CONFIDENTIAL UNTIL PUBLISHED

External Assessment Group Report commissioned by the NIHR Evidence Synthesis Programme on behalf of NICE

Tarlatamab for previously treated advanced small-cell lung cancer

---

Produced by	**Southampton Health Technology Assessments Centre (SHTAC)**
Authors	Karen Pickett, Senior Research Fellow, Evidence Synthesis Asyl Hawa, Research Fellow, Health Economics Keith Cooper, Senior Research Fellow, Health Economics Jo Picot, Senior Research Fellow, Evidence Synthesis David Alexander Scott, Principal Research Fellow, Statistics
Correspondence to	Dr Karen Pickett Southampton Health Technology Assessments Centre (SHTAC) School of Healthcare Enterprise and Innovation University of Southampton Alpha House, Enterprise Road, University of Southampton Science Park, Southampton SO16 7NS www.southampton.ac.uk/shtac
Date completed	19 July 2024

Source of Funding	This report was commissioned by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme as project number NIHR166739.

Acknowledgements

We thank the following for providing expert clinical advice to the project team and for reading and commenting on a draft of the report:

Dr Judith Cave, Medical Oncologist and Associate Director of Medical Education, University Hospital Southampton NHS Foundation Trust.

Dr Robin Young, Consultant Medical Oncologist, Sheffield Teaching Hospitals NHS Foundation Trust.

Dr Nicole Dorey, Consultant Clinical Oncologist, Torbay and South Devon NHS Foundation Trust.

We also thank Dr Jonathan Shepherd, Principal Research Fellow, and Dr Fay Chinnery, Research Fellow, both Southampton Health Technology Assessments Centre (SHTAC), for providing quality assurance reviews of the draft report. We additionally thank Lois Woods, Senior Research Assistant, SHTAC, for critically appraising the systematic literature review search strategies.

Declared competing interests of the authors and advisors

The authors, and Dr Young and Dr Dorey report none. Dr Cave reports that Amgen funded her to access the American Society of Clinical Oncology conference online in 2020. Dr Cave also reports that she expressed an interest in being a Principal Investigator for a study of tarlatamab but did not become an investigator following this.

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:

Pickett, K; Hawa, A; Cooper, K; Picot, J; Scott, DA. Tarlatamab for previously treated advanced small-cell lung cancer. A Single Technology Appraisal. Southampton Health Technology Assessments Centre, 2024.

Contributions of authors

Karen Pickett critically appraised the clinical effectiveness systematic review, the company's matching adjusted indirect comparison (MAIC), drafted the report, project managed the review and is the project guarantor; Asyl Hawa critically appraised the health economic systematic review and the economic evaluation, and drafted the report; Keith Cooper critically appraised the health economic systematic review and the economic evaluation, and drafted the report; Jo Picot critically appraised the clinical effectiveness systematic review and the company's MAIC, and drafted the report; David Alexander Scott critically appraised the company's MAIC, and drafted the report.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

AE	Adverse event
AIC	Akaike information criterion
BIC	Bayesian Information criterion
BICR	Blinded independent central review
BNF	British National Formulary
CAS	Cancer Analysis System
CAV	Cyclophosphamide, doxorubicin and vincristine
CI	Confidence interval
CIC	Commercial in confidence
CQ	Clarification question
CR	Complete response
CRS	Cytokine release syndrome
CS	Company submission
CSR	Clinical study report
CT	Computed tomography
DC	Disease control
DoDC	Duration of disease control
DOR	Duration of response
DSU	Decision Support Unit
EAG	External Assessment Group
ECOG	Eastern Cooperative Oncology Group
eMIT	Electronic market information tool
EORTC	European Organization for Research and Treatment of Cancer

EQ-5D-3L	European Quality of Life Working Group Health Status Measure 3 Dimensions, 3 Levels
EQ-5D-5L	European Quality of Life Working Group Health Status Measure 5 Dimensions, 5 Levels
ESS	Effective sample size
GP	General practitioner
HR	Hazard ratio
HRG	Healthcare Resource Group
HRQoL	Health-related quality of life
ICANS	Immune effector cell associated neurotoxicity syndrome
ICER	Incremental cost-effectiveness ratio
IV	Intravenous
KM	Kaplan-Meier
LSM	Least squares mean
LYG	Life years gained
MAIC	Matching-adjusted indirect comparison
MRI	Magnetic resonance imaging
MRU	Medical resource use
N	Number
N/A	Not applicable
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
OR	Objective response
ORR	Objective response rate
OS	Overall survival
PAS	Patient access scheme
PD-1	Programmed cell death protein 1
PD-L1	Programmed cell death ligand 1
PFS	Progression-free survival
PICOD	Patient, intervention, comparator, outcome, design
PR	Partial response
PS	Performance status
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit

Q2W	Once every other week (i.e. once every 2 weeks)
QALY	Quality-adjusted life year
QLQ-C30	Quality of life questionnaire core 30
QLQ-LC13	Quality of life questionnaire lung cancer module
QoL	Quality of life
r/r	Relapsed or refractory
RCT	Randomised controlled trial
RECIST	Response evaluation criteria in solid tumours
RR	Relative risk/risk ratio
SCLC	Small-cell lung cancer
SD	Standard deviation
SE	Standard error
SmPC	Summary of product characteristics
SOC	Standard of care
TA	Technology appraisal
TEAE	Treatment-emergent adverse event
TNM	Tumour, node, metastases
TSD	Technical Support Document
TTD	Time to treatment discontinuation
UK	United Kingdom
US	United States
VAS	Visual analogue scale

1 EXECUTIVE SUMMARY

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

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

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

1.1 Overview of the EAG's key issues

Table 1 Summary of key issues

ID	Summary of issue	Report sections
1	Uncertainty associated with the unanchored MAIC overall survival (OS) and progression-free survival (PFS) estimates	3.2.1, 3.2.2, 3.3.2, 3.4, 3.6, 5.2.2 and 6.3.1
2	Parametric curve used for modelling OS	4.2.6.1
3	Quality of life estimates used in the economic model	4.2.8.2
4	Exclusion of best supportive care as a comparator	2.3
5	Company submission (CS) evidence focuses on patients with an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1	2.2.3, 2.3, 3.2.1, 3.3 and 4.2.3

The key differences between the company's preferred assumptions and the EAG's preferred assumptions are the parametric distribution used for OS and the source of the quality-of-life values.

1.2 Overview of key model outcomes

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

Following their response to the clarification questions, the company updated their economic model. The company's revised base case deterministic cost-effectiveness results are shown in Table 2 with a confidential patient access scheme (PAS) discount applied for tarlatamab.

The ICER is £33,785 per QALY for tarlatamab versus standard of care (SOC), with a QALY gain of [redacted] and an additional cost of [redacted]. The results include a severity multiplier of 1.7 applied to the incremental QALYs.

Table 2 Company revised base case results with PAS for tarlatamab

Treatment	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY)
SOC	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Tarlatamab	[redacted]	[redacted]	[redacted]	[redacted]	£33,785

Source: Table 20 of company's clarification response, includes severity response
SOC, standard of care; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life year.

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

The EAG have not identified any key issues in relation to the company's decision problem.

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

Uncertainty associated with the unanchored MAIC OS and PFS estimates

Report section	3.2.1, 3.2.2, 3.3.2, 3.4, 3.5, 3.6, 5.2.2 and 6.3.1
Description of issue and why the EAG has identified it as important	No head-to-head randomised controlled trials (RCTs) of tarlatamab versus the company's selected standard of care comparators, carboplatin and etoposide chemotherapy (carboplatin + etoposide), cyclophosphamide in combination with doxorubicin and vincristine (CAV) and topotecan were identified in the company's decision problem population of patients receiving a third-line or later therapy (third-line therapy+). Therefore, an unanchored MAIC was conducted to compare the clinical efficacy of tarlatamab against a single basket standard of care comparator, incorporating these three treatments, in terms of OS and PFS. Data for tarlatamab came from a single-arm study (DeLLphi-301), and data for standard of care was derived from a UK real-world evidence study (the UK Cancer Analysis System [CAS] study). We consider the OS and PFS estimates from the unanchored MAIC base case uncertain, because: • Indirect comparative evidence is inferior to direct comparative data. Residual systematic error due to unobserved prognostic variables and effect modifiers cannot be entirely eliminated in an unanchored MAIC.[1] • An interim data-cut of the DeLLphi-301 trial, based on 97 participants, informed the data inputs for tarlatamab in the MAIC.

	• There was an uncertainty about the accuracy of the assessment of PFS in the DeLLphi-301 trial due to potential for unblinding of the Blinded Independent Central Review group. • There were differences between the participants in the DeLLphi-301 trial and UK CAS studies in the proportions who had received prior programmed cell death ligand 1 (PD-L1) inhibitor treatment. This was not included as a covariate in the MAIC analyses. It is unknown if this difference might affect the MAIC results. • The effective sample size was greatly reduced to just [redacted] participants, making the MAIC estimates potentially unreliable. This issue is important as the OS and PFS estimates from the MAIC were used in the company's economic model.
What alternative approach has the EAG suggested?	The EAG believe the unanchored MAIC was carried out appropriately and that there was no alternative analysis approach the company could have undertaken given the lack of head-to-head evidence. The EAG's clinical experts thought that the list of prognostic factors included in the MAIC base case was generally reasonable. The EAG asked the company to carry out two MAIC scenario analyses varying the prognostic factors to explore the impact on the results. These resulted in [redacted] effective sample sizes and [redacted] results for tarlatamab for PFS and [redacted] results for tarlatamab for OS than the base case analyses. The changes in the OS and PFS values were [redacted] but resulted in around a £10,000 to £12,000 reduction in the ICER compared to that obtained when using the MAIC base case results in the company's revised economic model base case. It would be useful to know why the CS and clarification response A22 state patients were included in the UK CAS study if they met diagnostic criteria between 1st January 2013 and 31st December 2020, while a confidential report of the MAIC states [redacted].[2] The EAG suspects this is an error, but an explanation would help further assess if there is a risk of selection bias.
What is the expected effect on the cost-effectiveness estimates?	Unknown. OS particularly drives the cost-effectiveness results, so any uncertainty in the precision of the OS estimate may alter the ICERs.
What additional evidence or analyses might help to resolve this key issue?	An explanation for the inconsistency in dates reported in the CS and the MAIC report regarding the diagnostic period in which participants were eligible for the UK CAS study. Provision of an unanchored MAIC analysis using more mature OS data from the DeLLphi-301 trial (if available) and a cost-effectiveness scenario analysis using the result of this may also help to address the concern above about the immaturity of the data.

