in collaboration with:

university of groningen

University Medical Center Groningen

Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]

Produced by Kleijnen Systematic Reviews (KSR) Ltd in collaboration with University Medical Center (UMC) Groningen, the Netherlands

Authors
Robert Wolff, Managing Director, KSR Ltd, United Kingdom (UK)
Thea van Asselt, Health Economist, UMC Groningen, the Netherlands
Sajad Emamipour, Health Economist, UMC Groningen, the Netherlands
Simon van der Pol, Health Economist, UMC Groningen, the Netherlands
Maarten Postma, Health Economist, UMC Groningen, the Netherlands
Annette Chalker, Systematic Reviewer, KSR Ltd, UK
Pawel Posadzki, Reviews Manager, KSR Ltd, UK
Charlotte Ahmadu, Health Economist, KSR Ltd, UK
Nigel Armstrong, Health Economics Manager, KSR Ltd, UK
Sean Harrison, Statistician, KSR Ltd, UK
Shelley de Kock, Information Specialist, KSR Ltd, UK
Jos Kleijnen, Director, KSR Ltd, UK

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

Date completed 03/09/2021

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number 13/51/34.

Declared competing interests of the authors:
None.

Acknowledgements
[redacted]

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:
Wolff R, van Asselt T, Emamipour S, van der Pol S, Postma M, Chalker A, Posadzki P, Ahmadu C, Armstrong N, Harrison S, de Kock S, Kleijnen J. Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: a Single Technology Assessment. York: Kleijnen Systematic Reviews Ltd, 2021.

Contributions of authors
Robert Wolff 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. Thea van Asselt acted as health economic project lead, critiqued the company's economic evaluation, and contributed to the writing of the report. Sajad Emamipour, Simon van der Pol, Maarten Postma, 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. Annette Chalker and Pawel Posadzki acted as systematic reviewers, critiqued the clinical effectiveness methods and evidence, and contributed to the writing of the report. Sean Harrison acted as statistician, critiqued the analyses in the company's submission and contributed to the writing of the report. Shelley de Kock critiqued the search methods in the submission and contributed to the writing of the report. 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 of Cancer Research
ACP	American College of Physicians
AE	Adverse event
AIC	Akaike Information Criterion
AiC	Academic in confidence
ALK	Anaplastic lymphoma kinase
ALT	Alanine aminotransferase
ANZCTR	Australian New Zealand Clinical Trials Registry
ASCO	American Society of Clinical Oncology
AST	Aspartate aminotransferase
ATE	Average treatment effect
ATT	Average treatment effect of the treated
BIC	Bayesian information criterion
BICR	Blinded independent central review
BNF	British National Formulary
BRAF	B-Raf Proto-oncogene
BSC	Best supportive care
CDF	Cancer Drugs Fund
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
CI	Confidence interval
CiC	Commercial in confidence
CMR	Cochrane Methodology Register
CRD	Centre for Reviews and Dissemination
CS	Company submission
CSR	Clinical study report
CT	Computerised tomography
CTCAE	Common Terminology Criteria for Adverse Events
DARE	Database of Abstracts of Reviews of Effects
DNA	Deoxyribonucleic acid
DSA	Deterministic sensitivity analysis
DSU	Decision Support Unit
ECOG	Eastern Cooperative Oncology Group
EED	Economic Evaluations Database
EGFR	Epidermal growth factor receptor
ELCC	European Lung Cancer Congress
eMIT	Electronic market information tool
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D	European Quality of Life-5 Dimensions
EQ-5D-5L	European Quality of Life-5 Dimensions-5 Levels
ERG	Evidence Review Group
ESMO	European Society for Medical Oncology
ESS	Effective sample size
EU CTR	European Clinical Trials Register
FDA	Food and Drug Administration
FE	Fixing errors
FV	Fixing violations
G12C	G12C amino acid substitution
GID	Guideline in development
HCHS	Hospital and Community Health Services
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment

IASLC	International Association for the Study of Lung Cancer
ICTRP	International Clinical Trials Registry Platform
ICER	Incremental cost effectiveness ratio
IPW	Inverse probability weighting
ISRCTN	International Standard Randomized Controlled Trial Number
ITC	Indirect treatment comparison
IV	Intravenous
KM	Kaplan-Meier
KRAS	Kirsten rat sarcoma viral oncogene homolog
KSR	Kleijnen Systematic Reviews
LYG	Life years gained
MAIC	Matching adjusted indirect comparison
MedDRA	Medical Dictionary for Regulatory Activities
MEK	Mitogen activated protein kinase
MeSH	Medical subject headings
MHRA	Medicines and Healthcare Products Regulatory Agency
MJ	Matters of judgement
MMRM	Mixed models with repeated measures
mRCT	metaRegister of Controlled
MRI	Magnetic resonance imaging
MTA	Multiple technology appraisal
N	Number of participants in the analysis set
n	Number of participants with observed data
N/A	Not applicable
NCI	National Cancer Institute
NHS	National Health Service
NHSCII	National Health Service Cost Inflation Index inflation indices
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NIHR	National Institute for Health Research
NMA	Network meta-analysis
NR	Not reported
NSCLC	Non-small cell lung cancer
NTRK	Neurotrophic tyrosine kinase
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PD-1	Programmed cell death-1
PD-L1	Programmed death-ligand 1
PF	Progression free
PFS	Progression-free survival
PP	Post progression
PPS	Post progression survival
PRO	Patient-reported outcome
PSA	Probabilistic sensitivity analysis
PSM	Partitioned survival model
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
PSWA	Propensity score weighted analysis
QALY	Quality adjusted life year
QLQ	Quality of Life Questionnaire
RA	Regression adjustment
RCT	Randomised controlled trial
RDI	Relative dose intensity
RECIST	Response Evaluation Criteria in Solid Tumours

RoB	Risk of bias
ROBINS-I	Risk Of Bias in Non-randomised Studies of Interventions
ROS	Proto-oncogene tyrosine-protein kinase
SAE	Serious adverse event
SD	Standard deviation
SLR	Systematic literature review
STM	State transition model
TA	Technology appraisal
TEAE	Treatment-emergent adverse events
TEW	Treatment effect waning
TRAE	Treatment related adverse events
TK	Tyrosine kinase
TSD	Technical support document
TTD	Time to treatment discontinuation
UK	United Kingdom
UKCCCR	United Kingdom Coordinating Committee on Cancer Research
UMC	University Medical Center
VAS	Visual analogue scale
VEGFR	Vascular endothelial growth factor receptor
WCLC	World Conference on Lung Cancer
WHO	World Health Organization
WTP	Willingness-to-pay

Table of Contents

Abbreviations **3**

Table of Tables **8**

Table of Figures **11**

1. EXECUTIVE SUMMARY **12**

1.1 Overview of the ERG's key issues 12
1.2 Overview of key model outcomes 13
1.3 The decision problem: summary of the ERG's key issues 13
1.4 The clinical effectiveness evidence: summary of the ERG's key issues 14
1.5 The cost effectiveness evidence: summary of the ERG's key issues 17
1.6 Summary of the ERG's view 20

2. CRITIQUE OF COMPANY'S DEFINITION OF DECISION PROBLEM **25**

2.1 Population 31
2.1.1 Previous treatment 31
2.1.2 Population in CodeBreaK100 31
2.1.3 Generalisability of trial population 32
2.2 Intervention 33
2.3 Comparators 33
2.4 Outcomes 34
2.5 Other relevant factors 34

3. CLINICAL EFFECTIVENESS **35**

3.1 Critique of the methods of review(s) 35
3.1.1 Searches 35
3.1.2 Inclusion criteria 38
3.1.3 Critique of data extraction 40
3.1.4 Quality assessment 40
3.1.5 Evidence synthesis 41
3.2 Critique of trials of the technology of interest, their analysis and interpretation (and any standard meta-analyses of these) 41
3.2.1 Design of CodeBreaK100 trial 41
3.2.2 Baseline characteristics of CodeBreaK100 trial 42
3.2.3 Quality of CodeBreaK100 trial 43
3.2.4. Results of CodeBreaK100 trial 45
3.3 Critique of trials identified and included in the indirect comparison and/or multiple treatment comparison 51
3.4 Critique of the indirect comparison and/or multiple treatment comparison 55
3.4.1 Matching of prognostic patient characteristics 55
3.4.2 PSWA using Flatiron 58
3.4.3 Results of indirect comparison 61
3.5 Additional work on clinical effectiveness undertaken by the ERG 64
3.6 Conclusions of the clinical effectiveness Section 64

4. COST EFFECTIVENESS **65**

4.1 ERG comment on company's review of cost effectiveness evidence 65

4.1.1 Searches performed for cost effectiveness Section 65
4.1.2 Inclusion/exclusion criteria 66
4.1.3 Conclusions of the cost effectiveness review 67
4.2 Summary and critique of company's submitted economic evaluation by the ERG 68
4.2.1 NICE reference case checklist 68
4.2.2 Model structure 69
4.2.3 Population 70
4.2.4 Interventions and comparators 71
4.2.5 Perspective, time horizon and discounting 73
4.2.6 Treatment effectiveness and extrapolation 73
4.2.7 Adverse events 78
4.2.8 Health-related quality of life 79
4.2.9 Resources and costs 83
5. COST EFFECTIVENESS RESULTS 87
5.1 Company's cost effectiveness results 87
5.2 Company's sensitivity analyses 87
5.3 Model validation and face validity check 89
6. EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES 91
6.1 Exploratory and sensitivity analyses undertaken by the ERG 91
6.1.1 Explanation of ERG base-case 91
6.1.2 ERG exploratory scenario analyses 92
6.1.3 ERG subgroup analyses 92
6.2 Impact on the ICER of additional clinical and economic analyses undertaken by the ERG 95
6.2.1 The results of deterministic ERG preferred base case scenario 95
6.2.2 The results of ERG sensitivity and scenario analyses 97
6.3 ERG's preferred assumptions 102
6.4 Conclusions of the cost effectiveness Section 102
7. END OF LIFE 104
8. REFERENCES 105

Table of Tables

Table 1.1: Summary of key issues 12

Table 1.2: Key issue 1. Population narrower than NICE scope 13

Table 1.3: Key issue 2. Generalisability / lack of UK participants 14

Table 1.4: Key issue 3. High risk of bias of CodeBreaK100 15

Table 1.5: Key issue 4. High number of serious adverse events observed in CodeBreaK100 15

Table 1.6: Key issue 5. Validity of ITC without a common comparator 16

Table 1.7: Key issue 6. Partitioned Survival Model structure not validated or justified 17

Table 1.8: Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in 2nd line 17

Table 1.9: Key issue 8. Docetaxel plus nintedanib modelling approach leading to worse survival 17

Table 1.10: Key issue 9. No waning of treatment effect 18

Table 1.11: Key issue 10. TTD modelling approach inconsistent with OS and PFS modelling 18

Table 1.12: Key issue 11. Time-to-death utilities do not seem well-informed 19

Table 1.13: Key issue 12. Disutility for IV administration not well justified 19

Table 1.14: Key issue 13. Relative dose intensity and wastage assumption not justified 19

Table 1.15: ERG base-case adjustments (comparator: docetaxel) 20

Table 1.16: ERG base-case adjustments (comparator: docetaxel + nintedanib) 21

Table 1.17: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel) 22

Table 1.18: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel + nintedanib) 23

Table 2.1: Statement of the decision problem (as presented by the company) 25

Table 2.2: Key issue 1. Population narrower than NICE scope 32

Table 2.3: Key issue 2. Generalisability / lack of UK participants 32

Table 3.1: A summary of sources searches to identify trial data 35

Table 3.2: Eligibility criteria used in search strategy for RCT and non-RCT evidence 38

Table 3.3: CodeBreaK100: study design 41

Table 3.4: Baseline characteristics of subjects in CodeBreak100, phase 2 42

Table 3.5: Quality assessment of CodeBreaK100 using the ROBINS-I tool 44

Table 3.6: Key issue 3. High risk of bias of CodeBreaK100 45

Table 3.7: Efficacy in CodeBreaK100 by adenocarcinoma histology (15 March 2021 data cut, post hoc analysis) 45

Table 3.8: Summary of overall adverse events in NSCLC subjects in CodeBreaK100 47

Table 3.9: Treatment-emergent adverse events of any severity occurring in >10% NSCLC patients in the CodeBreaK100 trial 48

Table 3.10: Treatment-related adverse events occurring in >5% of NSCLC subjects in CodeBreaK100 48

Table 3.11: Key issue 4. High number of serious adverse events observed in CodeBreaK100 49

Table 3.12: Overview of study designs of CodeBreaK100, SELECT-1 and LUME-Lung 1 52

Table 3.13: Comparison of baseline characteristics in CodeBreaK100, SELECT-1 and LUME-Lung 1 trials 53

Table 3.14: Key issue 5. Validity of ITC 60

Table 3.15: Results of MAIC for primary comparison of sotorasib vs docetaxel monotherapy 61

Table 3.16: Results of supplementary primary comparison using propensity score weighting analysis 63

Table 3.17: Results of secondary comparison implemented in the economic model 64

Table 4.1: A summary of the sources to identify cost effectiveness studies 65

Table 4.2: Eligibility criteria for the systematic literature reviews 66

Table 4.3: NICE reference case checklist 68

Table 4.4: Key issue 6. Partitioned Survival Model structure not validated or justified 70

Table 4.5: Key baseline patient characteristics of CodeBreaK100 used in the economic model 71

Table 4.6: Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in 2nd line 73

Table 4.7: Key issue 8. Docetaxel plus nintedanib modelling approach leading to worse survival 77

Table 4.8: Key issue 9. No waning of treatment effect 77

Table 4.9: Key issue 10. TTD modelling approach 78

Table 4.10: Datasets used for HRQoL analysis 79

Table 4.11: Health-state utility values - used in sensitivity analysis 80

Table 4.12: Time-to-death utilities - used in CS base case 80

Table 4.13: Key issue 11. Time-to-death utilities do not seem well-informed 82

Table 4.14: Key issue 12. Disutility for IV administration not well justified 82

Table 4.15: Unit drug costs 83

Table 4.16: Disease management and terminal care costs 84

Table 4.17: Subsequent treatment costs 85

Table 4.18: Treatment mix for 44 patients receiving subsequent treatments 85

Table 4.19: Key issue 13. Relative dose intensity and wastage assumption not justified 86

Table 5.1: Deterministic base-case results: sotorasib vs. docetaxel 87

Table 5.2: Disaggregated results by using Flatiron and CodeBreaK100 90

Table 6.1: Overview of key issues related to the cost effectiveness (conditional on fixing errors highlighted in Section 5.1) 93

Table 6.2: ERG base-case adjustments (comparator: docetaxel) 95

Table 6.3: ERG base-case adjustments (comparator: docetaxel + nintedanib) 96

Table 6.4: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel) 98

Table 6.5: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel + nintedanib) 98

Table 7.1: End of life criteria 104

Table of Figures

Figure 3.1: Kaplan-Meier plot of overall survival (safety analysis set) 46

Figure 3.2: Kaplan-Meier plot of progression-free survival 47

Figure 3.3: Waterfall plot of best tumour shrinkage 50

Figure 3.4: Swimmer plot of duration and time to response 51

Figure 3.5: Kaplan-Meier plot for primary MAIC analysis of OS for sotorasib and docetaxel monotherapy 62

Figure 3.6: Kaplan-Meier plot for primary MAIC analysis of PFS for sotorasib and docetaxel monotherapy 63

Figure 4.1: Model structure 69

Figure 4.2: Modelled OS curves taken from the economic model 76

Figure 4.3: Reported OS Kaplan-Meier plot from LUME-Lung-1 trial for nintedanib plus docetaxel (red line) versus placebo plus docetaxel (blue line) 76

Figure 5.1: The cost effectiveness acceptability curve for sotorasib versus docetaxel 89

Figure 6.1: Cost effectiveness plane for ERG base-case (Comparator: docetaxel) 100

Figure 6.2: Cost effectiveness acceptability curve for ERG base-case (comparator: docetaxel) 100

Figure 6.3: Cost effectiveness plane for ERG base-case (Comparator: docetaxel + nintedanib) 101

Figure 6.4: Cost effectiveness acceptability curve for ERG base-case (comparator: docetaxel + nintedanib) 101

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 relates to the clinical effectiveness, and Section 1.5 issues related 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 provides a summary of the key issues identified by the ERG.

Table 1.1: Summary of key issues

Issue	Summary of issue	Report Sections
1	Population narrower than NICE scope	2.1
2	Generalisability / lack of UK participants	2.1.3, 3.2.1
3	High risk of bias of CodeBreaK100	3.2.3
4	High number of serious adverse events observed in CodeBreaK100	3.2.4.5
5	Validity of ITC without a common comparator	3.3, 3.4
6	Partitioned Survival Model structure not validated or justified	4.2.2
7	Exclusion of platinum-based chemotherapy as a comparator in 2nd line	4.2.4
8	Docetaxel plus nintedanib modelling approach leading to worse survival	4.2.6
9	No waning of treatment effect	4.2.6
10	TTD modelling approach inconsistent with OS and PFS modelling	4.2.6
11	Time-to-death utilities do not seem well-informed	4.2.8
12	Disutility for IV administration not well justified	4.2.8
13	Relative dose intensity and wastage assumption not justified	4.2.9

ITC = indirect treatment comparison; IV = intravenous; NICE = National Institute for Health and Care Excellence; TTD = time to treatment discontinuation; UK = United Kingdom

The key differences between the company's preferred assumptions and the ERG's preferred assumptions are a different approach to estimating utility values, a different approach to estimating time to treatment discontinuation (TTD), the incorporation of treatment waning, and, specifically for the secondary comparison, assuming that docetaxel plus nintedanib cannot be worse than docetaxel in terms of overall survival (OS).

1.2 Overview of key model outcomes

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

Overall, the technology is modelled to affect QALYs by:

- Increasing survival, which accrues in progression-free survival (PFS; [redacted] vs. [redacted] months) as well as in post-progression survival (PPS; [redacted] vs. [redacted] months).
- Increasing health-related quality of life (HRQoL) because of longer survival (via time-to-death utilities) and because of the treatment-related disutility for docetaxel.

Overall, the technology is modelled to affect costs by:

- The higher cost of sotorasib compared to docetaxel ([redacted] vs. £17.95).
- Early treatment discontinuation for sotorasib compared to docetaxel.

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

- The hazard ratio applied to PFS to model sotorasib treatment duration (TTD).
- The time to death utility for >6 months prior to death.
- The OS hazard ratio for docetaxel plus nintedanib versus docetaxel (for the secondary comparison only).

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

The decision problem addressed in the company submission (CS) is slightly narrower than that specified in the final scope, see Table 1.2.

Table 1.2: Key issue 1. Population narrower than NICE scope

Report Section	**2.1**
Description of issue and why the ERG has identified it as important	There is a discrepancy of populations 1) defined in the NICE scope, 2) addressed in the CS decision problem, and 3) included in CodeBreaK100, providing the primary clinical trial evidence: 1. Adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC 2. Adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated 3. Adult patients with KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria, and had ECOG performance status of 0 or 1 Of note, the anticipated marketing authorisation is for the "treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemo-therapy and/or anti PD-1/PD-L1 immunotherapy, unless contra-indicated". The ERG would bring this issue to the attention of the committee as it potentially limits the population for which a decision is made.

Report Section	2.1
What alternative approach has the ERG suggested?	Further evidence should be gathered to cover the population defined in the NICE scope.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Further evidence should be gathered to cover the population defined in the NICE scope.
CS = company submission; ECOG = Eastern Cooperative Oncology Group; ERG = Evidence Review Group; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; PD-1 = programmed cell death-1; PD-L1 = programmed death-ligand 1; RECIST = Response Evaluation Criteria in Solid Tumours	

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

Generalisability to England and Wales is unclear due to the lack of centres in the United Kingdom, see Table 1.3.

The ERG assessed the risk of bias of the CodeBreaK100 study, the primary clinical trial evidence, using the ROBINS-I (Risk Of Bias in Non-randomised Studies of Interventions) tool and rated it the risk of bias to be "serious", see Table 1.4.

Furthermore, the ERG would like to highlight the high number of treatment-emergent adverse events (TEAEs) observed in the CodeBreaK100 study, see Table 1.5.

Finally, the ERG has concerns regarding the validity of indirect comparisons performed in the CS, see Table 1.6.

Table 1.3: Key issue 2. Generalisability / lack of UK participants

Report Section	2.1.3, 3.2.1
Description of issue and why the ERG has identified it as important	The participants of the CodeBreaK100 trial were included at 47 centres worldwide which did not include a centre in the UK. The generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of a high proportion of Asian participants (15.1% of the sample).
What alternative approach has the ERG suggested?	Further analyses of countries similar to the UK would be informative.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	Further analyses of countries similar to the UK would be informative.
ERG = Evidence Review Group; UK = United Kingdom	

Table 1.4: Key issue 3. High risk of bias of CodeBreaK100

Report Section	**3.2.3**
Description of issue and why the ERG has identified it as important	Using the ROBINS-I tool, the company rated overall risk of bias of CodeBreaK100 to "low to moderate". However, the ERG re-assessed the study and rated the risk of bias to be "serious". Specifically, domains relating to baseline confounding and measurement of ouctomes were rated as "serious" compared to "low" in the CS.
What alternative approach has the ERG suggested?	Further evidence should aim to minimise the risk of bias
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased
What additional evidence or analyses might help to resolve this key issue?	Further evidence should aim to minimise the risk of bias
CS = company submission; ERG = Evidence Review Group; ROBINS-I = Risk Of Bias in Non-randomised Studies of Interventions	

Table 1.5: Key issue 4. High number of serious adverse events observed in CodeBreaK100

Report Section	**3.2.4.5**
Description of issue and why the ERG has identified it as important	The ERG is concerned with the high number of treatment-emergent adverse events, i.e. 63 patients (50%) with NSCLC experienced serious AEs in the CodeBreaK100 trial. Twenty patients (15.9%) died.
What alternative approach has the ERG suggested?	None. The ERG wants to highlight the issue for the committee.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Potential guidance should reflect this issue.
AE = adverse event; ERG = ERG = Evidence Review Group; NSCLC = non-small cell lung cancer	

Table 1.6: Key issue 5. Validity of ITC without a common comparator

Report Section	3.3, 3.4
Description of issue and why the ERG has identified it as important	The ITC is unanchored i.e. no common comparator. Therefore, there are potentially relevant differences in prognostic factors between the studies included in the ITCs (CodeBreaK100, SELECT-1, LUME-Lung 1), e.g. regarding G12C KRAS mutation status, prior therapies, presence of brain metastases, and factors like sex and smoking history. It is not possible to match for all of these differences which might have an impact on the validity of the findings of any ITC. The company chose a MAIC for their primary analysis of the main comparison with docetaxel, which is particularly prone to bias given lack of identification of all relevant prognostic factors and clinical experts identified factors to be "very important", e.g. brain metastases and disease stage at baseline. However, these, alongside G12C mutation status, were not considered for the MAIC comparing CodeBreaK100 and SELECT 1. Also, because only summary statistics were available from SELECT-1, the CodeBreaK 100 had to be adjusted to match the SELECT-1 population. The company also conducted a supplementary analysis using the Flatiron study, which, using a method of adjustment, referred to as PSWA that appears to involve IPW, allowed the comparator data to match the CodeBreaK 100 population. A richer set of individual patient data also afforded a greater number of potential prognostic factors. In addition to the underlying uncertainty introduced by an indirect comparison of treatments (compared to a direct comparison), the differences between studies, the choice of baseline variables for matching, the choice of underlying data source and adjustment method can be questioned, and the ERG would have liked to see further analyses.
What alternative approach has the ERG suggested?	1. For the MAIC, an analysis with mutation status as covariate could be informative 2. For the PSWA, methods other than IPW, such as RA or doubly robust (RA plus IPW), could have been employed and so scenario analyses using these methods could be informative 3. For the PSWA, limiting to the docetaxel only population could be informative 4. In principle, evidence directly comparing treatments would provide more robust evidence.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	See suggestions above.

ERG = Evidence Review Group; G12C = G12C amino acid substitution; IPW = inverse probability weighting; ITC = indirect treatment comparison; KRAS = Kirsten rat sarcoma viral oncogene homolog; MAIC = matching adjusted indirect comparison; PSWA = propensity score weighted analysis; RA = regression adjustment

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 key issues in the cost effectiveness evidence are discussed in the Tables below.

