion criteria: • Prior chemotherapy, targeted cancer therapy, immunotherapy, or treatment with an investigational anti-cancer agent within 2 weeks or four half-lives whichever is longer, before the first administration of study drug • Untreated or active brain metastases • A history of malignancy other than the disease under study within 3 years before screening • A history of clinically significant cardiovascular disease • Known allergies, hypersensitivity, or intolerance to amivantamab or its excipients • Received an investigational drug (not including anti-cancer therapy) or used an invasive investigational medical device within 6 weeks before the planned first dose of study drug • Uncontrolled inter-current illness, including but not limited to poorly controlled hypertension or diabetes, ongoing or active infection, or psychiatric illness/social situation that would limit compliance with study requirements • Any specifically listed comorbidities such as leptomeningeal disease, human immunodeficiency virus (HIV), hepatitis B or C, and interstitial lung disease (ILD) • Any serious underlying medical or psychiatric condition
Study status	• Ongoing • Efficacy data cut-off date: 30th March 2021
Concomitant medication(s)	Permitted: Symptomatic treatment, prophylactic medications, localised limited radiotherapy of short duration (e.g., 5 days) for palliative purposes may be permitted but only after discussion with approval by the sponsor's medical monitor Disallowed: Any chemotherapy, anti-cancer therapy (other than study treatment(s)), or experimental therapy; radiotherapy to tumour lesions being assessed for tumour response prior to radiographic progression; use of live attenuated vaccines; use of phenytoin or phosphenytoin with carboplatin; nephrotoxic or ototoxic agents should be cautiously used with carboplatin; caution should be exercised when administering pemetrexed concurrently with a nonsteroidal anti-inflammatory drug to a participant whose creatinine clearance is <80 mL/min

RP2D	Amivantamab monotherapy, administered via IV infusion • 1,050 mg for patients with body weight <80 kg • 1,400 mg for patients with body weight ≥80 kg
Study outcome(s) (Part 2)	Primary outcome: • ORR Secondary outcomes: • CBR • DOR • PFS • OS • TTF • The best percentage change from baseline in SoD • HRQoL (exploratory descriptive analyses): PGIS, PGIC, NSCLC-SAQ and EQ-5D-5L VAS
Pre-planned subgroups	• Age: <65 versus ≥65 years and <75 versus ≥75 years • Sex: male versus female • Race: Asian versus non-Asian (patients with unknown race were not included in the subgroup analysis) • Baseline ECOG performance status: 0 versus ≥1 • History of smoking: yes, versus no • Prior immunotherapy: yes, versus no • Key EGFR Exon20ins variants (based on ctDNA analysis of pre-treatment samples). The change in SoD for target lesions was also described for these subgroups using a waterfall plot.

Based on Table 8. of the CS[4]
CBR = clinical benefit rate; CS = company submission; ctDNA = circulating tumour deoxyribonucleic acid; DOR = duration of response; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; EQ-5D-5L VAS = EuroQoL five-dimensions five-levels visual analogue scale; HRQoL = health-related quality of life; HIV = human immunodeficiency virus; ILD = interstitial lung disease; IV = intravenous; MET = mesenchymal epithelial transition; NSCLC = non-small-cell lung cancer; NSCLC-SAQ = Non-Small-Cell Lung Cancer Symptom Assessment Questionnaire; ORR = overall response rate; OS = overall survival; PFS = progression free survival; PGIC = Patient Global Impression of Change; PGIS = Patient Global Impression of Severity; PRO = patient-reported outcomes; RECIST = Response Evaluation Criteria in Solid Tumours; SoC = standard of care; SoD = sum of diameters; TKI = tyrosine kinase inhibitor; TTF = time to treatment failure; UK = United Kingdom

The number of patients coming from different cohorts to comprise Cohort D+ was not clear in the CS. After seeking clarification[30] the company provided these details, reported in Table 3.4. Although the CS stated that Cohort D+ "*consists largely of a subset of Cohort D and small number of patients in Cohort A*", which are both Cohorts of the study's Part 2, in their response they have now reported that five patients came from Part 1.

Table 3.4: Breakdown of patient numbers from CHRYSALIS; post platinum EGFR Exon20ins RP2D expanded efficacy analysis set (N=114)

Part and Cohort	Number of patients (N=114)
Part 1	[redacted]
Part 2 Cohort A	[redacted]
Part 2 Cohort D	[redacted]

Based on Table 17 in the clarification response[9]

ERG comments:

- In a clarification question the ERG inquired why since the decision problem did not specify performance status for the population, the evidence included in the CS was limited to patients with ECOG performance status 0 or 1.[9] The company has responded that "*…the NICE final scope is slightly broader than the marketing authorisation for amivantamab. As NICE appraise within the marketing authorisation, the marketing authorisation for amivantamab represents the population for the decision problem… the CHRYSALIS trial includes patients with an ECOG performance status of 0 or 1; i.e., a narrower population than the marketing authorisation. These data are the data upon which the marketing authorisation was granted and Janssen is requesting access for the licensed indication. That the CHRYSALIS trial, similar to most oncology trials, excludes some patients covered by the marketing authorisation does not mean that this submission is for a restricted population. The decision to treat patients above ECOG 1 is driven by the fitness of the patient and this would be based on the clinical assessment by the oncologist for treatment rather than mandated in the license. In alignment with this, a clinical expert consulted by Janssen during the development of this response document stated that clinicians would consider amivantamab as an option in some patients with ECOG >1*."[9]. The ERG points out that the population used for evidencing the efficacy and safety of the drug in question should match the characteristics of the perspective population under treatment. The standardised criteria which make out the assessment of ECOG status to measure the patient's performance status, are key to defining the population eligibility criteria for inclusion in the study as well as its progress within the study.
- The ERG inquired on the method that was used to identify EGFR Exon20ins mutations, for inclusion in the CHRYSALIS trial and if it was comparable (including with respect to specific mutations detected and limits of detection) with testing currently in place in routine practice in the UK[30]. The company has responded that in "*CHRYSALIS, EGFR Exon20ins mutations were assessed by local testing in the respective clinical trial centre locations or centrally using NGS testing for circulating tumour DNA (ctDNA), or tumour tissue where available. For central testing, Guardant was used for ctDNA while ThermoFischer was used for tumour tissue. The methods used are comparable to those available to patients in the UK as included on the NHS National Genomic Test Directory as part of the services provided by the Genomic Lab Hubs (GLHs)*"[9].
- According to the CHRYSALIS trial protocol[29], pemetrexed is included as concomitant medication to amivantamab, but it is also listed as a comparator. The ERG sought clarification on whether pemetrexed is a comparator, or part of SoC to be used alongside amivantamab, or both. The company has responded that data "*from CHRYSALIS presented in the submission are limited to patients enrolled and treated with amivantamab monotherapy in the dose escalation (Part 1) and dose expansion (Part 2) phases of the clinical trial. Thus, pemetrexed is not included in the intervention technology, and is listed appropriately as an example of treatments comprising "established clinical management without amivantamab" within the scope. The reference to pemetrexed in the CHRYSALIS protocol relates to a separate cohort which is not relevant for this submission. In one of the three cohorts in the dose escalation phase of the trial, patients were treated with amivantamab in combination with standard of care carboplatin and pemetrexed*."[9]. The ERG is satisfied that the company has clarified the use of pemetrexed in the trial and in the CS.
- Five patients in Cohort D+ came from Part 1 of the CHRYSALIS study as reported in Table 17 of the response to the clarification letter[9]. The CS stated that "*Patients enrolled to Part 1 were not required to meet any molecular eligibility requirements*."[4]. Nevertheless, in their response the company stated that Table 17 was a "*breakdown of the patient numbers comprising the efficacy analysis set N=114, patients with EGFR Exon20ins and post platinum chemotherapy who were treated at RP2D*"[9]. The CS also states that "*Part 1 was designed to determine the RP2D of*

amivantamab monotherapy in patients with advanced NSCLC based on safety, pharmacokinetic, pharmacodynamic, and anti-tumour activity data."[4]. The ERG is not entirely confident that the five patients included in Cohort D+, that came from Part 1 (dose escalation) of the study met the molecular eligibility requirements i.e., activating EGFR Exon20ins as well as the rest of the eligibility criteria that define Cohort D+.

- The ERG in its clarification letter asked the company to clarify if any of the patients in Cohort D+ received localised radiotherapy for palliative care, what criteria were used to select patients for this intervention, and what the recovery time between receipt of radiotherapy and amivantamab administration was. In response, the company stated that, *"During the on-treatment period, which was the time interval between the first dose of amivantamab and the end of treatment, 16 patients in the expanded efficacy analysis set (N=114) received palliative radiotherapy… 3 patients received palliative radiotherapy beyond the last dose date but before end-of-treatment, 1 patient received on-treatment salvage local therapy and 2 patients received on-treatment primary local therapy."*[9] They also stated that, *"There were no specific criteria for patient selection and the decision was based on investigator judgement,"*[9] and, *"Among the patients that received on treatment palliative radiotherapy and restarted treatment, treatment with amivantamab was re-started within 6–17 days after the end of radiotherapy."*[9] This did not include the three patients who did not restart amivantamab following palliative radiotherapy.

3.2.2 Statistical analyses of the CHRYSALIS trial

The population included in the CHRYSALIS trial differs from the one used for the CS. The details of the primary and supportive trial populations are presented in Table 3.5. Statistical analyses of the CHRYSALIS trial are summarized in Table 3.6.

The company, after the ERG sought clarification, confirmed that the primary population for safety results, reported in Table 3.3 "*(N=153) included only patients with EGFR Exon20ins NSCLC whose disease had progressed after platinum-based chemotherapy and had received at least one dose of the study drug, amivantamab*."[9]

Due to ambiguity in the CS, the ERG requested a clarification on the data-off dates used for the primary and the supportive clinical efficacy data. The company has now reported that "*the efficacy evidence for the N=114 efficacy population is derived from the 30th of March 2021 data cut-off*" and "*Supportive clinical efficacy data for the N=81 efficacy population is derived from the 8th October 2020 and 30th March 2021 data cut-offs*"[9].

Table 3.5: Trial populations used for the analysis of outcomes of CHRYSALIS

Analysis Set	Definition
Primary trial populations	
Efficacy results	
Post-platinum patients with EGFR Exon20ins RP2D expanded efficacy population (N=114)	Primary population for efficacy results: This population included all patients with EGFR Exon20ins NSCLC who received the RP2D prior to 4th June 2020 data cut-off with ≥3 disease assessments as of the 8th October 2020 data cut-off.
Safety results	
Post-platinum patients with EGFR Exon20ins RP2D safety population (N=153)	Primary population for safety results: This population included all patients with EGFR Exon20ins NSCLC who received prior chemotherapy at the RP2D prior to the 30th March 2021 data cut-off

Analysis Set	Definition
Supportive trial populations	
Efficacy results	
Post-platinum patients with EGFR Exon20ins RP2D initial efficacy population (N=81)	Supportive population for efficacy results: This population included all patients who received the first dose of amivantamab as monotherapy on or before 5th February 2020 and were response-evaluable with ≥3 disease assessments or discontinued treatment for any reason, including disease progression/death, prior to the 8th June 2020 data cut-off
Safety results	
All Treated at RP2D safety population (N=380)	Additional safety population: All patients enrolled in Part 1 (dose escalation) or Part 2 (dose expansion) irrespective of mutation status or prior chemotherapy, who received at least one dose of amivantamab monotherapy consistent with the RP2D (1,050 mg for body weight <80 kg and 1,400 mg for body weight ≥80 kg).
All Treated safety population (N=489)	Additional safety population: All patients enrolled in Part 1 or Part 2 who received at least one dose of amivantamab monotherapy at any dose (i.e., RP2D and non-RP2D).
Based on Tables 12 and 13 of the CS[4] CS = company submission; EGFR = epidermal growth factor receptor; NSCLC: non-small-cell lung cancer; RP2D: recommended Phase 2 dose **Note**: RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.	

Table 3.6: Summary of statistical analyses, CHRYSALIS trial

Hypothesis objective	The null hypothesis was that the ORR for amivantamab per RECIST v1.1 was ≤15%; the alternative hypothesis was that the ORR was ≥30%
Sample size, power calculation	The maximum total sample size at a RP2D for Part 2 was set to be approximately 460 patients, including approximately 40 patients in Cohort A, 20 patients in Cohort B, and up to 100 patients each if sufficient efficacy was observed in Cohorts C, D, MET-1, and MET-2 at a RP2D of amivantamab monotherapy. With a one-sided alpha of 2.5%, and a power of 87.5%, the total number of patients needed for each cohort was 86 response-evaluable patients. Assuming a non-evaluable rate of 15%, approximately 100 patients were to be enrolled within each cohort, although the number of patients was to be expanded beyond 100 patients (maximum of approximately 150) to further characterise activity for subpopulations within a cohort. The interim analysis was to be performed when approximately 30 patients were enrolled in each cohort and have sufficient data (i.e., post-baseline disease assessment) to be evaluable for response. Future enrolment into each cohort could have been terminated if it was determined during the first stage that the treatment was considered as ineffective as compared to other treatment options and/or not well tolerated. The sample size consideration for the subgroup in Cohort D who required to have had previous therapy with a combination platinum-doublet chemotherapy regimen was based on the null hypothesis of ORR ≤12%, and the alternative hypothesis of ORR >25%. To have a power of 80% to reject the null hypothesis with a one-sided alpha of 0.025, at least 60 patients were required to be enrolled in the subgroup; approximately 100 patients were targeted for enrolment to characterise the activity of amivantamab in this population.

Statistical analysis	Primary efficacy analysis of ORR with confirmed best overall responses was performed approximately 12 weeks after the last patient received the first infusion or at the end of study, whichever came first. The data cut-off was communicated to the sites. Any additional data were reported to the appropriate health authorities when all patients had finalised treatment with amivantamab. ORR was defined as the proportion of patients who achieved either a CR or PR in all treated analysis set (or response evaluable analysis set for interim monitoring) each expansion cohort (Part 2), as defined by investigator assessment using RECIST v1.1. Observed ORR along with their two-sided 95% exact CIs were presented for each cohort and dose level as appropriate. The null hypothesis for Cohort D was that the ORR was less than or equal to 15%, which was rejected if the lower bound of the 95% CI was greater than 15%. To control the overall type I error rate at 5% within each cohort, a sequential testing strategy was used. The hypotheses testing for subgroup within each cohort was only performed after null hypothesis for the whole cohort was rejected. The null hypothesis for the subgroup in Cohort D who require at least one prior line of platinum-containing chemotherapy is ORR ≤12%, which was rejected if the lower bound of the 95% CI was greater than 12% and was only tested after the null hypothesis for Cohort D (ORR ≤15%) was rejected.
Data management, patient withdrawals	A patient was withdrawn from the study for any of the following reasons: • Lost to follow-up • Withdrawal of consent for follow-up If a patient was lost to follow-up, every reasonable effort was made by the study site personnel to contact the patient and determine the reason for discontinuation/withdrawal. The measures taken to follow up were documented. In accordance with local regulations, information from public records were used to collect any missing survival data.
Based on Table 14 of the CS[4] CI = confidence intervals; CR = complete response; CS = company submission; ORR = overall response rate; PR = partial response; RECIST = response evaluation criteria in solid tumours; RP2D = recommended Phase 2 dose	

ERG comment: The ERG notes that there is a large difference between the efficacy and safety populations in terms of number of participants (N=114 versus N=153, or 34% more participants in the safety populations). This seems to be related to the fact that the efficacy population included only those who received the intervention prior to 4th June 2020 data cut-off with ≥3 disease assessments as of the 8th October 2020 data cut-off. The ERG does not know what the implications of this discrepancy are, but recommends the use of the ITT population for all analysis of all outcomes.

3.2.3 Baseline characteristics of the CHRYSALIS trial

Table 3.7 summarises the key baseline disease and demographic characteristics. The majority of patients were <75 years old (N=105, 92.1%), female (61.4%), Asian (51.8%), of normal weight (57%) and were non-smokers (57%). Most of the population had cancer Stage IV at initial diagnosis (78.9%) and an ECOG performance status 1 (70.2%). All patients had received platinum-based chemotherapy, as per inclusion criteria, while 43.9% had received IO agents. Only [redacted] UK patients were included in Cohort D+.

Table 3.7: Key patient baseline characteristics, CHRYSALIS trial (expanded efficacy population)

Baseline characteristic	Post-platinum patients with EGFR Exon20ins at RP2D (N=114)
Age (years)	
Mean (SD)	61.8 (10.0)
Median (range)	62.0 (36–84)
Age category, n (%)	
<65, n (%)	67 (58.8)
≥65, n (%)	47 (41.2)
<75, n (%)	105 (92.1)
≥75, n (%)	9 (7.9)
Sex, n (%)	
Male	44 (38.6)
Female	70 (61.4)
Race, n (%)	
Asian	59 (51.8)
Black or African American	3 (2.6)
White	42 (36.8)
Not reported	10 (8.8)
Weight, kg	
Mean (SD)	64.8 (15.8)
Median (range)	62.1 (35.4–115.0)
Body mass index, kg/m2	
Mean (SD)	24.1 (4.7)
Median (range)	23.5 (14.0–36.9)
Underweight (<18.5), n (%)	11 (9.6)
Normal (18.5–<25), n (%)	65 (57.0)
Overweight (25–<30), n (%)	25 (21.9)
Obese (≥30), n (%)	13 (11.4)
Initial diagnosis NSCLC subtype, n (%)	
Adenocarcinoma	109 (95.6)
Large cell carcinoma	0 (0)
Squamous cell carcinoma	3 (2.6)
Other	2 (1.8)
Histology grade at initial diagnosis, n (%)	
Moderately differentiated	23 (20.2)
Poorly differentiated	19 (16.7)
Well differentiated	7 (6.1)
Other	64 (56.1)
Not reported	1 (0.9)
Cancer stage at initial diagnosis, n (%)	
0	0 (0)
IA	7 (6.1)
IB	1 (0.9)
IIA	2 (1.8)

Baseline characteristic	Post-platinum patients with EGFR Exon20ins at RP2D (N=114)
IIB	4 (3.5)
IIIA	6 (5.3)
IIIB	4 (3.5)
IV	90 (78.9)
Location of metastasis, n (%)	
Bone	51 (44.7)
Liver	13 (11.4)
Brain	29 (25.4)
Lymph node	62 (54.4)
Adrenal gland	6 (5.3)
Other	62 (54.4)
Time from initial diagnosis of cancer to first dose, months	
Mean (SD)	22.3 (20.0)
Median (range)	17.5 (1.5–130.1)
Time from metastatic disease diagnosis to first dose, months	
Mean (SD)	18.3 (15.5)
Median (range)	15.5 (0.7–116.4)
Number of prior LOTs	
Mean (SD)	2.1 (1.3)
Median (range)	2 (1–7)
ECOG performance status, n (%)	
0	33 (28.9)
1	80 (70.2)
2	1 (0.9)
History of smoking, n (%)	
Yes	49 (43.0)
No	65 (57.0)
Prior systemic therapies of interest in ≥5% of patients, n (%)	
Platinum-based chemotherapy	[redacted]
EGFR TKI (1st generation)	[redacted]
EGFR TKI (2nd generation)	[redacted]
EGFR TKI (3rd generation)	[redacted]
IO agents	[redacted]

Based on Tables 9, 10 and 11 of the CS[4]

Note: RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg

CS = company submission; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; IO = immuno-oncology agent; LOT = lines of therapy; NSCLC = non-small-cell lung cancer; RP2D = recommended Phase 2 dose; SD = standard deviation; TKI = tyrosine kinase inhibitor

Generalisability to UK clinical practice

Given the large proportion of Asian patients in Cohort D+ (N=59, 51.8%), the ERG in its clarification letter[30] requested for the breakdown of the characteristics of those participants defined as Asian, as well as a discussion on the implications that this might have to the generalisability of the study population to the UK patient population. The baseline characteristics were provided by the company[9] and are now

presented in Tables 3.8 and 3.9. Regarding generalisability, the company stated that "*Clinical experts consulted by Janssen in the two advisory boards stated that the baseline characteristics of patients recruited to the CHRYSALIS trial broadly reflect those of patients seen in UK clinical practice. EGFR Exon20ins NSCLC is more prevalent in the Asian population than other races.*[14] *A clinical expert consulted by Janssen during the development of responses to this question stated that this was the case regardless of geographical location and that the proportion of Asian patients recruited to CHRYSALIS was broadly in line with what is seen in the UK.*[8] *Most patients with EGFR Exon20ins NSCLC are Stage IV at initial diagnosis.*[15] *The clinical expert also stated that the distribution of cancer stage at initial diagnosis seen in CHRYSALIS is reflective of clinical practice in the UK with most patients being Stage IV.*[9]

Table 3.8: Baseline demographic characteristics for patients defined as Asian (N=59) in CHRYSALIS expanded efficacy population (post-platinum patients with EGFR Exon20ins at RP2D, N=114)

Variable	Level / statistic	
Age	N	[redacted]
	Mean (SD)	[redacted]
	Median	[redacted]
	Range	[redacted]
Age (65 years threshold)	N	[redacted]
	<65	[redacted]
	≥65	[redacted]
Age (75 years threshold)	N	[redacted]
	<75	[redacted]
	≥75	[redacted]
Gender	N	[redacted]
	Male	[redacted]
	Female	[redacted]
Race	N	[redacted]
	Asian	[redacted]
Ethnicity	N	[redacted]
	Not Hispanic or Latino	[redacted]
Weight (kg)	N	[redacted]
	Mean (SD)	[redacted]
	Median	[redacted]
	Range	[redacted]
Height (cm)	N	[redacted]
	Mean (SD)	[redacted]
	Median	[redacted]
	Range	[redacted]
BMI (kg/m)	N	[redacted]
	Mean (SD)	[redacted]
	Median	[redacted]

Variable	Level / statistic	
	Range	[redacted]
BMI category	N	[redacted]
	Underweight (<18.5)	[redacted]
	Normal (18.5- <25)	[redacted]
	Overweight (25- <30)	[redacted]
	Obese (>30)	[redacted]