	The other limitations identified by the EAG likely cannot be addressed within this appraisal, given the nature of the available evidence.

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

Parametric curve used for modelling overall survival

Report section	4.2.6.1
Description of issue and why the EAG has identified it as important	The company model uses the exponential distribution for OS for the tarlatamab and SOC arm. The company fits the best fitting curve to the tarlatamab arm and then uses the same curve for the SOC arm.
What alternative approach has the EAG suggested?	The EAG prefers to use the gamma distribution for OS for the tarlatamab and SOC arms, based on the best fitting curve to the SOC arm. The EAG prefers to fit curves for OS to the SOC arm, rather than to the tarlatamab arm, as the SOC arm data are more mature and it has a larger sample size (n=540) for SOC vs (n=[redacted]) for tarlatamab arm.
What is the expected effect on the cost-effectiveness estimates?	Using the gamma distribution instead of the exponential distribution for both arms increases the ICER for the revised company base case from £33,785 to £39,074 per QALY.
What additional evidence or analyses might help to resolve this key issue?	Longer follow-up for the tarlatamab arm would improve the certainty of curve fitting for this arm.

Quality of life estimates used in the economic model

Report section	4.2.8.2
Description of issue and why the EAG has identified it as important	The company use the quality-of-life values from the DeLLphi-301 study for the progression-free and post-progression model health states, regardless of the treatment.
What alternative approach has the EAG suggested?	The EAG notes that the quality-of-life values are higher than would be expected and come from the full DeLLphi-301 dataset (n=97). We would have preferred to explore using quality-of-life values from the population used in the MAIC (n=[redacted]) which would better match the SOC population but whether these would also be higher than expected is unknown. We have used a study by Chouaid et al for the quality-of-life values in the EAG base case, although the population is third-line treatment for patients with non-small cell lung cancer (NSCLC), rather than small-cell lung cancer (SCLC).
What is the expected effect on the cost-effectiveness estimates?	Using the quality-of-life values from Chouaid et al. increases the ICER for the revised company base case from £33,785 to £43,521 per QALY.

What additional evidence or analyses might help to resolve this key issue?	Further clinical advice on whether quality-of-life data from a population of patients with NSCLC are generalisable to patients with SCLC.

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

Exclusion of best supportive care as a comparator

Report section	2.3
Description of issue and why the EAG has identified it as important	Best supportive care was included as a comparator in the NICE scope, but not in the company's decision problem. The company argue that patients who are candidates for best supportive care are a different population to those who are suitable for systemic therapy and will have worse ECOG performance status.
What alternative approach has the EAG suggested?	Clinical expert advice to the EAG is that the third-line therapy+ patients with SCLC who are likely to be treated with tarlatamab in clinical practice will be those who have an ECOG performance status of 0 or 1 and these patients currently tend to opt for systemic treatment. They considered the proportion of patients eligible for third-line treatment+ with an ECOG performance status of 0 or 1 who receive best supportive care to be low (one expert more specifically estimated <20%). One of the EAG's three clinical experts felt it was reasonable to not include best supportive care as a comparator, with the other two not commenting on this. The EAG suggests that obtaining further expert opinion about the relevance of this comparator would be beneficial, given that it was specified in the NICE scope.
What is the expected effect on the cost-effectiveness estimates?	The cost-effectiveness of tarlatamab versus best supportive care is currently unknown.
What additional evidence or analyses might help to resolve this key issue?	Further clinical expert opinion about whether or not best supportive care is a relevant comparator for tarlatamab.

CS evidence focuses on patients with an ECOG performance status of 0 or 1

Report section	2.2.3, 2.3, 3.2.1, 3.3 and 4.2.3
Description of issue and why the EAG has identified it as important	The clinical and cost effectiveness of tarlatamab versus standard of care is examined in the CS specifically among a population of patients with an ECOG performance status of 0 or 1 at third-line therapy+. The only identified trial of tarlatamab treatment at third-line therapy+ (DeLLphi-301) was conducted in this group of patients. This is a narrower population than that specified in the NICE scope and the company decision problem, [redacted]. The wording in the draft Summary of Product Characteristics (SmPC)

	[redacted] but the evidence from the trial may not be generalisable to patients with an ECOG performance status other than 0 and 1.
What alternative approach has the EAG suggested?	None. Clinical expert advice to the EAG is that only patients with an ECOG performance status of 0 or 1 at this stage will be suitable for tarlatamab. Therefore, we regard the company's focus on this population to be appropriate.
What is the expected effect on the cost-effectiveness estimates?	The company's model restricts the patient population to those with an ECOG performance status of 0 and 1, so the ICERs are only valid for this population. The cost-effectiveness for a wider group of patients, including those with a worse performance status has not been modelled.
What additional evidence or analyses might help to resolve this key issue?	Further discussion about the population of patients expected to be suitable for tarlatamab treatment in clinical practice.

1.7 Summary of EAG's preferred assumptions and resulting ICER

Based on the EAG's critique of the company's model (discussed in section 4), we have identified the following key aspects of the company base case with which we disagree. Our preferred model assumptions are the following:

- Overall survival: use gamma distribution for the tarlatamab and standard of care arms, instead of exponential distribution (section 4.2.6.1)
- Progression-free survival: use exponential distribution for tarlatamab arm, instead of lognormal for the tarlatamab arm (section 4.2.6.3)
- Health-related quality of life (HRQoL): use Chouaid et al. instead of DeLLphi-301 study (section 4.2.8.2),
- Adverse event costs: EAG recalculated values (section 4.2.9.3.2).

Table 2 shows the cumulative cost-effectiveness results of applying the EAG preferred model assumptions to the company's revised base case including the PAS discount for tarlatamab. Incorporating all the EAG assumptions, the ICER for tarlatamab vs BSC increases to £58,847 per QALY (including a severity multiplier of 1.7).

Table 3 EAG's preferred model assumptions: cumulative impact (deterministic) with PAS for tarlatamab

Preferred assumption	Treatment	Total costs	Total QALYs	Cumulative ICER £/QALY
	SOC	[redacted]	[redacted]	

Preferred assumption	Treatment	Total costs	Total QALYs	Cumulative ICER £/QALY
EAG corrected company revised base case	Tarlatamab	[redacted]	[redacted]	£34,958
+ OS: gamma for both arms	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£40,442
+ PFS: exponential for both arms	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£42,045
+ HRQoL: use Chouaid et al. (PFS 0.62; PD 0.47)	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£55,097
+ Adverse event costs: EAG recalculated.	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£58,847
EAG base case	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£58,847

EAG evidence assessment group; OS overall survival; PFS progression-free survival; HRQoL health-related quality of life; PD progressed disease.
Severity multiplier of 1.7 applied to QALYs.

The change that has the most significant impact on the cost-effectiveness results is using the gamma distribution for OS and using a different source for the HRQoL. Further details of the exploratory and sensitivity analyses done by the EAG are provided in section 6.3.

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

This report is a critique of the company's submission (CS) to the National Institute of Health and Care Excellence (NICE) from Amgen on the clinical effectiveness and cost effectiveness of tarlatamab for treating previously treated advanced small-cell lung cancer (SCLC). It identifies the strengths and weakness of the CS. Clinical experts were consulted to advise the external assessment group (EAG) and to help inform this report.

Clarification on some aspects of the CS was requested from the company by the EAG via NICE on 5th June 2024. A response from the company via NICE was received by the EAG on 20th June 2024 and this can be seen in the NICE committee papers for this appraisal.

2.2 Background

The company provides background information on lung cancer, and SCLC in particular, in CS section B.1.3.1 which provides an overview of lung cancer and describes clinical staging, epidemiology and prognosis of SCLC. CS section B.1.3.2 describes the symptom, health-related quality-of-life (HRQoL) and economic burdens of SCLC.

2.2.1 Background information on SCLC

SCLC is the rarer of the two main subtypes of lung cancer, the other main subtype being non-SCLC. Evidence indicates that the proportion of lung cases in England that are SCLC has been falling for at least two decades and in 2022 about 7% were SCLC.[3 4] People typically present with symptoms such as dyspnoea and a persistent cough and about two thirds of people with SCLC will already have metastatic disease (i.e. disease that has already spread to other parts of the body) when they are diagnosed.[5,6]

SCLC is classified as either limited stage or extensive stage disease. Our clinical experts advised us that people with limited stage SCLC will be treated with curative intent. Limited stage SCLC is where disease can be encompassed within a single radiotherapy field. These people receive multi-modality treatment usually involving combination chemotherapy and radiotherapy, although surgery is sometimes used. People with disease that cannot be encompassed within a radical radiotherapy field (which includes all those with metastatic SCLC) are considered to have extensive stage disease. One of our experts stated that 80% of patients with SCLC will present with extensive stage disease. If people with extensive stage disease have limited lymph node involvement, they may still be treated aggressively with curative intent, but more usually people are treated with palliative intent with

chemotherapy ± immunotherapy ± radiotherapy. People whose disease has relapsed after treatment for limited stage disease are usually considered to have extensive stage disease and one of our clinical experts advised are rarely salvageable with radical radiotherapy.

The focus of this NICE appraisal is previously treated advanced small-cell lung cancer. Our clinical experts advised us that advanced small-cell lung cancer corresponds to extensive-stage disease.

As the company state, lung cancer is the second most common cancer in England with approximately 40,000 new cases in England each year,[7] although we note that lung cancer diagnoses fell during the COVID-19 pandemic[8] and data for 2022 shows there were 36,886 lung cancer diagnoses in England.[4] The CS states that SCLC accounts for approximately 15-20% of lung cancer cases (CS section B.1.3.1) however, as noted above, the proportion of lung cases in England that are SCLC has been falling.[3] In 2022 there were 2,501 SCLC diagnoses among the total 36,886 lung cancer diagnoses in England which equates to 7%, in Wales the proportion was 9% among 2,211 lung cancer diagnoses.[4] The median age at SCLC diagnosis in England in 2019/2020 was 80 years whereas in Wales it was 70 years.[9]

SCLC is an aggressive disease which, coupled with the high likelihood of metastasis, means that it typically has a poor prognosis, particularly when extensive disease is present. A recent retrospective chart review of patients in France, Italy and the UK[10] found that among 231 first-line treated patients with limited SCLC median overall survival (OS) was 17.3 months whereas among the 308 first-line treated patients with extensive SCLC median OS was just 8.8 months. In patients with relapsed/refractory SCLC initiating second line treatment (n=225) median OS dropped to 6.6 months. Our clinical experts agreed that prognosis is poor. Some patients are not well enough to receive a second-line therapy whilst others choose not to receive one. There is no established third-line treatment and the proportion of patients who remain fit for third-line treatment is small.