Table 1.7: Key issue 6. Partitioned Survival Model structure not validated or justified

Report Section	**4.2.2**
Description of issue and why the ERG has identified it as important	The company used a partitioned survival model without elaborate justification and without an accompanying scenario implementing an STM to validate the results
What alternative approach has the ERG suggested?	The ERG did not suggest an alternative approach other than the STM
What is the expected effect on the cost effectiveness estimates?	The expected effect cannot be predicted
What additional evidence or analyses might help to resolve this key issue?	The ERG recognises that it is difficult and intensive to provide results from a model with an alternative structure.
ERG = Evidence Review Group; STM = state transition model	

Table 1.8: Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in 2nd line

Report Section	**4.2.4**
Description of issue and why the ERG has identified it as important	Compared to the final scope for this appraisal, platinum-based chemotherapy is excluded, while it is considered a relevant comparator in 2nd line for those that have received immunotherapy only in 1st line. According to clinical expert opinion, this concerns about 40% of the patient population in the scope: a very significant minority
What alternative approach has the ERG suggested?	The ERG has no alternative approach as adding the comparator to the model would require structural and substantial changes which are outside the scope of work for the ERG.
What is the expected effect on the cost effectiveness estimates?	Could potentially have a substantial impact on the cost effectiveness, direction unknown.
What additional evidence or analyses might help to resolve this key issue?	Implementing platinum-based chemotherapy in the model as an additional comparator would help to resolve the issue and reduce uncertainty.
ERG = Evidence Review Group	

Table 1.9: Key issue 8. Docetaxel plus nintedanib modelling approach leading to worse survival

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The indirect way of estimating OS and PFS for the secondary comparator docetaxel plus nintedanib leads to worse survival for

	docetaxel plus nintedanib compared to docetaxel plus placebo in the first six months of the OS curve.
What alternative approach has the ERG suggested?	The ERG prefers to assume that the HR for docetaxel plus nintedanib versus docetaxel plus placebo cannot go above 1.
What is the expected effect on the cost effectiveness estimates?	Lowering the HR for docetaxel plus nintedanib versus docetaxel plus placebo will increase the ICER for sotorasib versus docetaxel plus nintedanib.
What additional evidence or analyses might help to resolve this key issue?	Direct evidence for this comparison.
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio; OS = overall survival; PFS = progression-free survival	

Table 1.10: Key issue 9. No waning of treatment effect

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The company's assumption of continued effect of sotorasib does not seem justified and is difficult to maintain given immature evidence.
What alternative approach has the ERG suggested?	The ERG suggested to start waning of the treatment effect at the 2-year timepoint and have it gradually decreased to an HR of 1 over a period of 5 years (with exploratory scenario analyses for 3 and 7 years).
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on lasting treatment effect.
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio	

Table 1.11: Key issue 10. TTD modelling approach inconsistent with OS and PFS modelling

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The TTD was modelled by applying a hazard ratio to PFS from CodeBreaK100. The ERG feels it would have been more consistent to model the TTD in the same way that OS and PFS were modelled, fitting a parametric curve on TTD data using weights based on the MAIC.
What alternative approach has the ERG suggested?	The ERG suggested to use the company's alternative approach, based on the MAIC, in the base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on observed treatment duration in sotorasib and comparator arms

Report Section	4.2.6
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; MAIC = matching adjusted indirect comparison; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation	

Table 1.12: Key issue 11. Time-to-death utilities do not seem well-informed

Report Section	4.2.8
Description of issue and why the ERG has identified it as important	The time to death utilities which the company used in the base-case did not seem well-informed. The data underlying the estimates were sparse, and increasingly so for the closer to death states.
What alternative approach has the ERG suggested?	The ERG suggested to use utilities based on disease progression as base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Fully specified models using also AN02 dataset should be provided to see which approach is most appropriate. But given that even AN02 probably has many missing data this may still not be ideal.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio	

Table 1.13: Key issue 12. Disutility for IV administration not well justified

Report Section	4.2.8
Description of issue and why the ERG has identified it as important	A disutility for IV administration of docetaxel is applied without sufficient justification for the size of the disutility or the exclusion of the potential disutility for taking eight tablets of sotorasib daily.
What alternative approach has the ERG suggested?	The ERG suggested to exclude the IV disutility in the base-case
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Comparative evidence on (observed) health state utilities in sotorasib and comparator arms could resolve this
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratios; IV = intravenous	

Table 1.14: Key issue 13. Relative dose intensity and wastage assumption not justified

Report Section	4.2.9
Description of issue and why the ERG has identified it as important	In their base-case, the company assumed a lower RDI for sotorasib than for comparators, which was not justified. The company also assumed zero wastage for sotorasib, which the ERG also considered not justified.

Report Section	4.2.9
What alternative approach has the ERG suggested?	The ERG proposed to take the average RDI as base-case, and to include wastage based on opened packs.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	For the wastage, the company would have to make a convincing case that opened packs, when not used, would be returned for usage by other patients, i.e. a specific program would have to be in place.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratios; RDI = relative dose intensity	

1.6 Summary of the ERG's view

In conclusion, cost effectiveness estimates of sotorasib compared with docetaxel and with docetaxel plus nintedanib are subject to considerable uncertainty, mainly because of immaturity of data and lack of comparative evidence in various areas. Even when all the ERG preferred assumptions were implemented in the model, uncertainty remained on a number of issues, such as whether all relevant comparators were included in the analysis, treatment duration and long-term efficacy of sotorasib, and comparative HRQoL values. The comparison for docetaxel plus nintedanib is potentially more heavily biased even because of the indirectness of the two-step approach to model OS and PFS, see Tables 1.15 to 1.18.

Table 1.15: ERG base-case adjustments (comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	25,932	0.59	43,660
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	26,369	0.59	44,394
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	26,429	0.59	44,496
MJ 4: Including drug wastage (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	27,552	0.59	46,387
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	25,932	0.55	47,208
MJ 6: Subsequent treatments based on alternative distribution					
Docetaxel	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	26,031	0.59	43,825
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	25,932	0.58	44,339
MJ 8: gradual waning of treatment effect over 5 years, starting at 2-year timepoint (key issue 9)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	25,788	0.53	48,332
ERG base-case					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,466	0.49	58,415
Based on CS updated model CS = company submission; ERG = Evidence Review Group; FV = fixing violations; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation					

Table 1.16: ERG base-case adjustments (comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,699	0.47	33,628
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,297	0.47	34,909
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,195	0.47	34,692
MJ 4: Including drug wastage (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,186	0.47	34,673
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,699	0.44	35,990
MJ 6: Subsequent treatment based on alternative distribution					
Docetaxel + nintedanib	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	15,797	0.47	33,839
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,699	0.46	34,087
MJ 8: gradual waning of treatment effect over 5 years, starting at 2-year timepoint (key issue 9)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,697	0.47	33,618
MJ 9: Assuming HR of 1 for OS for nintedanib for the first period (key issue 8)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,386	0.34	44,969
ERG base-case					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,012	0.33	52,051
Based on CS updated model CS = company submission; ERG = Evidence Review Group; FV = fixing violations; HR = hazard ratio; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; OS = overall survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation					

Table 1.17: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	27,976	0.49	57,567
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,466	0.49	58,444
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,715	0.49	58,986
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	29,635	0.49	60,809

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,419	0.47	60,428
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,497	0.50	57,206

Based on CS updated model

AE = adverse event; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; PFS = progression free survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

Table 1.18: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,664	0.33	50,249
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,012	0.33	51,874
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,214	0.33	52,733
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,244	0.33	52,851
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,010	0.33	52,179
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,012	0.33	52,074

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG scenario 6: Assuming constant HR of OS and PFS for nintedanib from 2nd period onwards					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,059	0.34	49,664
Based on CS updated model AE = adverse event; CS = company submission; ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio; OS = overall survival; PFS = progression free survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year					

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 previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC	Adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immuno-therapy, unless contra-indicated	Patient population in the CodeBreaK100 trial included KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria, and had ECOG performance status of 0 or 1.	The population is slightly narrower than population outlined in NICE scope, see Section 2.1 for details.
Intervention	Sotorasib	Sotorasib (LUMYKRAS™) administered orally at a dose of 960 mg (given as 8x 120 mg tablets) once daily until disease progression or unacceptable toxicity	N/A – in line with the NICE final scope.	The intervention is in line with the NICE scope.
Comparator(s)	**Non-squamous NSCLC:** • pemetrexed with carboplatin with or without pemetrexed maintenance • other platinum doublet chemotherapy with or without pemetrexed maintenance • nintedanib with docetaxel (adenocarcinoma histology) • docetaxel monotherapy	**Primary comparator:** Docetaxel monotherapy **Secondary comparator:** Nintedanib + docetaxel	Docetaxel monotherapy, i.e. the primary comparator– is outside the final scope issued by NICE by not targeting people with KRAS p.G12C mutation	The NICE lung cancer pathway and international clinical guidelines recognise the increasing role of combination immunotherapy and chemotherapy in the first-line setting for NSCLC. It is unclear why results for other comparators are

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	• atezolizumab • nivolumab (subject to ongoing CDF review) • pembrolizumab (PD-L1-expressing tumours) • best supportive care **Squamous NSCLC:** • gemcitabine with carboplatin or cisplatin • vinorelbine with cisplatin or carboplatin • docetaxel monotherapy • pembrolizumab (PD-L1-expressing tumours) • atezolizumab • nivolumab • best supportive care **People with KRAS p.G12C mutation and another driver mutation (including EGFR-TK, ALK or ROS1):** Established clinical management without sotorasib, including: • atezolizumab combination (after EGFR-TK or ALK-targeted therapies) • lorlatinib (after ALK-targeted therapies			unavailable, see Section 2.3 for details.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	• brigatinib (after ALK-targeted therapies) • ceritinib (after ALK-targeted therapies) • osimertinib (EGFR T790M mutation-positive after EGFR-TK targeted therapies) • pemetrexed with carboplatin • platinum doublet chemotherapy with or without pemetrexed maintenance • nintedanib with docetaxel (adenocarcinoma histology) • nivolumab (subject to ongoing CDF review)			
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • time to treatment discontinuation • adverse effects of treatment • health-related quality of life	• overall survival • progression-free survival • response rates • duration of response • adverse effects of treatment • health-related quality of life	The outcomes reported are largely in line with the NICE scope	Time to treatment discontinuation is missing in the CS, see Section 2.4 for details.
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year.	[Not completed in the CS]	[Not completed in the CS]	The approach taken for the economic analysis is largely in line with the reference case.

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
	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 sotorasib is conditional on the presence of KRAS G12C mutation. The economic modelling should include the costs associated with diagnostic testing for KRAS G12C 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'.			No full incremental analysis was performed though, see Table 4.3. The costs associated with diagnostic testing for KRAS G12C mutation was not included in the economic modelling because KRAS testing is routinely commissioned by NHS in NSCLC.
Special considerations including issues related to equity or equality	N/A	• In contrast to NSCLC patients with other oncogenic mutations, patients with advanced or metastatic KRAS p.G12C-mutated NSCLC who have failed prior therapy currently have no targeted therapy options, and very few other effective therapy options. Their prognosis is very poor, with OS significantly less than 2 years.		N/A

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
		• Sotorasib is a highly innovative, first in class therapy for KRAS p.G12C-mutated NSCLC. It provides an effective and tolerable targeted treatment option where previously there was none. It has been designated as a Promising Innovative Medicine via the UK Early Access to Medicines Scheme, and was granted an Innovation Passport under the Innovative Licensing and Access Pathway. UK orphan designation is pending. • Subject to approval, sotorasib is anticipated to be granted conditional marketing authorisation by the MHRA via the Project Orbis regulatory route on the basis of the results of the phase 2 CodeBreaK100 single arm trial. • As sotorasib is the first KRASG12Cinhibitor to progress to licensing by any regulatory authority there is a lack of data specifically in patients with KRAS p.G12C mutated NSCLC for the relevant comparators, or any other agents. • Indirect comparative data using the most robust methods possible indicate that sotorasib is highly effective in achieving clinically meaningfully improvements in PFS and OS by >3 months compared with relevant comparators. • Based on these data, sotorasib provides a step change in therapy for patients with KRAS p.G12C mutated NSCLC and is highly likely to be cost effective under the NICE end of life policy. • Phase 3 data from the CodeBreaK200 RCT are anticipated within the next 2 years. • Sotorasib may therefore be a candidate for the CDF.		

Based on Table 1 of the CS[1]
ALK = anaplastic lymphoma kinase; CDF = Cancer Drugs Fund; CS = company submission; ECOG = Eastern Cooperative Oncology Group; EGFR = epidermal growth factor receptor; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; MHRA = Medicines and Healthcare Products Regulatory Agency; N/A = not applicable; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; OS =

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope	ERG Comment
overall survival; PD-1 = programmed cell death-1; PD-L1 = programmed death-ligand 1; PFS = progression-free survival; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours; ROS = proto-oncogene tyrosine-protein kinase; TK = tyrosine kinase; UK = United Kingdom				

2.1 Population

The NICE scope defined the population of interest as *"adults with previously treated Kirsten rat sarcoma viral oncogene homolog (KRAS) p.G12C mutated, locally advanced or metastatic non-small cell lung cancer (NSCLC)"*.[2]

The company submission (CS) defined the population of interest as *"adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated"*.[1]

ERG comment: The population addressed in the CS is narrower than the population defined in the NICE scope:

1. The CS only considered patients *"previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated"*.[1]
2. The population in CodeBreaK100, providing the primary clinical trial evidence for sotorasib in the CS, is even narrower than that specified in the NICE scope, namely *"KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per the RECIST* [Response Evaluation Criteria In Solid Tumours] *1.1 criteria, and had ECOG* [Eastern Cooperative Oncology Group] *performance status of 0 or 1"*.[1]

2.1.1 Previous treatment

It is unclear why only platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy are considered while other lines of therapy are not.

2.1.2 Population in CodeBreaK100

The CS did not define the 1-3 prior lines of anticancer therapy. Real-world data show there is a variety of first-line treatment strategies (checkpoint inhibitor ± chemotherapy, platinum + pemetrexed, platinum + taxanes, or other chemotherapy) and a variation in second-line treatment regimens while a proportion of patients also receive third-line treatment.[3]

It is unclear why patients with ECOG performance status of 0 to 2 (on a 5-point scale, with higher numbers indicating greater disability) were eligible in phase I of the CodeBreaK100 trial, whereas phase II of the trial only included phases 0 to 1 (less severe).[4, 5]

Of note, the population in CodeBreaK100 appears to be not only narrower than the NICE scope but also than the anticipated marketing authorisation, e.g. in regards to the ECOG status of included participants. According to the CS, an application for UK marketing authorisation for sotorasib was submitted to the Medicines and Healthcare Products Regulatory Agency (MHRA) in January 2021 with a proposed indication for use as monotherapy for treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated. A conditional licensing approval via the Project Orbis regulatory route in the UK is anticipated [redacted].[1]

In response to the request for clarification, the company stated that *"the exclusion of patients with ECOG PS* [performance status] *2 from the CodeBreaK100 trial should not preclude the use of sotorasib within its licensed indication in such patients in clinical practice. Sotorasib should be an option available to clinicians for use in patients with ECOG PS 2 when clinically relevant"*.[6] However, no evidence was provided to support this statement.

2.1.3 Generalisability of trial population

As discussed in Section 3.2.1, the participants of the CodeBreaK100 trial were included at 47 centres worldwide which did not include a centre in the United Kingdom (UK). The generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of a high proportion of Asian participants (15.1% of the sample; see Table 4 of the CS).[1]

Table 2.2: Key issue 1. Population narrower than NICE scope

Report Section	2.1
Description of issue and why the ERG has identified it as important	There is a discrepancy of populations 1) defined in the NICE scope, 2) addressed in the CS decision problem, and 3) included in CodeBreaK100, providing the primary clinical trial evidence: 1. Adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC 2. Adult patients with KRAS p.G12C mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated 3. Adult patients with KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria, and had ECOG performance status of zero or one Of note, the anticipated marketing authorisation is for the "treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemo-therapy and/or anti PD-1/PD-L1 immunotherapy, unless contra-indicated". The ERG would bring this issue to the attention of the committee as it potentially limits the population for which a decision is made.
What alternative approach has the ERG suggested?	Further evidence should be gathered to cover the population defined in the NICE scope.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Further evidence should be gathered to cover the population defined in the NICE scope.

CS = company submission; ECOG = Eastern Cooperative Oncology Group; ERG = Evidence Review Group; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; NICE = National Institute for Health and Care Excellence; NSCLC = non-small cell lung cancer; PD-1 = programmed cell death-1; PD-L1 = programmed death-ligand 1; RECIST = Response Evaluation Criteria in Solid Tumours

Table 2.3: Key issue 2. Generalisability / lack of UK participants

Report Section	2.1.3, 3.2.1
Description of issue and why the ERG has identified it as important	The participants of the CodeBreaK100 trial were included at 47 centres worldwide which did not include a centre in the UK. The generalisability of participants included in CodeBreaK100 to clinical practice in England and Wales is unclear, e.g. due to inclusion of a high proportion of Asian participants (15.1% of the sample).

Report Section	2.1.3, 3.2.1
What alternative approach has the ERG suggested?	Further analyses of countries similar to the UK would be informative.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	Further analyses of countries similar to the UK would be informative.
ERG = Evidence Review Group; UK = United Kingdom	

2.2 Intervention

The intervention (AMG 510/LUMYKRAS™) is in line with the scope.

Sotorasib is administered orally at a dose of 960 mg (given as 8x 120 mg tablets) once daily until disease progression, no further clinical benefit is expected, unacceptable toxicity, withdrawal of consent, or death.[1] Sotorasib is a small molecule that specifically inhibits KRAS G12C amino acid substitution (G12C) in advanced solid tumours through a unique interaction with the P2 pocket of the switch II region.[5]

ERG comment: Participants in the CodeBreaK100 trial used a combination arm with sotorasib and anti PD-1/L1 or midazolam at phase I.[7] It is not clear how these participants were handled in the analyses given that sotorasib was outlined as monotherapy as per NICE scope.[2] This might have an impact on the results of effectiveness as well as cost effectiveness analyses.

2.3 Comparators

The description of the comparators in the NICE scope includes eight different treatments for non-squamous NSCLC, seven treatments for squamous NSCLC; and atezolizumab combination, lorlatinib, brigatinib, ceritinib, osimertinib, pemetrexed with carboplatin, platinum doublet chemotherapy (with or without pemetrexed maintenance), and established clinical management without sotorasib for people with KRAS p.G12C mutation and another driver mutation (including EGFR-TK, ALK or ROS1), see Table 2.1 and NICE scope.[2]

The CS listed two comparators, docetaxel monotherapy as the primary comparator and nintedanib + docetaxel as the secondary comparator.[1]

In response to the request for clarification, the company confirmed that other comparators have not been considered to be relevant comparators for sotorasib.[6]

ERG comment: The primary comparator selected by the company, docetaxel monotherapy, was listed as a comparator for non-squamous NSCLC in the NICE scope.[2] However, it is outside the NICE scope for people with KRAS p.G12C mutation and another driver mutation (including EGFR-TK, ALK or ROS1).[2]

It should be noted that Peter Clark (The Clatterbridge Cancer Centre NHS Foundation Trust; NHS England Cancer Drugs Fund (CDF) clinical lead) highlighted that *"KRAS 12C mutations are mutually exclusive to other targetable mutations"*.[8]

The company selected nintedanib in combination with docetaxel as the secondary comparator which is in line with NICE technology appraisal (TA) 347 for patients with adenocarcinoma and in line with the NICE scope.[2, 9]

Following advice by Peter Clark, the ERG considers the main comparator to be second-line docetaxel monotherapy and would consider the secondary comparator, nintedanib + docetaxel as a scenario analysis.[8]

2.4 Outcomes

The NICE final scope lists the following outcome measures:

- Overall survival
- Progression-free survival
- Response rates
- Duration of response
- Adverse effects of treatment
- Health-related quality of life.

ERG comment: Most of these outcomes were included in the decision problem addressed in the CS as well as assessed in the CodeBreaK100 trial except time to treatment discontinuation (TTD).[1]

However, as stated in the response for the request for clarification, TTD was used to inform the economic model.[6] However, as discussed in Section 4.2.6 of the report, TTD was based on progression-free survival (PFS) using a hazard ratio (HR).

As detailed in Section 3.2.4.5, the ERG is concerned with the high number of treatment-emergent adverse events (TEAEs).

As detailed in Section 3.2.4.6, health-related quality of life (HRQoL) was only summarised descriptively; and changes from baseline using mixed effects models for repeated measures are tested.

2.5 Other relevant factors

According to the company, sotorasib is highly innovative and has been granted an Innovation Passport under the Innovative Licensing and Access Pathway; and addresses a significant unmet need in patients with *KRAS p.G12C* -mutated NSCLC (Section B.2.12 of the CS).[1] The drug also received accelerated approval by the United States Food and Drug Administration (FDA) on 28 May 2021 under its Real-Time Oncology Review (Section B.1.2 of the CS).[1]

Sotorasib is offered at an undiscounted price of [redacted] per patient per treatment (Table 2 of the CS).[1] The company highlighted that sotorasib may be a candidate for the CDF.

Sotorasib might fulfil the end of life criteria as specified by NICE. However, as discussed in Section 7, the ERG has concern regarding the validity of the data used to inform the second criterion, extension of life of ≥3 months.

According to the company, *"no specific equality considerations are anticipated"* (SectionSection B.1.4 of the CS).[1]

3. CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

3.1.1 Searches

Appendix D of the CS detailed a systematic literature review (SLR) conducted to identify trial data for systemic drug therapies used in the management of patients with KRAS mutant NSCLC.[1]

The SLR search strategy was based on a published SLR by Schulz et al. which was conducted in the pre-treated NSCLC population.[10] As this review set out to include publications reporting outcome data for a KRAS mutant population, any studies identified as relevant by Schulz et al. were included as well as all relevant studies published during or after 2015 as identified by replicating the Schulz et al. strategy.[10]

Searches were run in June 2020 and updated on 26 January 2021. In addition to a search for randomised controlled trials (RCTs), searches were also conducted to identify single arm trials with KRAS mutant NSCLC. These searches were undertaken on 24 July 2019 and updated on 10 March 2021. A summary of sources searched is provided in Table 3.1.

Table 3.1: A summary of sources searches to identify trial data

Resource	Host/Source	Date Ranges	Dates searched
RCT searches			
Electronic Databases			
Embase	Ovid	1980 – present	25 June 2020 26 January 2021
MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions®	Ovid	1946 – present	12 June 2020 26 January 2021
CDSR CENTRAL	Ovid		12 June 2020 26 January 2021
Conference proceedings			
ASCO	https://www.asco.org/	January 2017 – January 2021	
ESMO	http://www.esmo.org/		
IASLC World Congress on Lung Cancer	https://wclc2019.iaslc.org/		
AACR	https://www.aacr.org/Pages/Home.aspx		
Clinical trial registries			
Clinicaltrials.gov	www.clinicaltrials.gov		

Resource	Host/Source	Date Ranges	Dates searched
NCI clinical trial database	https://www.cancer.gov/	January 2017 – January 2021	
UKCCCR Register of Cancer Trials	http://www.ctu.mrc.ac.uk/ukcccr/		
ISRCTN Register	https://www.isrctn.com/		
EORTC	https://www.ukctg.nihr.ac.uk/		
UK Clinical Trials Gateway	https://www.ukctg.nihr.ac.uk/		
mRCT	http://www.isrctn.com/page/mrct		
Searches for single-arm trials			
Electronic databases			
Embase	Ovid	2014 - 2019	24 July 2019
		2019 - 2021	10 March 2021
MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, MEDLINE Daily and MEDLINE	Ovid	2014 – 2019	24 July 2019
		2019 – 2021	10 March 2021
CDSR	Ovid	2014 - 2019	24 July 2019
DARE		2019 – 2021	10 March 2021
CENTRAL			
NHS EED			
HTA Database			
ACP Journal Club			
Conference proceedings			
ASCO	https://www.asco.org/	2017 – 2021	24 July 2019 10 March 2021
ESMO	http://www.esmo.org/		
WCLC	https://wclc20190iaslc.org/		
ELCC	https://www.esmo.org/Conferences/ELCC-2019-European-Lung-Cancer-Congress		
Clinical trials registries			
ClinicalTrials.gov	https://clinicaltrials.gov	January 2017 – January 2021	
NIH	https://www.nih.gov/		
World Health Organization ICTRP	http://www.who.int/ictrp/en/		
ANZCTR	http://www.anzctr.org.au/		
EU CTR	https://www.clinicaltrialsregister.eu/		

AACR = American Association of Cancer Research; ACP = American College of Physicians; ANZCTR = Australian New Zealand Clinical Trials Registry; ASCO = American Society of Clinical Oncology; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials;

Resource	Host/Source	Date Ranges	Dates searched
DARE = Database of Abstracts of Reviews of Effects; EED = Economic Evaluations Database; ELCC = European Lung Cancer Congress; EORTC = European Organization for Research and Treatment of Cancer; ESMO = European Society for Medical Oncology; EU CTR = European Clinical Trials Register; HTA = Heath Technology Assessment; IASLC = International Association for the Study of Lung Cancer; ICTRP = International Clinical Trials Registry Platform; ISRCTN = International Standard Randomized Controlled Trial Number; mRCT = metaRegister of Controlled Trials; NCI = National Cancer Institute; NHS = National Health Service; NIH = National Institutes of Health; RCT = randomised controlled trial; UK = United Kingdom; UKCCCR = United Kingdom Coordinating Committee on Cancer Research; WCLC = World Conference on Lung Cancer			

ERG comment: The CS provided sufficient details for the ERG to appraise the literature searches. A range of databases, conference proceedings and clinical trials registries were searched. Both the original and update searches were overall well conducted and documented, making them transparent and reproducible. A date limit was applied to the searches but this was justified as a previous SLR on pre-treated NSCLC population had been undertaken by Schulz et al.[10] A separate search for single-arm studies was undertaken without an RCT filter to pick up adverse events to any treatments for NSCLC. Searches for single-arm studies were restricted to English language only.

In response to the request for clarification, the company explained that an English language limit had been applied for pragmatic reasons as most high-quality studies are generally published in English.[6] To avoid language bias and to increase precision, the Centre for Reviews and Dissemination (CRD) guidance recommends that English language limits should not be applied at the searching stage.[11] Study design filters were applied to RCT searches but were not appropriately referenced. In response to the request for clarification, the company confirmed that a validated search filter had not been used and that the SLR was built upon the one conducted by Schulz et al.[6, 10] The ERG believes a validated RCT filter would have increased the comprehensiveness of the searches.

The CS reported that searches were modified between databases to account for differences in syntax and thesaurus headings. However, the ERG noticed that the RCT filter applied to MEDLINE searches had not been modified and many of the terms in the RCT filter did not map across automatically. The ERG requested that the company re-run MEDLINE searches with the correct medical subject headings (MeSH) terms to ensure that nothing had been inadvertently missed which the company did.[6] An additional 13 records were identified and screened. Only the population was searched for in both RCT searches and searches for single-arm studies. This seemed appropriate considering the sparsity of the literature.