Based on Table 15 of the clarification letter[9]
Note: If race was not reported, then that subject is excluded from the race subgroup; Ns for each parameter reflect non-missing values.
BMI = body mass index; SD = standard deviation

Table 3.9: Baseline clinical and disease characteristics for patients defined as Asian (n=59) in CHRYSALIS expanded efficacy population (post-platinum patients with EGFR Exon20ins at RP2D, N=114)

Variable	Level / statistic	
Initial diagnosis NSCLC subtype	N	[redacted]
	Adenocarcinoma	[redacted]
	Squamous cell carcinoma	[redacted]
	Other	[redacted]
Histology grade at initial diagnosis	N	[redacted]
	Moderately differentiated	[redacted]
	Poorly differentiated	[redacted]
	Well differentiated	[redacted]
	Other	[redacted]
Cancer stage at initial diagnosis	N	[redacted]
	IA	[redacted]
	IB	[redacted]
	IIA	[redacted]
	IIB	[redacted]
	IIIA	[redacted]
	IIB	[redacted]
	IV	[redacted]
Bone metastasis	N	[redacted]
	No	[redacted]
	Yes	[redacted]
Liver metastasis	N	[redacted]
	No	[redacted]
	Yes	[redacted]
Brain metastasis	N	[redacted]
	No	[redacted]
	Yes	[redacted]

Variable	Level / statistic	
Lymph node metastasis	N	[redacted]
	No	[redacted]
	Yes	[redacted]
Adrenal gland metastasis	N	[redacted]
	No	[redacted]
	Yes	[redacted]
Other metastasis	N	[redacted]
	No	[redacted]
	Yes	[redacted]
Time from initial diagnosis of cancer to first dose	N	[redacted]
	Mean (SD)	[redacted]
	Median	[redacted]
	Range	[redacted]
Time from metastasis disease diagnosis to first dose	N	[redacted]
	Mean (SD)	[redacted]
	Median	[redacted]
	Range	[redacted]
Prior lines of treatment	N	[redacted]
	Mean (SD)	[redacted]
	Median	[redacted]
	Range	[redacted]
Prior lines of treatment (Category)	N	[redacted]
	1	[redacted]
	2	[redacted]
	3	[redacted]
	4	[redacted]
	5	[redacted]
	6	[redacted]
	7	[redacted]
ECOG	N	[redacted]
	ECOG 0	[redacted]
	ECOG 1+	[redacted]
Smoking history	N	[redacted]
	Yes	[redacted]
	No	[redacted]
Hepatic impairment at baseline	N	[redacted]
	Normal (Total bilirubin ≤ ULN and AST ≤ ULN)	[redacted]

Variable	Level / statistic	
	Mild (Total bilirubin ≤ ULN and AST > ULN) or (ULN < Total bilirubin ≤ 1.5 x ULN)	[redacted]
Renal impairment at baseline	N	[redacted]
	Normal (EGFR: ≥ 90 mL/min/1.73m2)	[redacted]
	Mild (EGFR: 60 to < 90 mL/min/1.73m2)	[redacted]
	Moderate (EGFR: 30 to < 60 mL/min/1.73m2)	[redacted]
Based on Table 16 of the clarification letter[9] ECOG = eastern cooperative oncology group; EGFR = epidermal growth factor receptor; NSCLC = non-small-cell lung cancer; SD = standard deviation		

ERG comment:

- In the CS it was not clear how many UK patients were included in Cohort D+. The company has now reported that there were only [redacted] whose *"…baseline demographic characteristics cannot be presented in order to avoid patient identification."*[9]. Although the company maintains that the generalisability of the study population is not affected by the race baseline characteristics, the subgroup analysis has detected differences, please see Section 3.2.5.7 of this report.
- The eligibility criteria stated that only patients with ECOG status 0 or 1 were to be included in the CHRYSALIS trial (Table 3.3), nevertheless, one patient with ECOG status 2 was included (Table 3.7). It is unclear why this patient was included, and how further baseline and clinical characteristics compare to other patients who might have been excluded due to ECOG status.

3.2.4 Risk of bias assessment of the CHRYSALIS trial

Table 3.10 presents the risk of bias assessment of the CHRYSALIS trial conducted using the ROBINS-I[31] tool for assessing risk of bias in non-randomised studies of interventions.Quality assessments were completed by one reviewer and verified by a second independent reviewer. The ERG undertook an independent risk of bias assessment using the same tool (ROBINS-I), whose results are reported in the same table.

Table 3.10: Quality assessment of the CHRYSALIS trial

	Risk of bias	
Source of bias	**CS**	**ERG**
Overall bias due to confounding	Low	Moderate
Overall bias in selection of participants into the study	Low	Low
Overall bias in classification of interventions	Low	Low
Overall bias due to deviations from intended interventions	Low	Low
Overall bias due to missing data	Low	Low
Overall bias in measurement of outcomes	Moderate	Moderate
Overall bias in selection of the reported results	Low	Low
Overall risk of bias	Moderate	Moderate
Based on Table 15 of the CS[4] CS = company submission; ERG = Evidence Review Group		

ERG comment: All parts of the systematic review, including the risk of bias assessment, should be undertaken by a team and not a single person to ensure errors are minimised. It is not clear in the CS whether more than one reviewer was involved in the risk of bias assessment. Nevertheless, the ERG largely agrees with the risk of bias assessment executed by the company. The only difference is the pre-intervention domain of 'bias due to confounding' as the study did not use a method to control for measured confounders. As a result, of only one domain rating change the overall risk of bias rating of the study was not altered.

3.2.5 Efficacy results of the CHRYSALIS trial

The company submitted efficacy results for one primary and several secondary outcomes as presented in Table 3.3. The expanded efficacy population was used (N=114) until the 30th March 2021 data cut-off. In addition, the supportive efficacy trial population (N=81) was used including data until the 30th March 2021. A summary of the outcomes for the expanded efficacy population is presented in Table 3.11. When applicable (all outcomes apart from TTF and OS) both INV and blinded independent committee review assessed (BICR) results were provided in the CS. Further details and critique are provided in the following Sections.

Table 3.11: Summary of key outcomes from the CHRYSALIS trial (30th March 2021 data cut-off)

Outcome	Result
ORR, n (%) [95% CI]	BICR: 49 (43.0) [33.7, 52.6] INV: 42 (36.8) [28.0, 46.4]
CBR, n (%) [95% CI]	BICR: 84 (73.7) [64.6, 81.5] INV: 86 (75.4) [66.5, 83.0]
Median DOR, months (95% CI)	BICR: 10.84 (6.90, 14.98) INV: 12.45 (6.54, 16.13)
Median PFS, months (95% CI)	BICR: 6.74 (5.45, 9.66) INV: 6.93 (5.55, 8.64)
Median TTF, months (95% CI)	8.08 (6.67, 10.64)
Median OS, months (95% CI)	22.77 (17.48, NE)

Based on Table 16 of the CS[4]
Note: DOR is calculated as the time from initial response (either complete or partial response) to PD or death; PFS is defined as the time from first infusion of amivantamab to PD or death; TTF is defined as the time from the first infusion of amivantamab to discontinuation of treatment for any reason, including disease progression, treatment toxicity and death; OS is defined as the time from first infusion of amivantamab to death due to any cause
BICR = blinded independent committee review; CBR = clinical benefit rate; CI = confidence interval; DOR = duration-of-response; INV = investigator assessed; NE = not evaluable; OS = overall survival; ORR = overall response rate; PD = progressive disease; PFS = progression-free survival; TTF = time to treatment failure.

3.2.5.1 Primary outcome: overall response rate (ORR)

The company defined ORR as, "*the proportion of patients with a best overall response of a confirmed CR or PR based on RECIST v1.1 criteria (best response as recorded in the CRF from the start of the amivantamab until disease progression, withdrawal of consent, or start of a subsequent anti-cancer therapy, whichever came first). ORR was based on investigator assessment and BICR assessment*." The results are provided in Table 3.12.

Table 3.12: Summary of best overall response based on RECIST v1.1 from the CHRYSALIS trial (30th March 2021 data cut-off)

	Post-platinum patients with EGFR Exon20ins at RP2D (N=114)	
	BICR	**INV**
Best overall response, n (%)		
CR	3 (2.6)	0 (0)
PR	46 (40.4)	42 (36.8)
SD	47 (41.2)	56 (49.1)
PD	15 (13.2)	14 (12.3)
Not evaluable/unknown	3 (2.6)	2 (1.8)
ORR, n (%) [95% CI]	49 (43.0) [33.7, 52.6]	42 (36.8) [28.0, 46.4]
CBR, n (%) [95% CI]	84 (73.7) [64.6, 81.5]	86 (75.4) [66.5, 83.0]

Based on Table 17 of the CS[4]
Note: CBR is defined as the percentage of patients achieving confirmed complete or partial response, or durable stable disease (duration of at least 11 weeks). RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.
BICR = blinded independent committee review; CBR = clinical benefit rate; CR = complete response; CS = company submission; CI = confidence interval; EGFR = epidermal growth factor receptor; INV = investigator assessed; ORR = overall response rate; PD = progressed disease; PR = partial response; RP2D = recommended Phase 2 dose; SD = stable disease.

In the CS, the company also compared the single arm results to SoC treatments for ORR, by separately comparing to SoC data from a US (RWE) and UK (PHE) Cohort (see Section 3.4 for further details).

Comparing BICR assessed data from CHRYSALIS with the US Cohort, based on an IPW-ATT approach for statistical adjustment (see Section 3.4), the adjusted OR for amivantamab versus SoC was [redacted]. Based on a multivariable proportional hazards regression model the adjusted OR for amivantamab versus SoC was [redacted].

Comparing INV data from CHRYSALIS with the US Cohort, based on an IPW-ATT approach for statistical adjustment (see Section 3.4), the adjusted OR for amivantamab versus SoC [redacted]. Based on a multivariable proportional hazards regression model the adjusted OR for amivantamab versus SoC was [redacted].

The ERG inquired whether the patients were still receiving treatment at the time of the evaluation of best overall response, as reported in Table 3.12. The company in its response to clarification[9] stated that "*Considering INV-assessed best overall response (BOR), all patients for whom a partial response or stable disease was their BOR achieved this whilst receiving treatment. Two patients were recorded as having a non-evaluable BOR since treatment was discontinued before the first disease evaluation.*

For BICR-assessed BOR, all patients with a BOR of complete response, partial response or stable disease achieved this whilst receiving treatment. Two patients were recorded as having a non-evaluable BOR since due to discontinuation of treatment before disease evaluation, and one patient had stable disease on Day 38, but this was not counted given that it did not meet the minimum window of 42 days for standard disease assessment as outlined by the CHRYSALIS trial protocol." The company has also

provided further details on the "*timing for the assessment of best overall response in relation to treatment*" which is reported in Table 3.13.

Table 3.13: Summary of best overall response based on RECIST v1.1 and timing of assessment; Post-platinum EGFR Exon20ins RP2D expanded efficacy population (N=114)

	Best overall response: post-platinum Exon20ins RP2D expanded efficacy population (N=114, 30th March 2021 data cut-off)			
	BICR		INV	
	n (%)	Timing of evaluation	n (%)	Timing of evaluation
CR	[redacted]	[redacted]	[redacted]	[redacted]
PR	[redacted]	[redacted]	[redacted]	[redacted]
SD	[redacted]	[redacted]	[redacted]	[redacted]
PD	[redacted]	[redacted]	[redacted]	[redacted]
Not evaluable/ unknown	[redacted]	[redacted]	[redacted]	[redacted]
ORR, n (%) [95% CI]	[redacted]		[redacted]	
CBR, n (%) [95% CI]	[redacted]		[redacted]	

Based on Table 17 of the CS[9]
AE = adverse event; BICR = blinded independent committee review; BOR = best overall response; CBR = clinical benefit rate; CR = complete response; CS = company submission; CI = confidence interval; EGFR = epidermal growth factor receptor; INV = investigator assessed; N/A = not applicable; ORR = overall response rate; PD = progressed disease; PR = partial response; RP2D = recommended Phase 2 dose; SD = stable disease.
Note: CBR is defined as the percentage of patients achieving confirmed complete or partial response, or durable stable disease (duration of at least 11 weeks). RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.

3.2.5.2 Secondary outcome: duration of response (DOR)

The company calculated DOR as "*time from initial response of CR or PR to PD or death due to underlying disease, whichever comes first, only for patients who achieve CR or PR*". The results are presented in Table 3.14. The Kaplan-Meier (KM) plots according to the BICR and INV assessments are illustrated in Figure 3.2 and Figure 3.3, respectively. The BICR identified a total of 49 responders while the INV identified 42. The respective median DOR were 10.84 months (95% confidence interval (CI): 6.90, 14.98) and 12.45 months (95% CI: 6.54, 16.13).

Table 3.14: Summary of duration of response (DOR) from the CHRYSALIS trial (30th March 2021 data cut-off)

	Post-platinum patients with EGFR Exon20ins at RP2D (N=114, 30th March 2021 data cut-off)	
	BICR	**INV**
Responders, n	49	42
Event, n (%)	27 (55.1)	21 (50.0)
Censored, n (%)	22 (44.9)	21 (50.0)
Time to event (months)		
25th percentile (95% CI)	5.13 (4.07, 8.21)	4.96 (4.14, 8.31)
Median (95% CI)	10.84 (6.90, 14.98)	12.45 (6.54, 16.13)
75th percentile (95% CI)	21.65 (11.04, NE)	16.13 (12.68, NE)
Range	1.1+, 21.7	1.1+, 19.0+
Duration of response ≥6 months, n (%)	27 (55.1)	27 (64.3)
Duration of study treatment (months)		
N	49	42
Mean (SD)	12.13 (5.77)	12.77 (5.09)
Median	13.37	13.59
Range	1.7, 23.9	2.3, 23.9

Based on Table 18 of the CS[4]

Note: RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.

BICR = blinded independent review; CI = confidence interval; CS = company submission; EGFR = epidermal growth factor receptor; INV = investigator; NE = not evaluable; RP2D = recommended Phase 2 Dose; SD = standard deviation.

Figure 3.2: Kaplan-Meier plot of DOR – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114) by BICR assessment

Source: Figure 7 of the CS[4]
CS = company submission; BICR = blinded independent review; DOR = duration of response

Figure 3.3: Kaplan-Meier plot of DOR – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114) by INV assessment

Source: Figure 8 of the CS[4]
CS = company submission; DOR = duration of response; INV = investigator

3.2.5.3 Secondary outcome: progression-free survival (PFS)

The company defined PFS as "*the time from first infusion of amivantamab to PD or death due to any cause*". The PFS data are provided in Table 3.15,while the KM curves are illustrated in Figure 3.4 and Figure 3.5 for the BICR and INV assessments, respectively. Median PFS was 6.74 months (95% CI: 5.45, 9.66) according to BICR and 6.93 months (95% CI: 5.55, 8.64) according to INV; while the median follow was [redacted] months (range: [redacted]).

Table 3.15: Summary of progression-free survival (PFS) from the CHRYSALIS trial (30th March 2021 data cut-off)

	Post-platinum patients with EGFR Exon20ins at RP2D (N=114, 30th March 2021 data cut-off)	
	BICR	**INV**
Event, n (%)	80 (70.2)	81 (71.1)
Censored, n (%)	34 (29.8)	33 (28.9)
Time to event (months)		
25th percentile (95% CI)	3.94 (2.66, 4.83)	3.71 (2.60, 4.34)
Median (95% CI)	6.74 (5.45, 9.66)	6.93 (5.55, 8.64)
75th percentile (95% CI)	12.45 (10.87, NE)	16.56 (12.58, NE)
Range	(0.0+, 23.3)	0.0+, 24.1
3-month event-free rate (95% CI)	0.78 (0.69, 0.85)	0.77 (0.68, 0.84)
6-month event-free rate (95% CI)	0.55 (0.45, 0.64)	0.55 (0.45, 0.64)
9-month event-free rate (95% CI)	0.41 (0.31, 0.50)	0.39 (0.30, 0.48)
12-month event-free rate (95% CI)	0.29 (0.21, 0.39)	0.35 (0.26, 0.44)
15-month event-free rate (95% CI)	0.22 (0.14, 0.31)	0.28 (0.19, 0.37)
18-month event-free rate (95% CI)	0.14 (0.06, 0.26)	0.18 (0.09, 0.30)
21-month event-free rate (95% CI)	0.14 (0.06, 0.26)	0.18 (0.09, 0.30)
24-month event-free rate (95% CI)	0 (NE, NE)	0.18 (0.09, 0.30)
27-month event-free rate (95% CI)	NR	0 (NE, NE)

Based on Table 19 of the CS[4]
Note: RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.
BICR = blinded independent review; CI = confidence interval; CS = company submission; EGFR = epidermal growth factor receptor; INV = investigator; PFS = progression-free survival; NE = not evaluable; NR = not reported; RP2D = recommended Phase 2 dose.

Figure 3.4: Kaplan-Meier plot of PFS – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114) by BICR assessment

Source: Figure 9 of the CS[4]
BICR = blinded independent committee review; CS = company submission; IN = investigator; PFS = progression-free survival; RP2D = recommended Phase 2 dose

Figure 3.5: Kaplan-Meier plot of PFS – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114) by INV assessment

Source: Figure 10 of the CS[4]
BICR = blinded independent committee review; CS = company submission; INV = investigator; PFS = progression-free survival; RP2D = recommended Phase 2 dose

The CS also compared the single arm results to SoC treatments for PFS, by separately comparing to data from a US and UK Cohort (see Section 3.4 for further details).

Comparing BICR data with the US SoC data, based on an IPW-ATT approach for statistical adjustment (see Section 3.4), the adjusted HR for amivantamab versus SoC was [redacted]. Based on a multivariable proportional hazards regression model the adjusted HR for amivantamab versus SoC was [redacted].

Comparing INV data with the US SoC data, based on an IPW-ATT approach for statistical adjustment (see Section 3.4), the adjusted HR for amivantamab versus SoC was [redacted]. Based on a multivariable proportional hazards regression model the adjusted HR for amivantamab versus SoC was [redacted].

3.2.5.4 Secondary outcome: time-to-treatment failure (TTF)

The TTF was defined by the company as "*the time from the first infusion of amivantamab to discontinuation of treatment for any reason, including disease progression, treatment toxicity and death*". The TTF results are presented in Table 3.16, and illustrated in a KM plot in Figure 3.6. The median TTF was [redacted] (95% CI: [redacted]) with [redacted] of patients censored.

Table 3.16: Summary of TTF from the CHRYSALIS trial (30th March 2021 data cut-off)

	Post-platinum patients with EGFR Exon20ins at RP2D (N=114, 30th March 2021 data cut-off)
Event, n (%)	[redacted]
Censored, n (%)	[redacted]
Time to event (months)	
25th percentile (95% CI)	[redacted]
Median (95% CI)	[redacted]
75th percentile (95% CI)	[redacted]
Range	[redacted]
3-month event-free rate (95% CI)	[redacted]
6-month event-free rate (95% CI)	[redacted]
9-month event-free rate (95% CI)	[redacted]
12-month event-free rate (95% CI)	[redacted]
15-month event-free rate (95% CI)	[redacted]
18-month event-free rate (95% CI)	[redacted]
21-month event-free rate (95% CI)	[redacted]
24-month event-free rate (95% CI)	[redacted]
27-month event-free rate (95% CI)	[redacted]

Based on Table 20 of the CS[4]
Note: RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.
CI = confidence interval; CS = company submission; EGFR = epidermal growth factor receptor; NE = not evaluable; RP2D = recommended Phase 2 dose; TTF = time-to-treatment failure

Figure 3.6: Kaplan-Meier plot of TTF – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114)

Source: Figure 11 of the CS[4]
CS = company submission; TTF = time-to-treatment failure; RP2D = recommended Phase 2 dose

3.2.5.5 Secondary outcome: overall survival (OS)

The OS was defined in the CS as "*the time from first infusion of amivantamab to death due to any cause*". The OS results are presented in Table 3.17 and a KM plot is illustrated in Figure 3.7. 64.9% of patients was censored and the median OS was 22.77 months (95% CI: 17.48, NE). On the 30th of March 2021 data cut-off (median follow-up of [redacted] [range: [redacted]]), [redacted] ([redacted]) had died.

Table 3.17: Summary of OS from the CHRYSALIS trial (30th March 2021 data cut-off)

	Post-platinum patients with EGFR Exon20ins at RP2D (N=114, 30th March 2021 data cut-off)
Event, n (%)	40 (35.1)
Censored, n (%)	74 (64.9)

	Post-platinum patients with EGFR Exon20ins at RP2D (N=114, 30th March 2021 data cut-off)
Time to event (months)	
25th percentile (95% CI)	9.95 (8.48, 14.59)
Median (95% CI)	22.77 (17.48, NE)
75th percentile (95% CI)	NE (23.00, NE)
Range	(0.2, 30.5+)
3-month event-free rate (95% CI)	0.95 (0.89, 0.98)
6-month event-free rate (95% CI)	0.90 (0.83, 0.94)
9-month event-free rate (95% CI)	0.79 (0.70, 0.86)
12-month event-free rate (95% CI)	0.73 (0.63, 0.80)
15-month event-free rate (95% CI)	0.66 (0.55, 0.75)
18-month event-free rate (95% CI)	0.61 (0.49, 0.71)
21-month event-free rate (95% CI)	0.53 (0.39, 0.66)
24-month event-free rate (95% CI)	0.40 (0.21, 0.58)
27-month event-free rate (95% CI)	0.40 (0.21, 0.58)
30-month event-free rate (95% CI)	0.40 (0.21, 0.58)

Based on Table 21 of the CS[4]

Note: RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.

CI = confidence interval; CS = company submission; EGFR = epidermal growth factor receptor; NE = not evaluable; RP2D: recommended Phase 2 dose; OS = overall survival

Figure 3.7: Kaplan-Meier plot of OS – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114)

Source: Figure 12 of the CS[4]
CS = company submission; TTF = time-to-treatment failure; RP2D = recommended Phase 2 dose

The CS also compared the single arm results to SoC treatments for OS, by separately comparing to data from a US and UK Cohort (see Section 3.4 for further details).