In CS section B.1.3.2 the company summarise the symptom burden of SCLC initially due to the effects of the disease itself and then the additional symptom burden of chemotherapy. Brain metastases, which have been found to be present in more than 10% of patients at diagnosis,[10,11] also develop after diagnosis in another 40-50% of SCLC patients.[12] Brain metastases are associated with a high burden of neurological symptoms which have been summarised by the company.

In discussing the HRQoL burden of SCLC, the company highlight the limited data available for this type of lung cancer. Evidence from a systematic review found that HRQoL is lower

for patients with SCLC than the general population and tends to remain stable when patients are receiving treatment.[13] The company also highlight the substantial financial burden of lung cancer which was estimated to be £307 million in England in 2010 to the National Health Service (NHS) and wider economy.[14]

2.2.2 Background information on tarlatamab

Tarlatamab (brand name [redacted]®) is a type of targeted immunotherapy known as a bispecific T-cell engager (BiTE).[15] Each of tarlatamab's two arms has a binding domain, one that binds to tumour cells, and one that binds to T-cells. When tarlatamab simultaneously binds to tumour cells and T-cells, the T-cell is activated which leads to breakdown of the tumour cell membrane and then disintegration of the tumour cell.[15,16] The company describe tarlatamab in CS section B.1.2, CS Table 2 and show its mechanism of action in CS Figure 1.

The company anticipate that tarlatamab will receive a [redacted] marketing authorisation[17] in [redacted]. The company have provided the draft Summary of Product Characteristics (SmPC) as part of their submission to NICE which states that tarlatamab is indicated for "[redacted]".

2.2.3 Tarlatamab is administered as

[redacted]

2.2.4 The position of tarlatamab in the treatment pathway

The company describes the clinical pathway of care for SCLC in CS section B.1.3.3, drawing on NICE NG122[18] and NICE Technology Appraisals (TAs) 184[19] and 638[20] to outline the treatment options available for patients at first-, second- and third-line or beyond (the latter is referred to as 'third-line therapy+' throughout this report). All three of the EAG's clinical experts agreed with the company's depiction of the treatment pathway, as shown in CS Figure 2, but it was noted that best supportive care – which the company did not include – is also an option.

We understand from our clinical experts that:

- At first-line therapy, as is indicated by the company in CS section B.1.3.3, most patients will receive a programmed cell death ligand 1 (PD-L1) inhibitor (atezolizumab) with carboplatin and etoposide, with or without radiotherapy, unless there is a major contraindication. CS Figure 2 shows that, as per NICE TA638,[20] this option is recommended for patients with extensive-stage disease only. The CS figure shows that cisplatin-based combination chemotherapy, with or without radiotherapy, is also available for patients with extensive- or limited-stage disease.

- The choice of second-line treatment for SCLC depends on the interval between first-line therapy and relapse. If patients have a chemotherapy-free interval of six or more

months, they are considered to still be platinum-sensitive and the generally preferred treatment option is to re-challenge with platinum-based chemotherapy (e.g. carboplatin/cisplatin plus etoposide chemotherapy (hereafter referred to as 'carboplatin + etoposide') or carboplatin alone). If patients have disease progression after three to six months, they are considered platinum-chemotherapy refractory and so topotecan (a topoisomerase 1 inhibitor chemotherapy) or cyclophosphamide in combination with doxorubicin and vincristine (CAV) might be considered. The use of topotecan may vary between individual centres. CS Figure 2 shows that these treatments can be used either alone or alongside radiotherapy.

- As stated in CS section B.1.3.3 and section 2.2.1, there are currently no established third-line therapy+ options available for SCLC. Our clinical experts also noted that the proportion of patients who remain fit enough for third-line treatment is small and therefore patients rarely receive third-line therapy. Most patients at this stage will have had re-treatment with platinum-based chemotherapy at second-line. The majority of patients go onto best supportive care at third-line after a trial of two chemotherapy regimens. However, if a patient's Eastern Cooperative Oncology Group (ECOG) performance status score remains within the range of 0 or 1 (indicating no or mild restriction in performing daily activities and self-care[21]) and systemic treatments are available, patients will likely accept these.

- Given the lack of established treatments at third-line therapy+, the systemic therapies that are currently available at this stage are those that are used at second-line: carboplatin + etoposide, CAV, and topotecan. We understand from our clinical experts that patients may be re-treated with a therapy they have had before, and most patients will receive either CAV or topotecan. The EAG's experts indicated that carboplatin only and cisplatin + etoposide (both with or without radiotherapy) – which are second-line therapy options shown in CS Figure 2 – are rarely used at the third-line stage. Two of the EAG's experts noted that there have been topotecan product supply issues affecting the NHS recently, but they indicated that this was a temporary problem.

Best supportive care is aimed at managing patients' symptoms and is considered palliative care, with patients discharged back to their general practitioner (GP) or community palliative care team. Best supportive care has a number of components, which may differ depending on needs of the patient, including nursing and physical care, provision of equipment (such as

a hospital bed at home), nutritional support, help with activities of daily living, and medical interventions such as oxygen, analgesia, and sometimes dexamethasone and palliative radiotherapy.

The company is positioning tarlatamab as a third-line therapy+ option in the SCLC clinical pathway (CS Figure 2). All three of the EAG's experts agreed with the company's proposed positioning. One expert commented that this was because the available clinical effectiveness evidence is in relation to tarlatamab used as a third-line therapy+, but that, in their opinion, it would also be desirable for tarlatamab to be available as a second-line treatment. This expert anticipated that if tarlatamab is approved, then it may change clinical practice as clinicians may try to move patients to third-line therapy more quickly, which could potentially increase the number of patients who might receive third-line treatment. The EAG's experts also indicated that it is expected that patients will need to be fit to receive tarlatamab, so it is anticipated that the patients who will be treated with tarlatamab in clinical practice will have an ECOG performance status of 0 or 1. This is due to its potential toxicity and the need for hospital admission to monitor for adverse events. (As outlined in section 2.2.2, the draft SmPC [redacted];[22] a cost for a hospitalisation period of 24 hours at tarlatamab treatment initiation is included in the company's economic model (CS Table 53).) Our experts said that in practice, most patients at the third-line therapy stage have performance statuses of 2 to 4, so only a small, select group of patients will be suitable for tarlatamab.

EAG comment on background information and treatment pathway

The company provides accurate background information and an accurate depiction of the treatment pathway for SCLC in the CS, but they do not include best supportive care, which is also an option for patients. The company's proposed positioning of tarlatamab as a third-line therapy+ is appropriate. The EAG's clinical experts noted that only patients with an ECOG performance status of 0 or 1 at this stage will be suitable for tarlatamab.

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

Table 4 summarises the company's decision problem in relation to the final scope issued by NICE and includes the EAG's comments. In their decision problem, the company focus on a third-line therapy+ population of patients with advanced SCLC rather than the second-line therapy+ population specified in the NICE scope. This is appropriate, as it reflects the patient

population the EAG's clinical experts expect to be treated with tarlatamab in clinical practice (please see Table 4 below and section 2.2.3 above).

In their decision problem, the company have selected the comparators of topotecan, CAV and carboplatin + etoposide and conceptualised these as a "*single basket*" standard of care (SOC) comparator. The company's selection of these comparators reflects those specified in the NICE scope and those used in clinical practice (confirmed by our experts; please see Table 4). We also consider that the single standard of care comparator approach is appropriate; please see Table 4 for our reasons. Best supportive care was also specified as a comparator in the NICE scope, but the company have excluded this from their decision problem without explanation. We asked the company for their reason in clarification question A1. The company responded that patients who are candidates for best supportive care are a different patient population to those who are suitable for systemic therapy and will have worse ECOG performance status. Our clinical experts informed us that best supportive care is a relevant comparator for a third-line therapy+ population. However, as is outlined in section 2.2.3, our experts expect that the third-line therapy+ patients who will likely be treated with tarlatamab in clinical practice will be those who have an ECOG performance status of 0 or 1. Our experts indicated that these fitter patients tend to currently opt for systemic treatment and the remaining proportion who receive best supportive care is low (one expert estimated less than 20%). In one of the EAG's expert's opinion, it is reasonable to not include best supportive care as a comparator, but the other two experts made no comment. Another expert made the point that the proportion of patients who will be treated with best supportive care will depend on the other treatment options available.

Table 4 Summary of the decision problem

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
Population	Adults with advanced small-cell lung cancer with disease progression on or after prior therapy	Adult patients with advanced SCLC after platinum-based chemotherapy and at least one other treatment	The population addressed in this submission is narrower than the NICE scope, requiring patients to have advanced disease after platinum-based chemotherapy and at least one other treatment. This narrower population reflects the clinical evidence available for tarlatamab in	The company focus on a third-line therapy+ population which is a subgroup of the NICE scope population [redacted]. All the EAG's clinical experts agreed with this proposed positioning of tarlatamab and agreed that the prior therapies that this group would have received reflect clinical practice.

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
			this indication.	
Intervention	Tarlatamab	Tarlatamab 10 mg Q2W	N/A – in line with NICE scope	The intervention matches the NICE scope and reflects [redacted].[17] See section 2.2.2 of this report for details of the specific draft SmPC-recommended dosing regimen for tarlatamab.
Comparators	Established clinical management without tarlatamab, which may include: • Chemotherapy, including anthracycline-containing or platinum-based regimen. • Oral topotecan (when re-treatment with the first-line regimen is not considered appropriate and the combination of cyclophosphamide, doxorubicin and vincristine is contraindicated)	Third-line treatment standard of care (SOC) has been modelled as a single comparator, comprising the following treatments: • Topotecan • Cyclophosphamide + doxorubicin + vincristine • Carboplatin + etoposide	Patients with advanced SCLC following two or more prior treatments do not have dedicated treatment options. Patients consequently face extremely poor outcomes, with expected survival of only a few months, regardless of treatment	The EAG's clinical experts agreed that the company's selected comparators are an appropriate reflection of the treatments patients with advanced SCLC who have already received platinum-based chemotherapy and at least one other treatment typically receive at third-line therapy+ in clinical practice. As detailed above, clinical expert advice to the EAG is that the proportion of patients with an ECOG performance status score of 0 or 1 who receive best supportive care (another comparator specified in the NICE scope) in practice is low, with one expert stating they believed it was reasonable for it not to have been included as a comparator in the CS. In clarification question A2 we asked the company to provide evidence for their statement that their three selected comparators have similar efficacy. The company provided information in response, which we have summarised in Appendix 1. Some differences were identified in the efficacy of the treatments. However, we do not perceive that to be a reason not to include the treatments in a single basket comparator as long as the treatment distributions reflect the extent to which they are used in clinical practice.