An RCT filter was applied to searches of CDSR and CENTRAL which are already pre-filtered databases and therefore the use of a filter is considered to be overly restrictive. In response to the request for clarification, the company argued that the additional use of study filters in their experience did not significantly increase the risk of relevant studies being excluded.[6] However, this is against the explicit recommendation of the Cochrane Handbook for Systematic Reviews of Interventions which states that CENTRAL *"aims to contain only reports with study designs possibly relevant for inclusion in Cochrane Reviews, so searches of CENTRAL should not use a trials 'filter' or be limited to human studies"*.[12]

A wide range of conference proceedings and clinical trials registries were searched. Search terms were not provided in the CS but were supplied in response to clarification questions.[6] The ERG was satisfied that the search terms were sufficient. The reference lists of included publications and relevant SLRs and network meta-analyses (NMAs) were scanned for further studies.

3.1.2 Inclusion criteria

The eligibility criteria for RCTs and non-RCTs is presented in Table 3.2. However, it was initially unclear if inclusion screening was completed in duplicate or how consensus was reached. The company clarified that this stage had been completed in duplicate.[6]

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

	Description	Justification
Inclusion criteria		
Population	• Subject had provided informed consent prior to initiation. • Men or women ≥18 years old. • Pathologically documented, locally-advanced or metastatic stage IIIB-IV NSCLC with, KRAS p.G12C mutation or any other KRAS mutation (KRASm)) identified through DNA sequencing. • Subjects must have received (at least) prior standard therapy appropriate for their tumour type and stage of disease, or in the opinion of the investigator would be unlikely to tolerate or derive clinically meaningful benefit from appropriate standard of care therapy. • Subjects were willing to provide archived tumour samples or willing to undergo pre-treatment tumour biopsy (Part 1 Dose Exploration). • Subjects were willing to undergo pre-treatment tumour biopsy. Subjects can be allowed to enrol without undergoing a tumour biopsy upon agreement with Investigator and the Medical Monitor if a tumour biopsy was not feasible. • Measurable or evaluable disease per RECIST 1.1 criteria. • ECOG performance status of ≤2 (phase 1) or ≤1 (phase 2).	N/A
Interventions	• Sotorasib • Any therapies licensed in the United States or European Union for the second or later line treatment of patients with NSCLC • Any anti-cancer drugs, any line of treatment or no treatment	Consistent with final scope
Comparator	Any or none	Consistent with final scope

	Description	Justification
Outcomes	• Objective response rate • Partial response • Complete response • Duration of response • Disease control rate or clinical benefit rate • Treatment duration and dosing • Disease control rate • Time to response • Progression free survival • Progression after next line of therapy (PFS2) • Time to progression • Time to next treatment • Event-free survival • Overall survival • Patient-reported outcomes • HRQoL • All-grade treatment-emergent AEs • Treatment related Grade 3 or 4 AEs • Treatment related SAEs • Tolerability: dose reductions and interruptions, discontinuation (any reason), discontinuation (due to AEs)	Consistent with final scope
Study design	• Prospective randomised controlled trials (for the RCT search) • Non-RCTs, i.e. experimental/interventional, not observational (for the non-RCT search)	Separate searches were conducted for RCTs and non-RCTs. ERG comment: It is unclear why phase I studies were excluded as they comprise useful for safety data. Also, unclear why non-randomised clinical trials were ineligible since the CodeBreaK100 was a non-randomised trial.
Language restrictions	English language only	To reduce number of hits and to identify studies in patient populations relevant to the UK setting

	Description	Justification
Exclusion criteria		
Population	• Subjects with active brain metastases from non-brain tumours • Paediatric and adolescent (<18 years) patients • Patients with cancers other than NSCLC • Early-stage NSCLC patients (Stage<IIIB) • Trials studying safety and efficacy of treatment administered in adjuvant setting • Treatment naïve patients	As specified by final scope
Interventions	• Treatments specifically targeting EGFR/ALK or ROS 1 mutations or other targetable mutation • Radiotherapy or surgery	Not relevant to final scope
Outcomes	Non-clinical outcomes	Not relevant to final scope
Study design	• Non-RCTs (for the RCT search) • RCTs (for the non-RCT search)	Separate searches were conducted for RCTs and non-RCTs
Language restrictions	Abstracts published in non-English language	To reduce number of hits and to identify studies in patient populations relevant to the UK setting

Based on Tables 1, 2, and 6 of Appendix D of the CS[13]
AE = adverse event; ALK = Anaplastic lymphoma kinase; CS = company submission; DNA = deoxyribonucleic acid; ECOG = Eastern Cooperative Oncology Group; EGFR = Epidermal Growth Factor Receptor; ERG = Evidence Review Group; G12C = G12C amino acid substitution; HRQoL = health-related quality of life; KRAS = Kirsten rat sarcoma viral oncogene homolog; N/A = not applicable; NSCLC = non-small cell lung cancer; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours; ROS = proto-oncogene tyrosine-protein kinase; SAE = serious adverse event; UK = United Kingdom

ERG comment: The inclusion criteria noted the exclusion of non-randomised trials, despite the CodeBreaK100 study being of a non-randomised design. In response to the request for clarification, the company stated that the inclusion criteria do not include searches of comparative trials that were not randomised, based on the assumption that few comparative studies are likely to be non-randomised.[6]

3.1.3 Critique of data extraction

Information provided in the CS regarding data extraction was limited. In the response to the request for clarification, the company stated that each stage of the systematic review process was completed in duplicate.[6]

3.1.4 Quality assessment

The critical appraisal of the non-randomised study was completed using the Risk Of Bias in Non-randomised Studies of Interventions (ROBINS-I) tool.[14] The SELECT-1 trial and the LUME-Lung 1 trial were reported to be assessed using the NICE single technology appraisal user guide. However, the

CS noted that aspects of the CRD guidance had been utilised.[11] , In response to the request for clarification, the company stated that this is in line with the NICE STA user guide.[6]

3.1.5 Evidence synthesis

According to Section B.2.8 of the CS, *"no meta analyses have been conducted" "as current efficacy data for sotorasib in the treatment of KRAS p.G12C-mutated NSCLC are based on a phase 2 single-arm trial"*.[1]

ERG comment: The ERG agrees that meta-analysis would not have been helpful.

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

3.2.1 Design of CodeBreaK100 trial

The CodeBreaK100 trial is an ongoing phase 1/2 study, in which the phase 2 portion is a multicentre, non-randomised, open-label study.[1]

The population was comprised of adults with confirmed KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per RECIST 1.1 criteria and had an ECOG performance status of 0 or 1. The trial locations were located in 47 centres, with none of these being based in the United Kingdom.

The intervention was comprised of 960 mg of sotorasib, which is meant to be administered orally once per day without interruption until either disease progression, intolerance, withdrawal of consent, or death. There was no listed comparator. Statistical analyses are shown in Table 3.3.

In response to the request for clarification regarding the *"blinded independent central review"*, the company noted that the blinded independent central review referred to the assessment of response per RECIST 1.1 criteria by central review, rather than investigators.[6]

Table 3.3: CodeBreaK100: study design

Study	CodeBreaK100 (NCT03600883)
Study Design (n)	Single-arm, phase 2 trial conducted in 47 centres (N=126)
Population	Adults with confirmed KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per Response Evaluation Criteria In Solid Tumours (RECIST) 1.1 criteria, and had ECOG performance status of 0 or 1.
Intervention	Sotorasib 960 mg administered orally once per day without interruption (i.e., no planned off-treatment days) until disease progression, intolerance, withdrawal of consent or death.
Comparator	None.
Reported outcomes specified in the decision problem	• Objective response rate assessed by blinded independent central review • Overall survival • Duration of response • Progression-free survival • Response rates • Time to treatment discontinuation • Adverse effects of treatment • Health-related quality of life (EORTC QLQ-C30, QLQ LC13, EQ-5D-5L)

Study	CodeBreaK100 (NCT03600883)
All other reported outcomes	• Disease control • Time to release • 6- and 12-month PFS • 12-month OS • Patient-reported outcomes (NSCLC SAQ, FACT-G, PRO-CTCAE) • PK parameters and biomarkers (not further discussed in the CS)
Duration of study and follow-up	The CodeBreaK100 trial is ongoing. A primary analysis of efficacy, safety and patient-reported outcomes (PROs) data was conducted in September 2020. An updated analysis of efficacy and safety data for regulatory purposes was conducted on 1 December 2020. A phase 3 randomised controlled trial (RCT) comparing sotorasib against standard of care docetaxel in patients with NSCLC is ongoing with first results anticipated in 2022.
Countries	47 centres in the United States, Australia, Austria, Belgium, Canada, France, Germany, Japan, South Korea, and Switzerland.

Based on Table 4 of the CS[1]
CS = company submission; ECOG = Eastern Cooperative Oncology Group; EORTC = European Organisation for Research and Treatment of Cancer; EQ-5D-5L = European Quality of Life-5 Dimensions-5 Levels; KRAS = Kirsten rat sarcoma viral oncogene homolog; NSCLC = non-small lung cancer; PFS = progression-free survival; PRO = patient-reported outcome; QLQ = Quality of Life Questionnaire; RCT = randomised controlled trial; RECIST = Response Evaluation Criteria in Solid Tumours

ERG comment: In response to the request for clarification regarding the generalisability of CodeBreaK100 to the clinical practice in England and Wales, the company stated that *"five UK clinical experts at an Amgen Advisory board considered that the population of patients enrolled in the CodeBreaK100 trial was reflective of patients in UK clinical practice who would meet the anticipated licensed indication"*.[6] However, the ERG wishes to emphasise that the CodeBreaK100 study did not include a single UK centre. Furthermore, at phase I, participants in the CodeBreaK100 trial used a combination of sotorasib and anti PD-1/L1 or midazolam. It is not clear how these participants were handled in the analyses and this can potentially impact on the results of effectiveness as well as cost effectiveness analyses.

3.2.2 Baseline characteristics of CodeBreaK100 trial

The baseline characteristics of the CodeBreaK100 trial are presented in Table 3.4. The participants in the phase 2 study were not randomised. The mean age of the participants was 62.9 years with a range of 37 to 80 years. The majority of the participants were white while they were evenly split among male and females. The CodeBreaK100 participants were largely comprised of people with advanced disease stages and who were either current or former smokers.

Table 3.4: Baseline characteristics of subjects in CodeBreak100, phase 2

Sotorasib 960 mg (N=126)	
Sex - n (%)	
Male	63 (50.0)
Female	63 (50.0)
Race - n (%)	
Asian	19 (15.1)
Black or African American	2 (1.6)

Sotorasib 960 mg (N=126)	
White	103 (81.7)
Other	2 (1.6)
Age (years)	
Mean	62.9
SD	9.3
Median	63.5
Min, Max	37,80
Smoking history - n (%)[a]	
Never	6 (4.8)
Current or former	117 (92.9)
NSCLC stage – n (%)	
III	5 (4.0)
IV	121 (96.0)
Metastases – n (%)	
Brain (non-active)	26 (20.6)
Liver	26 (20.6)
NSCLC histology – n (%)	
Non-squamous	125 (99.2)
adenocarcinoma	120 (95.2)
Squamous	1 (0.8)
ECOG performance status – n (%)	
0	38 (30.2)
1	88 (69.8)
Prior lines of systemic anticancer therapy – n (%)	
1	54 (42.9)
2	44 (34.9)
3	28 (22.2)
Types of prior systemic anticancer therapy[b] – n (%)	
Platinum-based chemotherapy	113 (89.7)
PD-1 or PD-L1 inhibitors	115 (91.3)
Platinum-based chemotherapy and PD1/L1 inhibitors	102 (81.0)

Based on Table 6 of the CS[1]
[a] smoking status missing for 3 participants; [b] prior systemic anticancer therapy also included targeted biologics (23.8%), targeted small molecules (7.1%), and other (0.8%)
CS = company submission; ECOG = Eastern Cooperative Oncology Group; N = number of participants in the analysis set; n = number of participants in the corresponding category; NSCLC = non-small cell lung cancer; PD-1 = programmed cell death protein 1; PD-L1 = programmed death-ligand 1; SD = standard deviation

3.2.3 Quality of CodeBreaK100 trial

The critical appraisal of this single-arm, non-randomised study was conducted utilising the ROBINS-I, tool, see Table 3.5.

Table 3.5: Quality assessment of CodeBreaK100 using the ROBINS-I tool

Domains of risk of bias assessment							
1. Bias due to confounding	2. Bias in selection of participants into the study	3. Bias in classification and intervention	4. Bias due to deviations from intended intervention	5. Bias due to missing data	6. Bias due to measurement of outcomes	7. Bias in selection of reported result	Overall bias
1.1.PY[a] 1.2.PN[b] 1.3.PN[b] 1.4.PN[b] 1.5.PN[b] 1.6.NI[b] 1.7.NI[b] 1.8.PN[b] RoB: Serious[g]	2.1.PN[c] 2.2.N/A[b] 2.3.N/A[b] 2.4.PY[d] 2.5.N/A[b] RoB: Low	3.1.PY[d] 3.2.Y 3.3.PN RoB: Low	4.1.PN[c] 4.2.N/A[b] 4.3.N/A 4.4.PY 4.5.PY 4.6.NI[b] RoB: Moderate[g]	5.1.PY[d] 5.2.PN[c] 5.3.PN[c] 5.4.N/A[b] 5.5.N/A[b] RoB: Low	6.1.PN 6.2.PY[d] 6.3.PY[e] 6.4.PY[a] RoB: Serious[g]	7.1.PN[c] 7.2.PN[c] 7.3.PN[c] RoB: Low	Serious[f]

ERG's own assessment (please also see Table 7 of Appendix D of the CS)[13]
Response categories: N = No; N/A = Not Applicable; PN = Probably No; PY = Probably Yes; Y = Yes; NI = no information
[a] Rated PN in CS; [b] Not rated in CS; [c] Rated N in CS; [d] Rated Y in CS; [e] Rated N/A in CS; [f] Rated as "low to moderate" in CS; [g] Rated as "low" in CS
Responses in Red indicate potential marker for a serious risk of bias
Responses in Green indicate potential markers for low risk of bias
Response in Moderate indicate potential markers for moderate risk of bias
RoB = risk of bias

ERG comment: The ERG considers that this tool has not been appropriately used as there were 14 missing entries to the signalling questions in the CS.[13] Specifically, domains relating to baseline confounding and measurement of ouctomes were rated as "serious" compared to "low" in the CS.[13]

Hence the ERG undertook its own assessment, concluding that there was a high risk of bias related to baseline confounding, i.e. lower ECOG performance status of 0-1 at baseline favoured sotorasib. The ERG also considers that there was a high risk of bias in classification of interventions. Furthermore, the ERG concluded that appropriate methods to control for confounders such as stratification, regression, or probability weighting were not employed. In addition, there was a serious risk of bias in measurement of outcomes, i.e. outcome assessors were probably aware of the intervention received by the participants in the CodeBreaK100 trial.

In summary, the study has some important limitations as it has been judged by the ERG to be at a serious risk of bias in two (out of seven) domains of the ROBINS-I assessment tool.[14]

Table 3.6: Key issue 3. High risk of bias of CodeBreaK100

Report Section	3.2.3
Description of issue and why the ERG has identified it as important	Using the ROBINS-I tool, the company rated overall risk of bias of CodeBreaK100 to "low to moderate". However, the ERG re-assessed the study and rated the risk of bias to be "serious". Specifically, domains relating to baseline confounding and measurement of ouctomes were rated as "serious" compared to "low" in the CS.
What alternative approach has the ERG suggested?	Further evidence should aim to minimise the risk of bias.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	Further evidence should aim to minimise the risk of bias.

CS = company submission; ERG = Evidence Review Group; ROBINS-I = Risk Of Bias in Non-randomised Studies of Interventions

3.2.4. Results of CodeBreaK100 trial

The results presented in the CS were reported from a primary analysis, in which the data cut off was 1 September 2020, along with updated analyses with data cuts of 1 December 2020 and 15 March 2021. In the response to a request for separate results for participants with and without adenocarcinoma, the company provided the information presented in Table 3.7.

Table 3.7: Efficacy in CodeBreaK100 by adenocarcinoma histology (15 March 2021 data cut, post hoc analysis)

ORR	PFS				OS			
Events/ Subjects (%) (95% CI)	**Events/ Subjects**	**Median (Months) (95% CI)**	**6 months KM Estimate (95% CI) (%)**	**12 months KM Estimate (95% CI) (%)**	**Events/ Subjects**	**Median (Months) (95% CI)**	**6 months KM Estimate (%) (95% CI)**	**12 months KM Estimate (%) (95% CI)**
Adenocarcinoma								
44/118 (37.3) (28.6 to 46.7)	82/118	6.8 (5.1 to 8.2)	52.2 (42.3 to 61.2)	26.9 (18.6 to 36.0)	62/120	12.0 (10.0 to NE)	74.2 (65.2 to 81.2)	50.5 (40.9 to 59.3)
No adenocarcinoma								
2/6 (33.3) (4.3 to 77.7)	5/6	6.2 (1.2 to NE)	50.0 (11.1 to 80.4)	33.3 (4.6 to 67.6)	2/6	NE (6.6 to NE)	100.0 (NE to NE)	66.7 (19.5 to 90.4)

Based on response to question A18 in response to the request for clarification[6]
CI = confidence interval; CS = company submission; KM = Kaplan-Meier; NE = not estimable; PFS = progression-free survival; ORR = objective response rate; OS = overall survival

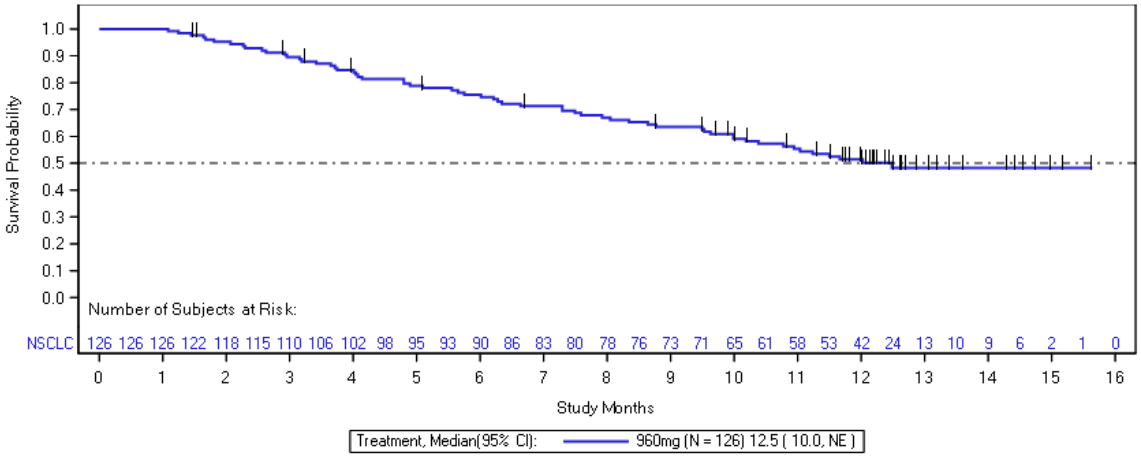
3.2.4.1 Objective response rate

ORR was the primary endpoint of the CodeBreaK100 study and was defined as the proportion of subjects with best overall response of complete response or partial response as assessed by RECIST 1.1.[1] The response was assessed by the blinded independent central review (BICR). The complete response and partial response required confirmatory computerised tomography (CT) or magnetic resonance imaging (MRI) repeat assessment at least 4 weeks after the first detection of response. According to the CS, clinical relevance was determined by the lower bound of the 95% CI excluding a prespecified benchmark of 23%.[1]

3.2.4.2 Overall survival (OS)

The median OS, as presented in the CS, was 12.5 (95% confidence interval (CI) 10.0 to not estimable) months.[1] The Kaplan-Meier (KM) estimate of survival was presented as 75.5% (95% CI 66.8 to 82.2) at 6 months and 51.4% (95% CI 41.9 to 60.1) at 12 months (Figure 3.1). Roughly half of the patients (46.8%) had experienced death at the time of the cut-off. The CS emphasises that the CodeBreaK100 study was not specifically powered for survival outcomes.[1]

Figure 3.1: Kaplan-Meier plot of overall survival (safety analysis set)



Snapshot date 21JAN2021. Phase 2 data cut-off date 01DEC2020.
Censor indicated by vertical bar |
NE = Not Estimable.
Death is an event.

Program: /userdata/stat/amg510/onc/20170543/analysis/eff_update_202101/figures/f-eff-km.sas
Output: f14n-04-003-001-eff-km-osm-nsclc-p2saf.rtf (Date Generated: 28JAN21:12:38:17) Source: adam.adsl, adam.adtte

Based on Figure 6 of the CS[1]

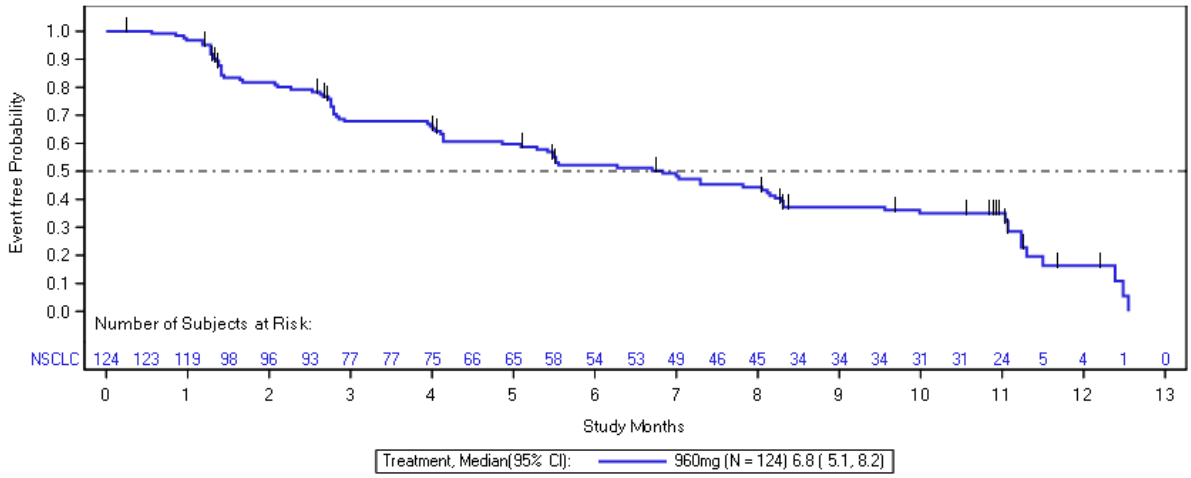
3.2.4.3 Duration of response

According to the CS, among the 46 responders who had NSCLC, the Kaplan-Meier estimate of median duration of response was 10 months (95% CI 6.9 to 11.1 months).[1] The company noted that 27 subjects (58.7%) were censored. The CS also stated that 20 of the 46 objective responders were still receiving treatment without disease progression.[1]

3.2.4.4 Progression-free survival (PFS)

The median PFS was reported to be 6.8 months (95% CI 5.1 to 8.2 months) at the time of the cut-off.[1] The KM estimate of survival was 52.2% (95% CI 42.6 to 60.9) at 6 months and 16.3% (95% CI 7.4 to 28.2) at 12 months. According to the CS, 56.5% of the patients had experienced disease progression, while 10.5% of patients experienced death.[1] Of note, 41 patients were censored, which comprised of 25 patients who were on the study without disease progression, seven who started new anticancer therapy, five who missed more than one consecutive assessment, and three who withdrew their consent.[1]

Figure 3.2: Kaplan-Meier plot of progression-free survival



Snapshot date 21JAN2021. Phase 2 data cut-off date 01DEC2020.
Censor indicated by vertical bar |
NE = Not Estimable.
Radiological Progression or Death (whichever occurs earlier) is an event.

Program: /userdata/stat/amg510/onc/20170543/analysis/eff_update_202101/figures/f-eff-km.sas
Output: f14n-04-002-001-eff-km-pfs-nsclc-p2fas.rtf (Date Generated: 28JAN21:12:38:15) Source: adam.adsl, adam.adtte

Based on Figure 5 of the CS[1]

3.2.4.5 Adverse effects of treatment

The CS provided the frequencies of treatment-emergent adverse events (TEAEs) experienced in the CodeBreaK100 study.[1] As presented in Table 3.8, nearly all participants in the CodeBreaK100 study (99.2%) experienced TEAEs.

Table 3.8: Summary of overall adverse events in NSCLC subjects in CodeBreaK100

Sotorasib 960 mg daily (N=126), n (%)	
All treatment-emergent adverse events	**125 (99.2)**
Grade ≥2	110 (87.3)
Grade ≥3	75 (59.5)
Grade ≥4	23 (18.3)
Serious adverse events	63 (50.0)
Leading to discontinuation of sotorasib	11 (8.7)
Serious	7 (5.6)
Non-serious	5 (4.0)
Fatal adverse events	20 (15.9)
Treatment-related treatment-emergent adverse events	**88 (69.8)**
Grade ≥2	49 (38.9)
Grade ≥3	26 (20.6)
Grade ≥4	1 (0.8)
Serious adverse events	10 (7.9)
Leading to discontinuation of sotorasib	9 (7.1)
Serious	4 (3.2)
Non-serious	5 (4.0)
Fatal adverse events	0 (0.0)

Sotorasib 960 mg daily (N=126), n (%)
Based on Table 17 of the CS[1] Coded using MedDRA version 23.1. Severity graded using CTCAE version 5.0 CS = company submission; CTCAE = Common Terminology Criteria for Adverse Events; MedDRA = Medical Dictionary for Regulatory Activities; N = number of participants in the analysis set, n = number of participants with observed data

The most commonly reported TEAE in ≥10% of NSCLC patients in the CodeBreaK100 trial included diarrhoea, nausea, fatigue, and elevations in alanine and aspartate aminotransferase, see Table 3.9 and Table 3.10 for TEAEs occurring in >5% of participants. The company noted that sotorasib appeared to be well tolerated and the adverse events were determined to be manageable.[1] As of the 01 December 2020 data cut-off, 37.3% of patients with NSCLC experienced events relating to hepatotoxicity or renal toxicity. However, this did not result in dose interruption or discontinuation.