Comparing BICR data with the US SoC data, based on an IPW-ATT approach for statistical adjustment (see Section 3.4), the adjusted HR for amivantamab versus SoC was [redacted]. Based on a multivariable proportional hazards regression model the adjusted HR for amivantamab versus SoC was [redacted].

Comparing BICR data with the UK PHE SoC data, based on a multivariable proportional hazards regression model the adjusted HR for amivantamab versus SoC was [redacted].

ERG comment:

3.2.5.6 Exploratory outcome: Health-related quality of life (HRQoL)

HRQoL consists of exploratory descriptive analyses that were meant to include four patient reported outcomes (PROs): PGIS, PGIC, NSCLC-SAQ and EQ-5D-5L VAS (see Table 3.3). PROs were not part of the original trial protocol but a later addition (protocol Amendment 7), which affected the data availability. Data were available for only a small subset of the population of interest (expanded efficacy population), n=[redacted] of 114 ([redacted]%).

The company opted to present results only for two of the outcomes, the ED-5D VAS and the NSCLC-SAQ results, as illustrated in Figure 3.8 and Figure 3.9, respectively.

NSCLC-SAQ is a 7-item questionnaire-based, PRO measure, used in advanced NSCLC clinical trials. It draws from a 7-day patient recall period and is based on verbal rating scales. The questionnaire assessed the patient reported symptoms of cough, pain, dyspnoea, fatigue and poor appetite. The total score can range from 0 to 20.

ED-5D-5L VAS is also a questionnaire-based PRO measure of health status, but it is designed to be used by the general population. It comprises the five dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The system includes five levels of severity for each of the five dimensions indicating no problem, slight problems, moderate problems, severe problems or extreme problems.

Figure 3.8: Change from baseline of NSCLC-SAQ total score over time – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114)

Source: Figure 13 of the CS[4]
CS = company submission; LS = least squares; NSCLC-SAQ = Non-Small-Cell Lung Cancer Symptom Assessment Questionnaire; TOT = time on treatment

Figure 3.9: Change of baseline of EQ-5D-5L VAS over time – the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114)

Source: Figure 14 of the CS[4]
CS = company submission; EQ-5D-5L = EuroQoL five-dimensions five-levels; LS = least squares; VAS = visual analogue scale

ERG comments:

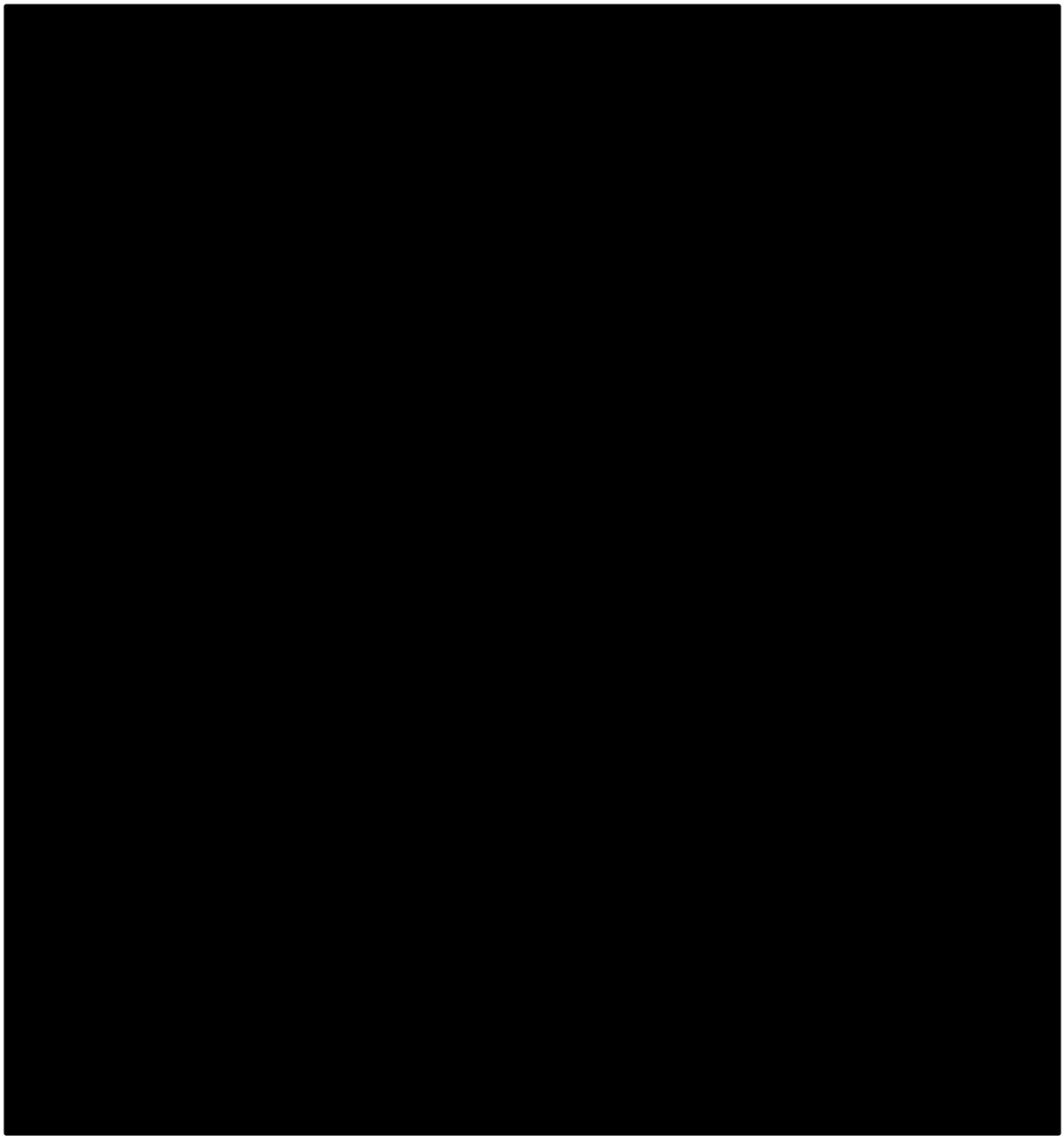
- The company has chosen to present only two out of the four PRO measures that were included in the CHRYSALIS for brevity. The results for PGIS and PGIC were not reported and a justification for their exclusion was not provided.
- In both Figure 3.8 (NSCLC-SAQ) and Figure 3.9 (ED-5D-5L) the included number of patients appears to be very small (n=[redacted]) and different from what was reported in the text (n=[redacted]).
- The number of patients available for this outcome is very small, and the estimates based on this small sample are uncertain. Further comments on the HRQoL outcomes are provided in the cost effectiveness part of this report.

3.2.5.7 Subgroup analysis

The CS presented an ORR subgroup analysis for the following demographic and clinical characteristics: age (four categories), sex, race (Asian versus non-Asian), ECOG status (0 versus ≥1), history of smoking and prior immunotherapy. Forest plots for BICR and INV assessments are illustrated in Figure 3.10 and Figure 3.11, respectively. From the 98 patients whose race could be determined, 59 were Asian (51.8%). The company argues that the results of the subgroup analysis regarding race, illustrate

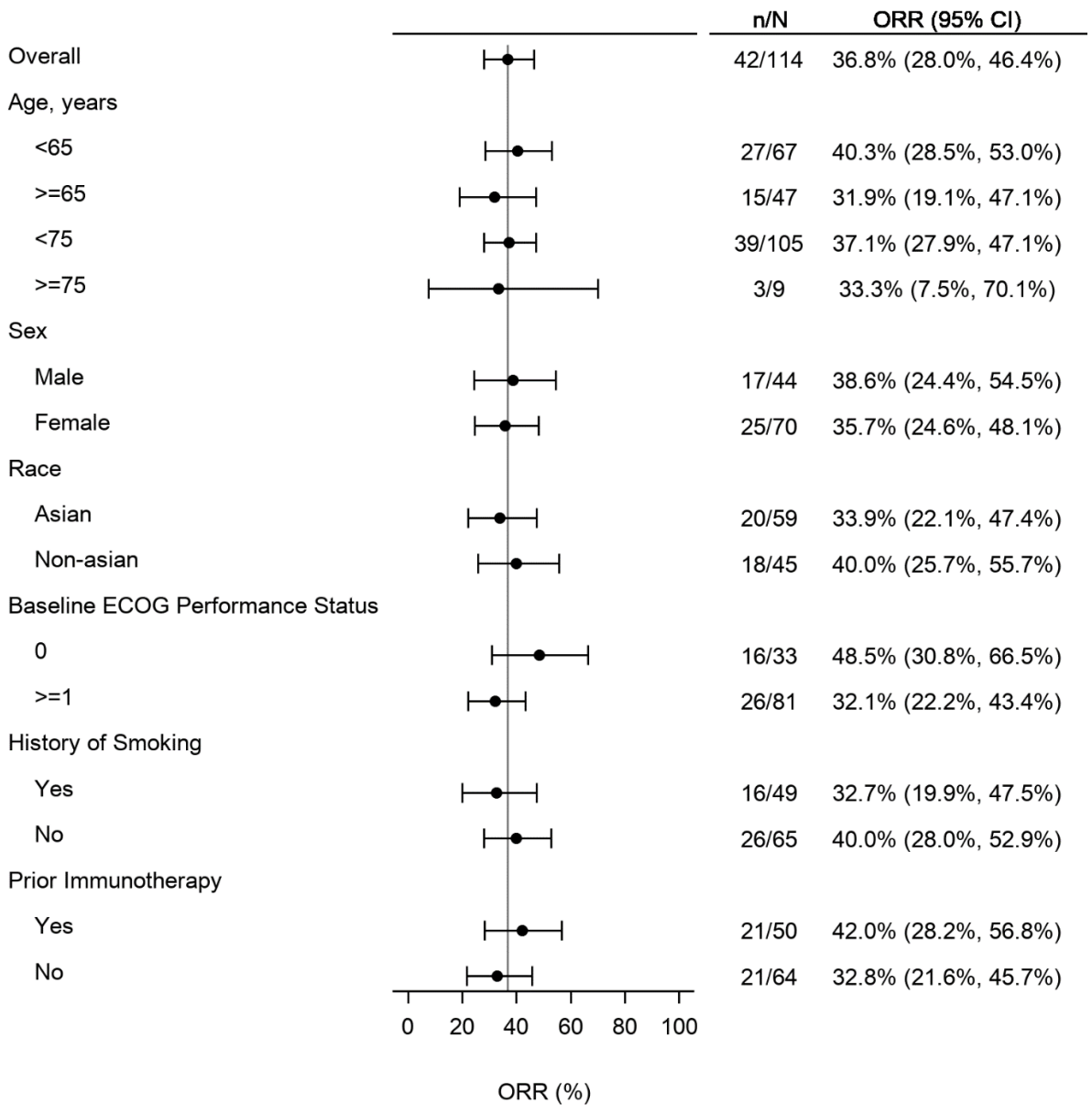
that the high proportion of Asian participants does not influence the generalisability of the efficacy results.

Figure 3.10: Forest plot of ORR based on RECIST v1.1– the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114) by BICR assessment



Source: Figure 15 of the CS[4]
Note: n = confirmed CR plus confirmed PR. If race was not reported, then that patient is excluded from the race subgroup. Chinese patients enrolled beyond the initial global cohort enrolment are excluded
CI = confidence interval; CR = complete response; CS = company submission; ECOG = Eastern Cooperative Oncology Group; ORR = overall response rate; PR = partial response

Figure 3.11: Forest plot of ORR based on RECIST v1.1– the CHRYSALIS trial (30th March 2021 data cut-off) - expanded efficacy population (N=114) by INV assessment



Source: Figure 16 of the CS[4]
Note: n = confirmed CR plus confirmed PR. If race was not reported, then that patient is excluded from the race subgroup. Chinese patients enrolled beyond the initial global cohort enrolment are excluded
CI = confidence interval; CR = complete response; CS = company submission; ECOG = Eastern Cooperative Oncology Group; ORR = overall response rate; PR = partial response

ERG comment:

- The results of the subgroup analyses regarding race (Asian versus non-Asian) illustrate that the ORRs vary. In the BICR assessment ORR for Asians is 45.8% (95% CI 32.7, 59.2) and for non-

Asians 40% (95% CI 25.7, 55.7) and in the INV assessment, ORR for Asians is 33.9% (95% CI 22.1, 47.4) and for non-Asians 40% (95% CI 25.7, 55.7).

- The effect of these differences on effectiveness and cost effectiveness, as far as the applicability to the UK population, is unknown.

3.2.6 Safety results of the CHRYSALIS trial

This Section reports on the safety results discussed in Section B.2.10 of the CS.

The CS reports safety results from the CHRYSALIS trial from the 8th October 2020 and 30th March 2021 data cut-offs. Results are presented for the post-platinum patients with Exon20ins at RP2D safety population (N=153) from the 30th March 2021 data cut-off. Additional data from the All Treated at RP2D safety population (N=380) and All Treated safety population (N=489) at the latest data cut-off are presented in Appendix F but are not summarised here.

ERG comment: In its clarification letter, the ERG asked the company to confirm if the safety population only included patients with EGFR Exon20ins NSCLC whose disease had progressed after platinum-based chemotherapy and had received at least one dose of the study drug, amivantamab. In response to clarification, the company stated that, *"Janssen can confirm that the safety population (N=153) included only patients with EGFR Exon20ins NSCLC whose disease had progressed after platinum-based chemotherapy and had received at least one dose of the study drug, amivantamab."*[9] The ERG is satisfied that the results presented in this Section are from a suitable analysis set.

3.2.6.1 Treatment duration and dosage

As of the latest data cut-off date (30th March 2021), from the EGFR Exon20ins RP2D safety population, the median follow up is stated to be [redacted] months. [redacted] of patients had completed the study, 62.1% (95/153) of patients were still in the study and [redacted] had prematurely terminated from study participation. The CS states that at this time 36.6% (56/153) were still receiving amivantamab while 63.4% (97/153) had discontinued treatment. When reviewing reasons for discontinuation 47.7% (73/153) of patients had progressive disease, 7.8% (12/153) had experienced AEs, 4.6% (7/153) were patient selected withdrawals, 1.3% (2/153) withdrew as a result of a physician decision and 2% (3/153) of patients expired. (see Table 3.18 below).

Table 3.18: Study and treatment disposition; post-platinum patients with EGFR Exon20ins at RP2D safety population (N=153)

Event, n (%)	Safety population (N=153, 30th March 2021 data cut-off)
Study disposition	
Patients ongoing	[redacted]
Completed study participation	[redacted]
Terminated study participation prematurely	[redacted]
Treatment disposition	
Patients ongoing	56 (36.6)
Discontinued study treatment	97 (63.4)
Reason for discontinuation	
Progressive disease	73 (47.7)
AE	12 (7.8)
Withdrawal by patient	7 (4.6)

Event, n (%)	Safety population (N=153, 30th March 2021 data cut-off)
Physician decision	2 (1.3)
Death	3 (2.0)

Based on table 26, CS[4]
Note: RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg. [a] Patient is considered to have completed the study if the patient died prior to the end of study.
AE = adverse event; CS = company submission; RP2D = recommended Phase 2 dose

The CS states that the median number of treatment cycles received in the safety population was seven, with 34.0% (52/153) subjects having received treatment for ≥10 cycles,[4] and the maximum number of treatment cycles was 27.46.4% (71/153) patients had received treatment for a period of ≥6 months with a median duration of treatment being 5.6 months. The maximum duration of treatment was 23.9% (see Table 3.19).

Table 3.19: Summary of treatment with amivantamab; post-platinum patients with EGFR Exon20ins at RP2D safety population (N=153)

Safety population (N=153, 30th March 2021 data cut-off)	
Duration of study treatment, months[a]	
Mean (SD)	7.28 (5.81)
Median	5.52
Range	(0.03; 23.89)
Duration of study treatment, n (%)	
<2 months	31 (20.3)
2 –<4 months	26 (17.0)
4 –<6 months	25 (16.3)
≥6 months	71 (46.4)
Total number of cycles[b]	
Mean (SD)	8.5 (6.2)
Median	7
Range	(1, 27)

Based on Table 27, CS[4]
Note: RP2D is defined as 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.
[a]Treatment duration is defined as the duration from the date of the first dose of amivantamab to the date of last dose of amivantamab+1 divided by 30.4375.
[b]A patient is considered as treated in a cycle if the patient received any non-zero dose of study agent in that cycle.
CS = company submission; SD = standard deviation; RP2D = recommended Phase 2 dose.

3.2.6.2 Summary of Treatment emergent adverse events (TEAEs)

The CS states that all patients experienced at least one treatment emergent adverse event (TEAE) while 98.0% had at least one TEAE reported by the investigator to be related to amivantamab. TEAEs at grade 3 or above were experienced by 41.8% of patients while 19.6% patients had TEAEs at grade 3 or above that were deemed to be related to amivantamab. Serious TEAEs were experienced by 28.8% of patients with 2.6% experienced grade 4 TEAEs while 7.2% of patients experienced a grade 5 event (fatal) and expired. Of those 28.8% of patients who experienced a serious TEAE, 8.5% were reported by the investigators. The CS states that all grade 5 fatal events were assessed as being unrelated to

amivantamab (see Table 3.20). The company did not provide a definition of 'serious' AEs within the CS document, however a review of the trial protocol clarified that a 'serious AE' would be based on 'I*CH and EU Guidelines on Pharmacovigilance for Medicinal Products for Human Use is any untoward medical occurrence that at any dose.*

- *Results in death*
- *Is life-threatening (The subject was at risk of death at the time of the event. It does not refer to an event that hypothetically might have caused death if it were more severe.)*
- *Requires inpatient hospitalization or prolongation of existing hospitalization*
- *Results in persistent or significant disability/incapacity*
- *Is a congenital anomaly/birth defect*
- *Is a suspected transmission of any infectious agent via a medicinal product*
- *Is Medically Important'*[29]

Table 3.20 provides data stating that 14.4% of patients experienced TEAEs that required dose reduction, while 11.8% of patients discontinued treatment as a consequence of AEs. Of the patients, 59.5% experienced a need for infusion modification and 35.9% of patients experiencing events that led to dose interruption. Investigators judged that all events (14.4%) that led to dose reduction were related to amivantamab while 5.2% of events that led to discontinued treatment, 58.8% of events that led to infusion modification and 20.9% of events that led to dose interruption were related to amivantamab.

Table 3.20: Overall summary of TEAEs; Post-platinum patients with EGFR Exon20ins at RP2D safety population (N=153)

Event, n (%)	Safety population (N=153, 30th March 2021 data cut-off)
Patients with ≥1 AE	153 (100.0)
Related AEs[a]	150 (98.0)
AEs leading to death[b]	11 (7.2)
Related AEs leading to death[a,b]	0
Serious AEs	44 (28.8)
Related serious AEs[a]	13 (8.5)
AEs leading to discontinuation of amivantamab	18 (11.8)
Related AEs leading to discontinuation of amivantamab[a]	8 (5.2)
AEs leading to dose reduction	22 (14.4)
Related AEs leading to dose reduction[a]	22 (14.4)
AEs leading to infusion modification[c]	91 (59.5)
Related AEs leading to infusion modification[a, c]	90 (58.8)
AEs leading to dose interruption[d]	55 (35.9)
Related AEs leading to dose interruption[a, d]	32 (20.9)
Grade ≥3 AEs	64 (41.8)
Related grade ≥3 AEs[a]	30 (19.6)
Grade 1	4 (2.6)

Grade 2	85 (55.6)
Grade 3	49 (32.0)
Grade 4	4 (2.6)
Grade 5	11 (7.2)

Based on Table 28, CS[4].
Note: RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg
[a]An AE is categorised as related if assessed by the investigator as possibly, probably, or very likely related to study agent
[b]AEs leading to death are based on AE outcome of fatal
[c]AEs leading to infusion modification of study agent are based on infusion interrupted, infusion rate decreased, and infusion aborted due to adverse event on the infusion eCRF page
[d]Excludes infusion related reactions
AE = adverse event; CS = company submission; RP2D = recommended Phase 2 dose

3.2.6.2.1 TEAEs occurring with a frequency of 10% or higher

The TEAEs which occurred with a frequency of 10% or higher in the EGFR Exon20ins at RP2D safety population (N=153) on 30th March 2021 data cut-off are summarised in Table 3.21. The more commonly reported TEAEs included infusion related reactions (63.4%), paronychia (52.9%), rash (43.1%), and dermatitis acneiform (39.2%). Along with stomatitis (22.2%), dry skin (13.7%) and diarrhoea (13.7%) these are stated in the CS to be common on-target events associated with EGFR inhibition. The CS also details that hypoalbuminemia (39.2%), constipation (23.5%) and peripheral oedema (22.9%) which are common on-target events associated with MET inhibition were also reported in >10% of patients in this population.