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
	• Best supportive care		received. The proposed treatment regimens, modelled as a single basket comparator, are considered an appropriate reflection of treatments typically prescribed at this stage of disease, given the similarly poor outcomes associated with each treatment. In line with discussion with the NICE technical	All three of the EAG's clinical experts considered that it was reasonable to treat the three comparators as a single basket standard of care comparator. One expert commented that it is likely there would be no meaningful data available to compare tarlatamab to the comparators individually. All of the EAG's experts agreed that topotecan, CAV and carboplatin + etoposide have similarly poor treatment outcomes. Overall, the single basket standard of care comparator approach appears reasonable.

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
			team and external assessment group (EAG), pairwise analyses against each individual comparator (topotecan, CAV and platinum-based chemotherapy) have also been conducted as scenario analyses.	
Outcomes	The outcome measures to be considered include: • overall survival • progression-free survival • response rates • adverse effects of treatment	• Response rates • PFS • OS • TEAEs • HRQoL	N/A – in line with NICE scope	The outcomes selected by the company match the NICE scope. Due to data availability, the company have used time-to treatment discontinuation as a proxy for PFS for the standard of care comparator (CS section B.2.9.3). Please see section 3.3.4 for a discussion about this.

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
	• health-related quality of life.			
Economic analysis	The reference case states the following requirements for cost-effectiveness analyses: costs assessed as cost per quality-adjusted life year (QALY), adequate time horizon, NHS and Personal Social Services perspective, commercial arrangements and managed access taken into account and availability and cost of biosimilar and generic products taken into account. (NICE scope wording abridged by EAG here for brevity.)	The economic analysis is expected to be in line with that described in the NICE decision problem	N/A – in line with NICE scope	The company's economic analysis matches NICE's reference case requirements. CS Table 2 includes details of a simple patient access scheme (PAS) discount for tarlatamab. The anticipated PAS price was used in the company's economic analysis (CS section B.3.5.1). The company used a multiplier for disease severity of 1.7 in their base case.
Subgroups	None specified	N/A	N/A	In practice the CS focuses on a sub-population of advanced SCLC patients who are receiving third-line therapy+. The CS

	Final scope issued by NICE	Company's decision problem	Rationale if different from the final NICE scope	EAG comments
				also presents subgroup analyses from the tarlatamab trial (the DeLLphi-301 trial) by various patient characteristics in CS section B.2.7.
Special considerations including issues related to equity or equality	None specified	N/A	N/A	The company have not identified any issues related to equity or equality. Two of the EAG's experts were not aware of any equity or equality issues potentially related to this appraisal. The other expert noted if the DeLLphi-301 trial protocol of admitting patients to hospital for monitoring of adverse events post-treatment after the first two infusions (see section 3.2.1.1) is adopted in clinical practice, this could create a resource issue and difficulties for patients who want to receive treatment outside major cancer centres. (The draft SmPC recommendations regarding monitoring are outlined in sections 2.2.2 and 2.2.3.)

Source: Partly reproduced from CS Table 1.
CAV, cyclophosphamide, doxorubicin and vincristine; CS, company submission; EAG, External Assessment Group; HRQoL, health-related quality of life; N/A, not applicable; NHS, National Health Service; NICE, National Institute for Health and Care Excellence; OS, overall survival; PAS, patient access scheme; PFS, progression-free survival; Q2W, once every other week; QALY, quality-adjusted life year; SCLC, small-cell lung cancer; SmPC, summary of product characteristics; TEAEs, treatment-emergent adverse events

3 CLINICAL EFFECTIVENESS

3.1 Critique of the methods of review(s)

The company conducted a systematic literature review to identify studies on the clinical efficacy (or effectiveness) and safety of treatments for patients with SCLC who have previously received two or more treatments and who have experienced disease progression on or after any prior therapy (CS section B.2.1 and CS Appendix D, section D.1). The searches for the systematic review (original search and an update search) covered a period from 2012 to 19th December 2023. As only approximately five months had elapsed since 19th December 2023 upon receiving the CS on 16th May 2024, we considered the searches to be sufficiently up to date. The search terms were appropriate, but we note that the company did not include terms for randomised controlled trials (RCTs) in their search strategies, despite this study design being listed in the eligibility criteria for the review (CS Appendix D, section D.1.2, Table 9). However, we do not believe that any relevant evidence has been omitted as a result. The EAG have not re-run the company's searches using a term for RCTs, but non-systematic searches we have conducted in the clinicaltrials.gov and PubMed databases have not identified any relevant RCTs of tarlatamab in the company's decision problem population. Our detailed assessment of the company's systematic review methods can be found in Appendix 1.

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

3.2.1 Included studies

The company's systematic literature review identified one study of tarlatamab relevant to their decision problem population. This was the ongoing DeLLphi-301 phase 2 trial[23] (CS section B.2.2).

The trial has been reported in a journal article, which the company supplied as part of their CS.[23] The company also provided the trial clinical study report (CSR).[24]

3.2.1.1 Study characteristics

The DeLLphi-301 trial is an open-label, registrational trial of tarlatamab examining the efficacy and safety of tarlatamab in people with relapsed or refractory SCLC after two or more previous lines of therapy. It does not compare tarlatamab to any other treatments. The trial was funded by the company (Amgen).[23] The CS states that the trial supported the marketing authorisation application for tarlatamab (CS section B.2.2).

The characteristics of the DeLLphi-301 trial are shown in Table 5. It consisted of three parts:

- In Part 1, participants were randomised to tarlatamab 10 mg every two weeks (Q2W) or 100 mg Q2W. During Part 1, an interim analysis determined which of the two initially examined tarlatamab doses would be the focus for Part 2 of the study.
- Part 2 was the dose expansion phase of the trial. From the interim analysis, the 10 mg Q2W dose was selected for Part 2 and additional participants were recruited to receive this dose only.
- In Part 3, a modified safety protocol was implemented for newly recruited participants with safety monitoring for potential Cytokine Release Syndrome (CRS) and/or neurological events changed from 48 hours (as it had been in Parts 1 and 2) to 24 hours.

The selected dose of 10 mg Q2W and the tarlatamab treatment regimen used in the DeLLphi-301 trial [redacted]. All our clinical experts confirmed that the permitted and disallowed medications and treatments in the DeLLphi-301 trial, as shown in Table 5, are reflective of how third-line therapy+ patients eligible for tarlatamab (if it is approved) would be treated in clinical practice.

All three of the EAG's clinical experts considered that the key inclusion criteria of the trial reflect the patients with SCLC who are expected to receive tarlatamab in clinical practice. However:

- One expert noted that the key inclusion criteria required participants to have measurable lesions within a 21-day period before taking their first dose of tarlatamab, and one expert noted that in practice, clinicians would not look at measurable lesions and patients would be treated regardless.
- The same expert also noted that the trial required re-biopsy, while in practice relapsed disease is not usually re-biopsied.
- Regarding the requirement for patients to have had one platinum-based regimen and at least one other prior line of therapy, another expert noted that 'one other prior line of therapy' would include re-treatment with chemotherapy.

One expert was also satisfied that the DeLLphi-301 trial key exclusion criteria were also representative of the patients expected to receive tarlatamab in clinical practice. The other two experts did not comment on this.

Our clinical experts noted that the DeLLphi-301 trial inclusion criteria represent a highly selected, small group of third-line therapy+ patients who will be fit enough and suitable to receive tarlatamab (i.e. people with an ECOG performance status of 0 or 1). Our experts said that in practice, most patients who reach third-line treatment will be more poorly and thus will not be suitable for tarlatamab.

In the trial, participants were treated until disease progression (CS section B.2.3.1). After this, participants could stay on tarlatamab if the investigator judged that they were receiving benefit and there were no unacceptable or significant co-morbidities. [redacted] Two of the EAG's three clinical experts do not expect that tarlatamab will be used post-progression in clinical practice, while one expected that it would be if it was clinically appropriate because this was how it was used in the trial and evidence suggests that tarlatamab is well-tolerated. Additionally, one expert commented that there can be doubt sometimes in practice about whether true progression is being observed and if there is some uncertainty about this, continuing treatment until the next assessment may be appropriate.

Interim results for the DeLLphi-301 trial are presented in the CS from a data-cut dated 27th June 2023. Of the participants eligible for the trial, 176 were randomised in Part 1 to either tarlatamab 100 mg (n = 88) or 10 mg (n = 88) (CS Figure 3), with one participant in each of these arms not receiving the allocated treatment. Another 12 participants received tarlatamab 10 mg in Part 2. Thus, 99 participants received tarlatamab in both Parts 1 and 2 of the trial. This population is referred to in the CS as the "*10 mg target group*" (CS section B.2.3.2) and is the focus of the evidence presented in the submission. In Part 3, 34 participants have so far received the 10 mg target dose, but due to immaturity of data, the company have not included either efficacy or safety results for these patients in the CS. However, results are available in the trial CSR[24] provided with the CS. In terms of presenting the clinical efficacy results in the CS, the EAG regards this as reasonable. However, in the context of the Part 1 and 2 10 mg participant group of only 99 participants, we believe that safety data for another 34 participants would be valuable and thus asked the company to provide a summary of the adverse event data for the Part 3 participants in clarification question A14 and the company accordingly provided these results.

The following outcomes from the DeLLphi-301 trial inform the company's economic analysis: OS, progression-free survival (PFS), time to treatment discontinuation (TTD), HRQoL (as measured by the European Quality of Life Working Group Health Status Measure 5

Dimensions, 5 levels [EQ-5D-5L]) and adverse event rates (specifically grade 1-2 CRS and immune effector cell-associated neurotoxicity syndrome (ICANS), and grade 3+ adverse events) (CS sections B.3.3 and B.3.4).