Table 3.9: Treatment-emergent adverse events of any severity occurring in ≥10% NSCLC patients in the CodeBreaK100 trial

Phase 2 NSCLC 960 mg daily (N = 126), n (%)	
Preferred Term	
Diarrhoea	62 (49.2)
Nausea	38 (30.2)
Fatigue	32 (25.4)
Aspartate aminotransferase increased	27 (21.4)
Alanine aminotransferase increased	26 (20.6)
Dyspnoea	24 (19.0)
Arthralgia	23 (18.3)
Vomiting	23 (18.3)
Constipation	22 (17.5)
Back pain	20 (15.9)
Anaemia	17 (13.5)
Blood alkaline phosphatase increased	17 (13.5)
Oedema peripheral	17 (13.5)
Cough	16 (12.7)
Decreased appetite	15 (11.9)
Pleural effusion	13 (10.3)
Based on Table 22 of Appendix F of the CS[13] Coded using MedDRA version 23.1; rows are sorted by preferred term in descending order of frequency CS = company submission; MedDRA = Medical Dictionary for Regulatory Activities; N = number of participants in the analysis set, n = number of participants with observed data	

Table 3.10: Treatment-related adverse events occurring in ≥5% of NSCLC subjects in CodeBreaK100

Treatment-related adverse events (TRAEs) occurring in > 5%, n (%)	Any Grade N = 126	Grade 3+ N = 126
Any event	**88 (69.8)**	**25 (19.8)**
Diarrhoea	39 (31.0)	5 (4.0)

Treatment-related adverse events (TRAEs) occurring in > 5%, n (%)	Any Grade N = 126	Grade 3+ N = 126
Nausea	24 (19.0)	0
ALT increase	19 (15.1)	8 (6.3)
AST increase	19 (15.1)	7 (5.6)
Fatigue	14 (11.1)	0
Vomiting	10 (7.9)	0
Blood alkaline phosphatase increase	9 (7.1)	1 (0.8)
Maculopapular rash	7 (5.6)	0
Based on Table 18 of the CS[1] ALT = alanine aminotransferase; AST = aspartate aminotransferase; CS = company submission; N = number of participants in the analysis set, TRAE = treatment-related adverse event		

Adverse events (AEs) of any grade, regardless of attribution, were observed in all but one patient (99.2%). The most common AEs included diarrhoea, nausea, fatigue, arthralgia (joint pain), increase in aspartate aminotransferase (ASP) or the alanine aminotransferase levels (ALT). Treatment-related AEs (TRAE) leading to dose modification (dose interruption, reduction, or both) happened in 28 patients (22.2%).[5]

ERG comment: The ERG is concerned with the high number of TEAEs, i.e. 63 patients (50%) with NSCLC experienced serious AEs in the CodeBreaK100 trial. Twenty patients (15.9%) died.

Table 3.11: Key issue 4. High number of serious adverse events observed in CodeBreaK100

Report Section	3.2.4.5
Description of issue and why the ERG has identified it as important	The ERG is concerned with the high number of treatment-emergent adverse events, i.e. 63 patients (50%) with NSCLC experienced serious AEs in the CodeBreaK100 trial. Twenty patients (15.9%) died.
What alternative approach has the ERG suggested?	None. The ERG wants to highlight the issue for the committee.
What is the expected effect on the cost effectiveness estimates?	Unclear.
What additional evidence or analyses might help to resolve this key issue?	Potential guidance should reflect this issue.
AE = adverse event; ERG = ERG = Evidence Review Group; NSCLC = non-small cell lung cancer	

3.2.4.6 Health-related quality of life (HRQoL)

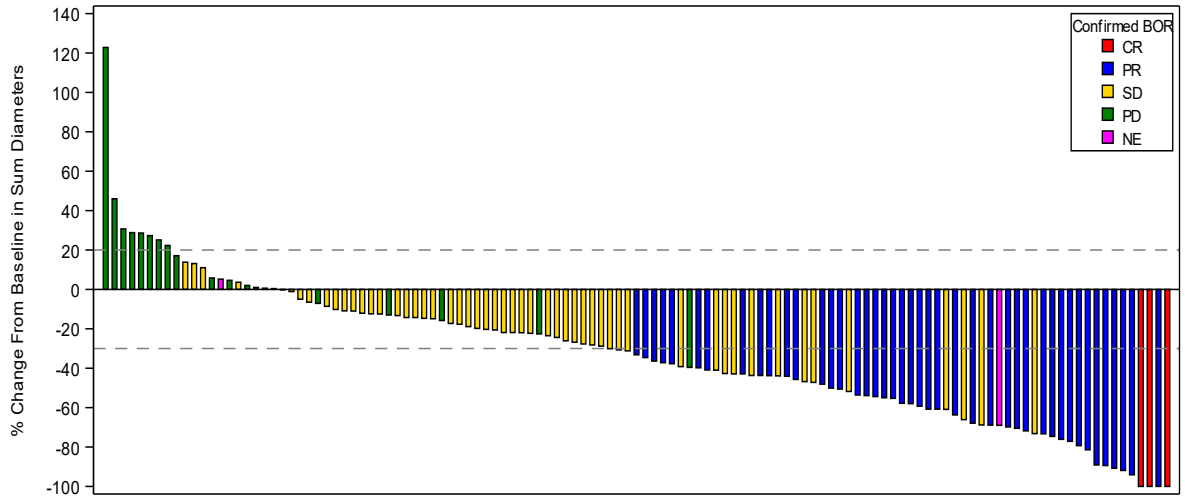
Information related to HRQoL was addressed as an exploratory analysis. For the purpose of the present CS, the European Quality of Life-5 Dimensions-5 Levels (EQ-5D-5L) was used to evaluate the generic health status of the participants.[1] At baseline, most of the participants either reported no problems or slight problems across the EQ-5D-5L health dimensions. However, 33% of participants who had reported either moderate or severe problems or were unable to perform the activity in the pain/ discomfort health dimension.

ERG comment: It is unclear why point estimates were unavailable for HRQoL in general (or PFS specific) for participants on chemotherapy that could be applied in scenario analyses.

3.2.4.7 Disease control rate

According to the CS, the disease control rate comprises of the complete response, partial response, or stable disease.[1] The disease control rate was determined to be high at 80.6% (95% CI 72.6 to 87.2). The CS noted that the percentage of subjects with stable disease was 43.5%. The company also emphasised that not all patients with advanced NSCLC have tumour shrinkage after cancer therapies.[1] Figure 3.3 depicts the tumour shrinkage by best overall response to sotorasib.

Figure 3.3: Waterfall plot of best tumour shrinkage



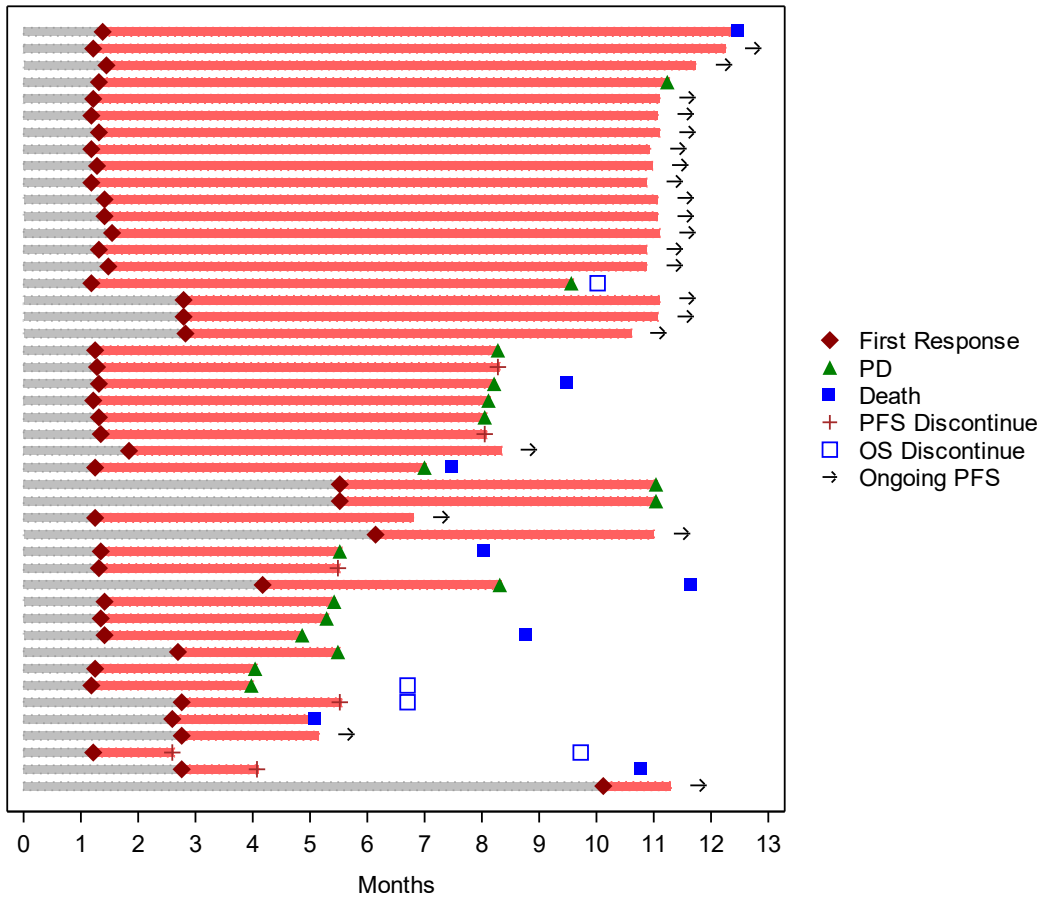
Phase 2 data cut-off date 01DEC2020.
Percent change from baseline in sum of diameters only considers tumor assessments prior to and include the 1st assessment where timepoint response is progressive disease, and prior to start of next anti-cancer therapy.
Two subjects without baseline target lesions and 3 subjects without post-baseline percent changes are not shown.

Based on Figure 3 of the CS[1]

3.4.2.8 Time to response

Among the 46 responders in the NSCLC group, the reported median time to response was 1.35 months within a range of 1.25 to 2.69 months.[1] Figure 3.4 depicts the duration and time to response. However, this is based on the December 2020 data cut-off.

Figure 3.4: Swimmer plot of duration and time to response



Phase 2 data cut-off date 01DEC2020.
'PFS Discontinue' indicates PFS censor due to no post-baseline assessment, withdrew consent, started of new anti-cancer therapy, missed two or more consecutive tumor assessments, off study due to sponsor decision, or lost to follow-up.
'OS Discontinue' indicate OS censor due to withdrew consent, completed study, off study due to sponsor decision, or lost to follow-up.

Based on Figure 3 of the CS[1]
CS = company submission; OS = overall survival; PD = progressive disease; PFS = progression-free survival

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

As detailed in Section 2.3, the CS considered docetaxel monotherapy as the primary comparator for sotorasib (referred to as the primary comparison) while docetaxel in combination with nintedanib was considered as a secondary comparator in patients with adenocarcinoma (referred to as the secondary comparison). As detailed in Section B.2.9 of the CS, the company expects the anticipated conditional approval of sotorasib to be based on the single-arm CodeBreaK100 trial, see Section 3.2 for details of the trial.[1]

The CS identified two studies, SELECT-1 and LUME-Lung 1, as relevant studies to inform an unanchored indirect treatment comparison (ITC) of sotorasib and docetaxel monotherapy and docetaxel combined with nintedanib, respectively, see Table 3.12 for details of the studies of the studies used for ITCs.[1]

Table 3.12: Overview of study designs of CodeBreaK100, SELECT-1 and LUME-Lung 1

Study characteristics	Sotorasib (CodeBreaK100)[15]	Docetaxel monotherapy (SELECT-1)[16]	Docetaxel + nintedanib (LUME-Lung 1)[17]
Blinding	Open label	Double-blinded	Double-blinded
Inclusion criteria	• Male or female patients (≥18 years) • Histologically confirmed locally advanced or metastatic NSCLC • KRAS p.G12C mutation identified through molecular testing • ECOG Performance Status 0 to 1 • ≥1 prior line of systemic anticancer therapy	• Male or female patients (≥18 years) • Histologically confirmed locally advanced or metastatic NSCLC • KRAS-mutation identified through molecular testing • WHO Performance Status 0 to 1 • 1 prior line of systemic anticancer therapy	• Male or female patients (>18 years) • Histologically confirmed locally advanced or metastatic NSCLC • ECOG Performance Status 0 to 1 • 1 prior line of systemic anticancer therapy
Key exclusion criteria	• Active brain metastases • Anti-tumour therapy including chemotherapy, antibody therapy, molecular targeted therapy, retinoid therapy within 28 days of study day 1	• Brain metastases • Received >1 prior anti-cancer drug regimen for advanced or metastatic NSCLC • Prior treatment with a MEK inhibitor or any docetaxel-containing regimen (prior treatment with paclitaxel is acceptable)	• Active brain metastases • Received >1 prior anti-cancer drug regimen for advanced or metastatic NSCLC • Prior treatment with a VEGFR inhibitor (other than bevacizumab) or docetaxel
Primary endpoint	Centrally assessed ORR	Investigator-assessed PFS	Centrally assessed PFS
Key secondary endpoints	• Centrally assessed PFS • Investigator-assessed PFS • OS	OS	OS

Based on Table 9 of the CS[1]
CS = company submission; ECOG = Eastern Cooperative Oncology Group; G12C = G12C amino acid substitution; KRAS = Kirsten rat sarcoma viral oncogene homolog; MEK = mitogen activated protein kinase; NSCLC = non-small cell lung cancer; ORR = objective response rate; OS = overall survival; PFS = progression-free survival; VEGFR = vascular endothelial growth factor receptor; WHO = World Health Organization

The CS summarised similarities and differences of these studies:

"CodeBreaK100, SELECT-1 and LUME-Lung 1 were all multicentre studies that recruited patients with confirmed locally advanced or metastatic NSCLC (stage IIIB to IV) who had failed prior therapy. CodeBreaK100 specifically enrolled patients with KRAS p.G12C mutations, whereas SELECT-1 enrolled patients with KRAS mutations at codon 12, 13 or 61.[16] LUME-Lung 1 did not specify KRAS mutations as an enrolment criterion and did not record KRAS mutations among the participants; however, in the subpopulation of interest (licensed population of patients with adenocarcinoma) the proportion of patients with KRAS p.G12C mutations is likely close to the prevalence of KRAS p.G12C mutations in the general non-squamous population (~13%). CodeBreaK100 enrolled patients with 1 to 3 prior therapies, whereas SELECT-1 and LUME-Lung 1 included patients with 1 prior therapy. All studies excluded subjects with active brain metastases, although CodeBreak100 and LUME-Lung 1 permitted inclusion of stable brain metastases.

All three studies reported PFS and OS as primary or secondary endpoints. PFS was assessed by investigators in SELECT-1, by both independent central review and by investigator in CodeBreaK100 and by independent central review in LUME-Lung 1".

Table 3.13 gives an overview of the baseline characteristics of these studies.

Table 3.13: Comparison of baseline characteristics in CodeBreaK100, SELECT-1 and LUME-Lung 1 trials

Baseline characteristics[a]	Sotorasib (CodeBreaK100) N=126[15]	Docetaxel monotherapy (SELECT-1) (N=256)[16]	Docetaxel + nintedanib (LUME-Lung 1) (N=322)[i][17]
Age	62.9 (mean)	60.9 (mean)	58.5 (median)
Gender (% female)	50%	43%	37%
Brain metastases (%)	21%	NR[c]	8%
Performance status (ECOG or WHO; % PS 1 [vs PS 0])	70%	59%	70%
Race (% white)	82%[d]	95%	NR[g]
% KRAS p.G12C-mutated	100%	42%[b]	NR[h]
Anti-PD-(L)1 in prior line(s)	91%	0%	0%
Number of prior lines (% with 1/2/3 prior lines)	43%/35%/22%	100%/0%/0%	Mostly 1 prior line[i]
Metastatic disease at baseline	96%	96%	90%
Histology (% non-squamous)	99%	95%	100%[j]
Smoking status (% ever smoker)	93%[e]	92%	64%
Other targetable mutations (EGFR, ALK, BRAF, ROS-1)	3%	NR[f]	NR
PD-L1 expression at baseline (<5% [vs >5%])	48%	58%	NR

Baseline characteristics[a]	Sotorasib (CodeBreaK100) N=126[15]	Docetaxel monotherapy (SELECT-1) (N=256)[16]	Docetaxel + nintedanib (LUME-Lung 1) (N=322)[i17]
Based on Table 10 of the CS[1] [a] all reported baseline characteristics in SELECT-1 and other key characteristics; [b] the rest of the population has KRAS mutations other than G12C; [c] not reported for SELECT-1. All studies had exclusion criteria for active brain metastases; [d] 15 percentage points of the 18% remaining correspond to Asian patients; [e] 2 percentage points of the remaining 7% are missing data; [f] probably very low due to KRAS mutant; [g] Race was not reported, the trial was non-US based and run mainly in Europe (71% of patients) as well as Asia; [h] LUME-Lung 1 did not enrol by or record genetic mutations; the % of KRAS p.G12C is likely close to the prevalence of KRAS p.G12C mutations in the general non-squamous population (~13%); [i] LUME-Lung 1 included patients with a prior platinum-based therapy and allowed adjuvant/neoadjuvant as line of therapy; [j] Based on the subpopulation of interest (adenocarcinoma) ALK = anaplastic lymphoma kinase; BRAF = B-Raf Proto-oncogene; CS = company submission; ECOG = European Co-operative Oncology Group; EGFR = epidermal growth factor receptor; KRAS = Kirsten rat sarcoma viral oncogene homolog; NR = not reported; PD-L1 = programmed death-ligand 1; PS = performance status; ROS = proto-oncogene tyrosine-protein kinase; WHO = World Health Organization			

The CS highlighted that *"as LUME-Lung 1 enrolled patients with mixed histology,[17] but nintedanib in combination with docetaxel is only licensed for use in patients with adenocarcinoma,[18] only the characteristics of the adenocarcinoma subpopulation of LUME-Lung 1 are considered"*.[1]

Overall, the distribution of patients between the three trials is similar in terms of age, disease stage and histology, and the majority of patients had ECOG/WHO performance status of 1.

However, the CS highlighted a few differences between the studies which arose from the different time at which these were conducted:[1]

1. G12C KRAS mutation status, i.e. 100% in CodeBreaK100, 42% in SELECT-1 (remaining patients had other KRAS mutations), and not reported for LUME-Lung 1.
2. CodeBreak100 included patients taking 1-3 prior therapies and a high proportion of patients who had prior use of PD(L)-1 inhibitors, reflecting the current treatment pathway for patients with KRAS p.G12C -mutated NSCLC in the UK. In contrast, the SELECT-1 and LUME-Lung 1 trials, which were both conducted before the evidence base supported front-line use of immunotherapy, included patients taking 1 prior therapy only and no PD(L)-1 inhibitors.
3. Based on inclusion criteria and/or a lack of recording, it is also not possible to compare for the presence of (non-active) brain metastases in SELECT-1, for the PD-1 expression in LUME-Lung 1, or for the presence of other targetable mutations in either of these comparator trials.
4. It is also of note that LUME-Lung 1 recruited fewer females, fewer prior smokers and patients with fewer brain metastases than CodeBreaK100.

According to the CS, UK clinical experts considered these were the best and most relevant sources of data available with which to make indirect comparisons for sotorasib in patients with KRAS p.G12C mutated NSCLC.[1, 19]

The company also used the Flatiron study as an alternative data source for the primary comparison.[20] The reason given for this being only supplementary was that docetaxel was only used in a minority of patients and that about 24% of patients had received prior first-line immunotherapy.

ERG comment: The ERG noted the differences between these studies. As discussed in more detail in Section 3.4, it is not possible to match for all of these differences which might have an impact on the validity of the findings of any ITC.

It is not entirely clear that SELECT-1 was a better data source than Flatiron. This is not least because individual participant data were available to perform what the company called a 'Propensity Score Weighted Analysis (PSWA)', which appears to be an Inverse Probability Weighting (IPW) analysis according to technical support document (TSD) 17, for the latter such that the comparator data could be adjusted to be more like the intervention population.[21] However, the size of reduction in any bias would depend on the degree to which prognostic factors could be identified, either from the CodeBreaK100 data for the MAIC or the Flatiron data for the PSWA.

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

3.4.1 Matching of prognostic patient characteristics

As noted in Section 3.3, there are differences between the three studies considered for indirect comparison, CodeBreaK100, SELECT-1 and LUME-Lung 1.

For the primary comparison (vs. docetaxel), *"an MAIC was used to compare changes in OS and PFS with sotorasib versus docetaxel monotherapy"* (see Section B.2.9.3.1 of the CS).[1]

The company rejected the use of a MAIC for the secondary analysis (vs. nintedanib plus docetaxel in adenocarcinoma). The reasons given was that: *"…the differences in patient characteristics and data availability for matching would present significant challenges, including reducing the effective sample size and precision for relative treatment effect estimates and introducing a further population that is less closely aligned with CodeBreak100 and not aligned with the SELECT-1 trial population to which sotorasib recipients in CodeBreaK100 had already been matched"* (see Section B.2.9.3.2 of the CS).[1]

According to the CS, *"further examination of the LUME-Lung 1 data indicated that a piecewise approach to hazard ratio estimation would be required, and estimation of the survival of patients with sotorasib vs nintedanib plus docetaxel could only be made following implementation of these data within the economic model"* (further details in Section B.3.3.5 of the CS).[1]

3.4.1.1 Choice of covariables for MAIC

In response to question A27 of the the request for clarification, the company gave details on the choice of covariables for the MAIC, stating that *"1) the populations being compared are defined similarly in terms of ECOG performance status at baseline and 2) adjustment based on ECOG score is performed to ensure the populations being compared are balanced. Other factors that were indicated by the physicians to be very important by a majority of physicians to assess the prognosis or response to treatment of patients included presence of brain metastases (ideally distinguishing between active and controlled brain metastases), disease stage at baseline (stage IIIb/c vs IV or IIIb-IVa vs IVc). A number of other factors were considered as being at least somewhat important for prognosis and should be considered when the information is available. Finally, age and gender, although not consistently considered as prognostic or predictive factors, were mentioned as key covariates to include in an adjusted comparative effectiveness analysis"*.[6]

Table 11 of the CS presented the starting list of prognostic covariates (five classified as "very important", 13 classified as "somewhat important", and three "additional covariates reported in other MAIC analyses").[1]

"Of these 21 potential covariates, 8 were selected for inclusion in the MAIC analysis based on data availability in SELECT-1, their prognostic importance for patients receiving sotorasib and docetaxel, and the feasibility of matching whilst preserving the effective sample size", as detailed in Table 12 of the CS, namely:

1. ECOG (% PS 1 [vs. PS 0])
2. Age (mean)
3. Metastatic disease stage at baseline
4. Smoking status (% ever smoker)
5. PD-L1 expression level
6. Gender (% female)
7. Histology (% Non-squamous)
8. Race (% white).[1]

ERG comment: It should be noted that some factors identified by the clinical experts, such as presence of brain metastases and disease stage at baseline, as well as other factors such as KRAS p.G12C mutation status were not considered for the MAIC comparing CodeBreaK100 and SELECT-1.

3.4.1.2 KRAS p.G12C mutation status

The CS stated that *"a propensity score weighted analysis approach such as MAIC requires the matching of prognostic patient characteristics to generate robust comparative treatment effect estimates. Due to missing data or other differences between the trials it would not be possible to match across all trials for KRAS p.G12C mutation status, brain metastases, prior lines of therapy or prior use of PD-L1 inhibitors. Given that PFS and OS outcomes are similar in the absence of targeted therapies, irrespective of KRAS status (see section B.2.9.2.1* [of the CS]*), the inability to match by specific KRAS status is unlikely to lead to biased estimates"*.[1]

In response to the request for clarification (question A24), the company replicates Table 3 of the CS to support the view that *"given the OS and PFS for All NSCLC patients are highly consistent to those for patients with KRAS mutations (and the KRAS mutation datasets are included in the All NSCLC dataset),it is reasonable to conclude that the results are consistent for those with and without KRAS mutations"*.[1, 6]

However, the company (in response to question A26b), highlighted that *"as targeted therapy for KRAS mutated NSCLC did not exist at the time of the LUME Lung 1 trial, and screening for KRAS mutant NSCLC was not routine practice, we have no means of knowing the KRAS mutant status of patients enrolled in the LUME Lung 1 trial. It was for this reason that it was not possible to match the LUME Lung 1 trial participants and CodeBreaK100 trial participants in an MAIC"*.[6]

ERG comment: Table 3 of the CS does show that median OS and PFS do appear to vary little between KRAS p.G12C-mutated and KRAS-mutated (non-p.G12C) NSCLC, e.g. first line: 12.0 (9.6, 15.3) vs. 12.2 (10.5, 14.4) and 5.0 (4.4, 5.8) vs. 5.6 (5.4, 6.0), respectively.[1, 6] However, despite the consequent increase in uncertainty by using only the 42% of patients with KRAS p.G12C mutation, an analysis with mutation status as covariate could be informative.