Table 3.21: TEAEs with a frequency of at least 10% by system organ class and preferred term; post-platinum patients with EGFR Exon20ins at RP2D safety population (N=153, 30th March 2021 cut-off)

Event	n (%)
Patients with one or more AEs	153 (100.0)
Skin and subcutaneous tissue disorders	136 (88.9)
Dermatitis acneiform	60 (39.2)
Rash	66 (43.1)
Pruritus	24 (15.7)
Dry skin	21 (13.7)
Gastrointestinal disorders	114 (74.5)
Constipation	36 (23.5)
Nausea	38 (24.8)
Stomatitis	34 (22.2)
Vomiting	21 (13.7)
Diarrhoea	21 (13.7)
Injury, poisoning and procedural complications	102 (66.7)
Infusion related reaction	97 (63.4)
Infections and infestations	107 (69.9)
Paronychia	81 (52.9)

Event	n (%)
Respiratory, thoracic and mediastinal disorders	88 (57.5)
Dyspnoea	30 (19.6)
Cough	26 (17.0)
General disorders and administration site conditions	96 (62.7)
Oedema peripheral	35 (22.9)
Fatigue	30 (19.6)
Pyrexia	26 (17.0)
Metabolism and nutrition disorders	92 (60.1)
Hypoalbuminaemia	60 (39.2)
Decreased appetite	27 (17.6)
Musculoskeletal and connective tissue disorders	73 (47.7)
Myalgia	18 (11.8)
Back pain	25 (16.3)
Nervous system disorders	50 (32.7)
Dizziness	18 (11.8)
Headache	11 (7.2)
Investigations	63 (41.2)
Alanine aminotransferase increased	34 (22.2)
Aspartate aminotransferase increased	25 (16.3)
Blood alkaline phosphatase increased	16 (10.5)
Psychiatric disorders	29 (19.0)
Insomnia	16 (10.5)

Based on Table 29, CS [4]
AEs = adverse events; CS = company submission; TEAE = treatment emergent adverse event
RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg. Patients are counted only once for any given event, regardless of the number of times they actually experienced the event

3.2.6.3 Grade ≥3 Treatment-emergent AEs

The CS provides data on TEAEs at grade ≥3 in the RP2D safety population (N=153) at the 30th March 2021 data cut-off (see Table 3.22 below) and highlights that these are the AEs considered in the cost effectiveness model informing this submission. There were [redacted] patients who experienced one or more grade ≥3 AEs with [redacted] ([redacted]) patients believed to be experiencing grade ≥3 TEAEs considered by the investigator to be related to amivantamab. The most common grade ≥3 AEs were pulmonary embolism and hypokalaemia, occurring in [redacted] and [redacted] patients, respectively. None of the AEs at grade 3 or higher occurred in ≥5% patients.

Table 3.22: Grade 3 or higher TEAE by preferred term: post-platinum patients with EGFR Exon20ins at RP2D safety population (N=153, 30th March cut-off

Event	n (%)
Subjects with one or more grade ≥3 AEs	[redacted]
Preferred term	
Pulmonary embolism	[redacted]

Event	n (%)
Hypokalaemia	[redacted]
Pneumonia	[redacted]
Dyspnoea	[redacted]
Hypoalbuminaemia	[redacted]
Paronychia	[redacted]
Diarrhoea	[redacted]
Infusion related reaction	[redacted]
Neutropenia	[redacted]
Hyponatraemia	[redacted]
Alanine aminotransferase increased	[redacted]
Hypophosphataemia	[redacted]
Hypotension	[redacted]
Gamma-glutamyl transferase increased	[redacted]
Rash	[redacted]
Respiratory failure	[redacted]
Anaemia	[redacted]
Respiratory tract infection	[redacted]
Sepsis	[redacted]
Acne	[redacted]
Cellulitis	[redacted]
Fatigue	[redacted]
Hypoxia	[redacted]
Pleural effusion	[redacted]
Pericardial effusion	[redacted]
Aspartate aminotransferase increased	[redacted]
Dermatitis acneiform	[redacted]
Headache	[redacted]
Hypertension	[redacted]
Oedema peripheral	[redacted]
Syncope	[redacted]
Abdominal pain	[redacted]
Atrial fibrillation	[redacted]
Blood alkaline phosphatase increased	[redacted]
Blood creatine phosphokinase increased	[redacted]
Decreased appetite	[redacted]
Lymphopenia	[redacted]
Mental status changes	[redacted]
Nausea	[redacted]
Pneumonia aspiration	[redacted]

Event	n (%)
Pneumonitis	[redacted]
Stomatitis	[redacted]
Vomiting	[redacted]
Aspiration	[redacted]
Hypocalcaemia	[redacted]
Infected dermal cyst	[redacted]
Insomnia	[redacted]
International normalised ratio increased	[redacted]
Muscular weakness	[redacted]
Pulmonary sepsis	[redacted]
Pulseless electrical activity	[redacted]
Rash papular	[redacted]
Renal vein thrombosis	[redacted]
Sudden death	[redacted]
Thrombocytopenia	[redacted]
Toxic epidermal necrolysis	[redacted]
Transitional cell carcinoma	[redacted]

Based on table 30, CS[4]

Note: Subjects are counted only once for any given event, regardless of the number of times they actually experienced the event. The event experienced by the subject with the worst toxicity is used. RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg.

AEs = adverse events; CS = company submission; RP2D = recommended Phase 2 dose; TEAE = treatment-emergent adverse event

3.2.6.4 Treatment related adverse events

The CS states that [redacted] (See Table 3.23 below) patients in the post-platinum patients with EGFR Exon20ins at RP2D safety population (N=153) had AEs reported by the investigator to be related to amivantamab. Skin and subcutaneous tissue disorders comprised the majority of AEs by System Organ Class, with 86.9% of patients affected. Infusion related reaction (IRR) was the most commonly reported with [redacted] of patients experiencing it. Paronychia was the second most reported AE with [redacted] of patients experiencing it. Rash and dermatitis acneiform were experienced by ([redacted]) and ([redacted]) of patients respectively. The CS clarifies that except for IRR, all treatment related AEs were comprised predominantly of on-target events associated with EGFR or MET inhibition and that on-target MET-associated events of hypoalbuminemia and peripheral oedema were reported as related to amivantamab in [redacted] and [redacted] of patients, respectively.

Table 3.23: Treatment-related AEs by system organ class and preferred term; Post-platinum patients with EGFR Exon20ins at RP2D safety population (N=153, 30th March 2021 cut-off)

Preferred term	n (%)
Patients with one or more related AEs	[redacted]
Skin and subcutaneous tissue disorders	[redacted]

Dermatitis acneiform	[redacted]
Rash	[redacted]
Pruritus	[redacted]
Dry skin	[redacted]
Injury, poisoning and procedural complications	[redacted]
Infusion related reaction	[redacted]
Gastrointestinal disorders	[redacted]
Stomatitis	[redacted]
Nausea	[redacted]
Infections and infestations	[redacted]
Paronychia	[redacted]
General disorders and administration site conditions	[redacted]
Fatigue	[redacted]
Oedema peripheral	[redacted]
Metabolism and nutrition disorders	[redacted]
Hypoalbuminaemia	[redacted]
Investigations	[redacted]
Alanine aminotransferase increased	[redacted]
Aspartate aminotransferase increased	[redacted]
Based on table 31, CS[4] AE = adverse event; CS = company submission; RP2D = recommended Phase 2 dose Note: RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg. Patients are counted only once for any given event, regardless of the number of times they experienced the event.	

3.2.6.5 Serious TEAEs

Serious TEAEs reported by the investigator for RP2D safety population (N=153) is summarised in Table 3.24 below. There were [redacted] patients that had TEAEs reported by the investigator to be serious. The most common serious TEAE being interstitial lung disease, reported in [redacted] patients ([redacted]).

Table 3.24: Serious TEAEs by system organ class, preferred term; RP2D safety population (N=153)

System organ class/preferred term	Safety population N=153, 30th March 2021 data cut-off, n (%)
Subjects with any serious TEAEs	[redacted]
Skin and subcutaneous tissue disorders	[redacted]
Rash	[redacted]
Toxic epidermal necrolysis	[redacted]
Injury, poisoning and procedural complications	[redacted]
Infusion related reaction	[redacted]
Gastrointestinal disorders	[redacted]
Diarrhoea	[redacted]

System organ class/preferred term	Safety population N=153, 30th March 2021 data cut-off, n (%)
Abdominal pain	[redacted]
Respiratory, thoracic and mediastinal disorders	[redacted]
Interstitial lung disease	[redacted]

Based on Table 32, CS[4]
RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg. Subjects are counted only once for any given event, regardless of the number of times they actually experienced the event. The event experienced by the subject with the worst toxicity is used.
ADR = adverse drug reaction; CS = company submission; RP2D = recommended Phase 2 dose; TEAEs = treatment emergent adverse events

ERG comments:

- The ERG notes that 41.8% of patients had experienced a grade 3 or higher AE, which according to the grading criteria of the National Cancer Institute common terminology criteria for adverse events (NCI CTCAE) is defined as being '*Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care activities of daily living*' at a minimum. According to the CS, 19.6% of these AEs were related to the administration of amivantamab when defined as '*assessed by the investigator as possibly, probably, or very likely related to study agent*'. However, events defined as 'serious' by the company (according to the criteria described in Section 3.2.6.2, include the definition 'medically important') occurred in 28.8% of patients, with 8.5% of serious events being related to amivantamab. The ERG would suggest that a grade 3 or above event which is 's*evere or medically significant*' is also 'medically important' and therefore could be defined as 'serious'. The ERG considers that there appears to be a lack of clarity and information leading to uncertainty regarding how 19.6.% of patients are experiencing grade 3 or above events related to amivantamab, yet only 8.5% of patients have experienced what is described as a serious AE**.**
- Concerning the statement in Section B.3.3.3 of the CS[4], *"safety profiles were considered and compared in the context of treatment classes rather than individual treatments, validating this approach,"* the ERG in its clarification letter asked the company to provide AEs specific to amivantamab rather than the class of treatments to which amivantamab belonged. In response, the company stated that, *"The text in the question refers to the approach taken to characterise the safety profile of UK SoC. AE incidence rates for the treatment classes included in the comparator basket were considered and compared in the context of treatment classes rather than individual treatments."*[9] Table 3.25 below reports the incidence of grade ≥3 AEs occurring in ≥5% of patients in the CHRYSALIS trial was also provided.
- Although a wide range of AEs were reported (table 3.22), the CS confirms that none of these were reported in more than 5% of the population. This data suggests that the likelihood of experiencing a severe AE is considerable (41.8%) in this population, and that less than 50% of these will be attributed to amivantamab (19.6%). While the more common AEs in this category included pulmonary embolism (4.6%), and hypokalaemia (3.9%), no incidence of specific or common severe or life-threatening AE's (as defined as grade ≥3 AEs in more than 5% of population) has been explicitly identified to be of concern. The ERG notes that this is based on a small sample and cautions that this should be considered in any interpretation.
- Table 3.24 states that only 7.2% of patients experienced any serious TEAEs. It is also apparent that in Table 3.24 interstitial lung disease is listed as a serious TEAE that has affected 2.6% of the

population, however there is no mention of this as a grade ≥3 TEAE in Table 3.22. We are unsure how a serious TEAE can be identified but yet not also be included in the data on grade 3 or above TEAE's. Furthermore, it is stated in the CS (Section B.2.10) that '*Forty-four patients (28.8%) had serious TEAEs'* however this does not appear to readily tally with the data included in Table 3.24 where it is stated that '*Subjects with any serious treatment-emergent AEs*' amounts to 11 patients (7.2%). The ERG considers that there appears to be a lack of clarity and consistency here on data reporting and defining.

Table 3.25: Incidence of Grade ≥3 AEs occurring in ≥5% of patients

AE, %	AMI	UK SoC				
		IO agents	EGFR TKIs	Pt-based chemotherapy	Non-Pt-based chemotherapy	Weighted average
Anaemia	[redacted]	0.5	0.0	11.8	3.8	3.2
Diarrhoea[a]	[redacted]	15.4	69.9	11.0	24.4	28.4
Fatigue	[redacted]	1.6	1.3	0.7	3.5	2.1
Febrile neutropenia	[redacted]	0.0	0.0	0.0	9.4	3.4
Neutropenia	[redacted]	0.5	0.0	11.8	14.6	7.2
Neutrophil count decreased	[redacted]	0.0	0.0	0.0	11.1	4.0
Rash	[redacted]	0.0	5.9	0.0	0.0	1.1
Thrombo-cytopaenia	[redacted]	0.0	0.0	7.4	0.0	1.1

Based on Table 14 of clarification letter response[9]
Note: [a] Due to its clinical relevance, the incidence of diarrhoea was considered at any grade.
AE = adverse event; AMI = amivantamab; EGFR = epidermal growth factor receptor; IO = immuno-oncology; Pt = platinum; SoC = standard of care; TKI = tyrosine kinase inhibitor; UK = United Kingdom

3.2.6.6 Mortality

The CS emphasises that '*OS is a secondary efficacy endpoint in this study, and survival data continues to be collected on all patients even after discontinuation of amivantamab during the Follow-up Period. In all cases of patient death, regardless of timing, the cause of death was separately reported. For all deaths that occurred during the Treatment Period (and up through 30 days after last dose), specific information regarding the cause of death was to be reported as a Grade 5 TEAE. Thus, patient deaths that are due to progressive disease, if occurring on treatment or within 30 days of the last dose, are also separately reported as an AE having an outcome of death'.*

Data is presented in Table 3.26 below to illustrate a summary of deaths that occurred at any time during the study in the RP2D safety population. The CS emphasises that the median follow-up was [redacted] months (range: [redacted]) and that these deaths were not reported as related to amivantamab by the investigator. Deaths were observed in [redacted] at any time on the study. Progressive disease was the most common cause of death [redacted] expired on treatment or within 30 days of the last dose of amivantamab. Of these, eight ([redacted]) patients died due to a TEAE, and [redacted] patients died due to progressive disease. The CS provides a summary of these deaths by preferred term and system organ class (see Table 34, CS). Briefly, on review of these data it is apparent that respiratory failure and dyspnoea accounted for the more common AEs that led to death with two patients (1.3%), dying of each.

Table 3.26: Summary of deaths during study; Post-platinum patients with RP2D safety population (N=153, 30th March 2021 cut-off)

Preferred term	n (%)
Deaths during study	[redacted]
PD	[redacted]
AE	[redacted]
Other	[redacted]
Deaths during treatment	[redacted]
AE	[redacted]
PD	[redacted]
Other	[redacted]

Based on table 33, CS[4]
RP2D: 1,050 mg if baseline weight <80 kg and 1,400 mg if baseline weight ≥80 kg. Deaths during treatment are presented for patients who died within 30 days of last amivantamab dose.
AE = adverse event; CS = company submission; PD = progressive disease; RP2D = recommended Phase 2 dose

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

The CS states that, in the absence of a direct head-to-head trial, and given that the SLR did not identify other relevant trials in this setting, that the two sources for the indirect treatment comparison would be the CHRYSALIS trial and RWE. Two RWE sources were included in the analyses:

- A US cohort that included pooled data from Flatiron Health Spotlight, ConcertAI and COTA data sources. This is referred to as US RWE.
- Data from PHE using routine population-level data available through PHE (now NHS Digital) National Cancer Registration and Analysis Service (NCRAS). These data are referred to as PHE.

A full critique of the CHRYSALIS trial is included in Section 3.2.

ERG comments:

- It was unclear to the ERG whether no other studies might have been suitable for a comparison with amivantamab. The ERG sought further information as to the means and rationale for the identification and selection of these two specific databases and in the request for clarification asked the company to provide insight. The company responded stating that '*The US RWE and the PHE cohort studies were initiated by Janssen with the objective of providing RWE data for patients with EGFR Exon20ins mutations previously treated with platinum-based chemotherapy to inform the external control arm for the CHRYSALIS trial.'* [9]
- The company also stated '*…the SLRs did not identify any studies reporting on clinical outcomes for patients with EGFR Exon20ins mutations positive NSCLC previously treated with platinum-based chemotherapy. As a result, individual patient level data derived from the US RWE and PHE studies were used as the only sources for these data for the adjusted comparison analyses*.' While we do not necessarily consider the data derived from these sources as inappropriate, the ERG expects that there must be a full, justified rationale with clear systematic and scientific robustness for the use of an evidence source.

- The PHE database includes data from a UK based population, while the RWE derived from the US included datasets from three specific databases, namely the Flatiron, COTA, and ConcertAI databases. We sought further information on the suitability of these databases as generalisable to the UK population. In the request for clarification, we asked that the demographic characteristics of the patients in these databases be provided with comparisons to a UK population. The company in its response provided tabulated data for each of the three RWE US databases,[9] as well as emphasising that this data was pooled and compared to the UK based PHE data in Section B.2.9 of the CS. The company also clarified that, '*UK-based clinical experts emphasised the high degree of alignment in the baseline characteristics of patients included in both of these RWE data sources and the CHRYSALIS trial, with the proportion of patients with brain metastases being the only characteristic highlighted as differing notably between them'.*
- While the demographic and patient data may be broadly similar, in the absence of a systematic approach to identifying and selecting this evidence, the impact of selection bias must be considered. The ERG addressed this in its clarification letter to the company and requested further information on how this was mitigated. The company responded acknowledging the presence of selection bias and explaining that this is difficult to avoid due to the rarity of the disease and that RWE cohorts are limited to patients with EGFR Exon20ins mutations for whom data are available. The company clarified that to counteract the impact of such bias, the US RWE data were adjusted to the CHRYSALIS population in terms of key prognostic variables and baseline characteristics. This included [redacted]. The company also stated that according to their clinical experts, '*the characteristics and outcomes broadly aligned with their expectations for the patient population in the UK, and that none of the baseline characteristics showed systematic differences that would confer a substantial selection bias'*[9]
- The ERG understands and appreciates that evidence sources for rare diseases may be difficult to obtain and be limited in their generalisability, however, the systematic approach to identification and selection of evidence must be robust and auditable. In this case we do not consider that this has been properly described. We do not necessarily deny the suitability of these RWE data sources, this is a separate issue, but there must be a clearly described, justified process and criteria, for why source X is identified and used over source Y. In this case the ERG does not feel that this has adequately occurred. Furthermore, while expert clinical opinion is a valuable tool, statements such as '*the characteristics and outcomes broadly aligned with their expectations for the patient population in the UK',* are secondary to the evidence which informs their expectations.

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

Because CHRYSALIS is a single arm trial, the company conducted an adjusted treatment comparison to inform the relative efficacy estimates for amivantamab versus a SoC utilising comparator data from RWE sources listed above in Section 3.3.

To account for differences in patient populations between CHRYSALIS and the RWE data sources, the comparisons adjusted for key prognostic variables, which were identified a priori by an SLR and validated by clinical experts. The following covariates were considered: [redacted]

[redacted]

Different statistical approaches were explored to conduct the adjusted comparisons, 1) inverse probability weighting (IPW) method, which uses the propensity score (probability of receiving the treatment) to estimate the average treatment effect on the treated (ATT) (by re-weighting only the comparator data), and 2) a multivariable regression approach with direct adjustment for covariates. Both methods were applied to the US RWE to estimate PFS, TTNT, OS and ORR. Only covariate adjustment was used for the PHE data, the reason given that: "IPW estimates were unstable due to the small sample size". Also, only for TTNT and OS were estimated, with the reason for lack of PFS as "Due to limitations in the data recorded in the PHE datasets, it was not possible to collect PFS for the PHE cohort." No reason was provided for not estimating ORR using the PHE. The US RWE was used in the base case, the reason given that the sample size was larger.

Baseline characteristics of the CHRYSALIS and US RWE cohorts are given in Table 3.27, and those for the UK PHE cohort in Tables 3.26 and 3.27.

The company also provided the results of tests of overlap in Appendix M i.e., plot of propensity scores and standardised mean differences (SMDs).

Table 3.27: Baseline characteristics of treatment lines for patients in CHRYSALIS and the US RWE cohort

Characteristic, n (%)	CHRYSALIS EAS	US RWE cohort	IPW ATT weighted US RWE cohort
N	114	[redacted]	[redacted]
Prior lines of treatment			
1	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]
3	[redacted]	[redacted]	[redacted]
4+	[redacted]	[redacted]	[redacted]
Brain metastasis			
No	85 (74.6)	[redacted]	[redacted]
Yes	29 (25.4)	[redacted]	[redacted]
Age			
<60	48 (42.1)	[redacted]	[redacted]
60–70	38 (33.3)	[redacted]	[redacted]
≥70	28 (24.6)	[redacted]	[redacted]
ECOG PS			
0	[redacted]	[redacted]	[redacted]
1	[redacted]	[redacted]	[redacted]
Number of metastatic locations			
1	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]
3	[redacted]	[redacted]	[redacted]
4	[redacted]	[redacted]	[redacted]

Characteristic, n (%)	CHRYSALIS EAS	US RWE cohort	IPW ATT weighted US RWE cohort
Missing	[redacted]	[redacted]	[redacted]
Haemoglobin			
Normal/high	[redacted]	[redacted]	[redacted]
Low	[redacted]	[redacted]	[redacted]
Sex			
Male	44 (38.6)	[redacted]	[redacted]
Female	70 (61.4)	[redacted]	[redacted]
Cancer stage at initial diagnosis			
I	[redacted]	[redacted]	[redacted]
II	[redacted]	[redacted]	[redacted]
IIIA	[redacted]	[redacted]	[redacted]
IIIB/IV	[redacted]	[redacted]	[redacted]

Based on Table 24 in CS.[4]
ATT = average treatment effect among the treated; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group performance score; EAS = efficacy analysis set; IPW = inverse probability weighting; RWE = real world evidence; US = United States

Table 3.28: Baseline characteristics of treatment lines for patients in CHRYSALIS and the PHE data source

Characteristic, n (%)	CHRYSALIS EAS	PHE Cohort[a]
N	114	[redacted]
Prior lines of treatment		
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3+	[redacted]	[redacted]
Brain metastasis		
No	85 (74.6)	[redacted]
Yes	29 (25.4)	[redacted]
Age		
≤55	[redacted]	[redacted]
55–≤60	[redacted]	[redacted]
> 60	[redacted]	[redacted]
ECOG PS		
0	33 (28.9)	[redacted]
1	80 (70.2)	[redacted]
Liver metastasis		
No	101 (88.6)	[redacted]
Yes	13 (11.4)	[redacted]
Sex		
Male	44 (38.6)	[redacted]

Characteristic, n (%)	CHRYSALIS EAS	PHE Cohort[a]
Female	70 (61.4)	[redacted]
BMI		
Underweight (<18.5)	11 (9.6)	[redacted]
Normal (18.5- <25)	65 (57.0)	[redacted]
Overweight (25- <30)	25 (21.9)	[redacted]
Obese (>30)	13 (11.4)	[redacted]

Based on Table 25, CS.[4]
Note: [a] Adjusted baseline characteristics are not available for the PHE cohort as only covariate adjustment was applied
BMI = body mass index; CS = company submission; ECOG PS = Eastern Cooperative Oncology Group performance score; EAS = efficacy analysis set; PHE = Public Health England

ERG comment:

- Because CHRYSALIS is non-comparative, an unanchored comparison is necessary. However, considerable potential for risk of bias is entailed in such an indirect comparison. Although methods for confounder adjustment appear robust, as evidenced by the adjusted baseline values in Table 3.27, these are limited by the covariates chosen, and it is highly likely that residual confounding will remain: as stated in TSD 17, the validity of the two methods of adjustment used by the company relies on the assumption of selection on observables.[32] Additionally, the UK data might have been preferred, but apparently this was not possible for PFS outcomes. However, the explanation given for not using the UK data was that the sample size for the US data set (n=206) is larger than the UK data set (n=16). This is a good reason why the data from the US might provide more precise estimates of effect, as well as more valid statistical adjustments, but does not mean the US data are more appropriate, *per se*, for modelling treatment responses for a UK population. The company explains that the US data were deemed relevant to the UK population on the basis of expert opinion, but the exact nature of this opinion was not described. No reason was provided in the cs for not estimating ORR using the PHE, but the FAC check stated that these data were not collected. The ERG agrees that, given the limitation in the UK data, the US RWE was probably more appropriate.
- The IPW method to estimate the ATT was also the most appropriate method, given a less stringent requirements for ignorability and overlap of covariates, essentially because only the comparator data need to be adjusted.[32] Also, there did seem to be sufficient adjustment given overlap in the distribution of propensity scores and SMDs, which were all below 0.25.[32] The ERG did also ask for a comparison with the IPW method to estimate the ATE, which showed very little difference in any outcome (PFS, OS or TTNT).[30] However, there remains doubt whether all appropriate data sources were found and so this constitutes a key issue.