Table 5 DeLLphi-301 study design and characteristics

Study characteristics	Details
Population	Patients with relapsed or refractory SCLC whose disease has progressed or recurred following one platinum-based regimen and at least one other line of therapy.
Intervention(s) and number of participants randomised or enrolled at each stage of the trial	• Part 1 (dose escalation phase): Tarlatamab 10 mg or 100 mg (participants randomised 1:1; N = 176). Participants received 1 mg of tarlatamab for the first dose (Cycle 1, day 1), and then step-up doses of 10 mg or 100 mg depending on their treatment arm. • Part 2 (dose expansion phase): Tarlatamab 10 mg (selected target dose following interim analysis on N=30/arm; N enrolled = 12) • Part 3 (modified safety protocol): Tarlatamab at selected target dose of 10 mg (N enrolled = 34). Safety monitoring reduced from 48 to 24 hours (please see row below).
Method of administration of tarlatamab and safety monitoring	Tarlatamab 1 mg was given on Cycle 1 Day 1. The full dose of either 10 mg or 100 mg was given on day 8 and 15 of Cycle 1, and Q2W after this. Administration was via an IV infusion for 60 minutes. An 8 mg dose of dexamethasone was administered IV before tarlatamab was given on Days 1 and 8 of Cycle 1, and prophylactic hydration (1 litre of normal saline) was administered IV after each dose in Cycle 1. Hospitalisation for safety monitoring was needed for 48 hours after the first two infusions of tarlatamab in Parts 1 and 2. This was reduced to 24 hours in Part 3.
Comparator	No comparator.
Key inclusion criteria	• Patient had provided informed consent/assent prior to initiation of any study specific activities/procedures. • Male and female patients ≥ 18 years of age (or legal adult age within country) at the time of signing the informed consent. • Histologically or cytologically confirmed r/r SCLC • Patients who progressed or recurred following 1 platinum-based regimen and at least 1 other prior line of therapy. • Patients willing to provide archived tumour tissue samples or willing to undergo pre-treatment tumour biopsy. • ECOG performance status of 0 or 1 • Minimum life expectancy of 12 weeks. • Measurable lesions as defined per RECIST 1.1 within 21 days prior to the first dose of tarlatamab. • Patients with treated brain metastases were eligible provided they met defined criteria.
Permitted and disallowed concomitant medications and treatments	Investigators were permitted to prescribe any concomitant medications or treatments deemed necessary for adequate supportive care, except for: • Other investigational agents

Study characteristics	Details
	• Concurrent experimental or approved anti-tumour therapies other than study drugs • Radiation therapy (with the exception of for symptom control) • Immunosuppressive agents
Study centres and locations	The study was conducted at 56 centres in Austria, Belgium, Denmark, France, Germany, Greece, Italy, Japan, Netherlands, Poland, Portugal, South Korea, Spain, Switzerland, Taiwan, United Kingdom (two study centres), and the United States. The company confirmed in clarification question response A5 that 5 patients were recruited from the two UK centres (1 at one centre, and 4 at the other).
Study start date and expected completion date	• The first patient was enrolled on 1st December 2021. • The estimated study completion date is 31st October 2025.
Follow-up	• Upon permanent discontinuation of tarlatamab, participants had a safety follow-up visit at around 42 (+5) days after their last tarlatamab dose. • After the safety follow-up visit, patients entered long-term follow-up, conducted every three months (+/- two weeks) for a year following the last dose of tarlatamab or five years following the first participant being enrolled (whichever occurred first), to measure OS and/or the initiation of subsequent cancer therapies. • At the interim data cut of 27th June 2023, the median duration of follow-up was 10.6 months for OS.
Primary outcomes	• OR (including CR and PR) • Incidence of TEAEs • Serum concentrations of tarlatamab
Other outcomes	DOR, PFS, OS, HRQoL, DC, DoDC, incidence of anti-tarlatamab antibody formation.

Source: Partly reproduced from CS Tables 3 and 4, CS section B.2.3 and clarification response A5. Estimated study completion date sourced from the clinicaltrials.gov DeLLphi-301 trial record.[25]
CR, complete response; DC, disease control; DoDC, duration of disease control; DOR, duration of response; ECOG, Eastern Cooperative Oncology Group; HRQoL, health-related quality of life; IV, intravenous; N, number; OR, objective response rate; OS, overall survival; PR, partial response; RECIST, response evaluation criteria in solid tumours; r/r, relapsed or refractory; SCLC, small-cell lung cancer; TEAE, treatment-emergent adverse event.

3.2.1.2 Patients' baseline characteristics

The company provide the baseline demographic and disease characteristics of the 99 participants in the DeLLphi-301 trial who received tarlatamab in both Parts 1 and 2 of the trial in CS section B.2.3.4. CS Table 6 shows the baseline demographic characteristics. Our clinical experts considered the characteristics presented to be representative of the patients seen in clinical practice, but two of the three experts noted that there was a higher proportion of Asian participants in the trial (41.4%) than observed in practice. The other expert noted that the proportion of patients who are Asian in practice will depend on geographical area.

The experts did not expect that race or ethnicity would impact on treatment response. The mean age of the trial participants was [redacted] years (standard deviation [SD] [redacted]). Two of the experts commented that the age of the participants in the trial is in keeping with the expectation that patients will only be suitable for tarlatamab if they have a performance status of 0 or 1.

CS Table 7 shows the baseline disease characteristics of the 99 DeLLphi-301 trial participants who received the 10mg target dose of tarlatamab in both Parts 1 and 2 of the trial. Overall, our experts thought these looked reasonable. Most of the patients had had two or more prior lines of therapy (n=97/99, 98.0%; percentage calculated by EAG), with the majority having two prior lines (65.7%). Of the 99 participants, 72.7% had received a prior programmed cell death protein 1 (PD-1) or PD-L1 inhibitor treatment, which is in line with clinical expert advice to the EAG that the majority of fit patients will have previously received atezolizumab (a PD-L1 inhibitor therapy). Other than the reporting of the proportion of patients who had received a prior PD-1 or PD-L1 inhibitor therapy, the EAG note that details of the specific therapies participants had previously received are not provided in Document B of the CS, so it is not possible to determine to what extent they reflect clinical practice. Two participants were protocol violators because they had not received two previous lines of therapy.

The proportions of patients who had an ECOG performance status of 0 or 1 at baseline were 26.3% and 73.7%, respectively (CS Table 7). One of our experts felt that this distribution reflected the patients seen in clinical practice (i.e. that the majority of third-line therapy+ patients who will be suitable for tarlatamab will have a performance status of 1). Another noted it was similar to the performance status distributions presented from a real-world study of UK patients with a performance status of either 0 or 1 used in the company's indirect comparison (see section 3.3). However, the remaining expert stated that they would not expect around a quarter of patients to have a score of 0 at third-line therapy. It is unclear if this expert's comment was in relation all third-line patients or just those with an ECOG performance status of 0 or 1.

Two of our experts thought the proportion of participants who had never smoked (8.1%; CS Table 7) was higher than is seen in practice. One commented that it would be expected that most or all SCLC patients would be current or former smokers. The other estimated that typically less than 5% of SCLC patients in practice have never smoked. Clinical expert advice to the EAG is that people who have never smoked are likely to be fitter and have

better outcomes, but all our experts felt the high proportion who had never smoked in the DeLLphi-301 trial was unlikely to affect the results.

3.2.1.3 Ongoing studies

As stated above, the DeLLphi-301 trial is ongoing and is estimated to complete in October 2025 (see Table 5). In the CS, the company states there is an ongoing trial of tarlatamab in a second-line therapy population, called the DeLLphi-304 trial (CS section B.2.11). We note, this is the phase 3 NCT05740566 trial that is comparing tarlatamab with standard of care in participants with relapsed SCLC after receiving platinum-based chemotherapy at first-line treatment.[26] The trial started on 31st May 2023 and is estimated to complete on 1st August 2027.[26] The company state that the indication is not relevant to the present submission. We agree that it is not relevant to the company's decision problem population, but it is relevant to the population specified in the NICE scope. It is unclear when interim results might become available from this trial.

EAG comment on included studies

The participants included in the DeLLphi-301 trial are generally representative of third-line patients seen in clinical practice and those expected to be treated with tarlatamab (i.e. patients with an ECOG performance status of 0 or 1). Limitations of the trial are: 1) it does not compare tarlatamab to any relevant comparators, and 2) only interim data are currently available.

3.2.2 Risk of bias assessment

The company did not state which source(s) of evidence they used to complete their assessment of the risk of bias (e.g. CS, CSR &/or journal publication by Ahn et al. 2023[23]). We have used the CS and journal publication by Ahn et al. 2023 supplemented with information from the trial CSR and protocol where necessary and we focus on the risk of bias for the 10mg target dose group treated as a single-arm study (i.e. participants in the 10mg target dose group in Part 1 and those from Part 2 of the study).

The company used the modified Downs and Black checklist for non-randomised trials[27] and state that the study scored as 'good', going on to say that this means the trial was associated with a low risk of bias. We find that that whilst the source paper cited by the company (Hooper et al.[27]) does ascribe 'quality levels' to score ranges we found no evidence in the paper to describe how they mapped the score ranges to the quality terms used (i.e. excellent, good, fair and poor). Additionally, the term 'bias' does not appear in the modified Downs and Black checklist so it is not clear what evidence the company has used to

determine that a 'good' quality level on this checklist aligns with a low risk of bias. In our view, a single-arm study is inherently likely to be at a high risk of bias because there is no comparator/control group to mitigate confounding.

The EAG's additional concerns about using the Downs and Black checklist are that the questions are equally weighted (scoring either 0 or 1) and then the scores are summed to give the overall 'quality' score. There is the potential for a study to achieve a relatively good score even if it is at a high risk of bias. Some of the questions asked by the checklist assume that there is a comparator arm (e.g. "Are the distributions of principal confounders in each group of patients to be compared clearly described?") which is not the case for the company's DeLLphi-301 study 10mg target dose group in Part 1 and 2 of the trial.

To compare the company's assessment and with our own we have completed the Downs and Black Checklist which can be seen in full in Appendix 3. As noted above, we are assessing the quality of the 10mg tarlatamab target dose as a 'single arm' of the DeLLphi-301 study (i.e. the participants receiving the target dose in both Parts 1 and 2 of the trial). This is important because for questions such as "Have all important adverse events that may be a consequence of the intervention been reported?" we agree that this is the case for the DeLLphi-301 10mg target dose group, but note that, in our view, adverse events have not been fully reported in the CS (e.g. data were not reported for the participants in Part 3 receiving the 10mg dose, but were provided in response to clarification question A14). Our assessment using the Downs and Black checklist differed from that of the company for five of the 26 questions and for one question we judged unclear (this was scored as zero). We do not agree with summing scores but for ease of comparison our assessment would be scored 15 (quality level 'fair') whereas the company assessment scored 20 (quality level 'good').