3.4.1.3 Brain metastases

The CS stated that not matching on brain metastases between CodeBreaK100 and SELECT-1 is *"unlikely to introduce significant bias"*.[1] In response to the request for clarification (question A28), the company elaborated on this point:[6]

"The proportion of patients with brain metastases was higher in CodeBreaK100 (21%) than in LUME Lung-1 (8%). The proportion with brain metastases in SELECT-1 was not reported. However, all three trials excluded patients with active (or symptomatic) brain metastases. (…) *As CodeBreaK100 enrolled a high proportion of patients with brain metastases, and somewhat higher than in patients recruited to LUME Lung 1, it is a reasonable assumption that SELECT-1 did not include a higher proportion of patients with non-active brain metastases than CodeBreaK100. Any negative influence on survival of the presence of brain metastases would therefore impact on the CodeBreak100 population to a greater extent than on the populations in LUME Lung 1 or SELECT-1. Therefore, the results of the comparison of sotorasib (from CodeBreaK100) vs nintedanib plus docetaxel (from LUME Lung 1) or docetaxel monotherapy (from SELECT-1) would favour the comparators. On this collective basis, our inability to match for brain metastases between CodeBreaK100 and SELECT-1 is unlikely to introduce bias in favour of sotorasib and is more likely to be conservative".*[6]

ERG comment: Although active brain metastases were excluded from all three trials, the presence of brain metastases did seem to affect prognosis as indicated by the subgroup analyses reported in Appendix E.[13]

In particular, median OS was not estimable for no metastases and percentage surviving to 12 months was 55.5 (44.8, 64.9) compared to 35.3 (23.4, 48.4) for presence of metastases. The company claim that, because the percentage was a lot higher for CodeBreaK100 than LUME-Lung 1 then it must also be higher than for SELECT-1, so that not adjusting for brain metastases is favourable to the comparator.

However, the ERG would regard this as speculation and therefore there is no way of knowing the effect of not adjusting for brain metastases on outcome.

3.4.1.4 Other baseline characteristics

In response to the request for clarification (questions A25a and A25b), the company confirmed that a number of factors, such as country of origin, socio-economic status, comorbidities, year of recruitment and number/severity of metastases as well as age, gender, smoking status, geographic region/ethnicity/race, body mass index/weight or history of alcohol abuse, were not considered as, based on a physician's assessment, none of these factors was found to be *"very important to consider"* (i.e. *"at least 4 of the 6"* physicians highlighting the importance).

The ERG noted differences in the smoking rate in LUME-Lung 1 study compared to CodeBreaK100 and SELECT-1 (64% versus 93% and 92%) and asked for clarification (question A26). In response, the company stated to *"not know the reasons for why the smoking history of patients enrolled in LUME Lung 1 was different to that in CodeBreaK100 and SELECT-1; however, it can be seen in the LUME Lung 1 trial results that progression-free survival (PFS) and overall survival (OS) were not significantly different between patients with or without a history of smoking (see Figure 4 in Reck 2014, which refers to the adenocarcinoma population)".*[6]

ERG comment: The ERG noted that a number of factors were not considered. While the clinical experts consulted by the company agreed with that approach, there remains uncertainty to the impact matching these factors would have had.

3.4.1.5 Standard of care

In response to a request to clarify whether the standard of care for NSCLC is likely to be equivalent between the studies (question A25c), the company stated that *"CodeBreaK100 subjects were more heavily pre-treated than SELECT-1 subjects, with SELECT-1 and LUME-Lung 1 patients receiving*

only 1 previous line of systemic anticancer therapy", concluding that *"this is a conservative limitation for the comparative effectiveness as a more heavily pre-treated population is generally associated with poorer clinical outcomes"*.[6] **ERG comment:** The assessment by the company is likely to be correct, however, this adds to the uncertainty linked to the ITCs.

3.4.2 PSWA using Flatiron

The company stated that they used a PSWA, which most closely resembles an IPW according to TSD 17.[21] The details of the method are reported in Appendix D.[13]

The weights were applied only to the comparator data, which effectively implies the estimation of average treatment effect of the treated (ATT) as opposed to average treatment effect (ATE), thus limiting applicability to the population of patients who received sotorasib as opposed to any in the index population.[21]

Firstly, Flatiron patients were selected to align with the CodeBreaK100 eligibility criteria:

- Diagnosis of advanced NSCLC between 01 January 2011 and index date
- First positive test for KRAS mutation no later than 21 days after index date (to avoid introducing immortal time bias in the analyses)
- Age 18 years or older at index date
- Started the selected line of treatment on/before 31 March 2020 (to allow sufficient opportunity for a follow-up time of at least 6 months)
- Structured electronic health record activity in the first 90 days after the date of advanced NSCLC diagnosis
- Previous treatment with at least one prior line of therapy containing anti-PD-1 or anti-PD-L1 immunotherapy and/or platinum-based chemotherapy
- Selected line of therapy does not contain a clinical study drug
- Selected line of therapy is not the patient's first line of treatment containing an anti-PD-(L)1 component
- Baseline ECOG performance status ≤1

In addition, the following selection rule was applied to determine which line of therapy was considered for the control cohort:

- If a patient had received between 2 and 4 (inclusive) lines of therapy on or before 31 March 2020, the latest line of therapy which met the inclusion criteria was selected.
- If a patient had received more than 4 lines of therapy on or before 31 March 2020, the 4th line was selected (unless that line of therapy did not meet the inclusion criteria, in which case the most recent eligible treatment line was included).
- If no line of treatment met the inclusion/exclusion criteria, the patient was not included in the analysis.

All platinum-based chemotherapy patients and not only those who had taken docetaxel monotherapy were included. As shown in Table 12, Appendix D, there were about 31% of the former and 10% (n=21) of the latter in the KRAS mutant population with about 29% and 13% (n=11) respectively in the KRAS p.G12C mutant population.[13]

The process of covariate selection started with the same set as for the MAIC, as shown in Table 11 of the CS:[1]

- 'Very important' covariates were included expect PD-L1 status due to 98.7% of values being missing in Flatiron.
- 'Somewhat important' covariates were included (except for eGFR, again due to missing data (38.7%)) on the basis of a "*stepwise variable selection algorithm*", which was not clearly explained.

The list of included covariates is shown in Table 10 of Appendix D, which showed the effect of adjustment.[13] Figure 6 showed the standardised differences in covariates between CodeBreaK100 and Flatiron.[13] This shows that adjustment reduced those differences to close to zero in the KRAS mutant population. However, the standardised differences remained above 0.1 for several covariates and above 0.2 for liver metastases, one prior line of therapy and two age groups in the KRAS p.G12C population. This and the small effective sample size were the reasons given for preferring the KRAS mutation population.

ERG comment: Estimation of ATT as opposed to ATE might be an issue depending on the degree of heterogeneity of treatment effect and applicability of the CodeBreaK100 trial. An analysis applying the propensity score weights to all patients could be informative. Also, there are methods other than IPW, such as regression adjustment (RA) or doubly robust (RA plus IPW), that could have been employed and so scenario analyses using these methods could also be informative.[21]

Selecting patients to align with the CodeBreaK100 trial is in principle a good idea. However, given that sotorasib is to be positioned for 2nd line or later, it is not clear to the ERG why patients only at 4th line were selected. Although patient numbers are small, it might have been informative to see results for the docetaxel monotherapy population.

The process of covariate selection was not entirely clear and would therefore benefit from further explanation. It did appear that better balance was achieved for the KRAS population and, as discussed above, it might be reasonable to consider the prognosis similar to the KRAS p.G12C population.

In conclusion, the ERG considers that there might be reasons to believe that the results of the PSWA (using Flatiron) are less biased than those of the MAIC (using SELECT-1) given that:

1. The PSWA adjusted the Flatiron data to make more comparable to the CodeBreaK100 population: the benefit of this lies in CodeBreaK100 being more applicable to the patients that might be treated in the UK with sotorasib, which is uncertain
2. Very little difference in effective sample size (104.8 for Flatiron in the KRAS population in the PSWA vs. OS/PFS 108.8/106.1 for CodeBreaK100 in MAIC primary analysis)
3. The MAIC primary analysis only adjusted for four covariates, which excluded brain metastasis, as opposed to 13 in the PSWA, which included brain metastasis

However, there remains considerable uncertainty in the effectiveness of sotorasib vs. docetaxel that might be to some extent addressed by further analysis using the Flatiron data as described above.

Table 3.14: Key issue 5. Validity of ITC

Report Section	3.3, 3.4
Description of issue and why the ERG has identified it as important	The ITC is unanchored i.e. no common comparator. Therefore, there are potentially relevant differences in prognostic factors between the studies included in the ITCs (CodeBreaK100, SELECT-1, LUME-Lung 1), e.g. regarding G12C KRAS mutation status, prior therapies, presence of brain metastases, and factors like sex and smoking history. It is not possible to match for all of these differences which might have an impact on the validity of the findings of any ITC. The company chose a MAIC for their primary analysis of the main comparison with docetaxel, which is particularly prone to bias given lack of identification of all relevant prognostic factors and clinical experts identified factors to be "very important", e.g. brain metastases and disease stage at baseline. However, these, alongside G12C mutation status, were not considered for the MAIC comparing CodeBreaK100 and SELECT 1. Also, because only summary statistics were available from SELECT-1, the CodeBreaK 100 had to be adjusted to match the SELECT-1 population. The company also conducted a supplementary analysis using the Flatiron study, which, using a method of adjustment, referred to as PSWA that appears to involve IPW allowed the comparator data to match the CodeBreaK100 population. A richer set of individual patient data also afforded a greater number of potential prognostic factors. In addition to the underlying uncertainty introduced by an indirect comparison of treatments (compared to a direct comparison), the differences between studies, the choice of baseline variables for matching, the choice of underlying data source and adjustment method can be questioned and the ERG would have liked to see further analyses.
What alternative approach has the ERG suggested?	1. For the MAIC, an analysis with mutation status as covariate could be informative 2. For the PSWA, methods other than IPW, such as RA or doubly robust (RA plus IPW), could have been employed and so scenario analyses using these methods could be informative 3. For the PSWA, limiting to the docetaxel only population could be informative 4. In principle, evidence directly comparing treatments would provide more robust evidence.
What is the expected effect on the cost effectiveness estimates?	The uncertainty is increased.
What additional evidence or analyses might help to resolve this key issue?	See suggestions above.

ERG = Evidence Review Group; G12C = G12C amino acid substitution; IPW = inverse probability weighting; ITC = indirect treatment comparison; KRAS = Kirsten rat sarcoma viral oncogene homolog; MAIC = matching adjusted indirect comparison; PSWA = propensity score weighted analysis; RA = regression adjustment

3.4.3 Results of indirect comparison

3.4.3.1 Primary analysis – MAIC using CodeBreaK100 and SELECT-1

Results for the primary analysis were reported in Section B.2.9.4.1.[1] Table 3.15 provides an overview of results while Figures 3.5 and 3.6 show Kaplan-Meier plots of the primary MAIC analysis for OS and PFS, respectively.

Table 3.15: Results of MAIC for primary comparison of sotorasib vs docetaxel monotherapy

Analyses	CodeBreaK100 N (OS / PFS)	CodeBreaK100 ESS (OS / PFS)	Median OS Sotorasib vs. Docetaxel	Median PFS Sotorasib vs. Docetaxel
Unadjusted	126	126	[redacted]	[redacted]
MAIC Model: "all variables of prognostic importance" (Primary analysis)	123/ 121	108.8/ 106.1	[redacted]	[redacted]
MAIC Model: "all available covariates" (sensitivity analysis)	98/ 96	53.3/ 53.1	[redacted]	[redacted]

Based on Table 14 of the CS[1]
* Median OS not reached, OS was 50.4% at 12.5 months; ¶ Median OS not reached, OS was 52.5% at 12.0 months
CI = confidence interval; CS = company submission; ESS = effective sample size; HR = hazard ratio; MAIC = matching adjusted indirect comparison; OS = overall survival; PFS = progression-free survival

Figure 3.5: Kaplan-Meier plot for primary MAIC analysis of OS for sotorasib and docetaxel monotherapy

Based on Figure 7 of the CS[1]
CI = confidence interval; CS = company submission; HR = hazard ratio; MAIC = matching adjusted indirect comparison; OS = overall survival

Figure 3.6: Kaplan-Meier plot for primary MAIC analysis of PFS for sotorasib and docetaxel monotherapy

Based on Figure 8 of the CS[1]
CI = confidence interval; CS = company submission; HR = hazard ratio; MAIC = matching adjusted indirect comparison; PFS = progression-free survival

3.4.3.2 Supplementary primary comparison – Propensity score weighting analysis using CodeBreaK100 and Amgen Flatiron Health real-world evidence study

As described in Section B.2.9.4.1, *"this supplementary analysis was undertaken to explore an alternative data source and method of estimating relative treatment effects for sotorasib vs docetaxel monotherapy (using the basket of standard of care chemotherapy regimens in the Amgen Flatiron real-world evidence cohort as a proxy for docetaxel monotherapy)"*.[1] Results are presented in Table 3.16.

Table 3.16: Results of supplementary primary comparison using propensity score weighting analysis

Outcome	Flatiron N before adjustment	KRAS mutant		KRAS-p.G12C mutant subgroup	
		ESS	Median HR (95% CI)	ESS	Median HR (95% CI)
Overall survival	206	104.8	[redacted]	17.8	[redacted]
Progression-free survival	206	104.8	[redacted]	17.8	[redacted]

Based on Table 15 of the CS[1]
CI = confidence interval; CS = company submission; ESS = effective sample size; HR = hazard ratio; KRAS = Kirsten rat sarcoma viral oncogene homolog

3.4.3.3 Secondary comparison implemented in the economic model

According to Section B.2.9.4.2 of the CS, an *"estimation of the survival of patients with sotorasib vs nintedanib plus docetaxel was implemented in the economic model"*.[1] Table 3.17 presents the results for the secondary comparison while Section 4.2.6.6 provides a critique of the approach.

Table 3.17: Results of secondary comparison implemented in the economic model

	Sotorasib	Nintedanib plus docetaxel	Increment
Mean OS (months)*	[redacted]	[redacted]	[redacted]
Mean PFS (months)*	[redacted]	[redacted]	[redacted]

Based on Table 16 of the CS[1]
* Derived from economic model with 20-year time horizon, undiscounted values (see Section B.3.3.5 of the CS for details on the implementation)
CS = company submission; OS = overall survival; PFS = progression-free survival

3.5 Additional work on clinical effectiveness undertaken by the ERG

As detailed in Section 3.2.3, the ERG re-assessed the risk of bias of the CodeBreaK100 study using ROBINS-I.[14]

3.6 Conclusions of the clinical effectiveness Section

As the clinical effectiveness searches were run in June 2020 and updated on 26th January 2021, the ERG considers it likely that all potentially relevant studies were included in the systematic review. However, the ERG remains concerned about the application of English language restrictions and a lack of validated search filter for RCTs which both could negatively impact on the comprehensiveness and precision of the company's clinical effectiveness review.

The ERG has identified some inconsistencies in the study selection process that potentially introduce bias. For instance, exclusion of non-RCTs or phase I trials is questionable and based on the company's assumption that the evidence-base is limited. The ERG did not identify any issues with regards to data extraction.

The CodeBreaK100 study was a single arm, multicentre, non-randomised, open-label, phase II study.[5] Therefore, due to the absence of a comparator arm, the interpretation of the results is problematic. The study did not include a single centre from the UK which indicates generalisability of the CodeBreaK100's findings into clinical practice in England and Wales. It is not clear how participants at phase I of the trial were handled in the analyses as they used a combination of sotorasib and anti PD-1/L1 or midazolam. The ERG also undertook its own risk of bias assessments and found some serious limitations in the CodeBreaK100 study.

As there was no comparative trial data, the only available analysis was an unanchored ITC between sotorasib and a) docetaxel monotherapy (SELECT-1) and b) docetaxel + nintedanib (LUME-Lung 1).[16, 17] The ERG highlighted a few dissimilarities between the studies and stressed that it is not possible to match for all of these differences which potentially impacts validity of the findings of any ITC. The ERG also believes that the results of the MAIC (using SELECT-1) are potentially more biased than an alternative approach using PSWA (based on Flatiron data).

4. COST EFFECTIVENESS

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

This Section pertains mainly to the review of cost effectiveness analysis studies. However, the search Section (5.1.1) also contains summaries and critiques of other searches related to cost effectiveness presented in the CS. Therefore, the following Section includes searches for the cost effectiveness analysis review, measurement and evaluation of health effects as well as for cost and healthcare resource identification, measurement and valuation.

4.1.1 Searches performed for cost effectiveness Section

The following paragraphs contain summaries and critiques of all searches related to cost effectiveness presented in the CS.

Searches for cost effectiveness analysis review

Appendix G of the CS detailed a SLR conducted to identify published studies evaluating cost effectiveness, costs and resource use and HRQoL for treatments in NSCLC.[13] Searches were undertaken on 20 February 2020 and updated on 29 January 2021. Searches for costs and healthcare resource use were restricted to 2009 onwards. An English language restriction was reported but this was not applied at the searching stage. A summary of the sources searched is provided in Table 4.1.

Table 4.1: A summary of the sources to identify cost effectiveness studies

Resource	Host/Source	Date Ranges	Dates searched
Electronic databases			
MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations	Ovid	1946 to present	20 February 2020 29 January 2021
MEDLINE Daily, MEDLINE and Versions			
Embase		1974 to present	20 February 2020 29 January 2021
• CDSR • DARE • CENTRAL • CMR • NHS EED • HTA Database			20 February 2020 29 January 2021

Resource	Host/Source	Date Ranges	Dates searched
• ACP			
Congress searches			
ASCO	https://www.asco.org/	2017 - 2020	
ESMO	http://www.esmo.org/		
WCLC	https://wclc2019.iaslc.org/		
AACR	https://www.aacr.org/Pages/Home.aspx		

AACR = American Association of Cancer Research; ACP = American College of Physicians; ASCO = American Society of Clinical Oncology; CDSR = Cochrane Database of Systematic Reviews; CENTRAL = Cochrane Central Register of Controlled Trials; CMR = Cochrane Methodology Register; DARE = Database of Abstracts of Reviews of Effects; EED = Economic Evaluations Database; ESMO = European Society for Medical Oncology; HTA = Heath Technology Assessment; NHS = National Health Service; WCLC = World Conference on Lung Cancer

ERG comment: Searches were undertaken for a SLR to identify all cost effectiveness, HRQoL and cost and resource use studies. The CS provided sufficient details for the ERG to appraise the literature searches. A range of databases and conference proceedings were searched as well as previous NICE submissions for disease management costs.

The search strategy for the population focused specifically on KRAS mutated NSCLC and may have been too narrow to identify all relevant studies for cost effectiveness, HRQoL and cost and resource use. A date limit of 2009 was applied to searches for health economics but this was considered appropriate. As for clinical effectiveness searches, the strategies between Embase, MEDLINE and the Cochrane Library were not modified in all cases to take account for differences in thesaurus headings. However, the ERG was satisfied that the sufficient use of free-text terms compensated for this failure.

The use of filters in NHS EED may have been overly restrictive as this database is topic specific. However, as NHS EED is no longer being updated, the ERG is satisfied that anything of relevance is unlikely to have been missed.

4.1.2 Inclusion/exclusion criteria

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

Table 4.2: Eligibility criteria for the systematic literature reviews

	Inclusion criteria	Exclusion criteria
Patient population	NSCLC patients with KRAS mutated (further specification not required) with a primary interest in $KRAS^{G12C}$	Known KRAS mutation-negative status
Intervention	Any	Drug targeted to ALK, BRAF, EGFR, NTRK, or ROS1 (unless a KRAS mutated NSCLC comparator group is included)
Comparator	Any or none	N/A
Outcomes(s)	- Health-related quality of life - Quality-adjusted life-years gained	Any other

	Inclusion criteria	Exclusion criteria
	- Progression-free life-years gained - Life-years gained - Treatment cost by stage of disease (e.g., pre-progression vs. post-progression), including healthcare resource use, cost of care, cost of illness - Health state utilities - Economic evaluations	
Study design	Any	Animal/in vitro studies, case studies, and case reports
Date restrictions	***Costs/healthcare resource use*** - 2009 to present ***HRQoL and economic evaluation*** - No limit	
Language restrictions	English language	
Publication type	All primary publications and systematic reviews	Non-systematic reviews, editorials, notes, and letters
Country	Not restricted	
Based on Appendix G of the CS[13] ALK = anaplastic lymphoma kinase; BRAF = B-Raf Proto-oncogene; CS = company submission; EGFR = epidermal growth factor receptor: HRQoL = health-related quality of life; KRAS = Kirsten rat sarcoma viral oncogene homolog; N/A = not applicable; NSCLC = non-small cell lung cancer; NTRK = neurotrophic tyrosine kinase; ROS = proto-oncogene tyrosine-protein kinase		

ERG comment: The eligibility criteria used by the company provided sufficient detail and appeared to be appropriate.

4.1.3 Conclusions of the cost effectiveness review

Searches were undertaken for a SLR to identify all cost effectiveness, HRQoL and cost and resource use studies. The CS provided sufficient details for the ERG to appraise the literature searches. The search strategy for the population focused specifically on KRAS mutated NSCLC and may have been too narrow to identify all relevant studies for cost effectiveness, HRQoL and cost and resource use. As for clinical effectiveness searches, the strategies between Embase, MEDLINE and the Cochrane Library were not modified in all cases to take account for differences in thesaurus headings.

No published economic studies were identified in the SLR which examined the cost effectiveness of interventions for the management of patients with KRAS p.G12C mutation-positive locally advanced or metastatic NSCLC or for KRAS mutation in general. Also, no relevant studies on HRQoL to inform the decision problem were identified.

The ERG agrees that the eligibility criteria are suitable to fulfil the company's objective to identify cost effectiveness studies. Eligibility criteria were suitable for the SLR performed.

According to the CS, the SLR identified 14 studies reporting costs and healthcare resources used in patients with NSCLC and a KRAS p.G12C mutation.[1] Of these, 13 were on costs associated with biomarker testing which was not considered relevant for this appraisal. The company concluded that

the studies identified in the SLR on costs and healthcare resource use did not provide adequate costs and resource use valuations which were useful to a UK clinical setting, although it was not clear from Appendix I and the CS why the 14th study was not relevant.[1, 13]

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 health technology assessment	Reference case	ERG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Direct health effects for patients included
Perspective on costs	NHS and PSS	NHS and PSS
Type of economic evaluation	Cost utility analysis with fully incremental analysis	Cost utility analysis, two separate analyses for two comparators – hence no full incremental analysis was performed, because the populations were not comparable.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The time horizon of 20 years is considered long enough to reflect all relevant differences in costs and outcomes.
Synthesis of evidence on health effects	Based on a systematic review	Systematic review conducted to identify additional evidence on health effects beyond trial data. However, none of the studies found pertained to the KRAS p.G12C mutation.
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of health-related quality of life in adults.	Health effects were expressed in QALYs. Quality of life was measured with EQ-5D-5L and mapped to the EQ-5D-3L.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Reported directly by patients.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Crosswalk – representative sample of the UK population.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No equity issues have been identified.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be	The model includes the costs that relate to NHS and PSS

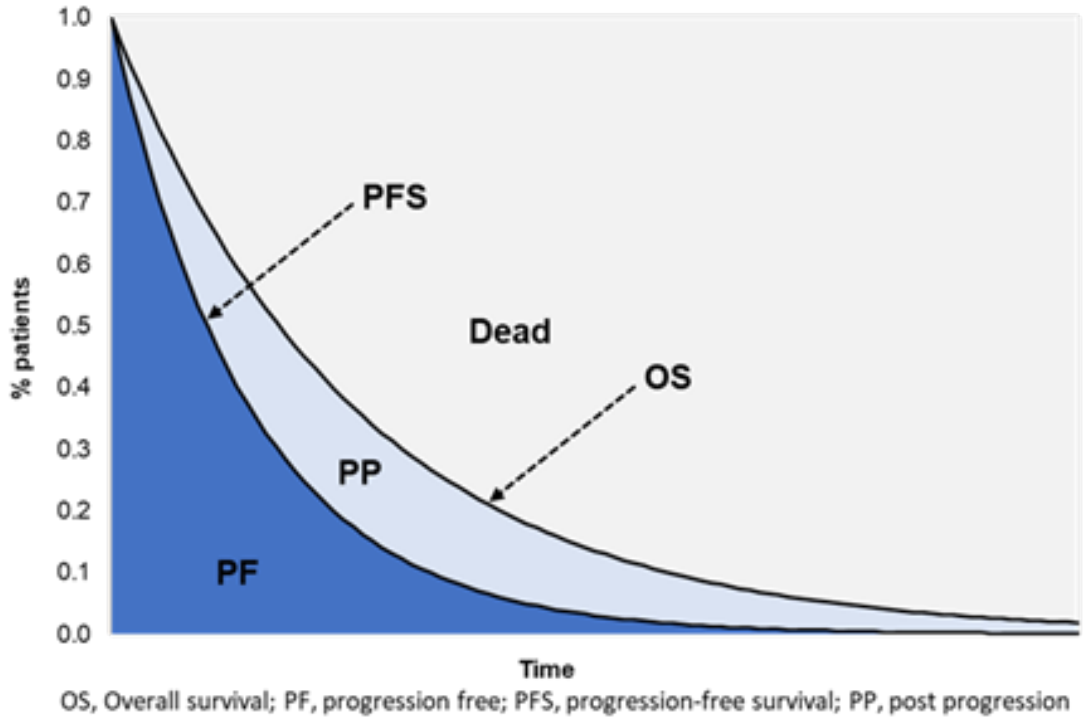
Element of health technology assessment	Reference case	ERG comment on company's submission
	valued using the prices relevant to the NHS and PSS	resources, valued using the prices relevant to the NHS and PSS.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Costs and health effects are discounted at 3.5%.
EQ-5D = European Quality of Life-5 Dimensions; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSS = Personal Social Services; QALY = quality-adjusted life year; UK = United Kingdom		

4.2.2 Model structure

The company developed a cohort-level partitioned survival model (PSM) in Microsoft Excel, with the following three health states: progression free (PF), post progression (PP) and death. The proportions of patients in each health state at the beginning of each model cycle are calculated from the PFS and OS curves from relevant clinical trials. In the model, all patients start in the progression free health state and on treatment.