Table 3.29: Comparison of HRs for overall population and subgroups by LOT. The HRs denote the relative effect between amivantamab and SoC (adjusted, based on US RWE).

HR (95% CI), ATT approach	OS	PFS (BICR)	TTNT
Base case (2L+)	[redacted]	[redacted]	[redacted]
2L subgroup	[redacted]	[redacted]	[redacted]
3L+ subgroup	[redacted]	[redacted]	[redacted]

Based on Table 5 in clarification response.[9]

HR (95% CI), ATT approach	OS	PFS (BICR)	TTNT
2L = second line; 3L+ = third line and beyond; ATT = average treatment effect among the treated; CI = confidence interval; HR = hazard ratio; BICR = blinded independent committee review; LOT = line of therapy; OS = overall survival; PFS = progression-free survival; TTNT = time to next treatment.			

The company also examined the effect of line of therapy using other methods, gaining qualitatively similar results. However, the company did not provide sub-group comparisons using only the comparators that would be standard of care for that particular PD-L1 sub-group as specifically requested in the ERG clarification question.

ERG comment:

- Adjustment of the US RWE resulted in a decrease in the treatment effect, albeit only slightly, due to better comparator outcomes. Using the PHE, the treatment effect on OS increased. Of course, it is impossible to know how much reduction in bias there was, but the choice of the US RWE does at least seem conservative relative to the PHE. In terms of the request for sub-grouping around PD-L1 status, the company's response was as follows: *"For the PD-L1 subgroup analyses, a test for PD-L1 status was performed for [redacted] patients in the CHRYSALIS population, and [redacted] tested positive. In the US cohort, [redacted] lines of therapy corresponded to patients who tested PD-L1 positive. Of these, only [redacted] lines of therapy consisted of nivolumab or pembrolizumab monotherapies. In the PHE cohort, [redacted] patient had a positive PD-L1 status and was not treated with nivolumab or pembrolizumab monotherapies. It is therefore not feasible to conduct a comparative analysis on this subgroup."*
- The ERG agrees that PD-L1 sub-group analyses would have been unfeasible for the reasons given.

3.5 Additional work on clinical effectiveness undertaken by the ERG

Not applicable.

3.6 Conclusions of the clinical effectiveness Section

The CS and response to clarification provided full details for the ERG to appraise the literature searches conducted to identify studies about clinical efficacy and safety outcomes in patients with advanced NSCLC with EGFR Exon 20 insertion mutations.[4, 9, 21] The searches were conducted in January 2021 and updated in September 2021. Searches were transparent and reproducible, and comprehensive search strategies were used. A good range of databases and grey literature resources were searched. Despite the use of a focused population facet of search terms, the literature searches were comprehensive, and it was unlikely that relevant studies were missed.

The CS presented the results of one study, the CHRYSALIS trial[4] a Phase 1b, single arm, first-in-human, open-label, multicentre, 2-part trial. The trial included 77 participants in Part 1 (to determine recommended dose, median [redacted] months) and 285 participants in Part 2 (to determine safety and pharmacokinetics, 9.9 months)

Detailed efficacy results are presented in Section 3.2.5 while detailed safety results are presented in Section 3.2.6. The results are summarised below for the cut-off date of 30th March 2021 (median follow up time [redacted] months):

- **ORR** rates were 43% (95% CI 33.7% to 52.6%) for BICR and 36.8% (95% CI 28.0% to 46.4%) for INV.
- **CBR** rates were 73.7% (95% CI 64.6% to 81.5%) for BICR and 75.4% (95% CI 66.5% to 83.0%)

- **DOR** (median) was 10.84 months for BICR (95% CI 6.90 to 14.98) and 12.45 months for INV (95% CI 6.54 to 16.13)
- **PFS** (median) was 6.74 months for BICR (95% CI 5.45 to 9.66) and 6.93 months for INV (95% CI 5.55 to 8.64)
- **OS** (median) was 22.77 months (95% CI 17.48 to 'not evaluable')
- **TTF** (median) was 8.08 months (95% CI 6.67 to 10.64)

HRQoL was also evaluated as an exploratory analysis. Four PROMs were meant to be included, but only two were reported - the ED-5D VAS and the NSCLC-SAQ results. In neither of these analyses was a significant change in QoL from baseline observed. It should be noted that the graphical data reported in the CS are both limited in size (n=26) and different from what was reported in the text (n=30). The small number of patients available for this outcome may explain the high levels of uncertainty observed.

Due to the single-arm nature of the CHRYSALIS trial, an adjusted treatment comparison was conducted to derive comparative efficacy for amivantamab versus SoC treatments – a basket of treatments comprising treatments currently used for this population. Using US SoC data, these additional analyses showed that amivantamab offers statistically significant benefits over SoC in terms of PFS [HR [redacted]] and OS [HR [redacted]]. Although methods for confounder adjustment appear robust, these analyses are inevitably limited by the covariates chosen. However, the biggest limitation is that only a subset of results based on different data sources and methods used have been reported. For example, results based on UK data should have been presented more fully, and this is believed to have increased the risk of reporting bias.

The ERG raised a number of concerns with the clinical effectiveness evidence, including issues with the choice of populations for efficacy and safety, comparators, short follow-up time, and the real-world data used to identify comparators (see Section 1).

4. COST EFFECTIVENESS

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

One set of systematic literature searches was performed to identify cost effectiveness studies, health-state utility values, and cost and healthcare resource use studies (CS, Appendix G, Appendix H and Appendix I).[21]

4.1.1 Searches performed for cost effectiveness Section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the CS.[4, 21] The CADTH evidence-based checklist for PRESS was used to inform this critique.[22, 23] The CS was checked against the STA specification for company/sponsor submission of evidence.[24]

Appendix G, Appendix H and Appendix I of the CS reported the literature searches used to identify cost effectiveness studies, health-state utility values, and cost and healthcare resource use studies.[21] Searches were conducted in May 2020, then updated in February 2021, and updated again in November 2021.

A summary of the resources searched is provided in Table 4.1.

Table 4.1: Resources searched for the cost effectiveness literature review (as reported in the CS)

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE, MEDLINE In-Process, MEDLINE Daily, Epub Ahead of Print	Ovid SP	Latest update: 1946 to November 01, 2021	04/05/20 04/02/21 02/11/21
Embase	Ovid SP	Latest update: Embase 1974 to November 1st, 2021	04/05/20 04/02/21 02/11/21
NHS EED	CRD website	NHS EED: Issue 2 of 4, April 2015	04/05/20
HTA Database	CRD website	Issue 4 of 4, October 2016	04/05/20
INAHTA HTA database	INAHTA website	Latest update: up to Nov 1 2021	04/02/21 02/11/21
Additional resources			
HERC Database of Mapping Studies	https://www.herc.ox.ac.uk/downloads/herc-database-of-mapping-studies	Latest update: up to November 1 2021	04/06/20 24/02/21 10/11/21
CEA Registry	http://healtheconomicsdev.tuftsmedicalcenter.org/cear2/search/search.aspx	Latest update: up to November 1 2021	04/06/20 24/02/21 10/11/21

Resource	Host/Source	Date Ranges	Dates searched
ScHARRHUD	http://www.scharrhud.org/	Latest update: up to November 1 2021	04/06/20 24/02/21 10/11/21
EQ-5D Publications Database	http://eq-5dpublications.euroqol.org/?noheader=true	Latest update: up to November 1 2021	04/06/20 24/02/21 10/11/21
Conference proceedings			
AACR annual meeting	Online abstract books	2018-2021	Not reported
ASCO annual meeting	https://meetinglibrary.asco.org/	2018-2021	Not reported
ESMO congress	Online abstract books	2018-2021	Not reported
ESMO ELCC	Online abstract books	2018-2021	Not reported
ISPOR annual international and European meetings	https://www.ispor.org/heor-resources/presentations-database/search	2018-2021	Not reported
HTA organisations			
AEMPS	https://www.aemps.gob.es/home.htm	Latest update: up to Nov 2021	08/06/20 01/03/21 11/11/21
AIFA	http://www.agenziafarmaco.gov.it	Latest update: up to Nov 2021	08/06/20 01/03/21 11/11/21
AWMSG	http://www.awmsg.org/	Latest update: up to Nov 2021	05/06/20 18/03/21 12/11/21
BAG	https://www.bag.admin.ch/bag/de/home.html	Latest update: up to Nov 2021	08/06/20 01/03/21 11/11/21
Danish Medicine Council	https://medicinraadet.dk/igangvaerende-vurderinger	Latest update: up to Nov 2021	08/06/20 01/03/21 12/11/21
FinCCHTA	https://www.ppshp.fi/Tutkimus-ja-opetus/FinCCHTA/Sivut/default.aspx	Latest update: up to Nov 2021	08/06/20 01/03/21 11/11/21
G-BA	https://www.g-ba.de/	Latest update: up to Nov 2021	08/06/20 01/03/21 12/11/21
HAS	https://www.has-sante.fr/portail/	Latest update: up to Nov 2021	08/06/20 01/03/21 11/11/21
MSCBS	http://www.mscbs.gob.es/	Latest update:	08/06/20

Resource	Host/Source	Date Ranges	Dates searched
	home.htm	up to Nov 2021	01/03/21 11/11/21
NCPE	http://www.ncpe.ie/	Latest update: up to Nov 2021	08/06/20 01/03/21 11/11/21
NICE	https://www.nice.org.uk/	Latest update: up to Nov 2021	08/06/20 01/03/21 12/11/21
NIPH	https://www.fhi.no/en/	Latest update: up to Nov 2021	08/06/20 01/03/21 12/11/21
SMC	https://www.scottishmedicines.org.uk/	Latest update: up to Nov 2021	08/06/20 01/03/21 12/11/21
SBU	https://www.sbu.se/en/	Latest update: up to Nov 2021	08/06/20 01/03/21 12/11/21
Zorginstituut Nederland	https://www.zorginstituut nederland.nl/	Latest update: up to Nov 2021	08/06/20 01/03/21 12/11/21

Additional resources: CEA = Cost-Effectiveness Analysis Registry; HERC = Health Economics Research Centre; ScHARRHUD = School of Health and Related Research Health Utilities Database
Conference proceedings: AACR = American Association for Cancer Research; ASCO = American Society of Clinical Oncology; ESMO = European Society for Medical Oncology; ELCC = European Lung Cancer Annual Congress; HTA = Health Technology Assessment; ISPOR = International Society for Pharmacoeconomics and Outcomes Research
HTA organisations: AEMPS = Agencia Española de Medicamentos y Productos Sanitarios; AIFA = Agenzia Italiana del Farmaco; AWMSG = All Wales Medicines Strategy Group; BAG = Bundesamt für Gesundheit; FinCCHTA = Finnish Coordinating Center for Health Technology Assessment; G-BA = Gemeinsamer Bundesausschuss; HAS = Haute Autorité de Santé; MSCBS = Ministerio de Sanidad, Consumo y Bienestar Social; NCPE = National Centre for Pharmacoeconomics; NICE = National Institute for Health and Care Excellence; NIPH = Norwegian Institute of Public Health; SBU = Swedish Agency for Health Technology Assessment and Assessment of Social Services; SMC = Scottish Medicines Consortium

ERG comment:

- The CS provided full details of the literature searches for the ERG to appraise.[4, 21]
- A comprehensive range of databases, supplementary resources, conference proceedings, and HTA organisation websites were searched.
- Full details of the database searches, including the database name, host platform, date range and date searched, were provided.
- Full details of the supplementary economic specific resources searched were provided, including url links, search terms used, date searched, and results.
- Full details of the conference proceeding searches were provided. The search terms used, url links, date range, and results, were reported.
- Full details of the comprehensive list of HTA organisation websites searched were provided, including the url links, search terms used, date searched, and results.

- The database search strategies were well structured, transparent and reproducible. They included truncation, proximity operators, synonyms, and subject headings (MeSH in MEDLINE and the CRD databases, and EMTREE in Embase). There were no language or date limits for the economic evaluation and health-state utility values elements of the searches. A 5-year date limit was included for the cost and resource use element of the searches in MEDLINE and Embase.
- It would have been preferable for the database search strategies to be presented exactly as run, rather than copied into a tabular format, as item eight of the PRISMA-S checklist recommends.[25] The Cochrane Handbook also recommends that *"...bibliographic database search strategies should be copied and pasted into an appendix exactly as run and in full, together with the search set numbers and the total number of records retrieved by each search strategy. The search strategies should not be re-typed, because this can introduce errors"*.[26]
- The population facet used for the cost effectiveness searches was much broader than that used for the clinical effectiveness searches: NSCLC plus advanced/metastatic. To further ensure sensitivity, the search strategies did not include a facet for interventions/comparators.
- The search strategies did not include the MeSH or EMTREE terms for NSCLC: *Carcinoma, Non-Small-Cell Lung/* or *exp non small cell lung cancer/*.
- The final line from the NHS EED/HTA database search strategy was missing. This was likely to be a reporting error rather than a searching error.
- Study design search filters for economic evaluations, utilities and HRQoL, and cost and resource use were included. The Scottish Intercollegiate Guidelines Network (SIGN) filter for economic studies was used, with additional terms derived from other sources.[33] It would have been helpful if the other sources of additional terms had been cited.[25]
- The update search results were de-duplicated against the original results, as limiting by publication date risks missing relevant studies.[34, 35]

4.1.2 Inclusion/exclusion criteria

In- and exclusion criteria for the review on cost effectiveness studies, utilities and costs and resource use are presented in Table 4.2.

Table 4.2: Eligibility criteria for the systematic literature reviews

Category	Inclusion criteria	Exclusion criteria
Population	Patients with metastatic or surgically unresectable NSCLC. Patients with stage IIIB, IIIC or IV disease. Studies with patients only specified as "stage 3" eligible only if stage 4 patients were also included within the study population .	Patients without metastatic or unresectable NSCLC or studies where outcomes were not presented separately for the patients of interest. Patients with locally advanced disease. Patients with stage 3 disease, if sub-stage b or c not specified.
Intervention	IOs as monotherapy or in combination with chemotherapy. Chemotherapy (platinum or non-platinum-based regimens). Nintedanib in combination with chemotherapy. TKIs[a]	Any other intervention.
Comparators	Any comparator (or none).	–

Category	Inclusion criteria	Exclusion criteria
Outcomes	Cost effectiveness outcomes, including but not limited to: ICERs Cost per clinical outcome Total QALYs Total LYGs Total costs Incremental costs and QALYs	Studies not presenting relevant outcomes for the population of interest.
Study design	Any of the following analysis types: Cost-utility Cost effectiveness Cost-consequence Cost-benefit Cost-minimisation	Any other types of study design.
Publication type	Original research studies (including economic evaluations, observational, interventional and real-world evidence studies). HTAs Congress abstracts published in or after 2018	Any other publication type, including studies not reporting any original research. Congress abstracts published before 2018.
	SLRs were included in the SLR at title/abstract for bibliography searching, these were then subsequently excluded for being an irrelevant study design at full-text review.	
Other considerations	Human subjects English language abstract/full text OECD countries	–

Based on Table 24 of Appendix G of CS.[4]
Note: [a] *Initially, due to the large volume of evidence in the field of NSCLC, the results were limited to publications relevant to OECD countries. However, due to the emerging real-world evidence that has identified TKIs as a constituent of the UK standard of care treatments deemed the relevant comparator for amivantamab, the scope was updated as part of the second SLR update to include economic evaluations reporting on TKIs. Due to the large number of additional economic evaluations included based on this expanded scope, evaluations conducted from a UK perspective were prioritised for extraction. For consistency, these prioritisation criteria were applied across all interventions in the economic evaluations stream. Economic evaluations from a non-UK perspective were still included but are presented as a list.*
CS = company submission; EGFR = epidermal growth factor receptor; HTA = health technology assessment; ICER = incremental cost-effectiveness ratio; IO = immune-oncology; LYG = life years gained; NSCLC = non-small-cell lung cancer; OECD = Organisation for Economic Co-operation and Development; QALY = quality-adjusted life years; SLRs = systematic literature reviews; TKI = tyrosine kinase inhibitor; UK = United Kingdom

ERG comment: The ERG agrees that the eligibility criteria are broadly suitable to fulfil the company's objective to identify cost effectiveness studies. However, the exclusion of non-English studies could have led to some relevant studies being missed.

In addition, there appeared to be some issues with the review methodology which potentially impinge on the ability of the review to ensure that the eligibility criteria were adhered to, including:

- The data extraction was not completed by two independent reviewers, which increases the risk of mistakes made at this stage. In the FAC, the company clarified two independent reviewers were used.

- It is unclear whether the quality assessment was conducted by independent reviewers, which makes the quality assessments less robust. In the FAC, it was clarified that quality assessments were completed by one reviewer and verified by a second independent reviewer.

4.1.3 Conclusions of the cost effectiveness review

The CS provides an overview of the included cost effectiveness, utility and resource use and costs studies, but no specific conclusion was formulated.

ERG comment:

- The CS and response to clarification provided full details for the ERG to appraise the literature searches conducted to identify economic, health-state utility values, and cost and healthcare resource use studies.[4, 9, 21] Searches were conducted in May 2020, then updated in February 2021, and updated again in November 2021. The searches were transparent and reproducible, and comprehensive search strategies were used. A good range of databases and grey literature resources were searched. Search strategies included validated study design search filters. Overall, the ERG has no concerns about the literature searches conducted.
- The eligibility criteria were broadly suitable for the SLR performed. However, the ERG raised several concerns, including about the exclusion of non-English studies, the comparators, and the review methodology (see Section 4.1.2).

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

4.2.1 NICE reference case checklist

Table 4.3: NICE reference case checklist

Element of HTA	Reference case	ERG comment on CS
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	In line with reference case
Perspective on costs	NHS and PSS	In line with reference case
Type of economic evaluation	Cost utility analysis with fully incremental analysis	Partly in line with reference case (i.e., no fully incremental analysis was performed)
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	In line with reference case
Synthesis of evidence on health effects	Based on systematic review	In line with reference case
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	In line with reference case
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Partly in line with reference case (QoL data from the CHRYSALIS trial was only used in a scenario analysis).

Element of HTA	Reference case	ERG comment on CS
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	In line with reference case
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	In line with reference case
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	In line with reference case
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	In line with reference case

CS = company submission; ERG = Evidence Review Group; HTA = health technology assessment; NHS = National Health Service; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom

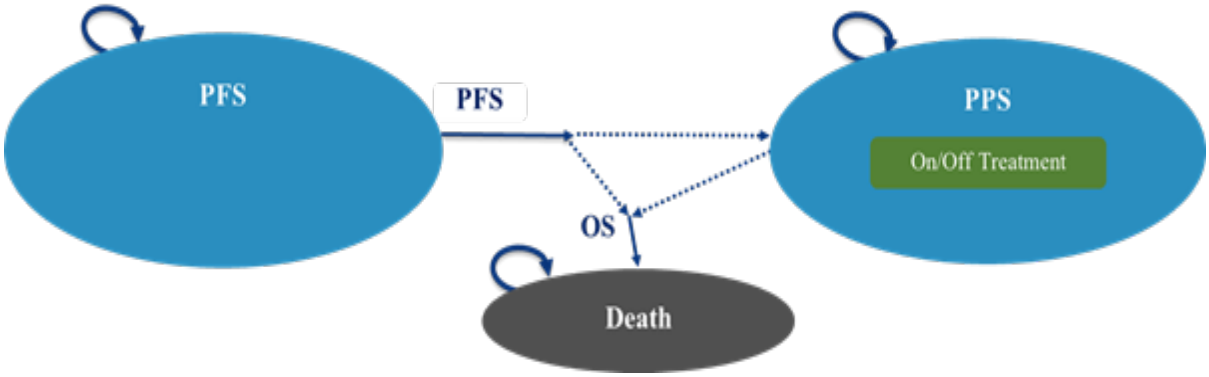
4.2.2 Model structure

A partitioned survival model (PSM) was developed including three health states: a PFS state, a PPS state, and death (Figure 4.1). The company stated that a partitioned survival analysis approach was chosen because it permits the use of outcome data from the adjusted treatment comparison presented in Section B.2.9 of the CS and permits the clinical benefits of amivantamab to be captured by reflecting the increased proportion of patients expected to be alive and/or progression-free over time. In addition, it was deemed in line with previous cost effectiveness models in metastatic NSCLC with EGFR. The model was developed in Microsoft Excel.