We asked the company to provide a critical appraisal of the DeLLphi-301 trial using the criteria for non-randomised and non-controlled studies as set out in section 2.5 of the NICE company evidence submission template guidance.[28] Process and methods' guide' (clarification question A7). The company's critical appraisal using these criteria is shown in Table 6 with our judgements using these criteria and comments added. Our judgements differ from those of the company regarding whether the outcome was accurately measured to minimise bias and whether the follow-up of patients was complete. In the latter case, this may be down to differences in how we have interpreted in the question. The company have focused on the thoroughness of follow-up procedures whereas we believe the question is

asking about the length of follow-up (sufficient to capture the outcomes of interest) and completeness of follow-up (i.e. have all patients completed the study).

Table 6 Critical appraisal of the Dellphi-301 trial using NICE's criteria for non-randomised and non-controlled studies.

Question	Company answer	Supporting information for answer	EAG answer and comments
Was the cohort recruited in an acceptable way?	Yes	The cohort was recruited based on strict inclusion criteria, including confirmed SCLC, previous treatments, and measurable lesions, ensuring a representative sample of the target population.	**Yes**, study enrolment criteria ensure the patients represent the relevant population, although only five patients were recruited from the UK.
Was the exposure accurately measured to minimise bias?	Yes	The administration of tarlatamab was consistent across the study with specific dosages (10 mg and 100 mg) and schedules, and adherence was monitored closely. Adherence challenges were noted with 6 patients in the 10 mg group and 14 patients in the 100 mg group dying before the post-baseline scan.	**Yes**. CS section B.2.10.2 provides reporting on exposure at the data cut-off date (27th June 2023).
Was the outcome accurately measured to minimise bias?	Yes	Outcomes were measured using objective criteria (RECIST 1.1) and assessed by blinded independent central review. Objective	**Uncertain**. Objective criteria assessed by BICR, however we have concerns that blinding could have been compromised if the BICR body [redacted]

Question	Company answer	Supporting information for answer	EAG answer and comments
		response rates were 40% (97.5% CI: 29-52) for the 10 mg group and 32% (97.5% CI: 21-44) for the 100 mg group.	[redacted]. In their response to clarification question A8, the company did not address this point.
Have the authors identified all important confounding factors?	Yes	The study design and statistical analysis accounted for several potential confounders, including prior treatments and patient baseline characteristics.	**Yes**. The CSR lists a number of covariates that were included, where appropriate, in subgroup and multivariate analyses. Covariates included platinum sensitivity, and brain and liver metastases at baseline.
Have the authors taken account of the confounding factors in the design and/or analysis?	Yes	Confounders were considered in the study design and analyses, as seen in the stratification of patients and adjustment in the statistical models.	**Yes**. The CSR states that [redacted].
Was the follow-up of patients complete?	Yes	Follow-up was thorough, with scheduled imaging assessments and safety follow-ups detailed in the study protocol.	**No**. Study is ongoing and of the 99 patients who received tarlatamab in the 10mg target group (Part 1 and Part 2 of the study) 29 (29.3%) were still continuing treatment at the interim analysis.
How precise (for example, in terms of confidence interval and p values) are the results?	N/A	Results of all primary and secondary efficacy analyses were reported with appropriate confidence interval margins.	Confidence intervals are reported where appropriate but, because the study is still ongoing, some values are not estimable yet.

Source: Based on the table in company clarification response A7 supplemented with EAG answers and comments.

BICR, blinded independent central review; CS, company submission; CSR, clinical study report; EAG, External Assessment Group; MAIC, matching-adjusted indirect comparison; N/A, not applicable; RECIST, response evaluation criteria in solid tumours.

EAG comment on risk of bias assessment

Overall, we find the company's DeLLphi-301 trial to be a well conducted single-arm study which is subject to the risk of bias that is associated with this study design. We have particular concerns about the potential for blinding of the independent central review body to have been undermined if it was [redacted]. The study is ongoing so final results are not yet available.

3.2.3 Outcomes assessment

The CS lists the outcomes measures included in the DeLLphi-301 trial in CS Table 4 (three primary and eight 'key' secondary).The EAG notes that two of the three listed primary outcomes in CS Table 4 also appear in the list of key secondary outcomes in Table 4 (treatment emergent adverse events and serum concentrations of tarlatamab). CS Table 5 provides outcome definitions for a single primary outcome, objective response rate, and five secondary endpoints. The outcomes presented in CS Table 4 and CS Table 5 include all those listed in the NICE scope and company decision problem except for HRQoL. We note however that CS section B.2.6.3 presents results for patient reported outcomes obtained using the European Organization for Research and Treatment of Cancer (EORTC) quality of life questionnaire lung cancer module (QLQ-LC13), subscale global health status/quality of life (QoL) and the physical functioning scale of the EORTC quality of life questionnaire core 30 (QLQ-C30). These HRQoL outcomes are described as 'exploratory endpoints' and in response to clarification question A12 the company confirmed that the EQ-5D-5L questionnaire (including visual analogue scale [VAS] scores) was also used and symptom burden (i.e. cough, chest pain and dyspnoea) was measured by the Patient Global Impression of Severity/Change (PGI-S and PGI-C). As stated in section 3.2.1, the data cut-off date for the outcomes presented in the CS was 27th June 2023 (median length of treatment was [redacted] weeks (range [redacted] to [redacted] weeks).

The outcomes that inform the economic model are:

- OS
- PFS

- TTD (note this outcome was not reported in CS section B.2.6 so we requested the results in clarification question A11)
- EQ-5D utilities (as no EQ-5D results were reported in CS section B.2.6 we requested a summary of the VAS scores in clarification question A13)
- Adverse events (Grade 3+ events and Grade 1-2 CRS and ICANS).

3.2.3.1 Efficacy outcome(s)

In this section we focus on the primary outcome of objective response and the efficacy outcomes that inform the economic model: OS, PFS and TTD. The CS does not state the frequency of outcome assessments but the trial publication[23] states that imaging assessments were scheduled every six weeks for the first year and every 12 weeks thereafter.

CS Table 5 states that the primary outcome of objective response rate was defined as the proportion of patients with best overall response of complete response or partial response assessed using the RECIST (Response Evaluation Criteria in Solid Tumours) criteria version 1.1. The primary analysis, which is presented in the CS, is based on the assessment by blinded independent central review (BICR). We asked the company to confirm how disease progression was assessed by the BICR body (clarification question A8) and the company responded that [redacted]. It is not clear from the company's answer whether some images were sent to the BICR body from unscheduled visits when no progression was detected nor whether the BICR body was aware that at least some, if not all, of the images sent from unscheduled visits [redacted]. Our concern is that if the BICR body knew that images from unscheduled visits [redacted] that this could lead to unblinding of the BICR body as noted in our risk of bias assessment in section 3.2.2. The CS presents results from the BICR Full Analysis Set (CS Table 8) which includes all patients from Part 1 or Part 2 of the trial who received at least one dose of tarlatamab and who had one or more measurable lesions at baseline that could be assessed by BICR using the RECIST 1.1 criteria.

OS was calculated as "*the time from the first dose of tarlatamab to death due to any cause. Any patients known to be alive were censored either at the date last known to be alive or at the analysis data cut-off date, whichever was earlier*".

PFS was calculated as the time from the first dose of tarlatamab to the earliest event of progressive disease assessed by RECISTS 1.1 or death due to any cause. Three groups of patients were censored in the PFS analysis (these are described in CS Table 5) and the EAG does not have any concerns about them.

TTD was not included as an outcome in the clinical effectiveness results section of the CS (CS B.2.6) but we asked the company to provide these data (clarification question A11). We requested the TTD results (with and without adjustment for post-progression tarlatamab use) because TTD from the real-world UK Cancer Analysis System (CAS) Registry dataset was used as a proxy for PFS estimates for the standard of care in the matching-adjusted indirect comparison (MAIC) (see section 3.3 of this report for our critique of the MAIC). In the unadjusted TTD analyses (which include the use of tarlatamab after progression) participants still in receipt of tarlatamab at the data cut-off were censored and their censoring time was the time of their last tarlatamab exposure before the data cut-off. Discontinuation of tarlatamab by participants prior to the data cut-off was counted as an event and the event time was the time of the participants last tarlatamab exposure. For the adjusted TTD analyses (which aim to provide an estimate of the TTD that would have been observed if there had been no post-progression use of tarlatamab) those participants who had continued to receive tarlatamab after progression had their event times altered so that their TTD was set to the time of progression (i.e. they were treated as having discontinued tarlatamab at the time of progression). Participants who did not have post-progression use of tarlatamab were treated in the same way as in the unadjusted analysis.

3.2.3.2 HRQoL outcomes

HRQoL outcomes are not described in CS section B.2.3 which summarises the methodology of the relevant clinical effectiveness evidence (although the abbreviation HRQoL is listed at the foot of CS Table 4 the term HRQoL does not appear in that table). However, some results are presented for patient-reported outcomes in CS section B.2.6.3 'Exploratory endpoints'. We asked the company to describe these endpoints (clarification question A12). In response to clarification question A12 the company stated that the following measures were used:

- The EORTC 30-item Quality of Life Questionnaire (QLQ-C30),[29]
- The 13-item lung cancer module (QLQ-LC13)
- The EQ-5D-5L questionnaire (including VAS scores)
- The Patient Global Impression of Severity/Change (PGI-S and PGI-C)

- The GP5 question of the Functional Assessment of Cancer Therapy – General Form (FACT-G).[30]

The company did not provide details of the score ranges or direction of scoring for the above measures.

3.2.3.3 Safety outcomes

The Safety Analysis Set was defined as all participants who received at least one dose of tarlatamab (CS Table 8). However, the company explain in CS section B.2.10 that adverse events are only reported in the CS for the subgroup of participants from the Safety Analysis Set who received the [redacted] 10 mg dose of tarlatamab in either Part 1 or part 2 of the trial. Participants who received 10 mg of tarlatamab in Part 3 of the trial were not included. The company provided the data for the Part 3 participants in response to Clarification question A14. In response to clarification question A12 the company state that patient reported adverse events were measured using selected questions from the Patient reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE).

Adverse events were defined as treatment-emergent when they occurred after the first dose of tarlatamab was administered and up to and including 47 days after the last dose of tarlatamab was administered, or the end of study date (whichever was earlier). The study publication[23] states that:

- adverse events were graded using the CTCAE version 5.0, which incorporates some elements of the Medical Dictionary for Regulatory Activities (MedDRA) version 26.0 terminology.
- CRS events and ICANS events were identified on the basis of a narrow search and a broad search respectively for preferred terms in the MedDRA version 26.0.
- the severity of CRS events and ICANS events were graded according to the American Society for Transplantation and Cellular Therapy 2019 consensus guidelines.