Figure 4.1 shows the model structure of the partitioned survival model.

Figure 4.1: Model structure



Based on Figure 9 of the CS[1]
CS = company submission

ERG comment: The main concern of the ERG relates to the use of a PSM without a state transition model (STM) alongside it to validate the model structure. The company stated that the model structure applied was fully aligned with the primary objectives of treatment in oncology and NSCLC, namely avoiding disease progression and prolonging life, and that all relevant health states were included.[6]

The ERG considers this to be not an exclusive feature of a PSM, an STM would have aligned fully with these objectives as well and could have included the same health states. Therefore, the ERG requested the company to provide an STM as a scenario for validation purposes, as recommended in NICE Decision Support Unit (DSU) TSD 19.[22] In response to the request for clarification, the company stated that they considered it to be a recommendation in TSD 19 that an STM should be accompanying a PSM for validation.[6] The company also stated they believed an STM would not overcome the potential downsides of a PSM and that the scenarios provided would explore these sufficiently.

The ERG acknowledges that every model approach has its limitations but is still concerned that the consequences of choice of model structure may not be fully overseen because all choices and scenarios implemented follow this chosen structure. Size and direction of bias (if any) associated with choice of model structure cannot be estimated in the absence of alternative approaches.

Table 4.4: Key issue 6. Partitioned Survival Model structure not validated or justified

Report Section	**4.2.2**
Description of issue and why the ERG has identified it as important	The company used a partitioned survival model without elaborate justification and without an accompanying scenario implementing an STM to validate the results
What alternative approach has the ERG suggested?	The ERG did not suggest an alternative approach other than the STM
What is the expected effect on the cost effectiveness estimates?	The expected effect cannot be predicted
What additional evidence or analyses might help to resolve this key issue?	The ERG recognises that it is difficult and intensive to provide results from a model with an alternative structure.
ERG = Evidence Review Group; STM = state transition model	

4.2.3 Population

Consistent with the NICE scope, the population considered in the CS (Table 1 of the CS) was adults with previously treated KRAS p.G12C mutated, locally advanced or metastatic NSCLC.[1] The anticipated licensed indication of sotorasib is: for the treatment of adult patients with KRAS p.G12C-mutated locally advanced or metastatic NSCLC previously treated with platinum-based chemotherapy and/or anti PD-1/PD-L1 immunotherapy, unless contraindicated.[23]

The phase 2 trial evidence for sotorasib, i.e. the single-arm CodeBreaK100 study, focused on safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy. The population in the CodeBreaK100 study is defined as: adults with confirmed KRAS p.G12C-mutated NSCLC who had progressed after receiving 1-3 prior lines of anticancer therapy, had measurable disease per Response Evaluation Criteria In Solid Tumors (RECIST) 1.1 criteria, and had ECOG performance status of 0 or 1.[1]

Subgroup analyses were not included in the cost effectiveness analysis given there were no subgroups observed with substantially different efficacy compared to the whole population. Further, the relatively small population enrolled in CodeBreaK100 (N=126) would limit the sample size available and interpretability of any subgroup analyses.

The key baseline patient characteristics in the economic model are listed in Table 4.5 below and were obtained directly from CodeBreaK100.

Table 4.5: Key baseline patient characteristics of CodeBreaK100 used in the economic model

Patient characteristic	Mean / %	Source
Age at baseline (years)	62.9	CodeBreaK100 CSR, Table 9.2[24]
Gender (female)	50%	CodeBreaK100 CSR, Table 9.2[24]
Weight (kg)	71.1	CodeBreaK100 CSR, Section 9.3[24]
Body Surface Area (BSA, m^2)	1.81	Calculation - Mosteller formula[24]
Based on Table 22 of the CS BSA = body surface area; CS = company submission; CSR = clinical study report; SD = standard deviation		

ERG comment: As already mentioned in Section 2.1.2, the population in CodeBreaK100 and therefore also the population in the economic model, appears to be narrower than that defined in the NICE scope. In addition, the population in the secondary comparison (docetaxel plus nintedanib) may be different from the population in the primary comparison.

The company did not perform a full incremental analysis to compare all three treatment strategies in this appraisal. In their response to the question of the ERG in the clarification phase whether this was because of non-matching populations, the company stated that *"a minority who are eligible for docetaxel will have an add-on nintedanib"* and that there was no easy way to produce a relative treatment effect between sotorasib and nintedanib plus docetaxel.[6]

The ERG considers that the absence of a full incremental analysis for the three treatment options negatively impacts the validity of the comparison and the generalisability of results to UK clinical practice.

4.2.4 Interventions and comparators

The intervention considered in the CS was sotorasib, a $KRAS^{G12C}$ inhibitor. Sotorasib is administered once daily as oral monotherapy, at a dose of 96 0mg (8x 120 mg tablets). The comparators considered were docetaxel monotherapy, or nintedanib for patients with adenocarcinoma. As discussed in Section 2.3, the NICE scope listed the following comparators:

Non-squamous NSCLC:

- pemetrexed with carboplatin
  with or without pemetrexed maintenance
- other platinum doublet chemotherapy
  with or without pemetrexed maintenance
- nintedanib with docetaxel (adenocarcinoma histology)
- docetaxel monotherapy
- atezolizumab
- nivolumab (subject to ongoing CDF review)
- pembrolizumab (PD-L1-expressing tumours)
- best supportive care

Squamous NSCLC:

- gemcitabine with carboplatin or cisplatin
- vinorelbine with cisplatin or carboplatin
- docetaxel monotherapy
- pembrolizumab (PD-L1-expressing tumours)

- atezolizumab
- nivolumab
- best supportive care

People with *KRAS p.G12C* mutation and another driver mutation (including EGFR-TK, ALK or ROS1):

- Established clinical management without sotorasib, including:
    - atezolizumab combination (after EGFR-TK or ALK-targeted therapies)
    - lorlatinib (after ALK-targeted therapies)
    - brigatinib (after ALK-targeted therapies)
    - ceritinib (after ALK-targeted therapies)
    - osimertinib (EGFR T790M mutation-positive after EGFR-TK targeted therapies)
    - pemetrexed with carboplatin
    - platinum doublet chemotherapy
    - with or without pemetrexed maintenance
    - nintedanib with docetaxel (adenocarcinoma histology)
    - nivolumab (subject to ongoing CDF review)

The company justified the limited number of comparators as follows:

- For immunotherapy and combination radiotherapy: re-challenge is not routine clinical practice according to clinical expert opinion obtained from a UK advisory board.
- Co-occurrence of KRAS p.G12C next to another driver mutations, is very rare (<1%).[25]
- Docetaxel monotherapy is considered a key second- and subsequent-line option in NSCLC.[26, 27]
- For adenocarcinoma patients eligible for docetaxel, a combination of nintedanib and docetaxel may be administered in some regions in the UK.

Additionally, the CS states that the use of docetaxel monotherapy as the comparator was agreed upon in scientific advice from NICE and EUnetHTA.[1]

In the company's response to clarification, the company mentioned that the PSWA of chemotherapy regimens from the Flatiron database compared to sotorasib, could be used as a proxy of using platinum doublet chemotherapy as a comparator.[6] These cost effectiveness results were explored in scenario analyses (Section B.3.7.3.1, Table 46 of the CS).[1]

Sotorasib dose reductions are recommended in case of adverse reactions. The first reduction brings the total dosage to 480 mg (four tablets) and the second reduction to 240 mg (two tablets), taken once daily. If patients are unable to tolerate 240 mg daily, treatment should be discontinued. Dose modifications related to adverse events are displayed in the draft SmPC provided by the manufacturer.[23]

ERG comment:

a) The number of comparators included in the cost effectiveness analysis is limited compared to the initial scope set out by NICE. Importantly, platinum-based chemotherapy is excluded, while it is considered a relevant comparator in 2nd line for those that have received immunotherapy only in 1st line. According to clinical expert opinion, this concerns about 40% of the patient population in the scope: a very significant minority.[8]
b) The ERG does not consider the suggestion made by the company in their response to clarification that Table 46 in the CS (the analysis using Flatiron data) could be used as a pragmatic reflection of sotorasib versus platinum-based chemotherapy, to be supported by the information presented in the

CS.[1, 6] No conclusions regarding the cost effectiveness of sotorasib versus platinum doublet therapy should be drawn from the analysis presented by the company.

Table 4.6: Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in 2nd line

Report Section	4.2.4
Description of issue and why the ERG has identified it as important	Compared to the final scope for this appraisal, platinum-based chemotherapy is excluded, while it is considered a relevant comparator in 2nd line for those that have received immunotherapy only in 1st line. According to clinical expert opinion, this concerns about 40% of the patient population in the scope: a very significant minority
What alternative approach has the ERG suggested?	The ERG has no alternative approach as adding the comparator to the model would require structural and substantial changes which are outside the scope of work for the ERG.
What is the expected effect on the cost effectiveness estimates?	Could potentially have a substantial impact on the cost effectiveness, direction unknown.
What additional evidence or analyses might help to resolve this key issue?	Implementing platinum-based chemotherapy in the model as an additional comparator would help to resolve the issue and reduce uncertainty.
ERG = Evidence Review Group	

4.2.5 Perspective, time horizon and discounting

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

ERG comment: In the CS, the company states a 20-year time horizon was used, at what point <1% of the patients is expected to be alive.[1] This was considered to represent a lifetime time horizon. The approach is in concordance with the NICE reference case.

4.2.6 Treatment effectiveness and extrapolation

The main source of evidence on treatment effectiveness for sotorasib is the CodeBreaK100 trial (updated from the initial submission to include data up to 15 Match 2021).[28]

Treatment effectiveness of the comparators is derived from the SELECT-1 trial for docetaxel and from the LUME-Lung 1 trial (adenocarcinoma subgroup) for nintedanib plus docetaxel. An additional analysis was provided using real-world data from the Flatiron cohort, in which a basket of standard-of-care chemotherapy was used. As no head-to-head trial was performed comparing sotorasib to its comparators, all analyses are indirect analyses, the methods of which are described in Section 3.4.

4.2.6.1 Sotorasib versus docetaxel

For the base-case estimation of the OS for sotorasib versus docetaxel, an HR of [redacted] was derived from the MAIC indirect analysis and an HR of [redacted] for PFS. Several parametric distributions were fit to the data from the CodeBreaK100 trial and the adjusted data from the SELECT-1 trial. In the CS a joint-fit restricted lognormal model fit best to the OS data, considering the Akaike information criterion (AIC) and Bayesian information criterion (BIC) and sensitivity analyses with the other models

were also provided. For the PFS, a restricted joint fit using a generalised gamma model fit best to the data when the AIC was considered and a lognormal fit best to the data when the BIC was considered.

4.2.6.2 Sotorasib versus nintedanib plus docetaxel

For the nintedanib plus docetaxel comparator, no patient-level data were available, instead, pseudo-patient level data was generated from the published results of the LUME-Lung 1 trial, which compared nintedanib plus docetaxel to placebo plus docetaxel. No MAIC was performed, as the patient population in the LUME-Lung 1 trial was deemed to differ too much from the CodeBreaK100 trial population. Nintedanib was modelled by applying time-dependent HRs to the data for docetaxel patients from SELECT-1. For the OS comparison between docetaxel plus nintedanib versus docetaxel plus placebo, piecewise HRs were used: for 0-6 months [redacted], for 6- to 26 months [redacted], and for 26 months and over [redacted]. Piecewise HRs were also considered for the PFS: for 0-2 months [redacted], for 2-6 months [redacted], and for 6 months and over [redacted].

4.2.6.3 Flatiron real-world data

An alternative analysis was provided in the CS using the Flatiron real-world dataset. In this analysis, the sotorasib data from the CodeBreaK100 trial was compared to a basket of standard-of-care chemotherapy: 21 out of 206 patients in this dataset were on docetaxel monotherapy. Also 85 out of 206 participants had a KRAS p.G12C mutation. Using a propensity score analysis described in Section 3.4, the HR for OS was estimated at [redacted] for the KRAS p.G12C mutant subgroup and the HR for the PFS was estimated at [redacted]. For both the OS and PFS a restricted joint fit lognormal model provided the best fit considering the AIC and BIC.

4.2.6.4 Waning of treatment effect

In the base-case of the CS, sotorasib was extrapolated for the full time horizon of the analysis.[1] In the original CS, a scenario analysis was provided to limit the treatment effect of sotorasib to 5 years and in the company's response to the ERG clarification questions, seven additional scenario analyses were provided.[1, 6] Two methods were used to incorporate treatment effect waning (TEW): gradual TEW and immediate TEW.

In the gradual TEW, the sotorasib effects gradually decrease for 5 years starting in year 2, 3, 4 or 5; for the immediate TEW, the sotorasib HRs were immediately set to 1 (meaning no benefit compared to the comparator) from year 2, 3, 4 and 5. In response to the request for clarification, the company noted that *"TEW is a very blunt tool and in an ideal world its use should be limited to cases where there is no (or very little) available external data to compare or adjust long term extrapolations with"* and provides several reasons why TEW should not be applied in this case.[6]

4.2.6.5 Treatment duration

TTD for sotorasib was estimated by applying an HR of [redacted] to the PFS curve. A sensitivity analysis was included where treatment discontinuation was modelled using separate parametric models, which is an approach in line with the methods used to model OS and PFS for sotorasib and docetaxel. According to the CS, for docetaxel no robust data were available, and TTD was assumed to be equal to PFS. For nintedanib plus docetaxel, treatment duration was also set to be equal to PFS, which is a conservative estimate according to the CS, as in a previous NICE submission (TA347) the PFS rate was higher than the discontinuation rate.[9]

4.2.6.6 ERG comment

a) The methods to extrapolate the treatment effect of sotorasib versus docetaxel using the CodeBreaK100 and SELECT-1 trial data are well explained in the CS and the decisions made are clear. It should be noted however, that the decision for a specific parametric model remains somewhat arbitrary and can have a major influence on the model outcomes. As provided in the company's response to clarification questions, the deterministic incremental cost effectiveness ratio (ICER) ranges from £30,112 to £62,123 per QALY depending on the chosen PFS and OS functions.

   As the presented OS and PFS curves are the results of an indirect analysis, the ERG expects additional uncertainty regarding the chosen model, mainly for the comparator. Since the generalised gamma distribution provided the best fit for PFS considering the AIC, the ERG considers the use of this distribution to be an important scenario to include; next to the base-case in which the lognormal distribution was used which provided the best fit considering the BIC.

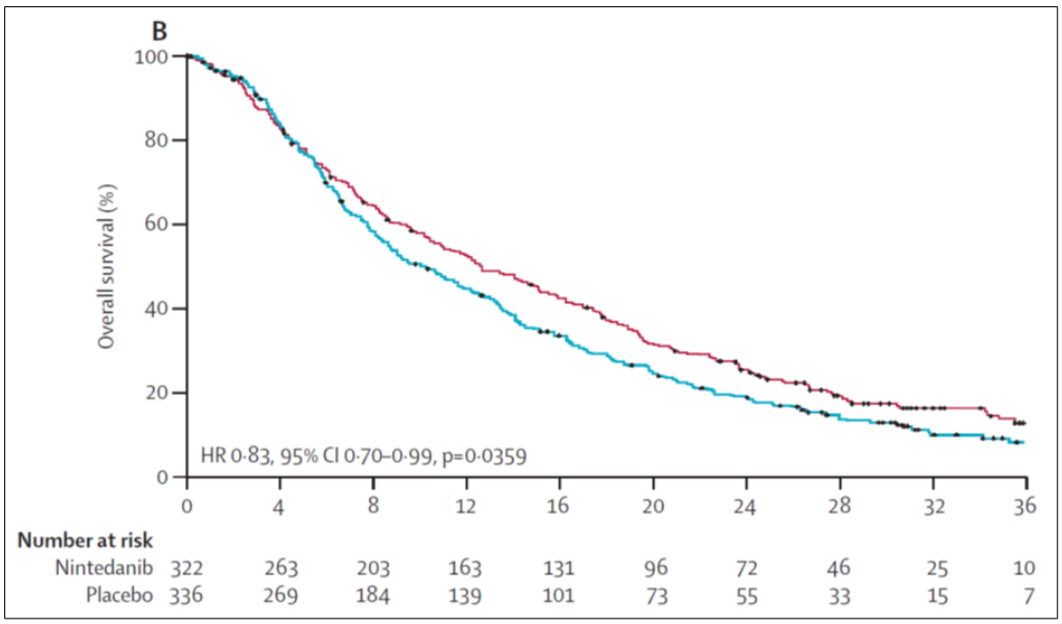
b) The modelling of nintedanib plus docetaxel is subject to considerable uncertainty: first, the LUME-Lung 1 trial data is used to compare docetaxel plus nintedanib to docetaxel plus placebo, then the resulting HRs are applied to the SELECT-1 data, which are then used for the indirect analysis using the same methods as for the sotorasib versus docetaxel comparison.

   In addition to the uncertainty introduced by this method, the ERG has major concerns regarding the clinical plausibility of the resulting OS curve. First of all, the patient populations of the SELECT-1 trial and LUME-Lung-1 trial differ mainly in terms of smoking and performance status (Table 10 of CS) and the CS does not report any adjustments for these differences.[1] Additionally, the resulting HR of [redacted] for the first 6 months results in a major rise in mortality (see Figure 4.2, copied from the economic model provided by the company). The ERG finds it implausible that adding nintedanib to docetaxel treatment would result in a major rise in mortality and does not consider the resulting OS curve to be in line with the Kaplan-Meier-curve reported in the LUME-lung-1 trial (see Figure 4.3). There was no expert opinion provided in the CS to support Figure 4.2. Note that the titles of Tables 30 and 31 in the CS contain an error, as the HRs provided are for docetaxel plus nintedanib versus docetaxel plus placebo; not for nintedanib plus docetaxel versus sotorasib.[1]

Figure 4.2: Modelled OS curves taken from the economic model

[redacted]

Figure 4.3: Reported OS Kaplan-Meier plot from LUME-Lung-1 trial for nintedanib plus docetaxel (red line) versus placebo plus docetaxel (blue line)



On visual inspection of the OS and PFS survival curves provided in the company's response to clarification questions, none of the fitted curves have a particularly good fit. A piecewise analysis was used, with two cut-off points, for the OS at 6 months and at 26 months. The ERG does not agree with the cut-off point at 26 months and the company failed to justify this approach both in the initial CS and in the company's response to clarification questions. The ERG suggests reducing the number of cut-off points to one at month 6.

c) The company did not consider any waning of the treatment effect, and in their response to the clarification questions, the company noted that *"TEW is a very blunt tool and in an ideal world its*

use should be limited to cases where there is no (or very little) available external data to compare or adjust long term extrapolations with".[6] According to the ERG, there is no external data available in this case, as the only data regarding the treatment effects of sotorasib come from the CodeBreaK100 trial, with a limited follow-up time and no comparators.

The ERG does agree with some of the points made in the company's response, e.g. that the impact of discontinuation is already somewhat "baked" into the model. On the other hand, it may not be reasonable to expect that patients continue to benefit from the treatment indefinitely, even after they have stopped treatment. Considering that only 18 months of CodeBreaK100 trial data have been collected and there is no additional information available for the sotorasib treatment effects beyond this, the ERG thinks it is a feasible approach to introduce a gradual TWE after 24 months, for which a waning period of 5 years can be used; the period is suggested in the company's response to clarification questions.

d) The TTD was modelled by applying an HR to PFS from CodeBreaK100. The company explored an alternative approach in a sensitivity analysis where the weights generated from the MAIC analysis were applied to the CodeBreaK100 discontinuation data and parametric models were fitted to extrapolate the treatment duration. However, the company considered this approach to be more complex and ultimately dependent on the variable selection in the MAIC analysis. The ERG feels it would have been more consistent to model the TTD in the same way that OS and PFS were modelled, i.e. based on MAIC. Also, by basing the TTD on the PFS, TTD would still be, via PFS, ultimately dependent on the variable selection in the MAIC. Moreover, in Figure 36 of the CS the company presents Kaplan-Meier data for TTD alongside the modelled curve and the ERG believes this to be a poor fit, potentially underestimating true TTD in the long run.

Table 4.7: Key issue 8. Docetaxel plus nintedanib modelling approach leading to worse survival

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The indirect way of estimating OS and PFS for the secondary comparator docetaxel plus nintedanib leads to worse survival for docetaxel plus nintedanib compared to docetaxel plus placebo in the first 6 months of the OS curve.
What alternative approach has the ERG suggested?	The ERG prefers to assume that the HR for docetaxel plus nintedanib versus docetaxel plus placebo cannot go above 1
What is the expected effect on the cost effectiveness estimates?	Lowering the HR for docetaxel plus nintedanib versus docetaxel plus placebo will increase the ICER for sotorasib versus docetaxel plus nintedanib
What additional evidence or analyses might help to resolve this key issue?	Direct evidence for this comparison
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio; OS = overall survival; PFS = progression-free survival	

Table 4.8: Key issue 9. No waning of treatment effect

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The company's assumption of continued effect of sotorasib does not seem justified and is difficult to maintain given immature evidence.

What alternative approach has the ERG suggested?	The ERG suggested to start waning of the treatment effect at the 2-year timepoint and have it gradually decreased to an HR of 1 over a period of 5 years (with exploratory scenario analyses for 3 and 7 years).
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on lasting treatment effect.
ERG = Evidence Review Group; HR = hazard ratio; ICER = incremental cost effectiveness ratio	

Table 4.9: Key issue 10. TTD modelling approach

Report Section	**4.2.6**
Description of issue and why the ERG has identified it as important	The TTD was modelled by applying a hazard ratio to PFS from CodeBreaK100. The ERG feels it would have been more consistent to model the TTD in the same way that OS and PFS were modelled, fitting a parametric curve on TTD data using weights based on the MAIC.
What alternative approach has the ERG suggested?	The ERG suggested to use the company's alternative approach, based on the MAIC, in the base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Mature data on observed treatment duration in sotorasib and comparator arms
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; MAIC = matching adjusted indirect comparison; OS = overall survival; PFS = progression-free survival; TTD = time to treatment discontinuation	

4.2.7 Adverse events

The company included grade 3+ adverse events with an incidence of ≥5% in any of the comparator arms in the analysis, considering data from the CodeBreaK100, SELECT-1 and LIME-Lung 1 trials.[1] Only TRAEs were included in the analyses, as only these were available from the LUME-Lung 1 trial. Treatment-emergent adverse events were included in a scenario analysis for the comparison of sotorasib to docetaxel.

Disutilities related to adverse events were included in the analysis, the values of which are provided in Table 36 of the CS.[1] If no disutility value could be identified, this was assumed to be 0. This is the case for: decreased neutrophils, increased AST, and pleural effusion.

ERG comment:

a) The inclusion of only TRAEs could negatively impact the validity of the assessment, as the quality of life of patients may not be captured well if TEAEs are excluded. However, the company provided

a scenario analysis including TEAEs for the comparison of sotorasib to docetaxel which increased the ICER from £43,660 to £44,116 per QALY, which the ERG considers a minor impact.

b) Disutilities were assumed to be 0 if no disutility value could be identified. The CS states that: *"This assumption could potentially be conservative given the generally increased frequency of these AEs in the comparator arms versus sotorasib"*.[1]
The ERG does not agree with this statement, although it seems to be reasonable for the nintedanib plus docetaxel comparison, it is not reflected by the data for docetaxel monotherapy. Within the sotorasib arm, the incidence of decreased neutrophils was 0.8% and the increased AST incidence was 5.6%, while this was 0.0% and 0.0% respectively for the docetaxel arm. As no disutility was applied to these adverse events, this is expected to favour the cost effectiveness of sotorasib compared to docetaxel. In contrast, as decreased neutrophils are more prevalent in the nintedanib plus docetaxel comparison, it may negatively impact the cost effectiveness of this comparator.

4.2.8 Health-related quality of life

The utility values were estimated for the following health states: progression-free, and post-progression, via a disutility subtracted from the progression-free utility. Notably, these health state utilities were only used in a sensitivity analysis as the approach taken in the base-case was to use time to death utilities.

4.2.8.1 Utility values

In the absence of studies from the SLR (see Section 4.1.3), the primary source of HRQoL values in the model was CodeBreaK100.[1] HRQoL was collected in CodeBreak100 using the EuroQoL-5D-5L instrument.[29] This instrument was completed on the first day of cycle 1, on every first day of subsequent cycles until cycle 7, and then on the first day of every second cycle until end of treatment. The company defined various datasets, see Table 4.10 for details. Using mixed models with repeated measures (MMRM), utilities were estimated using two approaches: time to death and health states. The analysis included several combinations of datasets and covariates.