The allocation of patients into health states was directly based on treatment-specific PFS and OS functions. The model considers up to two distinct lines of treatment: current-line treatment while in the PFS state, and a subsequent line while in the PPS state. Time on treatment was assumed to be equal to progression. Upon disease progression patients could receive a basket of subsequent treatments. The proportion of patients receiving these treatments and the composition of the subsequent treatment basket was based on US RWE pooled data. Only costs of subsequent treatments were considered in the model, as it was assumed that efficacy was implicitly captured in OS extrapolations.

A lifetime horizon (i.e., 15 years) with a cycle length of 4 weeks (including half-cycle correction) was applied to ensure all costs and QALYs were captured. This was considered appropriate given that the mean starting age of the patients (61.75 years) and their poor prognosis.

Figure 4.1: Model structure



Source: Figure 23 of the CS

CS = company submission; PFS = progression-free survival

ERG comment: The main concern of the ERG relates to the use of a PSM without exploring a state transition model (STM) alongside it. The NICE Decision Support Unit (DSU) TSD 19 recommends the use of STMs alongside PSMs to verify the plausibility of PSM extrapolations and to explore key clinical uncertainties in the extrapolation period. In response to clarification question B2, the company stated that although over- or underestimation of long-term outcomes is a potential limitation of a PSM, the CHRYSALIS trial data were relatively mature and the risk of long-term over- or under-estimation of outcomes with a PSM was therefore likely limited. In addition, the company validated their approach based on literature comparing PSM and STM approaches and other NSCLC NICE submissions. Although the ERG ideally would have liked to see a STM to verify the PSM results, the ERG agrees the company's arguments are reasonable.

4.2.3 Population

The population considered in the CS (CS, Table 1) was adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon20ins, whose disease has progressed on or after platinum-based chemotherapy, which is different from the population defined in the final NICE scope and may not be generalisable to the England and Wales NHS population.

The modelled baseline patient characteristics were presented in Table 43 of the CS. These have been taken from the patients in the CHRYSALIS trial, as clinical experts indicated that they were largely generalisable to the patient population in the UK.

ERG comment: The main concern of the ERG relates to the population considered by the company being slightly narrower than the population defined in final NICE scope. The narrower population may not be generalisable to the England and Wales NHS population and may for example have led to an underestimation of AEs. More details regarding this issue are provided in Sections 2.1 and 3.2.

4.2.4 Interventions and comparators

The intervention considered in the CS was amivantamab monotherapy. Amivantamab was administered via IV infusion at 1,050 mg for patients with body weight <80 kg and 1,400 mg for patients with body weight ≥80 kg once weekly for the first 4 weeks and then once every 2 weeks starting at week 5, consistent with the regimen used in the CHRYSALIS trial and the SmPC for amivantamab. Although the protocol of the CHRYSALIS trial allowed patients to continue to receive treatment following disease progression, UK clinical experts considered this does not reflect clinical practice and treatment discontinuation was therefore assumed upon disease progression.

The NICE scope listed the following comparators: established clinical management without amivantamab, including but not limited to atezolizumab, nivolumab (subject to an ongoing NICE appraisal), pembrolizumab (for disease with PD-L1 >1%) and chemotherapy such as docetaxel alone or with nintedanib, pemetrexed and carboplatin. As the CHRYSALIS trial is a single arm study, data informing comparator efficacy in the economic model were derived from pooled US RWE data. According to clinical experts, there is no established standard treatment pathway for patients with EGFR Exon20ins mutated NSCLC in the UK and amivantamab was therefore compared to a basket of treatments termed UK SoC within the model. The treatment classes included in this basket were IO agents ([redacted]), EGFR TKIs ([redacted]), platinum-based chemotherapy regimens ([redacted]), non-platinum-based chemotherapy regimens ([redacted]) and other ([redacted]), as reported in Table 5 of the CS (transposition of the values for IO agents and EGFR TKIs corrected by the ERG). After redistribution of the 9% in the 'other' category, the four treatment classes included in this basket were IO agents ([redacted]), EGFR TKIs ([redacted]), platinum-based chemotherapy regimens ([redacted]) and non-platinum-based chemotherapy regimens ([redacted]), as reported in Table 38, CS. For costing purposes, the individual treatments considered in each of these four treatment classes were as follows:

- IO agents: atezolizumab (45%), pembrolizumab (45%) and nivolumab (10%)
- EGFR TKIs: afatinib (100%)
- Platinum-based chemotherapy: carboplatin plus gemcitabine (33.3%), carboplatin plus pemetrexed (33.3%) and carboplatin plus vinorelbine (33.3%)
- Non-platinum-based chemotherapy: docetaxel plus nintedanib (75%) and docetaxel monotherapy (25%)

Scenario analyses were performed to assess the impact of varying the treatments and treatment proportions implemented in the model.

The composition of the basket for subsequent treatments received following amivantamab or UK SoC was sourced from the subsequent treatment distribution of patients receiving third-line or later therapy in the pooled US RWE database and are presented in Tables 39 and 40 of the CS. In line with this study, [redacted] of patients are modelled to receive subsequent treatments (calculated from the proportion of second line patients receiving a third-line treatment upon progression), with the remaining [redacted] of patients receiving no active treatment and assumed to receive best supportive care (BSC). A scenario analysis was explored in which the subsequent treatment composition for patients following amivantamab was sourced from the subsequent treatment distribution of patients receiving third-line or later therapy in the CHRYSALIS trial.

ERG comment: The main concern of the ERG relates to the effectiveness of the comparator basket being representative of UK clinical practice.

Due to considerable heterogeneity in treatments due to lack of specifically recommended treatments in the UK, data informing comparator efficacy were derived from a basket of treatments from a US RWE database study. The comparator effectiveness and costs are therefore based on the average clinical effectiveness and weighted average costs across all the treatments included in the comparator basket. As reported in Table 38 of the CS, the company assumed [redacted] of the comparator basket to exist of EGFR TKIs. It is, however, unclear to the ERG whether this is consistent with UK clinical practice, especially given that, as reported on page 23 of the CS, Exon20ins mutations have been associated with resistance to EGFR TKIs. In addition, the results of the indirect treatment comparison excluding TKIs in response to clarification question A6c show that the HRs are slightly higher than the base case HRs, indicating that the effectiveness of EGFR TKIs for Exon20ins mutations may indeed be questionable. Therefore,

the inclusion of the substantial proportion of EGFR TKI in the US RWD is considered as a source of uncertainty by the ERG, potentially underestimating outcomes for the comparator basket. This means that ICERs might be under-estimated. The ERG would like to see an analysis where EGFR TKI therapies are excluded from the US RWD informing the comparator basket. In addition, although the ERG acknowledges the limitation of small sample sizes of patients receiving individual treatments in the RWE sources, a fully incremental analysis of all relevant comparators in the comparator basket would be informative (as was requested in the clarification letter, but not provided) to address the uncertainty of assuming average effectiveness and costs of a basket of treatments.

4.2.5 Perspective, time horizon and discounting

The analysis is performed from the NHS and PSS perspective. Discount rates of 3.5% are applied to both costs and benefits. The model cycle length is 4 weeks with a lifetime time horizon (15 years).

ERG comment: The approach is in concordance with the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness used for amivantamab and SoC are the CHRYSALIS trial and US RWE data respectively. The baseline characteristics of the modelled population were based on the CHRYSALIS trial. To account for differences in the treatment populations of CHRYSALIS and the US RWE used to inform comparator efficacy, the treatment comparisons were adjusted for differences in key prognostic variables at baseline (identified a priori by an SLR and validated by clinical experts). For the US RWE an ATT approach (IPW) was used while for the scenario analysis using the PHE data a covariate adjustment approach was used.

The main outcomes regarding treatment effectiveness were OS and PFS. The company stated that the criteria that were used to decide on the best parametric fit were 1) visual fit to the observed KM curve, 2) statistical fit based on AIC and BIC statistics, and 3) face validity based on expert opinion.

4.2.6.1 Company's base case

The company selected the Weibull model in its base case for the extrapolation of OS in the amivantamab arm. For amivantamab PFS, the company selected the generalised gamma model. Progression in the base case was assessed with a BICR. For more details regarding the company's survival curve selection see Table 4.4 (criteria based on NICE DSU TSD 14).

For both OS and PFS for patients receiving SoC the KM curve was directly used rather than selecting a parametric model. The company argued that extrapolation of OS and PFS for SoC was not necessary, as the KM data was based on a 'robust' population size (n=206) and all patients had reached the end point for both outcomes.

Table 4.4: Criteria for choice of survival curves

Criteria for choice of survival curve	OS	PFS	TTD
General considerations	SoC Extrapolation of the US RWE data informing efficacy for UK SoC was not deemed necessary	SoC Extrapolation of the US RWE data informing efficacy for UK SoC was not deemed necessary	Amivantamab The company assumed that time on treatment was equal to PFS. SoC

Criteria for choice of survival curve	OS	PFS	TTD
	due to the maturity of the available data.	due to the maturity of the available data.	It was assumed that SoC time on treatment is equal to SoC PFS.
Reporting of log-cumulative hazard plots, quantile-quantile plots or suitable residual plots to allow initial selection of appropriate models	Log-cumulative hazard plots, Schoenfeld residuals were provided. Proportional hazards assumption does not hold.*	Log-cumulative hazard plots, Schoenfeld residuals were provided Proportional hazards assumption does not hold.*	*Not reported by the company*
Fit to the observed data based on AIC and BIC	Amivantamab Lowest AIC: Weibull Lowest BIC: Exponential SoC (US RWE - scenario) Lowest AIC & BIC: Weibull	Amivantamab Lowest AIC: Log-logistic Lowest BIC: Log-logistic SoC (US RWE - scenario) Lowest AIC & BIC: Log-logistic	*Not reported by the company*
Fit to the observed data based on visual comparison with the Kaplan-Meier curves	Plots including KM curve and all parametric curves were provided for amivantamab and SoC. No further comment was made based on their visual fit.	Plots including KM curve and all parametric curves were provided for amivantamab and SoC. No further comment was made based on their visual fit.	*Not reported by the company*
Clinical plausibility of the extrapolation based on comparison with data	*Not reported by the company*	*Not reported by the company*	*Not reported by the company*
Clinical plausibility of the extrapolation based on clinical expert opinion	5-year OS expectation of 7-8%.	5-year PFS expectation less than 1%. 2-year PFS expectation about 10%.	Assumption that PFS equals TT
Base case approach	Based on expert opinion and best fit with AIC the Weibull curve was chosen. KM curves were considered directly for SoC in the CS base case.	Generalised gamma curve was selected based on expectation of 2-year and 5-year PFS. KM curves were considered directly	TTD was set equal to PFS based on expert opinion

Criteria for choice of survival curve	OS	PFS	TTD
		for SoC in the CS base case.	
Based on CS Section 3.3 CS = company submission; KM = Kaplan-Meier; OS = overall survival; PFS = progression=free survival; RWE = real world evidence; SoC = standard of care; TTD = time to treatment discontinuation; UK = United Kingdom; US = United States; *provided in response to clarification letter			

4.2.6.2 Scenario analyses

To explore the impact of alternative assumptions the company conducted several scenario analyses:

- For the amivantamab treatment population the impact of using IA progression instead of BICR-assessed progression was explored. A log-normal model was selected for the scenario based on AIC and BIC fit.
- For SoC OS, based on the US RWE, the impact of using a Weibull model (based on statistical fit) and a generalised gamma model (based on expert expectations of survival) were explored. Further the UK PHE data was explored as an alternative source of data. Here, the KM curve was directly implemented in the model.
- For SoC PFS, based on the US RWE, the impact of using a log-logistic model (based on statistical fit) was explored. Again, using the impact of the UK PHE data was explored by implementing the KM curve directly into the model. In this case TTNT was used as a proxy as progression data was unavailable.

4.2.6.3 Time to treatment discontinuation

To calculate treatment costs (i.e., drug acquisition and drug administration costs), TTD was implemented in the model. While the median treatment duration ([redacted] months) in the CHRYSALIS trial was longer than the median PFS ([redacted] months), clinical experts stated that time to discontinuation would usually be the same as time to progression. Therefore, the company base case assumed that time-on-treatment was equal to PFS.

ERG comment: The main concerns of the ERG relate to: a) using KM data for SoC survival analyses of PFS and OS; b) assumption that treatment discontinuation is equal to PFS; c) lack of transparency and choice of curve for the modelled treatment discontinuation; d) adherence of the company to the NICE DSU TSD 14[36]; e) a lack of exploration around uncertainty of the parametric survival curves; f) no inclusion of treatment waning in the model; g) alternative methods to perform indirect treatment comparison; and h) external validation of parametric curves.

a) For survival analyses of OS and PFS in the SoC arm, the company argued that due to the maturity of the data and all patients reaching the specified end point or being censored within the timeframe of data collection, KM data could be directly implemented rather than fitting a parametric survival model. However, this is not necessarily in line with NICE DSU TSD 14, which states that "*parametric models are likely to represent the preferred method for incorporating survival data into health economic models in the majority of cases*". For example, the 'stepped' nature of KM curves, resulting from follow-up only occurring at pre-specified time intervals, means that events are only observed to have occurred at specific intervals which could create bias in survival analysis results. Moreover, the implementation of KM data may introduce overfitting of the modelled survival outcomes. Implementing KM curves biases the SoC treatment effectiveness as patients do not transition smoothly. Instead at each measurement point all patients who have died or progressed

will leave the health state at once, which is not valid. Hence, the ERG requested a scenario analysis in which the most appropriate parametric models were selected for OS and PFS in the SoC arm including a PSA with 5,000 iterations. The company complied with this request, implementing a Weibull model for OS and a log-logistic model for PFS. The ICER of the resulting probabilistic analysis including the PAS price was £40,353. The ERG therefore chose to implement these parametric models into its base case.

b) For the estimation of TTD, the company assumed that treatment would be discontinued when a patient progresses, setting TTD equal to PFS. The CHRYSALIS trial, however, allowed patients to remain on treatment after disease progression and median TTD ([redacted]) was substantially longer than median PFS ([redacted]). The ERG questions the company's approach. The assumption reduces the estimated treatment costs of amivantamab without reducing the estimated effectiveness after progression of amivantamab. The ICER is therefore likely underestimated. The ERG therefore requested a scenario analysis in which TTD would be informed by the CHRYSALIS trial protocol for amivantamab, which increased the ICER to £50,549 per QALY gained. An additional scenario analysis was conducted in which TTNT was used as a proxy for TTD in the SoC arm, decreasing the ICER to £33,708 per QALY gained. The company argued that the second scenario analysis was more valid, as in this case the assumptions made for each treatment arm would be in line with each other. The ERG disagrees with this judgement. TTNT likely overestimates TTD, as the time to the start of a next treatment is per definition longer than the time to discontinue treatment. Additionally, no compelling evidence was provided by the company to demonstrate the TTNT as a good approximation to TTD. While acknowledging that this approach may be conservative, the ERG therefore chose to implement parametric survival curves for TTD in the amivantamab arm and take PFS as a proxy for TTD in the SoC arm.

c) Upon request, the company conducted two scenario analyses using parametric survival curves to reflect treatment discontinuation (as described in critique b)). The choice of survival curves for these analyses was not transparent (i.e., lacked details regarding the NICE DSU TSD 14 criteria). For amivantamab a Gompertz model was implemented, while the KM-curve was used for SoC. The limited indicators that are available to the ERG showed that the Gompertz model had the fourth best statistical fit (exponential, Weibull and log-logistic models all had a better fit) and did not clearly have the best visual fit. The Gompertz model distinguished itself from other models by being the most pessimistic curve (i.e., resulting in the lowest number of patients on-treatment over time). For SoC, the generalised gamma model had the best statistical fit. The ERG therefore implemented an exponential model for amivantamab, which had the best statistical fit and was in between the most optimistic and pessimistic curves, hence not presenting an extreme of early discontinuation or late discontinuation. For a scenario analysis exploring TTNT as a proxy for TTD in the SoC arm, the generalised gamma model was chosen.

d) In the initial CS, there was substantial uncertainty surrounding the adherence of the company to the NICE DSU TSD 14[36]. Upon request for clarification, log-cumulative hazard plots and Schoenfeld residual plots were submitted by the company. However, for other additional analyses conducted for other clarification requests, NICE DSU TSD 14 details were again not submitted. The ERG could therefore only judge the new analyses on statistical measures of fit and visual fit, rather than all relevant NICE DSU TSD 14 criteria.

e) For the modelling of PFS in amivantamab even though AIC and BIC indicated that a log-logistic curve would be the best fit, a generalised gamma curve was implemented based on the fit to expected progression-free rates based on expert opinion. The resulting uncertainty was not explored. Upon clarification, the company elaborated that while log-logistic curves had a better statistical fit and the log-logistic curves would be consistent with a decreasing hazard, log-logistic curves had a long tail, which did not seem like a valid assumption to the analysts. The ERG has

looked into the impact of assuming the log-logistic curve for PFS in the amivantamab arm and this did not seem to have a large impact on the ICER.

f) The ERG considered that the assumption of a lifelong treatment effect may not be warranted and requested the company to explore treatment waning in the model. Upon request to do so, the company refused with the arguments that 1) treatment waning would be implicitly captured in the selected curves, 2) due to the poor prognosis patients receive treatment for a relatively short amount of time, and 3) amivantamab is a treat to progression treatment. It is unclear to the ERG whether this assumption holds true in clinical practice as there is limited evidence provided on treatment waning by the company. The follow-up of the CHRYSALIS study is notably shorter than the time horizon in the economic model. Hence, it is unclear to the ERG whether the benefits of amivantamab could be assumed to last over the full-time horizon. This has also been acknowledged in other STAs. For example, in TA520, the appraisal committee concluded that a lifetime treatment effect was implausible. The ERG would like to see an updated economic model in which the company explores treatment waning scenarios. Additional evidence to support the company's statement that treatment waning would be implicitly captured in the selected curves would also be informative to address this issue.

g) The comparative effectiveness of amivantamab versus SoC was explored via covariate adjustment and IPW. However, alternative approaches to address confounding in the indirect treatment comparison are possible. Hence, the ERG requested the company to implement matching instead of IPW to examine the potential uncertainty introduced by different methodological choices. In response to clarification question B4, the company performed a PSM analysis in which SoC patients from the US RWE and those from CHRYSALIS have been matched to estimate the relative efficacy of amivantamab versus UK SoC. This resulted in an ICER of £45,092 per QALY gained. The ERG acknowledges the concerns of the company that the matching results in a smaller sample size and that the IPW results therefore might be slightly more robust. However, the ERG implemented the results of the PSM analysis as second ERG base case.

h) In response to clarification question B5, the company provided an overview of the validity of the extrapolated OS and PFS rates beyond the trial data for both amivantamab and SoC. The company stated that, to this extent, "*clinicians were presented with both KM data and curve extrapolation options for OS and PFS for both amivantamab and UK SoC (as informed by US RWE or PHE cohort data). The clinicians were then asked whether the KM curves and the available extrapolations broadly aligned with their clinical expectations for EGFR Exon20ins mutated NSCLC patients in UK clinical practice receiving either amivantamab or UK SoC after the failure of platinum-based chemotherapy*". The resulting estimates are presented in Table 29 of the company's response to clarification. In Table 30 of the company's response to clarification, the corresponding modelled long-term OS and PFS rates assumed in the base case economic analysis are presented. Although the modelled results seem to be in line with clinical expectations, the ERG would like to emphasise that rates of OS and PFS in the model seem to be slightly underestimated for SoC and overestimated for amivantamab compared to estimations made by the clinicians.

4.2.7 Adverse events

The economic model included grade ≥3 AEs that were reported in more than 5% of patients in key trials, except for incidence of diarrhoea, which was considered at any grade due to its clinical relevance (see Table 49 of the CS). In the CS, it was stated that "*clinical expert opinion received by Janssen supports that these AEs are relevant for inclusion and that no relevant events expected to affect more than 5% of patients have been omitted*".[4] AEs were only considered for current-line treatments, and AEs associated with subsequent-line treatments were not included. The main sources of evidence on treatment AEs used for intervention and comparators were clinical trials (CHRYSALIS for

amivantamab, AURA3 for platinum-based chemotherapy (as per TA653) and LUX-Lung-8 for EGFR TKIs) or previous NICE appraisals (TA520 for IO agents and non-platinum-based chemotherapy).[10, 37, 38]

The consequences of AEs were modelled in terms of the accrual of associated management costs and disutilities. The percentage of patients who experienced AEs was calculated at the start of the model and one-off costs and disutilities were incurred at this stage.

ERG comment: In the CS base case, disutilities associated with grade ≥3 AEs were based on a weighted average based on the treatment class proportions in the US RWE. Considering that SoC is a basket of treatments including IO agents, EGFR TKIs and platinum and non-platinum-based chemotherapies, it is uncertain whether this basket is representative of UK clinical practice (see Section 4.2.4). Hence, the ERG would have liked to see a scenario analysis where EGFR TKI therapies are excluded from the US RWD informing the comparator basket AEs disutilities.

4.2.8 Health-related quality of life

The company stated that EQ-5D-5L data were collected in CHRYSALIS at day 1 of each cycle, at the end of treatment and during post-treatment follow-up. However, in the CS, the company states that "*the number of responses to the EQ-5D-5L questionnaire was low at the time of data cut-off and were therefore not used in the model*" [4].

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

According to the CS, the SLR identified 50 articles reporting on 47 unique studies. Although an appendix was provided with more details, the company did not summarise in the CS whether any of these studies could be used in the economic model.

4.2.8.2 Health state utility values

Health state utility values used in the economic model have been sourced from TA484/TA713, a previous NICE appraisal in advanced non-squamous NSCLC after chemotherapy[11, 39]. In the CS, the company stated that "*this was considered a suitable source for utility data given the similarity of this population to the population of interest in this submission*". Furthermore, the company stated that UK clinical experts consulted as part of this appraisal confirmed that the utility values used are appropriate[4].