A safety follow-up visit occurred approximately 42 (+5) days after the last dose of tarlatamab (CS section B.2.3.1) even if participants were receiving a subsequent anti-cancer therapy.

EAG comment on outcomes assessment

We consider that the outcome measures reported for the DeLLphi-301 trial are appropriate. The company include those listed in the NICE scope and the company decision problem. We have no concerns about the outcome definitions, but we note that

the company did not provide details of the score ranges or direction of scoring for the exploratory HRQoL measures used in the DeLLphi-301 trial.

3.2.4 Statistical methods of the included studies

The EAG's critique of the statistical methods used in the DeLLphi-301 trial is presented in Table 7.

Table 7 EAG critique of the statistical methods of the DeLLphi-301 trial

DeLLphi-301
Analysis populations
Brief description CS Table 8 defines the DeLLphi-301 analysis populations, two of which are presented in the CS: the BICR Full Analysis Set (primary analysis set) and the Safety Analysis Set. BICR Full Analysis Set defined as all patients who were randomised (Part 1) or enrolled (Part 2), who received at least one dose of tarlatamab, and had one or more measurable lesions at baseline as assessed by BICR using the RECIST 1.1 criteria. Safety Analysis Set defined as all patients who received at least one dose of tarlatamab.
EAG comment: The BICR Full Analysis Set does not contain any patients from Part 3 of the trial. Additionally, although CS Table 8 states that this population is presented in the CS, in terms of results, the CS only presents results for those participants who received the 10mg tarlatamab dose, because this is the [redacted] dose. CS Figure 3 shows that 99 of a possible 100 participants due to receive 10mg tarlatamab were included in the CS BICR Full Analysis Set results because one participant did not receive any treatment. The CS also presents data for the Safety Analysis Set only for those patients who received the [redacted] dose of 10mg tarlatamab in Parts 1 and 2 of the trial. The EAG views it as acceptable to have limited reporting to those participants who received the [redacted] dose of tarlatamab. However, we note that 34 participants enrolled in Part 3 of the trial received the 10mg dose, yet safety results were not presented for these participants in the CS. We believe that the safety results from these participants would be valuable, so we therefore requested these results in clarification question A14 and the company provided them in response.
Sample size calculations
Brief description

DeLLphi-301
It was not clear to the EAG if a formal method had been used to determine sample size. In response to clarification question A18 the company explain the rationale for their chosen sample size, expanding on the information provided in CS Table 9. The company planned to enrol approximately 100 participants in total at the selected target dose for Part 1 and Part 2 of the trial. Initially patients would be enrolled and randomised in Part 1 to one of two doses (tarlatamab 100mg or 10mg, approximately 90 participants expected to be enrolled at each dose). An interim analysis took place to select a target dose for expansion after 30 patients had been recruited to each of the Part 1 doses, with recruitment continuing to Part 1 as the interim analysis took place. When the target dose had been selected the final approximately 10 participants were enrolled to Part 2 at the selected dose. The results from two open-label studies, one of pembrolizumab[31] and one of nivolumab,[32] which reported objective response rates (ORRs) of 19% and 12% respectively for patients with SCLC who had received two or more previous lines of therapy (i.e. the current therapy was third line or later) informed the company's determination of the appropriate sample size. For Parts 1 and 2 of the study, which are the focus of the CS, a sample size of 100 participants receiving the 10mg target tarlatamab dose was estimated to provide a 63% probability of observing at least one adverse event with a true 1% incidence rate, and a 99% probability of observing at least one adverse event with a true 5% incidence rate. For Part 3 of the study a sample size of 30 receiving the 10mg target tarlatamab dose was estimated to provide a 26% probability of observing at least one adverse event with a true 1% incidence rate, and a 79% probability of observing at least one adverse event with a true 5% incidence rate.
EAG comment: The sample size of 100 participants in total at the selected target dose for Part 1 and Part 2 of the trial appears to have been calculated appropriately and 99 participants are included in the BICR full analysis set for Parts 1 and 2 at the 10mg target dose.
Methods to account for multiplicity
Brief description To adjust for multiplicity a 97.5% two-sided confidence interval was used for the ORR whereas for other outcomes at 95% two-sided confidence interval was used. In response to clarification question A20 the company stated their adjustment for multiplicity was based on the Bonferroni method.
EAG comment: The company had an appropriate procedure in place to reduce the risk of statistically significant effects being detected by chance.

DeLLphi-301
Analysis of outcomes
Brief description Time to event variables (duration of response, duration of disease control, PFS, OS, and TTD, the latter provided in response to clarification question A11): reported as Kaplan-Meier estimates of the median time to event and percentiles with associated two-sided confidence intervals [97.5% confidence intervals (CIs) for ORR and 95% CIs for other outcomes and the subgroup analyses of ORR]. Confidence intervals for Kaplan-Meier estimates of median time to event were calculated using the Brookmeyer and Crowley method. Landmarks for time to event endpoints (e.g. 1-year OS): estimated using Kaplan-Meier methods. The Greenwood formula was used to estimate the standard error subsequently used in the calculation of the confidence interval. Proportions: Confidence intervals were estimated using the Clopper-Pearson method[33] Censoring rules: Provided in CS Table 5 for PFS and OS and in the response to clarification question A11 for TTD.
EAG comment: Appropriate analysis methods have been used for the primary outcome and the time to event outcomes, including those that inform the economic model. However, we note that CS section B.3.3.1 states that [redacted] participants continued tarlatamab post-disease progression in total in the study; whereas, CS section B.3.3 states that this figure was [redacted] participants. Patients continuing tarlatamab post-progression do not appear to have been censored from the analyses of the participants receiving the 10mg target dose in Parts 1 and 2 of the study, the results of which are presented in CS section B.2.6. Given this, some of the results, depending on the outcome, may not fully reflect those that might occur in clinical practice if tarlatamab is approved, because, as is stated in section 3.2.1.1, [redacted].
Handling of missing data
Brief description Table 9 states that imputation for missing or incomplete data was performed if required. However, in response to clarification question A16 the company states that no imputation was performed for efficacy endpoints or PRO endpoints. If dates were missing or incomplete for adverse events and concomitant medication use, imputation was carried out as per the statistical analysis plan. In response to clarification question A15 the company state that reasons for missing PRO data (other than disease progression or death) were not collected in DeLLphi-301 and they provide details of the completion (for all randomised participants) and compliance (for participants still in the study who would

DeLLphi-301
be expected to be able to complete the PRO instrument at that visit) rates for the QLQ-C30, QLQ-LC13 and EQ-5D-5L scales at study visits from baseline to Cycle 15 Day 1.
EAG comment: For the efficacy and PRO endpoints no imputation for missing data was used. For the primary outcome, as patients without a post-baseline tumour assessment (n=[redacted], CS Table 11) were considered non-responders (CS Table 5), this would lead to a conservative estimate of treatment effect.
Sensitivity & post-hoc analyses
Brief description Few sensitivity analyses were described in the CS so we sought clarification (clarification question A17). In response to clarification question A17 the company confirmed that sensitivity analyses were conducted using an alternative censoring rule for DOR, DoDC and PFS and they provide the results of these analyses which are similar to the primary analysis.
EAG comment: Appropriate sensitivity analyses have been conducted to test the use of an alternative censoring rule.

Source: Table compiled by the EAG
BICR, blinded independent central review; CI, confidence interval; CS, company submission; DOR, duration of response; DoDC, duration of disease control; EAG, external assessment group; MAIC, matching-adjusted indirect comparison; OR, objective response; ORR, objective response rate; OS, overall survival; PFS, progression-free survival; RECIST, response evaluation criteria in solid tumours; SCLC, small-cell lung cancer; TTD, time to treatment discontinuation.

EAG comment on study statistical methods

The statistical methods used in the DeLLphi-301 study appear to be appropriate and the EAG has no concerns, but it should be noted that the trial results presented in CS section B.2.6 include the participants who received tarlatamab post-progression. We note a lack of clarity in the CS about the number of participants who received tarlatamab post-progression.

3.2.5 Efficacy results of the intervention studies

In this section we report on the primary outcome of the DeLLphi-301 trial (objective response), the outcomes that inform the economic model (OS, PFS and TTD), HRQoL outcomes, subgroup analyses for the primary outcome and safety outcomes. The CS additionally provides data for the secondary outcomes of duration of response and disease control (CS section B.2.6.2) which we do not include here.

3.2.5.1 Objective response (Primary outcome)

The objective response rate (best overall response of either complete response or partial response) for the 99 participants assigned to the 10mg dose of tarlatamab in either Part 1 or Part 2 of the DeLLphi-301 study was 40.4% (97.5% confidence interval [CI] 29.4% to 52.2%). Only one (1.0%) participant had a confirmed complete response and 39 (39.4%) had a confirmed partial response. Response outcomes do not contribute data to the economic model. CS Table 11 provides a summary of objective response.

3.2.5.2 OS

Median OS for the 99 participants who received at least one dose of tarlatamab after assignment to the 10mg dose in either Part 1 or Part 2 of the DeLLphi-301 trial was 14.3 months (95% CI: 10.8 months to not estimable) with a median follow-up time of 10.6 months (Table 8 and Figure 1). Thirty-five participants (35.4%) had experienced an event at the time of analysis, and, of the remaining 64 participants who were censored, 57 were alive at the last follow-up and continuing in the study, six had withdrawn consent from the study and one had been lost to follow-up.