Table 4.10: Datasets used for HRQoL analysis

	Original	AN01	AN02
Safety analysis set	N=126	N=122[#]	N=86
Full analysis set	N=123	N=119[*]	N=84

Based on pages 111 and 112 of the CS[1]
AN01 = patients who completed at least one EQ-5D-5L questionnaire in line with study protocol with all fields of the questionnaire completed; AN02 = patients who completed EQ-5D-5L at baseline visit per protocol and at least one other completed EQ-5D visit
[#] used for time to death utilities analysis in the model; [*] used for health state utilities analysis in the model
CS = company submission; EQ-5D-5L = European Quality of Life-5 Dimensions-5 Levels; HRQoL = health-related quality of life

Health-state utility values

For the health-state utilities, the CS presented results for both the AN01 and the AN02 full dataset.[1] Although the company did find that a model including both progression status and baseline utility score as covariates fitted best, this would require the use of the AN02 dataset since in the AN01 dataset not all subjects had completed the baseline questionnaire. And so, *"to account for all information available"*, as the company stated, the MMRM with only progression status based on AN01 was used to inform the model.[1]

Time-to-death utility values

For the time-to-death utility analysis, the AN01 dataset was used as well, but based on the safety analysis set instead of full analysis set, which implied a few additional subjects were included compared to the health state utility values, see Table 4.11.

The company provided mean utility scores visually in Figure 40 of the CS and the final time-to-death utility scores used to inform the model in Table 35 of the CS, which were updated in the addendum accompanying the response to clarification to reflect the latest data cut-off (15 March 2021).[1, 30] These updated time-to-death utilities were used in the company's base-case, and preferred over the health-state utility scores. The company stated this to be, amongst other reasons, because the time-to-death approach reflects the findings of studies which have shown NSCLC patients to have markedly decreased utilities towards the end of life.[1, 6]

A summary of all utility values used in the cost effectiveness analysis is provided in Tables 4.11 and 4.12.

Table 4.11: Health-state utility values - used in sensitivity analysis

Health state	Utility value (mean and 95% CI)	Reference
Progression-free	0.734 (0.700 to 0.769)	CodeBreaK100[30, 31a] and UK crosswalk tariffs[32]
Disutility in progressed disease	0.064 (0.097 to 0.031)	
Post-progression	0.670	Calculation

Based on addendum to clarification response, Table 7[30]

[a] Obtained from CodeBreaK100 Clinical Study Report, Tables 14n-4.7.701, 14n-47.702 and subsequent analyses

CI = confidence interval; UK = United Kingdom

Table 4.12: Time-to-death utilities - used in CS base case

Health state	Utility value (mean and 95% CI)	Reference
Utility more than 6 months to death	0.762 (0.698, 0.767)	CodeBreaK100[30, 31a] and UK crosswalk tariffs[32]
Disutility between 3 and 6 months to death (versus more than 6 months)	0.047 (0.090, 0.004)	
Disutility between 1 and 3 months to death (versus more than 6 months)	0.125 (0.176, 0.074)	
Disutility less than 1 month to death (versus more than 6 months)	0.233 (0.312, 0.153)	
Utility between 3 and 6 months to death	0.715	Calculated
Utility between 1 and 3 months to death	0.637	Calculated
Utility in last month of life	0.529	Calculated

Based on addendum to clarification response, Table 8[30]

[a] Obtained from CodeBreaK100 Clinical Study Report, Tables 14n-4.7.701, 14n-47.702 and subsequent analyses

CI = confidence interval; CS = company submission; UK = United Kingdom

4.2.8.2 Disutility values

In the absence of reported utility data for the comparators, the company included a disutility to express the implications of a hospital-based intravenous (IV) administration and increased cytotoxicity of docetaxel and nintedanib plus docetaxel.[1] The utility decrement was set at 0.025 (per cycle on treatment), based on a previous study in advanced NSCLC and disutility associated with IV administration.[33] This previous study (published in 2010) was on the cost effectiveness of erlotinib versus docetaxel and reported utilities of 0.451 and 0.426 for oral therapy and IV therapy respectively, in the progression free health state. These utilities were determined by having 154 members of the general population from four UK sites filling out a visual analogue scale (VAS).[33]

Disutilities of adverse events were discussed in Section 4.2.7.

The utilities in the economic model were not adjusted for age and sex. In response for a scenario analysis including age related decrements, the company added a scenario applying an adjustment to utilities based on the sex-matched general population utilities, to ensure that the estimated patient utilities never exceed that of the general population.[6] The company did however not apply an age-related decrement in this scenario since the TTD utility values were considered to already account for aging.

ERG comment: The main concerns of the ERG relate to a) the choice of TTD utilities over health state utilities as the TTD utilities seem less well informed; b) the treatment disutility applied for the comparators; and c) the absence of an age-related decrement:

a) Although using a time to death approach to utility scores may be justified in this population, the ERG has concerns about the data underlying the estimates used in the model. Firstly, by relying on the AN01 dataset, all patients that at least filled out one EQ-5D questionnaire were included in the analysis. Given that there were at maximum 14 timepoints available at which patients could have completed a questionnaire (see Figure 38 of the CS), one questionnaire seems the bare minimum and this raises questions about representativity of the sample.[1]
   In the AN02 dataset, patients had to have completed at least 2 questionnaires, one of which at the baseline visit. Using the AN02 dataset may have been more valid and stable, but the company discarded the AN02 *"to account for all information available"* even though the mixed model including baseline utility score as a covariate had a better fit than the model with progression status alone.
   Then, for the time to death analysis, the AN01 dataset was again preferred over the AN02 dataset, seemingly because the company wanted to align with the health state utility analysis (but nevertheless did decide to use the safety analysis dataset here instead of the full analysis dataset). Although for the health state utility approach some results of the AN02 dataset were presented, for the TTD approach no information on AN02 analyses were provided. In addition, the TTD utility scores presented in Table 35 of the CS and the final TTD utilities used in the model (see Table 4.12) above do not seem to match very well with the visual representation of mean utilities shown in Figure 40 of the CS.[1]
   Also apparent from Figure 40 of the CS is that numbers of distinct patient underlying the scores were quite small, i.e.., 86, 30, 31 and 12 for the more than 6 months, 3 months to 6 months, 1 month to 3 months, and less than 1 month to death categories.
   Altogether, the ERG considers the TTD utilities not reliable and therefore prefers the utilities by health state approach.
b) The disutility applied for IV administration of docetaxel. In the clarification phase, the ERG asked whether the disutility applied was appropriate for use in this case, given that the value for this disutility was derived by using a VAS instrument in a general population (so not EQ-5D, and

therefore not officially utilities) and the values obtained for the progression free health states in this study were vastly lower than observed in CodeBreaK100 (i.e. 0.426 and 0.451 compared to 0.74, respectively).[34] The company, in their response, made a case for treatment-specific utilities when comparing targeted therapy vs. chemotherapy.[6]

The ERG agrees that the use of treatment-specific utilities may be justified but considers the source used for disutility in the company base-case to be a questionable one. Apart from this, the company did not provide a response to the questions of the ERG how one day of IV infusion per 3 weeks would compare, in terms of quality of life, to taking eight tablets every day, as is the case for sotorasib treatment and for which no disutility was applied.. The company provided two alternative scenarios, which both effectively increase the disutility compared to the company base-case.[1, 6]

c) The ERG considers the fact that utilities in the model were not adjusted for age to be a potential source of bias. Although the company did provide a scenario where utilities in the model could not exceed the sex-adjusted utilities in the general population, the utilities in the model could then in theory still exceed the age-adjusted utilities in the general population, even though TTD utilities would decrease over time. The ERG would have liked to see a scenario as requested, including age-related utility decrements, to estimate the impact of such a scenario.

Table 4.13: Key issue 11. Time-to-death utilities do not seem well-informed

Report Section	**4.2.8**
Description of issue and why the ERG has identified it as important	The time to death utilities which the company used in the base-case did not seem well-informed. The data underlying the estimates were sparse, and increasingly so for the closer to death states.
What alternative approach has the ERG suggested?	The ERG suggested to use utilities based on disease progression as base-case.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	Fully specified models using also AN02 dataset should be provided to see which approach is most appropriate. But given that even AN02 probably has many missing data this may still not be ideal.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio	

Table 4.14: Key issue 12. Disutility for IV administration not well justified

Report Section	**4.2.8**
Description of issue and why the ERG has identified it as important	A disutility for IV administration of docetaxel is applied without sufficient justification for the size of the disutility or the exclusion of the potential disutility for taking eight tablets of sotorasib daily.
What alternative approach has the ERG suggested?	The ERG suggested to exclude the IV disutility in the base-case
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses	Comparative evidence on (observed) health state utilities in sotorasib and comparator arms could resolve this

Report Section	4.2.8
might help to resolve this key issue?	
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; IV = intravenous	

4.2.9 Resources and costs

The cost categories included in the model were treatment acquisition costs, medical costs (treatment administration, monitoring and disease management, subsequent treatments, and terminal care), and costs of managing adverse events.

Unit prices were based on the NHS reference prices, British National Formulary (BNF), Personal Social Services Research Unit (PSSRU) and Electronic Market Information Tool (eMIT).[35-38] All costs, where necessary were inflated to the 2018/2019 cost year to remain consistent with the latest available NHS Reference Costs using the PSSRU Hospital and Community Health Services (HCHS) and the NHS Cost Inflation Index inflation indices (NHSCII).[37]

4.2.9.1 Treatment costs (with patient access scheme (PAS))

Drug acquisition costs for the intervention and comparators are presented in Table 4.15. The sotorasib dose of 960 mg per day is consistent with the anticipated license and the dosing regimen in CodeBreaK100.[5, 39, 40] Dosage for docetaxel and docetaxel plus nintedanib is aligned with UK clinical practice and informed by NHS treatment protocols.[41]

Estimates of relative dose intensity (RDI) as observed in respective clinical trial programmes were applied to calculate total monthly costs.[9, 15, 39, 41, 42] RDI for sotorasib was slightly lower at 89% compared to docetaxel and nintedanib (90.3% and 921.1% respectively). In response to the request for clarification the company stated that there would be no reason to assume that RDI is truly lower for sotorasib, and that the differences in these observations may reflect random sampling error.[6]

Drug wastage was not discussed as such in the CS but from Table 48 in the CS it is apparent that the base-case assumption was zero wastage and that a scenario was run to test the impact of potential drug wastage in clinical practice by estimating drug acquisition costs based on total packs as opposed to treatments received.[1] In their response to the request for clarification, the company stated to maintain their base-case assumption of zero wastage.[6] This was justified with arguments on the ability to implement dose reductions and the single strength formulation of sotorasib, which would allow the pharmacist to optimise the dose without wastage and provide the appropriate supply of drugs to patients until disease progression is recorded. The company stated that they believe the scenario analysis including wastage would significantly overestimate the true drug utilisation.[6]

Table 4.15: Unit drug costs

Drug	Unit	Unit cost (£)	Reference	Dose	RDI	Cost per month (£)
Sotorasib	240x 120 mg tablets	[redacted]	[redacted]	960 mg per day	89%	[redacted]
Docetaxel	160 mg per vial	17.95	eMIT[38]	75 mg/m^2 on day of treatment	90.3%	19.93

Drug	Unit	Unit cost (£)	Reference	Dose	RDI	Cost per month (£)
Nintedanib	120x 100 mg tablets	2,151.10	BNF[36]	400 mg per day (21-day cycle)[a]	92.1%	1,926.28

Based on Table 38 of the CS[1]
[a] Nintedanib administered on days when docetaxel is not taken, i.e. 20 days per 21 day cycle
BNF British National Formulary; CS = company submission; eMIT = electronic Market Information Tool, RDI = relative dose intensity

Treatment administration costs were assumed to be zero for sotorasib and nintedanib as these are both taken orally. For docetaxel, administration costs were based on NHS reference costs for the administration of simple parenteral chemotherapy.[35] See also Table 38 in the CS.[1]

4.2.9.2 Health state and event costs

Costs of monitoring and disease management were largely based on assumptions used and accepted in previous NICE STAs, in particular NICE TA347 on nintedanib plus docetaxel.[9] Apart from a per-cycle cost per health state, a one-off cost was applied at treatment initiation and at progression. A one-off cost was also applied at the moment of dying to reflect the cost of terminal care, based on the values used in the NICE multiple technology appraisal (MTA) for erlotinib and gefitinib, see Table 4.16 for an overview.[43]

Table 4.16: Disease management and terminal care costs

		Source
Health state	**Cost per cycle (£)**	
Progression-free	77.04	NHS reference costs 2018/2019[35]; PSSRU[37]; aligned with NICE TA347[9] and TA428[44]
Post-progression	39.98	
Event	**Cost (£)**	
At treatment initiation	834.25	NHS reference costs 2018/2019[35]; PSSRU[37]; aligned with NICE TA347[9] and TA428[44]
At progression	116.53	
Terminal care	3,759.73	Appendix L of CS[13]

Based on Tables 39 and 42 of the CS[1]
CS = company submission; NHS = National Health Service; NICE = National Institute for Health and Care Excellence; PSSRU = Personal Social Services Research Unit; TA = Technology Appraisal

4.2.9.3 Adverse event costs

The unit costs related to the management of adverse events were mainly derived from a previous NICE MTA for erlotinib and gefitinib.[43]

4.2.9.4 Subsequent treatment costs

The costs of subsequent treatment were included in the economic model as a one-off cost at disease progression, see Table 4.17. The distribution of subsequent treatments for docetaxel and docetaxel plus nintedanib was informed by previous STAs as was treatment duration. The distribution of subsequent treatments for patients who progress on sotorasib was informed by UK clinical experts. In response to the request for clarification, the company also provided data on observed treatment mix in 44 patients

receiving subsequent treatments in CodeBreaK100, which revealed that patients would often receive more than one subsequent treatment, i.e. the 44 patients in the sample altogether received [redacted] subsequent treatments, see Table 4.18.[6]

Table 4.17: Subsequent treatment costs

Subsequent treatment / Original treatment	BSC	Platinum-based	Docetaxel	Source
Sotorasib (%)	50%	10%	40%	Assumption based on clinical expert feedback
Docetaxel (%)	70%	30%	0%	NICE TA 347 – assumption[9]
Nintedanib + docetaxel (%)	70%	30%	0%	NICE TA 347 – assumption[9]
Treatment duration (weeks)	14	14	14	NICE TA347, TA428[9, 44]
Cost of subsequent treatment (£)	0	2,835	1,219	Calculation – appendix L[13]

Based on Table 41 of the CS[1]
BSC = best supportive care; CS = company submission; NICE = National Institute for Health and Care Excellence; TA = Technology Appraisal

Table 4.18: Treatment mix for 44 patients receiving subsequent treatments

Treatment	N	Proportion of 77 treatments (%)	Proportion of 44 patients (%)
Pemetrexed or docetaxel	[redacted]	[redacted]	[redacted]
Platinum based chemotherapy	[redacted]	[redacted]	[redacted]
Others* or non-interventional therapy	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

Based on page 67 of the response to the request for clarification[6]
* Other includes novel treatments assessed in clinical trial settings and other treatments not relevant UK clinical practice or unknown
UK = United Kingdom

ERG comment: The main concerns of the ERG relate to:

a) The RDI of sotorasib being lower than for the comparators while the company have stated in their response to clarification that there is no reason to assume this. The ERG believes it would be reasonable to set the RDI for sotorasib, docetaxel, and docetaxel plus nintedanib at 90.5% which is the average of the observed RDIs for the three interventions considered.
b) The assumption of zero wastage, which the ERG considers to be overly optimistic. Although an oral drug at a fixed dose will be associated with less wastage than IV treatment which is dosed based on BSA, the ERG does not believe it to be likely that packs of sotorasib, once delivered to the patient and opened, will be returned and later administered to other patients. Hence, some wastage will always occur, no matter how precise and short-term the dosing.
c) Subsequent treatment costs for sotorasib are likely underestimated by assuming that patients would receive only one subsequent treatment while data from CodeBreaK100 suggests otherwise. The ERG considers the percentage of actual patients receiving docetaxel and platinum-based

chemotherapies is more relevant here than the mix between therapies. Notably, the percentage of patients receiving platinum-based chemotherapies may be underestimated in the model.

Table 4.19: Key issue 13. Relative dose intensity and wastage assumption not justified

Report Section	**4.2.9**
Description of issue and why the ERG has identified it as important	In their base-case, the company assumed a lower RDI for sotorasib than for comparators, which was not justified. The company also assumed zero wastage for sotorasib, which the ERG also considered not justified.
What alternative approach has the ERG suggested?	The ERG proposed to take the average RDI as base-case, and to include wastage based on opened packs.
What is the expected effect on the cost effectiveness estimates?	The ICER for sotorasib will increase
What additional evidence or analyses might help to resolve this key issue?	For the wastage, the company would have to make a convincing case that opened packs, when not used, would be returned for usage by other patients, i.e. a specific program would have to be in place.
ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; RDI = relative dose intensity	

5. COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The CS base-case cost effectiveness results from the updated model indicated that sotorasib is both more effective and more costly than docetaxel, which resulted in an ICER of £43,660 per QALY gained (Table 5.1).[1] When comparing sotorasib to the secondary comparator, docetaxel plus nintedanib, the deterministic ICER was £33,628 per QALY gained (with additional costs of £15,599, incremental QALYs 0.47 and life years gained (LYG) 0.61).

It should be noted that in the original CS, the ICER for sotorasib vs. docetaxel was £47,176 per QALY gained.[1] After updating the model with the new data cut-off point of 15 March 2021 for CodeBreaK100 (the original CS was submitted with data cut-off of 01 December 2020), the ICER decreased by 7.5%.[15] The increase in OS of sotorasib compared to docetaxel was the main driver for the lowered ICER compared to the original submission.

Table 5.1: Deterministic base-case results: sotorasib vs. docetaxel

	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	[redacted]				
Docetaxel	[redacted]	[redacted]	[redacted]	25,932	0.77	0.59	43,660
Based on updated company model ICER = incremental cost effectiveness ratio; LYG = life years gain; QALY = quality-adjusted life years							

Overall, the technology is modelled to affect QALYs by:

- Increasing survival, which accrues in PF ([redacted] vs [redacted] months) as well as in PP ([redacted] vs [redacted] months).
- Increased QoL because of the longer survival, and because of treatment-related disutility for docetaxel.

Overall, the technology is modelled to affect costs by:

- The higher cost of sotorasib compared to docetaxel ([redacted] vs. £17.95).
- Early treatment discontinuation for sotorasib compared to docetaxel.

ERG comment: To test the effect of extreme values on the model, the weight was set to zero and there were small changes on ICER. The reason was that the treatment of sotorasib was not dependant on weight. In the updated model after clarification, weight was removed.

5.2 Company's sensitivity analyses

The company performed and presented the results of probabilistic sensitivity analyses (PSAs), deterministic sensitivity analyses (DSAs) as well as scenario analyses.[1] The PSA included probabilistic parameters that were used to estimate QALYs and costs. The PSA was run for 1,000 iterations. The 2.5th and 97.5th percentiles for the probabilistic incremental costs and QALYs were £21,121 to £32,110 and 0.28 to 0.98, respectively. The PSA shows consistency with the deterministic results with an ICER of £43,183 per QALY gained. The probability of sotorasib being cost effective against docetaxel is [redacted] (Figure 5.1).

The PSA for the secondary comparator (docetaxel plus nintedanib) is more favourable towards sotorasib with a probability of being cost effective of [redacted]. The 2.5th and 97.5th percentiles for the probabilistic incremental costs and QALYs were £9,455 to £22,437 and -0.07 to 0.96, respectively. The probabilistic ICER of £33,368 per QALY gained is consistent with the deterministic results.

Based on the DSA of sotorasib versus docetaxel, the parameters that have the greatest effect on the ICER are the following:

- the HR applied to PFS to model sotorasib treatment duration (TTD)
- the time to death utility for >6 months prior to death
- disease management costs per week in the progression free health state

Based on the DSA of sotorasib versus docetaxel plus nintedanib, the parameters that have the greatest effect on the ICER are the following:

- HR of OS in the third period (from week 113 until week 261)
- HR of OS in the first period (up to week 26)
- HR of OS in the second period (from week 26 until week 113)

Consistently, modelling assumptions that relate to these parameters likely have the greatest effect on the ICER. This is illustrated by the following CS scenarios that have a substantial impact on the ICER (Table 48 of the CS):[1]

- Limiting treatment effects to 5 years (ICER: £46,684 per QALY gained)
- Applying health state utilities by progression status (ICER: £47,208 per QALY gained)
- Including drug wastage (ICER: £46,387 per QALY gained)
- Excluding RDI (ICER: £48,944 per QALY gained)
- MAIC-adjusted TTD curve from CodeBreaK100 (ICER: £44,496 per QALY gained)
- Generalised gamma distribution selected to estimate long-term PFS (ICER: £45,123 per QALY gained)
- Joint (unrestricted) lognormal distribution selected to estimate long-term PFS (ICER: £47,917 per QALY gained)

ERG comment:

a) Patient characteristics (age, sex, BSA) should not be included in PSA.
b) A scenario assuming TTD for sotorasib was equal to PFS (like for the comparators) was not included.

Figure 5.1: The cost effectiveness acceptability curve for sotorasib versus docetaxel

QALY = quality-adjusted life year; WTP = willingness to pay

5.3 Model validation and face validity check

Some aspects of validation were discussed by the company in the validation Section of the CS (Section B.3.9).[1] The clinical plausibility of the parametric models used was evaluated by comparing modelled median PFS and OS to the reported medians in the MAIC adjusted CodeBreaK100 trial and the docetaxel arm of the SELECT-1 study (CS Section B.3.9.1).[1] Also, the predicted OS landmark results at the 1 year, 5 year and 10 year points for the various parametric models were evaluated based on clinical expert opinion. The base-case jointly fitted (restricted) log-normal distribution was considered to be clinically valid for the population under consideration.

The real-world Flatiron Health database was used to test the robustness of the results generated by the MAIC. Using this data, the ICER of the base-case scenario would be £38,279 per QALY gained which is 12.3% less than the ICER of the base-case using CodeBreaK100 data.[1] The main difference was caused by the longer OS and PFS when using Flatiron instead of CodeBreaK100 (see Table 5.2). The company considered these results to be consistent with the conclusion of the MAIC analysis and underlining the robustness of the analyses presented.[1]

Lastly, quality control of the economic model was performed by systematic examination of calculations, extreme value analysis and tracing of calculations. The company used a verification checklist to guide this, but details of this checklist were not made available.

Table 5.2: Disaggregated results by using Flatiron and CodeBreaK100

	Flatiron		CodeBreaK100		Difference between Flatiron and CodeBrak100	
	Sotorasib	**Docetaxel**	**Sotorasib**	**Docetaxel**	**Sotorasib**	**Docetaxel**
PFS, mean (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
OS, mean (months)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LYG in PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LYG in OS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Costs (£)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
ICER(£/QALY)	38,279		43,660		-5,381	

Based on updated company model
ICER = incremental cost effectiveness ratio; LYG = life year gained; OS = overall survival; PFS = progression-free survival; QALY = quality-adjusted life year

ERG comment: The ERG considers the validation as described by the company to be minimal. As discussed in Section 4.2.2, the company did not provide a scenario with a state transition model (STM) and so validating the model structure in this way was not possible. The Flatiron analysis shows some rather distinct changes (in PFS, OS) for mainly the docetaxel arm compared to the CodeBreaK100 analysis.

In the absence of suitable clarification for these differences, the ERG does not agree with the company that the results from the Flatiron analysis underline the robustness of the analyses presented. However, if there is a lack of correspondence between the results based on the MAIC (using SELECT-1) and the PSWA (using Flatiron), this might be because the latter provides estimates that are less biased, although, as discussed in Section 3.4.2, this is very uncertain.

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:[45]

- 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 & 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):[46]

- 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 Explanation of ERG base-case

Adjustments made by the ERG, to derive the ERG base-case (using the CS base-case as a starting point) are listed below. Section 6.2 shows the impact of each adjustment plus the combined effect of all abovementioned adjustments simultaneously, in the deterministic, probabilistic and scenarios analyses.

6.1.1.1 Fixing errors

No errors were found in the CS model.

6.1.1.2 Fixing violations

1. Patient characteristics included in the PSA (Section 5.2).
   The ERG corrected this.

6.1.1.3 Matters of judgement

2. Key issue 10 (Section 4.2.6)
   TTD modelling approach for sotorasib: the ERG used the approach based on the MAIC fitting parametric models instead of HR applied to PFS.

3. Key issue 11 (Section 4.2.8)
   Method for health state utilities; the ERG used utilities based on disease progression instead of time to death.
4. Key issue 13 (Section 4.2.9)
   Relative dose intensity (RDI); the ERG assumed these to be equal (at average) for all interventions.
5. Key issue 13 (Section 4.2.9)
   Method to calculate treatment costs: the ERG preferred to calculate treatment costs on a per-opened-pack basis.
6. Distribution of subsequent treatments (Section 4.2.9)
   The ERG changed the distribution of subsequent treatments based on total patients receiving these.
7. Key issue 12 (Section 4.2.8)
   The ERG excluded the utility decrement for IV infusion.
8. Key issue 9 (Section 4.2.6)
   The ERG implemented a limit to the treatment effect at the 2 year timepoint, with a subsequent gradual waning of the effect over 5 years.
9. Key issue 8 (Section 4.2.6)
   For the secondary comparison (docetaxel plus nintedanib), the ERG assumed OS for docetaxel plus nintedanib could not be worse than OS for docetaxel plus placebo (i.e. where HR exceeded 1 it would be set equal to 1).

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. The main concern for the ERG was extrapolating the effectiveness of sotorasib and the comparators.