Utilities were not age-adjusted, which the company justified by stating that the time horizon of the economic model is relatively short, and the impact of age-adjustment on the results is therefore likely to be marginal.

The company stated in the CS that the standard error for utilities was assumed to be ±10% of the mean.

A summary of all utility values used in the cost effectiveness analysis is provided in Table 4.5.

Table 4.5: Health state utility values

Health state	Utility value	Standard error
Progression-free survival	0.713	0.0713
Post-progression survival	0.569	0.0569

Based on TA484/TA713.93
Based on CS, Table 51
CS = company submission

4.2.8.3 Disutility values

The company implemented one-off disutilities for AEs, sourced from TA520, TA484/TA713 and the published literature (see Table 4.6).

Table 4.6: Summary of AE disutilities applied in the cost effectiveness model

AE	Disutility (SE)	Source
Anaemia	−0.073 (0.018)	Nafees et al. (2008) as per TA484/TA713 and TA52094[10, 39, 40]
Diarrhoea	−0.047 (0.016)	Nafees et al. (2008) as per TA484/TA71394[39, 40]
Fatigue	−0.073 (0.018)	Nafees et al. (2008) as per TA484/TA713 and TA52094[10, 39, 40]
Febrile neutropenia	−0.090 (0.016)	Nafees et al. (2008) as per TA484/TA713 and TA52094[10, 39, 40]
Neutropenia	−0.090 (0.015)	Nafees et al. (2008) as per TA484/TA713 and TA52094[10, 39, 40]
Neutrophil count decreased	0	TA484/TA713 and TA52094[10, 39, 40]
Rash	−0.032 (0.012)	Nafees et al. (2008)[40]
Thrombocytopaenia	−0.108 (0.011)	Tolley et al. (2013)[41]
Based on CS, Table 50 AEs = adverse events; CS = company submission		

ERG comment: The main concerns of the ERG relate to a) exclusion of age-adjustment to the health state utilities; and b) source of health state utilities.

a) The company stated that given the relatively short time horizon of the model, the impact of age-adjustment on results is likely to be marginal and as such, utilities were not age-adjusted. In response to clarification question B12, the company provided an updated model which included the possibility to run the model with age-adjusted utilities, which slightly increased the ICER to £40,293 per QALY gained. This adjustment was included in the ERG base case.
b) Although EQ-5D-5L data were collected in CHRYSALIS, health state utilities in the economic model were sourced from TA484/TA713 as the number of EQ-5D-5L responses from the CHRYSALIS trial was low at the time of data cut-off. In response to clarification question B11, the company provided a scenario analysis informing health state utilities based on the collected HRQoL data in CHRYSALIS. This resulted in a slight increase in the ICER (£42,117 per QALY gained compared to £39,764 per QALY gained in its base case). Given the small sample from which utilities were collected, the ERG is not necessarily against the use of utilities from TA484/TA713. In response to clarification question B11, the company presented scenario analyses investigating the effect of using health state utilities from TA428 and TA347. This resulted in an ICER of £35,617 per QALY gained and £38,086 per QALY gained. The ERG acknowledges the limitations of the HRQoL data in CHRYSALIS and is satisfied with the additional analyses the company provided, which only had a minor impact on the ICER.

4.2.9 Resource use and costs

The cost categories included in the model were drug acquisition costs, drug administration costs, costs of subsequent treatments, medical & monitoring costs (i.e., liver function test, renal function test, full blood test, outpatient oncologist visit, CT scan (chest), General Practitioner (GP) surgery visit, GP home

visit, non-admitted monitoring consultation, and palliative care), costs of managing AEs, and end-of-life costs[4].

Unit prices were based on the NHS reference prices, British National Formulary (BNF), Personal Social Services Research Unit (PSSRU) and pharmaceutical electronic market information tool (eMIT).

4.2.9.1 Resource use and costs data identified in the review

According to the CS, the SLR identified seven articles reporting on seven unique studies in patients with lung cancer. The company stated that no studies reporting on cost and healthcare resource use were conducted in the population considered in this submission (adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon20ins, whose disease has progressed on or after platinum-based chemotherapy).

4.2.9.2 Drug acquisition and administration costs (with PAS)

Drug acquisition costs for every 4-week model cycle were calculated for each treatment based on the dosing schedule and the UK list price of each pack or vial. The company stated that, in the base case, no vial sharing is assumed given the small patient population.

All drugs administered orally or via IV infusion were assumed to be administered in an outpatient setting. The administration-related costs were derived according to data available from the NHS Reference Costs 2019/20.

Dosing regimens and cost per model cycle of intervention and comparators, including amivantamab PAS discount can be found in Table 4.7. A summary of drug costs, administration costs, AE management costs, disease management costs, and subsequent treatment costs per cycle can be found in Table 4.8.

Table 4.7: Dosing regimens and cost per model cycle of intervention and comparators, inclusive of amivantamab PAS discount

Treatment	Dosing regimen	Stopping rule	Cost per dose	Admins per cycle	Cost per treatment cycle	Weeks per cycle	Cost per model cycle
Amivantamab (1,050 mg)	1,050 mg or 1,400 mg (weight dependent) weekly for 4 weeks and bi-weekly thereafter	Treat to progression	£[redacted]	Initial cycle: 4 Subsequent cycles: 2	Initial cycle: £[redacted] Subsequent cycles: £[redacted]	4	Initial cycle: £[redacted] Subsequent cycles: £[redacted]
Amivantamab (1,400 mg)		Treat to progression	£[redacted]		Initial cycle: £[redacted] Subsequent cycles: £[redacted]	4	Initial cycle: £[redacted] Subsequent cycles: £[redacted]
EGFR TKIs (note: in the CS base case, only the costs for afatinib were assumed)							
Afatinib	Oral, 40 mg daily	Treat to progression	£72.26	28	£2,023.28	4	£ 2,023.28
Osimertinib	Oral, 80 mg daily	Treat to progression	£192.33	28	£5,385.33	4	£5,385.33
IO agents							
Atezolizumab	1,200 mg every 3 weeks	Treat to progression	£3,807.69	1	£3,807.69	3	£5,076.92
Pembrolizumab	200 mg every 3 weeks	Treat to progression	£5,260.00	1	£5,260.00	3	£7,013.33
Nivolumab	240 mg every 2 weeks	Treat to progression	£3,291.00	1	£3,291.00	2	£6,582.00
Platinum-based chemotherapy regimens							
Carboplatin + gemcitabine							Initial cycle: £84.92 Subsequent cycles: £0

Treatment	Dosing regimen	Stopping rule	Cost per dose	Admins per cycle	Cost per treatment cycle	Weeks per cycle	Cost per model cycle
Carboplatin	Area under curve 6 mg/mL per minute administered every 3 weeks	Four treatment cycles or progression	£27.03	1	£108.10	12	Initial cycle: £36.03 Subsequent cycles: £0
Gemcitabine	1,250 mg/m^2 on days 1 and 8 every 3 weeks		£18.33	2	£146.65	12	Initial cycle: £48.88 Subsequent cycles: £0
Carboplatin + vinorelbine							Initial cycle: £76.74 Subsequent cycles: £0
Carboplatin	Area under curve 5 mg/mL per minute administered every 3 weeks	Four treatment cycles or progression	£25.67	1	£102.66	12	Initial cycle: £34.22 Subsequent cycles: £0
Vinorelbine	25 mg/m^2 on days 1 and 8 every 3 weeks		£15.95	2	£127.56	12	Initial cycle: £42.52 Subsequent cycles: £0
Carboplatin + pemetrexed							Initial cycle: £1,459.22 Subsequent cycles: £0
Carboplatin	Area under curve 5 mg/mL per minute administered every 3 weeks	Four treatment cycles or progression	£25.67	1	£102.66	12	Initial cycle: £34.22 Subsequent cycles: £0
Pemetrexed	500 mg/m^2 on day 1 every 3 weeks		£1,068.75	1	£4,275.00	12	Initial cycle: £1,425.00 Subsequent cycles: £0
Non-platinum-based chemotherapy regimens							
Docetaxel + nintedanib							First six cycles: £1,935.83 Subsequent cycles: £1,912.09
Docetaxel	75 mg/m^2 repeat cycle every 3 weeks	Fixed duration (six cycles)	£17.81	1	£18.26	3	£24.35
Nintedanib	200 mg twice daily on days 2–21 of cycle	Treat to progression	£35.85	40	£1,434.07	3	£1,912.09

Treatment	Dosing regimen	Stopping rule	Cost per dose	Admins per cycle	Cost per treatment cycle	Weeks per cycle	Cost per model cycle
Docetaxel	75 mg/m^2 repeat cycle every 3 weeks	Treat to progression	£17.81	1	£18.26	3	£24.35
Based on CS, Table 53 CS = company submission; EGFR = epidermal growth factor receptor; IO = immune-oncology; TKIs = tyrosine kinase inhibitor							

4.2.9 Resource use & monitoring costs

The types of resource use incorporated in the model were based on TA520[10]. The company stated that "*this was considered to be a suitable source for healthcare resource use given that it is a relatively recent NICE appraisal that considered a patient population analogous to that of this submission*".

4.2.10 Adverse reaction unit costs and resource use

The cost of managing AEs experienced by patients receiving treatments was included as a one-off cost in the economic model. The company stated that the costs per event were based on NHS Reference Costs 2019–20 as per TA653[42].

4.2.11 End-of-life costs

A one-off cost representing the cost of terminal care was applied in the model in the first cycle post-death. The cost applied in the model (£3,803.36) was derived as per the assumptions in TA520, using costs from the NHS Reference Costs (2019/20) and PSSRU (2021) [10].

Table 4.8: Summary of drug costs, administration costs, AE management costs, disease management costs, and subsequent treatment costs per cycle

Drug costs, initial cycle		**Measurement of uncertainty (distribution)**
Amivantamab	£13,780.99	Assumed to be ±10% of the mean (Gamma)
IO agents	£6,098.81	
EGFR TKIs	£2,023.28	
Pt-based chemotherapy	£540.29	
Non-Pt-based chemotherapy	£1,457.81	
Drug costs, subsequent cycles		**Measurement of uncertainty (distribution)**
Amivantamab	£6,890.49	Assumed to be ±10% of the mean (Gamma)
IO agents	£6,098.81	
EGFR TKIs	£2,023.28	
Pt-based chemotherapy	£0.00	-
Non-Pt-based chemotherapy	£1,440.00	Assumed to be ±10% of the mean (Gamma)
Administration costs, initial cycle		
Amivantamab	£885.39	Assumed to be ±10% of the mean (Gamma)
IO agents	£309.89	
EGFR TKIs	£207.79	
Pt-based chemotherapy	£666.41	
Non-Pt-based chemotherapy	£295.13	
Administration costs, subsequent cycles		
Amivantamab	£442.70	Assumed to be ±10% of the mean (Gamma)
IO agents	£309.89	
EGFR TKIs	£0.00	-
Pt-based chemotherapy	£0.00	-
Non-Pt-based chemotherapy	£73.78	Assumed to be ±10% of the mean (Gamma)
AE management costs		
Amivantamab	£242.43	Assumed to be ±10% of the mean (Gamma)

UK SoC	£628.82	
Disease management costs, progression-free		
Amivantamab	£648.19	Assumed to be ±10% of the mean (Gamma)
UK SoC	£823.35	
Disease management costs, post-progression		
Amivantamab	£536.28	Assumed to be ±10% of the mean (Gamma)
UK SoC	£536.28	
Disease management costs, one-off cost		
Mortality	£3,803.36	Assumed to be ±10% of the mean (Gamma)
Subsequent treatment costs		
Amivantamab	£8,200.12	Assumed to be ±10% of the mean (Gamma)
UK SoC	£8,469.41	
Based on CS Table 60 AEs = adverse events; CS = company submission; EGFR = epidermal growth factor receptor; IO = immuno-oncology; Pt = platinum; SoC = standard of care; TKIs = tyrosine kinase inhibitor; UK = United Kingdom		

ERG comment: The main concerns of the ERG relate to a) treatment costs for EGFR TKIs solely being based on afatinib and b) exclusion of costs for diagnostic testing for EGFR in people with NSCLC.

a) In the CS base case, treatment costs for EGFR TKIs are solely based on afatinib (e.g., excluding osimertinib) rather than calculating this based on the proportion of patients per EGFR TKI in the US RWE. This is likely not in line with UK clinical practice (see Section 4.2.4). Furthermore, the company provided a scenario analysis in which the costs of EGFR TKIs were solely based on osimertinib, which decreased the ICER to £31,224 per QALY gained. Although the ERG prefers EGFR TKIs to be removed from the model (Section 4.2.4), if the company decides to include them, the EGFR TKI treatment costs should be based on proportions in line with clinical evidence.
b) In the final scope issues by NICE, it is stated that "*The use of amivantamab is conditional on the presence of an EGFR mutation. The economic modelling should therefore include the costs associated with diagnostic testing for EGFR in people with NSCLC who would not otherwise have been tested. A sensitivity analysis should be provided without the cost of the diagnostic test*". However, in response to clarification question B13, the company argued: "*EGFR Exon20ins mutations can be tested as part of the EGFR test conducted at diagnosis for all NSCLC patients. As such, Janssen, considers there are no additional costs likely to be incurred by the NHS over and above the current standard of care EGFR testing requirements for all NSCLC patients*". The ERG is satisfied with this justification.

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

5.1.1 Company's deterministic base case results

In the company's probabilistic cost effectiveness results (probabilistic) indicated that amivantamab is both more costly (additional costs of [redacted]) and more effective (incremental QALYs of [redacted]) UK SoC, amounting to an ICER of £40,246 per QALY gained (see Table 5.1). The probability of amivantamab being cost effective at a threshold of £50,000 per QALY gained was around 68% (i.e., due to variation in the PSA results when running the model multiple times).

Overall, the technology is modelled to affect QALYs by (deterministic):

- Increased PPS, with an increment of 0.526 years (63% of total incremental LYs) in the amivantamab arm (1.349 years) compared with UK SoC (0.823 years)
- Increasing PFS, with an increment of 0.314 years (37% of total incremental LYs) in the amivantamab arm (0.818 years) compared with UK SoC (0.504 years)

Overall, the technology is modelled to affect costs by (deterministic):

- The higher drug costs (additional cost of [redacted], [redacted] of total incremental costs), administration costs (additional cost of [redacted], [redacted] of total incremental costs) and post-progression disease management costs (additional cost of [redacted], [redacted] of total incremental costs)

Table 5.1: Company's probabilistic base case results (with PAS)

Technology	Total costs	Total LYs	Total QALYs	Incremental costs	Incremental LYs	Incremental QALYs	ICER (£/QALY)
UK SoC	[redacted]	1.32	[redacted]	-	-	-	-
Amivantamab	[redacted]	2.21	[redacted]	[redacted]	0.88	[redacted]	£40,246

Sources: CS Table 64 and Table 65[4]

CS = company submission; ICER = incremental cost-effectiveness ratio; LYs = life years; QALY = quality-adjusted life year; SoC = standard of care; UK = United Kingdom

ERG comment: The main concern of the ERG relates to the lack of a fully incremental analysis for all relevant comparators in the comparator basket. Although the ERG acknowledges the limitation of small sample sizes of patients receiving individual treatments in the RWE sources, a fully incremental analysis of all relevant comparators in the comparator basket would be informative (as was requested in the clarification letter, but not provided) to address the uncertainty of assuming average effectiveness and costs of a basket of treatments.

5.2 Company's sensitivity analyses

The company performed and presented the results of PSA, DSA as well as scenario analyses. The parameters that had the greatest effect on the ICER based on the company's DSA were:

- PFS KM curve for the UK SoC arm
- Drug costs in subsequent cycles for the amivantamab arm
- Health state utilities for PFS and PPS

The CS scenarios that have the greatest impact on the ICER (not including scenarios related to discount rates and time horizon) were:

- UK SoC efficacy based on PHE data (decreased ICER to £25,865)
- Using osimertinib to represent EGFR TKIs (decreased ICER to £31,224)
- Using INV as a measure of progression (increased ICER to £42,249)

ERG comment: The main concern of the ERG related to the fact that the majority ([redacted]) of the incremental QALY gain was accrued post-progression. Upon a request for justification, the company argued that this was in line with the submitted evidence. The company added that UK clinical experts agreed with this judgement as amivantamab offered another line of treatment leading to the list of available treatments becoming exhausted later. The ERG is satisfied with this response.

5.3 Model validation and face validity check

5.3.1 Face validity assessment

The company states that expert clinical input was sought during the development of the cost effectiveness model to ensure that the inputs and assumptions used in the analysis were relevant to UK clinical practice and to validate the clinical plausibility of the outcomes predicted by the model. Moreover, feedback was obtained in two advisory boards and in total, input was gathered from seven UK clinical experts. The CS provides limited information on these clinical experts or advisory boards (i.e., how issues were presented, what topics were discussed, whether there was disagreement).

5.3.2 Technical verification

In the CS, it is stated that the model programming was checked by an analyst who was not involved in the original development of the model. Moreover, the company reports to have held a model challenge session with health economic experts to gain insights and advice regarding the most appropriate assumptions and inputs to consider for the cost effectiveness model. In the CS, it is mentioned that the model was validated "using a validation checklist similar that reported in the published literature". This checklist was not provided in the CS. In response to clarification question B18b, the company indicated that this checklist was based on the TECH-VER checklist. Furthermore, in response to clarification question B18a, the company provided additional information on the stress test checklist used to validate the model.

5.3.3 Comparisons with other technology appraisals

In the CS base case, no cross-validation with other technology appraisals was performed by the company regarding the modelled outcomes (e.g., comparisons of extrapolated PFS or OS curves, QALY gains, or total cost estimates).

In response to clarification question B19, the company provided comparisons with other relevant NICE TAs focused on similar, potentially relevant, diseases. To this extent the company provided a summary of key previous appraisals as per the NICE final scope and NG122 (TA347, TA428, TA484/TA713, TA520 and TA653).

5.3.4 Comparison with external data used to develop the economic model

No external data was used to validate outcomes in the CS base case model. In the CS, it is stated that parametric distributions were selected based on clinical expert input. This selection process did not involve external data.

5.3.5 Comparison with external data not used to develop the economic model

Not performed.

ERG comment: The main concern of the ERG relates to differences between the probabilistic results when running the same model multiple times (without changing model settings). This is likely due to the lack of a fixed random seed in the model PSA, which results in slightly different random draws each time the model runs. When running the model multiple times, the ERG estimates the ICER to fluctuate roughly with £500 to £1,000 per QALY gained.

6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

Table 6.1 summarises the key issues related to the cost effectiveness categorised according to the sources of uncertainty as defined by Grimm et al. 2020[43]:

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

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

Based on all considerations in the preceding Sections of this ERG report, the ERG defined a new base case. This base case included multiple adjustments to the original base case presented in the previous Sections. These adjustments made by the ERG form the ERG base case and were subdivided into three categories (derived from Kaltenthaler 2016):[44]

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

6.1.1 ERG base case

Adjustments made by the ERG, to derive the ERG base case (using the CS base case as starting point) are listed below. Table 6.2 shows how individual adjustments impact the results plus the combined effect of all abovementioned adjustments simultaneously, resulting in the ERG base case. The 'fixing error' adjustments were combined, and the other ERG analyses were performed also incorporating these 'fixing error' adjustments given the ERG considered that the 'fixing error' adjustments corrected unequivocally wrong issues.

6.1.1.1 Fixing errors

There were no errors identified by the ERG.

6.1.1.2 Fixing violations

1. Exclusion of age-adjustment to the health state utilities (Section 4.2.8): In the CS base case, the company did not include an age-adjustment to the health state utilities given the relatively short time horizon of the model. However, the ERG decided to include age-adjustments as it is in line with good modelling practice.

6.1.1.3 Matters of judgement

2. Indirect treatment comparison approach for the comparative effectiveness of amivantamab versus SoC (Section 4.2.6): The comparative effectiveness was explored via covariate adjustment and IPW and propensity score matching (PSM). The ERG decided to opt for two ERG base cases because it remains undecided regarding the best way to determine the comparative effectiveness of amivantamab versus SoC. Hence, the ERG opted for two separate ERG base cases in which ERG base case one was based on the IPW approach and ERG base case two was based on the propensity score matching approach.
3. Implementation of parametric survival curves in SoC arm (Section 4.2.6): In line with the company's scenario analyses, the ERG implemented a Weibull curve for OS and a log-logistic curve for PFS.
4. TTD for amivantamab was informed by the CHRYSALIS trial protocol instead of assuming TTD is equal to PFS (Section 4.2.6): Instead of assuming TTD being equal to PFS, the ERG implemented TTD using an exponential curve informed by CHRYSALIS trial data.

6.1.2 ERG exploratory scenario analyses

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

6.1.2.1 Exploratory scenario analyses

5. Informing health state utilities based on the collected HRQoL data in CHRYSALIS (Section 4.2.8): health state utilities in the economic model were sourced from TA484/TA713 as the number of EQ-5D-5L responses from the CHRYSALIS trial was low at the time of data cut-off. Nevertheless, an ERG scenario informing utilities based on CHRYSALIS data was conducted to assess the impact on the ICER.
6. Assuming TTNT as a proxy for treatment discontinuation in the SoC population (Section 4.2.6): For this scenario, TTNT estimates were used as a proxy for TTD in the SoC arm. For this analysis the generalised gamma model was chosen. For amivantamab, the ERG implemented the exponential model for TTD in its base case (see ERG change 4).

6.1.3 ERG subgroup analyses

No subgroup analyses were performed by the ERG.