6.1.2.1 Exploratory scenario analyses

1. AE disutilities (Section 4.2.7)
   For AEs where disutility was zero, a disutility of 0.05 was assumed.
2. Treatment-emergent vs. treatment-related AEs (Section 4.2.7)
3. PFS distribution (Section 4.2.6)
   Assuming a generalised gamma distribution instead of lognormal distribution for PFS.
4. Gradual treatment waning (Section 4.2.6)
   Assuming gradual waning of treatment effect over 3 years (instead of 5 years).
5. Gradual treatment waning (Section 4.2.6)
   Assuming gradual waning of treatment effect over 7 years (instead of 5 years)
6. Piecewise HR for docetaxel plus nintedanib vs. docetaxel (Section 4.2.6)
   Assuming constant HR of OS and PFS for nintedanib from the second period (from week 113) onwards

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
Key issue 6. Model structure	4.2.2	Methods	State transition model to validate current PSM results	+/-	No	No
Key issue 7. Exclusion of platinum-based chemotherapy as a comparator in the 2nd line	4.2.4	Methods	Amend model to include platinum-based comparator	+/-	No	Yes
Key issue 8. Docetaxel plus nintedanib modelling approach	4.2.6	Bias & indirectness	Assumed that HR of docetaxel plus nintedanib versus docetaxel cannot exceed 1	+	Partly	Yes
Key issue 9. Treatment waning	4.2.6	Unavailability – immature data	Assumed gradual waning of treatment effect over 5 years, starting at 2-year timepoint	+	Partly	Yes
Key issue 10. TTD modelling approach	4.2.6	Bias & indirectness	Assumed alternative approach using MAIC and parametric distributions	+	Partly	Yes
Key issue 11. Health-related quality of life approach	4.2.8	Unavailability/missing data/small sample sizes	Assumed utilities based on disease progression	+	Partly	Yes
Key issue 12. Disutility for IV infusion	4.2.8	Unavailability of comparative HRQoL data	Excluded disutility	+	Partly	Yes
Key issue 13. RDI and wastage assumption	4.2.9	Unavailability of evidence for the company's assumption	Equal RDI and costs based on opened packs	+	Partly	Yes

[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
ERG = Evidence Review Group; HR = hazard ratio; HRQoL = health-related quality of life; ICER = incremental cost effectiveness ratio; IV = intravenous; RDI = relative dose intensity; TTD = time to treatment discontinuation						

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

6.2.1 The results of deterministic ERG preferred base case scenario

In Section 6.1, the ERG base-case was presented, 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 for the primary comparator (docetaxel). The largest impact on the ICER was caused by limiting the treatment effect of sotorasib at 2 years with a gradual waning effect over 5 years after (MJ 8), which resulted in a 10.7% increase of the ICER compared to the CS base-case (£48,332 per QALY vs. £43,660 per QALY), mainly due to a decrease in LYG. The ERG base-base, combining all proposed adjustments, was 33.8% higher than the CS base-case (£58,415 per QALY vs. £43,660 per QALY). The main reasons for this difference were higher drug acquisition costs for sotorasib ([redacted] vs. [redacted]) and lower LYG in the post-progression health state ([redacted]).

The impact of each individual change and the combined effect of all changes simultaneously for the secondary comparator (docetaxel + nintedanib) was presented in Table 6.3. Changing the HR for OS (MJ 9) had the largest impact on the ICER, increasing it with 33.7% compared to the CS base-case (£44,969 per QALY vs. £33,628 per QALY). The ERG base-case ICER was 54.8% higher than the CS base-case (£52,051 per QALY vs. £33,628 per QALY). The main reasons for this difference were higher drug acquisition costs for sotorasib ([redacted] vs [redacted]) and lower LYG gained in the post-progression health state ([redacted]) for docetaxel + nintedanib.

Table 6.2: ERG base-case adjustments (comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	25,932	0.59	43,660
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	26,369	0.59	44,394
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	26,429	0.59	44,496
MJ 4: Including drug wastage (key issue 13)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	27,552	0.59	46,387
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	25,932	0.55	47,208
MJ 6: Subsequent treatments based on alternative distribution					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	26,031	0.59	43,825
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	25,932	0.58	44,339
MJ 8: gradual waning of treatment effect over 5 yrs, starting at 2-year timepoint (key issue 9)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	25,788	0.53	48,332
ERG base-case					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,466	0.49	58,415
Based on CS updated model CS = company submission; ERG = Evidence Review Group; FV = fixing violations; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation					

Table 6.3: ERG base-case adjustments (comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
FV 1: Excluding patients' characteristics from PSA					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,699	0.47	33,628
MJ 2: Assuming equal RDI (90.5%) for all technologies (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,297	0.47	34,909
MJ 3: Assuming parametric distribution for TTD of sotorasib (key issue 10)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,195	0.47	34,692
MJ 4: Including drug wastage (key issue 13)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,186	0.47	34,673
MJ 5: Using health state utilities instead of time to death category (key issue 11)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,699	0.44	35,990
MJ 6: Subsequent treatment based on alternative distribution					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,797	0.47	33,839
MJ 7: Exclude utility decrement for IV infusion (key issue 12)					
Docetaxel + nintedanib	[redacted]	[redacted]			

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
Sotorasib	[redacted]	[redacted]	15,699	0.46	34,087
MJ 8: gradual waning of treatment effect over 5 yrs, starting at 2-year timepoint (key issue 9)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,697	0.47	33,618
MJ 9: Assuming HR of 1 for OS for nintedanib for the first period (key issue 8)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	15,386	0.34	44,969
ERG base-case					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,012	0.33	52,051

Based on CS updated model

CS = company submission; ERG = Evidence Review Group; FV = fixing violations; HR = hazard ratio; ICER = incremental cost effectiveness ratio; IV = intravenous; MJ = matter of judgment; OS = overall survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year; RDI =relative dose intensity; TTD = time to treatment discontinuation

6.2.2 The results of ERG sensitivity and scenario analyses

The sensitivity and scenario analyses were performed based on the ERG base-case. The results of the PSA for the ERG base-case were in line with deterministic results for both primary and secondary comparator (Tables 6.4 and 6.5). The probability of sotorasib being cost effective against docetaxel and docetaxel + nintedanib was [redacted], respectively. Figures 6.1 to 6.4 show the cost effectiveness plane and cost effectiveness acceptability curve for the docetaxel and docetaxel + nintedanib comparisons.

- The first ERG scenario had a small impact on ICER compared to ERG base-case for both comparators. The reason was that adding disutility of "decreased neutrophils" and "increased aspartate aminotransferase", led to a very minor decrease in incremental QALYs.
- The second ERG scenario with assuming treatment-emergent instead of TRAEs, resulted in slightly higher ICERs compared to the ERG base-case.
- The third ERG scenario slightly increased the ICER for both comparators. The incremental QALYs did not change in this scenario compared to the ERG base-case, however, people spend more time on treatment which led to an increase in costs.
- The fourth and fifth ERG scenario explored different periods for the waning effect of sotorasib. The ICER of sotorasib vs. docetaxel increases when shortening the waning period to 3 years and decreases when applying a waning effect over 7 years. For the secondary comparator however (sotorasib vs. docetaxel plus nintedanib) both the 3 year and the 7 year scenario result in an increase in the ICER. The reason for this is that in the company model, and also in the ERG analysis, the waning effect is applied to nintedanib as well.
- The sixth ERG scenario was explored only for the secondary comparator (docetaxel plus nintedanib). Since in this scenario the HR for OS of docetaxel plus nintedanib versus docetaxel is on average higher than the ERG base case ([redacted] vs. [redacted]), the QALYs decreased slightly for docetaxel plus nintedanib which made the ICER just below £50,000 per QALY.

Table 6.4: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	27,976	0.49	57,567
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,466	0.49	58,444
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,715	0.49	58,986
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	29,635	0.49	60,809
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,419	0.47	60,428
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	28,497	0.50	57,206

Based on CS updated model

AE = adverse event; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; PFS = progression free survival; PSA = probabilistic sensitivity analysis; QALY = quality-adjusted life year

Table 6.5: Probabilistic sensitivity analysis (PSA) and deterministic scenario analyses (conditional on ERG base-case, comparator: docetaxel + nintedanib)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG base-case (PSA)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	16,664	0.33	50,249
ERG scenario 1: Disutility of 0.05 for "decreased neutrophils" and "increased aspartate aminotransferase" for AEs with disutility of zero					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,012	0.33	51,874

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER (£/QALY)
ERG scenario 2: Treatment emergent AEs (instead of treatment-related)					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,214	0.33	52,733
ERG scenario 3: Assuming generalised gamma distribution instead of lognormal distribution for PFS					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,244	0.33	52,851
ERG scenario 4: Assuming gradual waning of treatment effect (after 2 years) over 3 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,010	0.33	52,179
ERG scenario 5: Assuming gradual waning of treatment effect (after 2 years) over 7 years					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,012	0.33	52,074
ERG scenario 6: Assuming constant HR of OS and PFS for nintedanib from 2nd period onwards					
Docetaxel + nintedanib	[redacted]	[redacted]			
Sotorasib	[redacted]	[redacted]	17,059	0.34	49,664

Based on CS updated model

AE = adverse event; CS = company submission; ERG = Evidence Review Group; ICER = incremental cost effectiveness ratio; PFS = progression free survival; QALY = quality-adjusted life year

Figure 6.1: Cost effectiveness plane for ERG base-case (Comparator: docetaxel)

ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.2: Cost effectiveness acceptability curve for ERG base-case (comparator: docetaxel)

Figure 6.3: Cost effectiveness plane for ERG base-case (Comparator: docetaxel + nintedanib)

ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

Figure 6.4: Cost effectiveness acceptability curve for ERG base-case (comparator: docetaxel + nintedanib)

ERG = Evidence Review Group; QALY = quality-adjusted life year; WTP = willingness-to-pay

6.3 ERG's preferred assumptions

In Section 6.2, the results based on the ERG preferred assumptions were presented. The estimated ERG base-case ICERs were £58,415 and £52,051 per QALY gained for sotorasib versus docetaxel and docetaxel + nintedanib, respectively which was 33.8% and 54.8% higher than the CS base-case. The probabilistic ERG base-case analyses indicated that the probability of sotorasib being cost effective was [redacted] against docetaxel and docetaxel + nintedanib, respectively, at a willingness-to-pay (WTP) threshold of £50,000 per QALY gained. Comparing sotorasib to docetaxel, the most influential adjustment in the ERG base-case was limiting the treatment effect to 2 years with a waning effect over 5 years. Comparing sotorasib to docetaxel + nintedanib, the most influential adjustment was setting the HR of OS to 1 for docetaxel plus nintedanib versus docetaxel. Concerning exploratory scenarios, using a generalised gamma distribution for PFS was the most influential scenario, driving the ICER upwards, for both comparisons.

6.4 Conclusions of the cost effectiveness Section

As discussed in Section 4.1.1, the search strategy for the population focused specifically on KRAS mutated NSCLC and may have been too narrow to identify all relevant studies for cost effectiveness, HRQoL and cost and resource use.

Separate sets of searches were conducted to identify cost effectiveness studies, HRQoL studies and healthcare resource use evidence. The eligibility criteria used by the company provided sufficient detail and were suitable to fulfil the company's objective to identify cost effectiveness studies.

The CS was largely in line with the NICE reference case. The CS partly deviated from the scope, however, where it concerned the comparators modelled. More specifically, platinum-based chemotherapy in the 2nd line was excluded as a comparator, while expert opinion indicated that it is a relevant treatment option for a substantial part of the population. Also, the company did not perform a full incremental analysis but instead presented two pairwise comparisons.

Although the ERG agreed that a partitioned survival model seemed appropriate for the decision problem, they would have liked to see a state transition model as a scenario to validate the results of the company's partitioned survival model.

The ERG considered the absence of any waning of the treatment effect in the company model not well justified. Data from the CodeBreaK100 trial are not sufficiently mature to assume a continuous effect of sotorasib. Given the available follow-up in CodeBreaK100 of 18 months (with many patients censored) the ERG believes that implementing a gradual waning of the treatment effect over 5 years, starting from the 2 year point, is a fair and maybe even already optimistic scenario.

The ERG was concerned about the approach taken to estimate treatment duration. Instead of taking a similar approach as for OS and PFS, TTD was linked to PFS via a HR. The ERG was not convinced by the rationale of the company to choose this approach and felt it more consistent to take the same approach for TTD as for OS and PFS, which is to fit parametric models to CodeBreaK100 discontinuation data (weighted based on the MAIC).

A major concern of the ERG was the validity of the modelling approach in the secondary comparison, sotorasib versus docetaxel plus nintedanib. The two-step approach taken potentially introduces bias, of which the fact that modelled OS for docetaxel plus nintedanib was initially below OS for docetaxel may be only one symptom. The ERG believes that the docetaxel plus nintedanib comparison is subject to large uncertainty, beyond what the ERG was able to take into account in their ERG base-case analysis.

Some comments on the incorporation of adverse events in the economic model were made by the ERG, but these could be resolved in the ERG analyses and were of minor importance for overall cost effectiveness results.

With respect to the implementation of health state utility values in the mode, the ERG had some major concerns. Firstly, the datasets used (AN01 and AN02) contained a much smaller number of EQ-5D observations than could have been expected based on the sample size and number of timepoints available for collecting these data, and so the mixed models were based on a sample that may not be representative of the whole population. Furthermore, the time-to-death approach based on AN01 preferred by the company was not justified by statistical arguments, while results for alternative approaches (with AN02 data for instance) were not presented. And because of the preferred time-to-death utilities, the company considered it not necessary to apply an age-related decrement. The company then also proposed to apply a disutility for IV infusion of docetaxel but did not discuss the potential disutility of having to take eight tablets daily for sotorasib. The ERG considered this approach altogether not well justified and feels that substantially more evidence on comparative HRQoL is necessary to be able to resolve these issues.

The ERG considered the company's assumption of no wastage for sotorasib to be unrealistic. Without a specific program in place that would guarantee that opened packs could be returned by the patient and then used by another patient, the cost calculation based on opened packs seems closest to daily practice. The values for RDI and subsequent treatments were deemed to slightly favour sotorasib while not entirely justified, so the ERG adjusted these to be more conservative. For a reliable estimate of subsequent treatments provided after sotorasib, more evidence is warranted.

The ERG made various adjustments to the company base-case. The probabilistic ERG base-case ICER for sotorasib versus docetaxel was [redacted] per QALY gained (based on 1,000 iterations). For sotorasib versus docetaxel plus nintedanib, the ICER was [redacted]. The most influential scenario for both comparators was where the generalised gamma distribution for PFS was used instead of the lognormal distribution, driving the ICER upwards.

In conclusion, cost effectiveness estimates of sotorasib compared with docetaxel and with docetaxel plus nintedanib are subject to considerable uncertainty, mainly because of immaturity of data and lack of comparative evidence in various areas. Even when all the ERG preferred assumptions were implemented in the model, uncertainty remained on a number of issues, such as whether all relevant comparators were included in the analysis, treatment duration and long-term efficacy of sotorasib, and comparative HRQoL values. The comparison for docetaxel plus nintedanib is potentially more heavily biased even because of the indirectness of the two-step approach to model OS and PFS.

7. END OF LIFE

According to the CS, sotorasib in its full anticipated licensed indication as a second- or subsequent line therapy meets the NICE criteria for an end of life medicine, see Table 7.1.[1]

Table 7.1: End of life criteria

Criterion	Data available	Reference in CS (Section and page number)
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	Large real world evidence studies indicate that that OS with non-targeted 2nd line therapies is <10 months, and with 3rd line therapies is <7 months. OS with 2nd line docetaxel monotherapy in the SELECT-1 study was 7.9 months.[16] OS with 2nd line nintedanib plus docetaxel in the LUME-Lung 1 study was 10.9 months.[17]	Section B.1.3.1.2, pages 19-21
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	A robust MAIC indicates sotorasib provides at least an additional [redacted] in median OS compared with docetaxel monotherapy based on available trial data. The economic model estimates that sotorasib plausibly provides an additional undiscounted mean OS of [redacted] months compared with docetaxel monotherapy and [redacted] months compared with nintedanib plus docetaxel*.	Section B.2.9.4.1, page 57 Section B.2.9.4.2, page 60

Based on Table 20 of the CS[1]
* Derived from economic model with 20-year time horizon, values undiscounted (see Section B.3.3.5 of the CS for how comparison of sotorasib vs nintedanib plus docetaxel is implemented)
CS = company submission; MAIC = matching adjusted indirect comparison; NHS = National Health Service; OS = overall survival

ERG comment: The ERG considers the first criterion, life expectancy less than 24 months, to be met.

Regarding the second criterion, extension of life of ≥3 months, the ERG agrees that, based on the data cited by the company, the criterion has been met. However, as discussed in Section 3.3 and 3.4, the ERG has concerns regarding the validity of the indirect comparisons referred to by the company, see key issue 5.

8. REFERENCES

[1] Amgen Ltd. *Sotorasib for previously treated KRAS p.G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: submission to National Institute of Health and Care Excellence. Single technology appraisal (STA)*. Cambridge: Amgen, 2021. 140p.

[2] National Institute for Health and Care Excellence. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer. Final scope (Appendix B) [Internet]*. London: NICE, 2021 [accessed 20.7.21]. 6p. Available from: https://www.nice.org.uk/guidance/gid-ta10639/documents/final-scope

[3] Sebastian M, Eberhardt WEE, Hoffknecht P, Metzenmacher M, Wehler T, Kokowski K, et al. KRAS G12C-mutated advanced non-small cell lung cancer: a real-world cohort from the German prospective, observational, nation-wide CRISP Registry (AIO-TRK-0315). *Lung Cancer* 2021;154:51-61.

[4] Hong DS, Fakih MG, Strickler JH, Desai J, Durm GA, Shapiro GI, et al. KRAS(G12C) Inhibition with Sotorasib in Advanced Solid Tumors. *N Engl J Med* 2020;383(13):1207-1217.

[5] Skoulidis F, Li BT, Dy GK, Price TJ, Falchook GS, Wolf J, et al. Sotorasib for lung cancers with KRAS p.G12C mutation. *N Engl J Med* 2021;384(25):2371-81.

[6] Amgen Ltd. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: response to request for clarification from the ERG*: Amgen, 2021. 77p.

[7] Amgen. A phase 1/2, study evaluating the safety, tolerability, PK, and efficacy of AMG 510 in subjects with solid tumors with a specific KRAS mutation (CodeBreaK 100). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2018- [cited 2021 Sep 1]. Available from: http://clinicaltrials.gov/show/NCT03600883. NLM Identifier: NCT03600883

[8] Clark P. *ID3780 Sotorasib for previously treated KRAS G12C NSCLC - comparator query. Clinical expert opinion [Personal communication]*.

[9] 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 TA347 [Internet]*. London: NICE, 2015 [accessed 13.8.21]. 49p. Available from: www.nice.org.uk/guidance/ta347

[10] Schulz C, Gandara D, Berardo CG, Rosenthal R, Foo J, Morel C, et al. Comparative Efficacy of Second- and Subsequent-line Treatments for Metastatic NSCLC: A Fractional Polynomials Network Meta-analysis of Cancer Immunotherapies. *Clin Lung Cancer* 2019;20(6):451-460 e5.

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

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

[13] Amgen Ltd. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: appendices*. Cambridge: Amgen, 2021. 156p.

[14] 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.

[15] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC With KRAS p.G12C mutation (CodeBreaK 100). Clinical study report CodeBreak100 efficacy data (01 December data cut). Data on file [PDF provided by the company]*. Thousand Oaks (CA): Amgen, 17 February 2021. 105p.

[16] Janne PA, van den Heuvel MM, Barlesi F, Cobo M, Mazieres J, Crino L, et al. Selumetinib Plus Docetaxel Compared With Docetaxel Alone and Progression-Free Survival in Patients With KRAS-Mutant Advanced Non-Small Cell Lung Cancer: The SELECT-1 Randomized Clinical Trial. *JAMA* 2017;317(18):1844-1853.

[17] Reck M, Kaiser R, Mellemgaard A, Douillard J-Y, Orlov S, Krzakowski M, et al. Docetaxel plus nintedanib versus docetaxel plus placebo in patients with previously treated non-small-cell lung cancer (LUME-Lung 1): a phase 3, double-blind, randomised controlled trial. *Lancet Oncol* 2014;15(2):143-155.

[18] Boehringer Ingelheim Limited. Vargatef 100 mg soft capsules [Internet]. electronic Medicines Compendium (eMC), 2021 [accessed 1.9.21]. Available from: https://www.medicines.org.uk/emc/product/3647

[19] Amgen Inc. *Data on File: Amgen NSCLC UK Advisory Board - 11th February 2021*, 2021

[20] Amgen Inc. *Data on file: a retrospective cohort study of patient characteristics and treatment outcomes among patients with advanced non-small cell lung cancer in the Flatiron Health-Foundation Medicine Clinico-Genomics Database (Study 20200132)*, 2020

[21] 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, 2015 [accessed 1.9.21] Available from: http://www.nicedsu.org.uk

[22] Woods B, Sideris E, Palmer S, Latimer N, Soares M. *NICE DSU technical support document 19. Partitioned survival analysis for decision modelling in health care: a critical review [Internet]*. Sheffield: Decision Support Unit, 2017 [accessed 9.7.21]. 72p. Available from: http://nicedsu.org.uk/wp-content/uploads/2017/06/Partitioned-Survival-Analysis-final-report.pdf

[23] Amgen Inc. *Draft Summary of Product Characteristics: LUMYKRAS (Sotorasib)*, 2021

[24] Amgen Inc. *Data on File: CodeBreaK100 Safety data (01 December 2020 data cut)*. Thousand Oaks (CA): Amgen, 2021

[25] Aggarwal S, Hsu H, Tu H, Carrigan G, Wang X, Ngarmchamnanrith G, et al. Clinicopathological characteristics, treatment patterns, and outcomes in patients with KRAS p.G12C mutant advanced non-small cell lung cancer in the Flatiron Health-Foundation Medicine Clinico-Genomic Database. Poster presented at the European Society for Medical Oncology (ESMO) Virtual Congress; 19-21 September 2020. 2020.

[26] National Institute for Health and Care Excellence. *Lung cancer: diagnosis and management. NICE guideline NG122 [Internet]*. London: NICE, 2019 [accessed 20.7.21]. 41p. Available from: www.nice.org.uk/gidance/ng122

[27] National Institute for Health and Care Excellence. Lung cancer overview [Internet]. London: NICE, 2021 [accessed 13.8.21]. Available from: https://pathways.nice.org.uk/pathways/lung-cancer

[28] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Clinical study report CodeBreak100 full primary analysis Mar 2021 data cut. Data on file* Thousand Oaks (CA): Amgen, 15 March 2021

[29] EuroQoL. EuroQol Instruments - EQ-5D-3L, EQ-5D-5L, EQ-5D-Y [Internet]. 2021 [accessed 30.4.21]. Available from: https://euroqol.org/

[30] Amgen Ltd. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: addendum. March 2021 data cut updates*. Cambridge: Amgen, 2021. 25p.

[31] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Clinical study report supplementary PROs primary analysis (September 2020 data cut). Data on file [PDF provided by the company]*. Thousand Oaks (CA): Amgen, 12 November 2020. 81p.

[32] van Hout B, Janssen MF, Feng YS, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health* 2012;15(5):708-15.

[33] Lewis G, Peake M, Aultman R, Gyldmark M, Morlotti L, Creeden J, et al. Cost-effectiveness of erlotinib versus docetaxel for second-line treatment of advanced non-small-cell lung cancer in the United Kingdom. *J Int Med Res* 2010;38(1):9-21.

[34] National Institute for Health and Care Excellence. *Sotorasib for previously treated KRAS G12C mutated, locally advanced or metastatic non-small-cell lung cancer [ID3780]: Clarification letter*. London: NICE, 2021

[35] NHS. *National Cost Collection for the NHS [Internet]*. London: NHS, 2021 [accessed 13.8.21] Available from: https://www.england.nhs.uk/national-cost-collection/#ncc1819

[36] Medicines Complete. British National Formulary [Internet]. London: Medicines Complete, 2021 [accessed 13.8.21]. Available from: https://about.medicinescomplete.com/publication/british-national-formulary/

[37] Personal Social Services Research Unit. *Unit costs of health and social care 2019 [Internet]*. Canterbury: PSSRU, University of Kent, 2019 [accessed 25.8.21] Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2019/

[38] Department of Health and Social Care. Drugs and pharmaceutical electronic market infrmationo tool (eMIT) [Internet]. London: Department of Health and Social Care, 2021 [accessed 13.8.21]. Available from: https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit

[39] Amgen Ltd. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Clinical study report CodeBreak100 full primary analysis Sept 2020 data cut. Data on file [PDF provided by the company]*. Thousand Oaks (CA): Amgen, 10 November 2020. 107p.

[40] Amgen Inc. *A phase 1/2, open-label study evaluating the safety, tolerability, pharmacokinetics, pharmacodynamics, and efficacy of AMG 510 monotherapy in subjects with advanced solid tumors with KRAS p.G12C mutation and AMG 510 combination therapy in subjects with advanced NSCLC with KRAS p.G12C mutation (CodeBreak 100). Amgen study 20170543 protocol*. Thousand Oaks (CA): Amgen, 2018

[41] Northern Cancer Alliance. Lung chemotherapy protocols and prescriptions [Internet]. Newcastle Upon Tyne: Norther Cancer Alliance, 2021 [accessed 13.8.21]. Available from: https://northerncanceralliance.nhs.uk/advisory_group/lung-expert-advisory-group/lung-chemotherapy-protocols-and-prescriptions/#

[42] National Institute for Health and Care Excellence. *Single Technology Appraisal: company evidence submission template [Internet]*. London: NICE, 2015 [accessed 11.6.15]. 28p. 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

[43] National Institute for Health and Care Excellence. *Erlotinib and gefitinib for treating non-small-cell lung cancer that has progressed after prior chemotherapy: technology appraisal guidance TA374 [Internet]*. London NICE, 2015 [accessed 20.7.21]. 67p. Available from: www.nice.org.uk/guidance/ta374

[44] National Institute for Health and Care Excellence. *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy: technology appraisal guidance TA428 [Internet]*. London: NICE, 2017 [accessed 20.7.21]. 28p. Available from: www.nice.org.u/guidance/ta428

[45] 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):205-216.

[46] 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.