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

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base case[b]	Required additional evidence or analyses
Representativeness of the comparator basket effectiveness to UK clinical practice	4.2.4	Bias and indirectness	Exclude EGFR TKIs from comparator basket.	+/-	No	Updated economic model excluding the costs and effects of EGFR TKIs.
Implementation of parametric survival curves instead of KM curves for SoC	4.2.6	Methods	Implement parametric models for survival analyses of OS and PFS in the SoC arm.	+	Yes	N/A
TTD assumed equal to PFS	4.2.6	Methods	Apply parametric survival model to TTD based on CHRYSALIS evidence.	+	Partly	Details of NICE DSU TSD 14 criteria assessment to support TTD curve selection.
Treatment waning	4.2.6	Bias and indirectness	Updated economic model including treatment waning scenarios. Additional evidence that treatment waning would be implicitly captured in the selected curves.	+/-	No	Updated economic model including treatment waning scenarios. Additional evidence that treatment waning would be implicitly captured in the selected curves
Exclusion of age-adjusted health state utilities in the CS base case	4.2.8	Methods	Include age-adjusted health state utilities	+	Yes	N/A
Lack of a fully incremental analysis for all relevant comparators in the comparator basket	5.1	Methods	Fully incremental analysis of all relevant comparators in the comparator basket.	+/-	No	Fully incremental analysis of all relevant comparators in the comparator basket.
Lack of a fixed random seed in model PSA	5.3	Imprecision	Implement fixed random seed to model PSA.	+/-	No	Implement fixed random seed to model PSA.

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

Key issue	Section	Source of uncertainty	Alternative approaches	Expected impact on ICER[a]	Resolved in ERG base case[b]	Required additional evidence or analyses
CS = company submission; DSU = Decision Support Unit; EGFR = epidermal growth factor receptor; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; KM = Kaplan-Meier; N/A = not applicable; NICE = National Institute for Health and Care Excellence; PFS = progression-free survival; PSA = probabilistic sensitivity analysis; SoC = standard of care; TKI = tyrosine kinase inhibitor; TSD = Technical Support Document; TTD = time to treatment discontinuation; UK = United Kingdom						

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

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

Table 6.2: ERG base case 1 (IPW approach) and base case 2 (PSM approach) (with PAS)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
CS deterministic base case					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	39,764
UK SoC	[redacted]	[redacted]			
Fixing violation (1-Exclusion of age-adjustment to the health state utilities)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	40,293
UK SoC	[redacted]	[redacted]			
Matter of judgement (2-Use of PSM approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	45,790
UK SoC	[redacted]	[redacted]			
Matter of judgement (3-Implementation of parametric survival curves in SoC arm)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	41,401
UK SoC	[redacted]	[redacted]			
Matter of judgement (4-Time to treatment discontinuation informed by the CHRYSALIS trial protocol)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	55,695
UK SoC	[redacted]	[redacted]			
Deterministic ERG base case 1 (IPW approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	56,799
UK SoC	[redacted]	[redacted]			
Probabilistic ERG base case 1 (IPW approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	54,418
UK SoC	[redacted]	[redacted]			
Deterministic ERG base case 2 (PSM approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	52,185
UK SoC	[redacted]	[redacted]			
Probabilistic ERG base case 2 (PSM approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	49,880
UK SoC	[redacted]	[redacted]			

CS = company submission, ERG = Evidence Review Group, ICER = incremental cost-effectiveness ratio, IPW = inverse probability weighting, PSM = propensity score matching, QALY = quality-adjusted life year, SoC = standard of care; UK = United Kingdom

Table 6.3: Probabilistic scenario analyses (conditional on ERG base case) (with PAS)

Technologies	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER (£/QALY)
ERG base case 1 (IPW approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	54,418
UK SoC	[redacted]	[redacted]			
Scenario analysis base case 1 (5-Health state utilities based on CHRYSALIS HRQoL data)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	58,764
UK SoC	[redacted]	[redacted]			
Scenario analysis base case 1(6-Assuming TTNT as proxy for treatment discontinuation in SoC)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	39,567
UK SoC	[redacted]	[redacted]			
ERG base case 2 (PSM approach)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	49,880
UK SoC	[redacted]	[redacted]			
Scenario analysis base case 2 (5-Health state utilities based on CHRYSALIS HRQoL data)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	53,390
UK SoC	[redacted]	[redacted]			
Scenario analysis base case 2 (6-Assuming TTNT as proxy for treatment discontinuation in SoC)					
Amivantamab	[redacted]	[redacted]	[redacted]	[redacted]	36,169
UK SoC	[redacted]	[redacted]			

ERG = Evidence Review Group; HRQoL = health-related quality of life; ICER = incremental cost-effectiveness ratio; IPW = inverse probability weighting; PSM = propensity score matching; QALY = quality-adjusted life year; TTNT = time to next treatment; SoC standard of care; UK = United Kingdom

6.3 ERG's preferred assumptions

The estimated ERG base case ICERs (probabilistic), based on the ERG preferred assumptions highlighted in Section 5.1, were £54,418 per QALY gained for ERG base case 1 and £49,880 per QALY gained for ERG base case 2. The probabilistic ERG base case 1 and ERG base case 2 analyses indicated cost effectiveness probabilities of 38% and 47% at a willingness to pay threshold of £50,000 per QALY gained. The most influential adjustments were implementing TTD using an exponential curve informed by CHRYSALIS trial data and selecting the PSM indirect treatment comparison approach for the comparative effectiveness of amivantamab versus SoC. The ICER increased most in the scenario analysis assuming TTNT (generalised gamma curve) as a proxy for treatment discontinuation in the SoC population.

6.4 Conclusions of the cost effectiveness Section

The company's cost effectiveness model partly complied with the NICE reference case. Deviations from the NICE reference case related to the exclusion of a fully incremental analysis which would include all UK SoC comparators separately (rather than a "basket" of comparators). The most prominent issues highlighted by the ERG were 1) the representativeness of the comparator basket effectiveness to UK clinical practice; 2) the assumption that treatment would be discontinued when a patient progresses (i.e. assuming TTD equal to PFS in the model); 3) using the KM curves to inform survival analyses for

UK SoC; 4) the exclusion of age-adjustment to the health state utilities; and 5) the company's assumption of a lifelong treatment effect. As a general source of uncertainty, the ERG was undecided regarding the best way to determine the comparative effectiveness of amivantamab versus SoC (i.e., IPW or PSM approach). To this extent, the ERG opted for two ERG base cases in its ERG analyses.

First, due to considerable heterogeneity in treatments due to lack of specifically recommended treatments in the UK, data informing comparator efficacy were derived from a basket of treatments from a US RWE database study. The comparator effectiveness and costs are therefore based on the average clinical effectiveness and weighted average costs across all the treatments included in the comparator basket. It is, however, unclear to the ERG whether this is consistent with UK clinical practice. This is especially important as Exon20ins mutations have been associated with resistance to EGFR TKIs, which are now included in the CS base case. In addition, the results of the indirect treatment comparison excluding TKIs show that the HRs are slightly higher than the base case HRs, indicating that the effectiveness of EGFR TKIs for Exon20ins mutations may indeed be questionable. An updated economic model excluding EGFR TKI therapies from the US RWD could resolve this issue. Moreover, although the ERG acknowledges the limitation of small sample sizes of patients receiving individual treatments in the RWE sources, a fully incremental analysis of all relevant comparators in the comparator basket would be informative to address the uncertainty of assuming average effectiveness of a basket of treatments.

Second, for the estimation of TTD the company assumed that treatment would be discontinued when a patient progresses, setting TTD equal to PFS. The CHRYSALIS trial, however, allowed patients to remain on treatment after disease progression and median TTD ([redacted]) was significantly longer than median PFS([redacted]). This assumption reduces the estimated cost of amivantamab without reducing the estimated effectiveness after progression of amivantamab. Upon request the company implemented a scenario examining the impact of separate TTD curves (i.e., assuming PFS is not necessarily equal to PFS). In its base case, the ERG implemented an exponential curve to model TTD for amivantamab.

Third, OS and PFS in the SoC arm were modelled based on the KM data. The company argued that due to the maturity of the data and all patients reaching the specified end point or being censored within the timeframe of data collection, KM data could be directly implemented rather than fitting a parametric survival model. However, this is not necessarily in line with NICE DSU TSD 14, which states that "*parametric models are likely to represent the preferred method for incorporating survival data into health economic models in the majority of cases*". The ERG decided that the implementation of KM data may introduce overfitting of the modelled survival outcomes. Implementing KM curves biases the SoC treatment effectiveness as patients do not transition smoothly. The ERG therefore implemented parametric models to inform survival analysis of OS and PFS for SoC in its base case.

Fourth, in the CS base case, the company did not include an age-adjustment to the health state utilities. It was argued that given the relatively short time horizon of the model, the impact of age-adjustment on the model results was likely to be marginal and as such, utilities were not age-adjusted. However, in line with good modelling practice, the ERG decided to include age-adjustments in its base case.

Fifth, the ERG considered that the assumption of a lifelong treatment effect may not be warranted and requested that the company explored treatment waning in the model, which the company did not implement. It is unclear to the ERG whether the assumption of a lifelong treatment effect holds true in clinical practice as there is limited evidence provided on the presence (or absence) of treatment waning by the company.

Finally, the ERG decided to opt for two ERG base cases because it remained undecided regarding the most appropriate approach to determine the comparative effectiveness of amivantamab versus SoC. The comparative effectiveness was explored via IPW and PSM approaches. Hence, the ERG opted for two separate ERG base cases in which one was based on the IPW approach (ERG base case 1) and the other one based on the PSM approach (ERG base case 2).

The CS base case probabilistic and deterministic ICERs were £40,246 and £39,764 per QALY gained, respectively. According to the company's model amivantamab is set to influence cost effectiveness by 1) increased PPS, with an increment of 0.526 years (63% of total incremental LYs) in the amivantamab arm (1.349 years) compared with UK SoC (0.823 years); 2) increased PFS, with an increment of 0.314 years (37% of total incremental LYs) in the amivantamab arm (0.818 years) compared with UK SoC (0.504 years); and 3) the higher drug costs, administration costs and post-progression disease management costs.

The two (probabilistic) ERG base case analyses resulted in ICERs of £55,043 per QALY gained (when assuming all ERG changes and the IPW approach to determine comparative effectiveness) and £49,273 per QALY gained (when assuming all ERG changes and the PSM approach to determine comparative effectiveness). The TTD informed by parametric curves based on the CHRYSALIS trial protocol had the biggest impact in the ICER compared to the CS base case. The ICER increased most in the scenario analysis in which health state utilities were based on CHRYSALIS HRQoL data. The ICER decreased most when assuming TTNT as a proxy for treatment discontinuation in SoC. It should be noted that the latter scenario assumes that TTNT is a good approximation to TTD, which is questionable according to the ERG (as discussed in Section 4.2.6. of this report).

In conclusion, there remains uncertainty about the effectiveness and relative effectiveness of amivantamab, which can be at least partly resolved by the company by conducting further analyses (e.g., incorporate the results of the indirect treatment comparison excluding EGFR TKIs in the model, perform a fully incremental analysis, and explore treatment waning). Moreover, the current assessment does not provide an appropriate estimation of the comparators listed in the scope. Therefore, the ERG believes that the CS nor the ERG report contains an unbiased ICER of amivantamab compared with relevant comparators.

7. END OF LIFE

The company states that amivantamab fulfils the first NICE end of life criteria (that the population's life expectancy is less than 24 months) and the second (that the survival benefit of amivantamab exceeds 3 months), see Table 7.1.

Table 7.1: End of life criteria

Criterion	Data available			Section in Document B of the CS
	Comparator	Median OS	Mean undis-counted life years	
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	UK SoC	US RWE: [redacted] CEM: [redacted]	1.38 LYs	B.2.9 (62), B.3.3 (101)
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	Amivantamab	CHRYSALIS: 22.77 (17.48, NE) CEM: [redacted]	2.31 LYs	B.2.6 (48), B.3.3 (101)
	Difference versus amivantamab	US RWE: [redacted] CEM: [redacted]	0.93 LYs	

Based on Table 36 of CS[4]
[a] Median OS is presented based on adjusted comparison with US data (US RWE), unadjusted comparison with UK data (PHE), the output of the cost effectiveness model (CEM) or the CHRYSALIS trial (CHRYSALIS).
CEM = cost effectiveness model; CS = company submission; NE = not evaluable; NHS = National Health Service; OS = overall survival; RWE = real-world evidence; SoC = standard of care; UK = United Kingdom

In Section 5.1 above, the ERG reports figures that also suggest that amivantamab satisfy both end of life criteria. Specifically, the ERG found that the life expectancy of patients without the treatment (SoC) is 1.33 LYs. On this basis, the ERG analysis confirms that criteria that patients do not survive more than 24 months is met. Relatedly, the ERG calculated that patients taking amivantamab have an additional 0.84 LYs, so the second criteria also appears to be met.

ERG comment:

- The ERG confirms that amivantamab fulfils the first NICE end of life criterion (that the population's life expectancy is less than 24 months).
- The ERG notes that there is uncertainty regarding the estimates of clinical effectiveness, and also that the reported values appear to be well over 3 months). Therefore, the ERG considers the 2nd end-of-life also to have been met.

8. REFERENCES

[1] National Institute for Health and Care Excellence. *Dacomitinib for untreated EGFR mutation-positive non-small-cell lung cancer. Technology appraisal guidance 595 [Internet]*. London: National Institute for Health and Care Excellence, 2019 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta595

[2] National Institute for Health and Care Excellence. *Brigatinib for ALK-positive advanced non-small-cell lung cancer that has not been previously treated with an ALK inhibitor. Technology appraisal guidance 670 [Internet]*. London: National Institute for Health and Care Excellence, 2021 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta670

[3] National Institute for Health and Care Excellence. *Entrectinib for treating ROS1-positive advanced non-small-cell lung cancer. Technology appraisal guidance 643 [Internet]*. London: National Institute for Health and Care Excellence, 2020 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta643

[4] Janssen-Cilag. *Amivantamab as a monotherapy for treatment of adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon 20 insertion mutations, whose disease has progressed on or after platinum-based chemotherapy [ID3836]. Document B. Company evidence submission. Single technology appraisal (STA)*: Janssen-Cilag, 2022

[5] National Institute for Health and Care Excellence. *Amivantamab for treating EGFR Exon 20 insertion-positive non-small-cell lung cancer after platinum-based chemotherapy. Final scope [Internet]*. London: National Institute for Health and Care Excellence, 2021 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/gid-ta10729/documents/final-scope-2

[6] US Food & Drug Administration. FDA approves first targeted therapy for subset of non-small cell lung cancer. 2021 [accessed 20.1.22]. Available from: https://www.fda.gov/news-events/press-announcements/fda-approves-first-targeted-therapy-subset-non-small-cell-lung-cancer

[7] European Medicines Agency. *Rybrevant. International non-proprietary name: amivantamab. Procedure No. EMEA/H/C/005454/0000. EPAR - Assessment report [Internet]*. Amsterdam: European Medicines Agency, 2021 [accessed 11.2.22] Available from: https://www.ema.europa.eu/en/documents/assessment-report/rybrevant-epar-public-assessment-report_en.pdf

[8] Janssen-Cilag Ltd. Summary of product characteristics. Rybrevant (amivantamab) 50mg/mL concentrate for solution for infusion [Internet]. electronic medicines compendium (emc), 2021 [accessed 14.2.22]. Available from: https://www.medicines.org.uk/emc/product/13084

[9] Janssen-Cilag. *Amivantamab for treating EGFR Exon 20 insertion-positive non-small-cell lung cancer after platinum-based chemotherapy [ID3836]. Response to request for clarification from the ERG*: Janssen-Cilag, 2022

[10] National Institute for Health and Care Excellence. *Atezolizumab for treating locally advanced or metastatic non-small-cell lung cancer after chemotherapy. Technology appraisal guidance 520 [Internet]*. London: National Institute for Health and Care Excellence, 2018 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta520

[11] National Institute for Health and Care Excellence. *Nivolumab for advanced non-squamous non-small-cell lung cancer after chemotherapy. Technology appraisal guidance 713 [Internet]*. London: National Institute for Health and Care Excellence, 2021 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta713

[12] National Institute for Health and Care Excellence. *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy. Technology appraisal guidance 428 [Internet]*. London: National Institute for Health and Care Excellence, 2017 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta428

[13] National Institute for Health and Care Excellence. *Nintedanib for previously treated locally advanced, metastatic, or locally recurrent non-small-cell lung cancer. Technology appraisal guidance 347 [Internet]*. London: National Institute for Health and Care Excellence, 2015 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta347

[14] Janssen. *Data on File. Clinical expert statement: amivantamab for EGFR Exon20ins NSCLC*: Janssen, 2022

[15] Girard N, Bazhenova L, Minchom A, Ou S, Gadgeel S, Trigo J, et al. MA04.07 Comparative clinical outcomes for patients with NSCLC harboring EGFR exon 20 insertion mutations and common EGFR mutations. *J Thorac Oncol* 2021;16(3):S145-S146.

[16] Remon J, Hendriks LEL, Cardona AF, Besse B. EGFR exon 20 insertions in advanced non-small cell lung cancer: a new history begins. *Cancer Treat Rev* 2020;90:102105.

[17] Morita C, Yoshida T, Shirasawa M, Masuda K, Matsumoto Y, Shinno Y, et al. Clinical characteristics of advanced non-small cell lung cancer patients with EGFR exon 20 insertions. *Sci Rep* 2021;11(1):18762.

[18] Takeda Oncology. *EXA GB LC 0002_Market Access Advisory Board December 2021 [CIC] [Data on file]*: Takeda, 2021

[19] Medicines & Healthcare Products Regulatory Agency. *Approval of application for an Innovation Passport as part of the Innovative Licensing and Access Pathway (ILAP) for amivantamab*: Medicines & Healthcare Products Regulatory Agency, 2021

[20] Vyse S, Huang PH. Targeting EGFR exon 20 insertion mutations in non-small cell lung cancer. *Signal Transduct Target Ther* 2019;4:5.

[21] Janssen-Cilag. *Amivantamab as a monotherapy for treatment of adult patients with locally advanced or metastatic NSCLC with activating EGFR Exon 20 insertion mutations, whose disease has progressed on or after platinum-based chemotherapy [ID3836]. Appendices. Company evidence submission summary for committee. Single technology appraisal (STA)*: Janssen-Cilag, 2022

[22] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline statement. *J Clin Epidemiol* 2016;75:40-6.

[23] Canadian Agency for Drugs and Technologies in Health. *PRESS - Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration (PRESS E&E) [Internet]*. Ottawa: CADTH, 2016 [accessed 11.2.22] Available from: https://www.cadth.ca/resources/finding-evidence/press

[24] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 11.2.22] Available from: http://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx

[25] Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev* 2021;10(1):39.

[26] Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. *Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021) [Internet]*: Cochrane, 2021 [accessed 11.2.22] Available from: https://training.cochrane.org/handbook

[27] Golder S, Peryer G, Loke YK. Overview: comprehensive and carefully constructed strategies are required when conducting searches for adverse effects data. *J Clin Epidemiol* 2019;113:36-43.

[28] Janssen. *Prognostic patient and disease characteristic SLR. Data on File*: Janssen

[29] Janssen Research & Development, LLC. *Clinical Protocol. A phase 1, first-in-human, open-label, dose escalation study of JNJ-61186372, a human bispecific EGFR and cMet antibody, in subjects with advanced non-small cell lung cancer. Protocol 61186372EDI1001; Phase 1. AMENDMENT 9. JNJ-61186372*: Janssen Research & Development, LLC, 2020

[30] National Institute for Health and Care Excellence. *Amivantamab for treating EGFR Exon 20 insertion-positive non-small-cell lung cancer after platinum-based chemotherapy [ID3836]. Clarification letter*. London: National Institute for Health and Care Excellence, 2022

[31] Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.

[32] Faria R, Hernandez Alava M, Manca A, Wailoo AJ. *NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data*. Sheffield: Decision Support Unit, ScHARR, 2015 Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/TSD17-DSU-Observational-data-FINAL.pdf

[33] Scottish Intercollegiate Guidelines Network (SIGN). SIGN search filters [Internet]. Edinburgh: SIGN, [accessed 21.1.22]. Available from: https://www.sign.ac.uk/what-we-do/methodology/search-filters/

[34] Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York, 2009 [accessed 11.2.22] Available from: http://www.york.ac.uk/inst/crd/SysRev/!SSL!/WebHelp/SysRev3.htm

[35] EUnetHTA JA3WP6B2-2 Authoring Team. *Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness. Methodological Guidelines. Version 2.0 [Internet]*. Diemen, The Netherlands: EUnetHTA, 2019 [accessed 22.2.22] Available from: https://www.eunethta.eu/wp-content/uploads/2020/01/EUnetHTA_Guideline_Information_Retrieval_v2-0.pdf

[36] Latimer N. *NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.* Sheffield: Decision Support Unit, ScHARR, 2011 Available from: http://nicedsu.org.uk/wp-content/uploads/2016/03/NICE-DSU-TSD-Survival-analysis.updated-March-2013.v2.pdf

[37] Goss GD, Cobo M, Lu S, Syrigos K, Lee KH, Göker E, et al. Afatinib versus erlotinib as second-line treatment of patients (pts) with advanced lung squamous cell carcinoma (SCC): Final analysis of the global phase III LUX-Lung 8 (LL8) trial. *Ann Oncol* 2018;29.

[38] Mok TS, Wu YL, Ahn MJ, Garassino MC, Kim HR, Ramalingam SS, et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N Engl J Med* 2017;376(7):629-40.

[39] National Institute for Health and Care Excellence. *Nivolumab for previously treated non-squamous non-small-cell lung cancer. Technology appraisal guidance 484 [Internet]*. London: National Institute for Health and Care Excellence, 2017 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta484

[40] Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84.

[41] Tolley K, Goad C, Yi Y, Maroudas P, Haiderali A, Thompson G. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. *Eur J Health Econ* 2013;14(5):749-59.

[42] National Institute for Health and Care Excellence. *Osimertinib for treating EGFR T790M mutation-positive advanced non-small-cell lung cancer. Technology appraisal guidance 653 [Internet]*. London: National Institute for Health and Care Excellence, 2020 [accessed 11.2.22] Available from: https://www.nice.org.uk/guidance/ta653

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

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