Table 8 Analysis of OS (safety analysis set)

	Participants, N
Number of patients who received at least 1 dose (1-10mg[a]) of tarlatamab	99
Patient status	Participants, N (%)
Events	35 (35.4)
- Death	35 (35.4)
Censored	64 (64.6)
- Alive at last follow-up	57 (57.6)
- Withdrawal of consent from study	6 (6.1)
- Decision by sponsor	0 (0.0)
- Lost to follow-up	1 (1.0)
- Completed study without death	0 (0.0)
OS (KM) (months)[b]	
25th percentile (95% CI)	5.7 (4.7, 10.5)
Median (95% CI)	14.3 (10.8, NE)
75th percentile (95% CI)	NE (NE, NE)
Min, Max (+ for censored)	0.3+, 15.2+
Follow-up time of OS (KM) (months)[b]	
25th percentile (95% CI)	[redacted]
Median (95% CI)	10.6 (9.2, 11.5)
75th percentile (95% CI)	[redacted]
Min, Max (+ for censored)	[redacted]

	Participants, N
Kaplan-Meier estimate (%) (95% CI)[c]	
At 3 months	88.7 (80.5, 93.6)
At 6 months	73.4 (63.2, 81.2)
At 9 months	68.0 (57.1, 76.6)
At 12 months	[redacted]

Source: Reproduced from CS Table 15 with the addition of footnote 'a' by the EAG and minor formatting alterations.
BICR, blinded independent central review; CI, confidence interval; EAG, External Assessment Group; KM, Kaplan-Meier; Max, maximum; Min, minimum; N, number; OS, overall survival.
[a] Participants enrolled in Part 1 of the study and randomised to receive 10mg tarlatamab and participants due to receive the selected 10mg dose in Part 2 of the study, received 1mg of tarlatamab for the first dose (Cycle 1, Day 1) before receiving step-up doses of 10mg tarlatamab. This was to mitigate the risk of cytokine release syndrome.
[b] Median and quantiles were estimated using Kaplan-Meier method and 95% CI of median were estimated using log-log transformation of KM survival estimate by Brookmeyer and Crowley (1982) method[34]
[c] 95% Cis were estimated using Kalbfleisch and Prentice (1980) method[35]

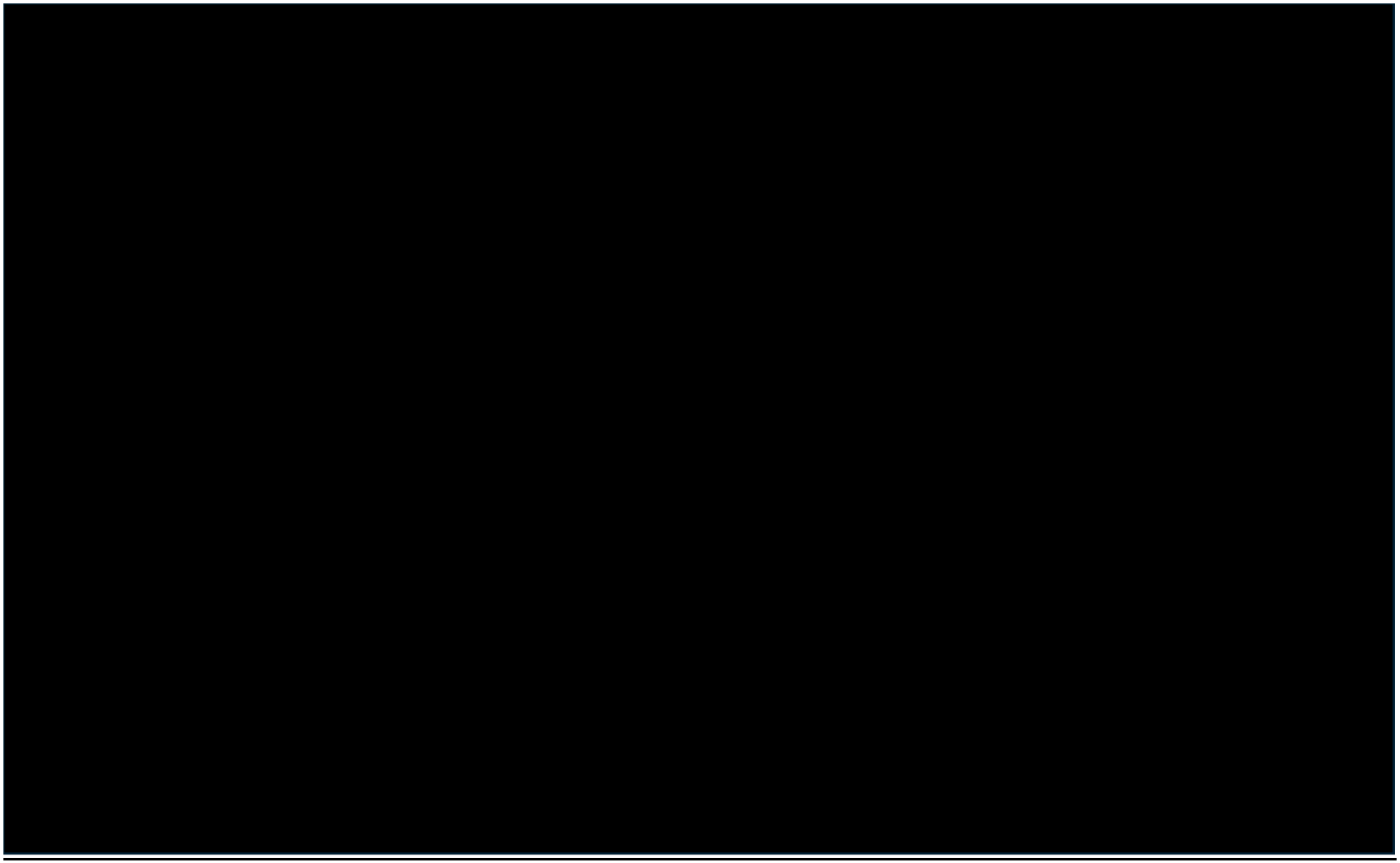

Figure 1 KM plot for overall survival (Safety analysis set)

Source: Reproduction of CS Figure 7
KM, Kaplan-Meier

3.2.5.3 PFS

Median PFS for the 99 participants who received at least one dose of tarlatamab after assignment to the 10mg dose in either Part 1 or Part 2 of the DeLLphi-301 as assessed by BICR was 4.9 months (95% CI: 2.9 to 6.7 months) with a median follow-up time of [redacted] months (Table 9 and Figure 2). There were eight (8.1%) deaths, 56 (56.6%) participants with disease progression and the remaining 35 (35.4) participants were censored, primarily because they remained in the study without disease progression or death (25 participants).

Table 9 PFS as assessed by BICR (BICR Full Analysis set for Part 1 and 2; 10 mg dose)

	Participants, N
Number of patients who received at least 1 dose (1-10mg[a]) of tarlatamab	99
Patient status	Participants, N (%)
Events	64 (64.6)
- Death	8 (8.1)
- Disease progression	56 (56.6)
Censored	35 (35.4)
- On study without disease progression or death	25 (25.3)
- No evaluable post-baseline disease assessment	2 (2.0)
- Missed 2 or more consecutive assessments	3 (3.0)
- Started new anti-cancer therapy	2 (2.0)
- Withdrawal of consent from study	3 (3.0)
- Decision by sponsor	0 (0.0)
- Lost to follow-up	0 (0.0)
- Completed study without disease progression or death	0 (0.0)
PFS (KM) (months)[b]	
25th percentile (95% CI)	2.4 (1.4, 2.8)
Median (95% CI)	4.9 (2.9, 6.7)
75th percentile (95% CI)	NE (7.1, NE)
Min, Max (+ for censored)	0.0+, 13.7+
Follow-up time of PFS (KM) (months)[b]	
25th percentile (95% CI)	[redacted]
Median (95% CI)	[redacted]
75th percentile (95% CI)	[redacted]
Min, Max (+ for censored)	[redacted]
Kaplan-Meier estimate (%) (95% CI)[c]	
At 3 months	58.8 (48.1, 68.1)
At 6 months	40.8 (30.6, 50.7)
At 9 months	28.5 (19.2, 38.6)

	Participants, N
At 12 months	[redacted]

Source: Reproduced from CS Table 14 with the addition of footnote 'a' by the EAG and minor formatting alterations.
BICR, blinded independent central review; CI, confidence interval; EAG, External Assessment Group; KM, Kaplan-Meier; Max, maximum; Min, minimum; N, number; PFS, progression-free survival.
[a] Participants enrolled in Part 1 of the study and randomised to receive 10mg tarlatamab and participants due to receive the selected 10mg dose in Part 2 of the study, received 1mg of tarlatamab for the first dose (Cycle 1, Day 1) before receiving step-up doses of 10mg tarlatamab. This was to mitigate the risk of cytokine release syndrome.
[b] Median and quantiles were estimated using Kaplan-Meier method and 95% CI of median were estimated using log-log transformation of KM survival estimate by Brookmeyer and Crowley (1982) method[34]
[c] 95% Cis were estimated using Kalbfleisch and Prentice (1980) method[35]

Figure 2 KM plot for PFS as assessed by BICR (BICR Full Analysis Set for Part 1 and Part 2)

Source: Reproduction of CS Figure 6
BICR, blinded independent central review; KM, Kaplan-Meier; PFS, progression-free survival.

3.2.5.4 TTD

In response to clarification question A11 the company provided data on TTD for the 99 participants who received at least one dose of tarlatamab in the 10mg target dose group. Results are provided including data for participants who continued to receive tarlatamab after BICR assessed progression (the unadjusted results) whose data were censored at the time of their last tarlatamab exposure before the data cut-off. Adjusted results are also provided in which the patients who continued to receive tarlatamab after progression were instead treated as if they had discontinued tarlatamab at the time of progression. Median TTD in the unadjusted analysis was [redacted] months and in the adjusted analysis was [redacted] months (Table 10). The Kaplan-Meier plot showing the unadjusted and adjusted curves together is provided in Figure 3 (figures showing the unadjusted and adjusted Kaplan-Meier plots in separate figures are available in clarification response A11 Figure 1 and Figure 2).

Table 10 TTD

	Participants, N	
Number of patients who received at least 1 dose (1-10mg[a]) of tarlatamab	99	
Patient status	Participants, N (%)	
	Unadjusted	Adjusted
Event (discontinuation), n (%)	[redacted]	[redacted]
Censored (still on treatment), n (%)	[redacted]	[redacted]
TTD (KM) (months)[b]		
25th percentile (95% CI)	[redacted]	[redacted]
Median (95% CI)	[redacted]	[redacted]
75th percentile (95% CI)	[redacted]	[redacted]
Min, Max (+ for censored)	[redacted]	[redacted]
Follow-up time (months)b		
25th percentile (95% CI)	[redacted]	[redacted]
Median (95% CI)	[redacted]	[redacted]
75th percentile (95% CI)	[redacted]	[redacted]
Min, Max (+ for censored)	[redacted]	[redacted]
Kaplan-Meier estimate (%) (95% CI)c		
At 3 months	[redacted]	[redacted]
At 6 months	[redacted]	[redacted]
At 9 months	[redacted]	[redacted]
At 12 months	[redacted]	[redacted]

Source: Reproduced from company response to clarification question A11 with minor formatting alterations by the EAG.
CI, confidence interval; EAG, External Assessment Group; KM, Kaplan-Meier; Max, maximum; Min, minimum; N, number; TTD, time to treatment discontinuation

[a] Participants enrolled in Part 1 of the study and randomised to receive 10mg tarlatamab and participants due to receive the selected 10mg dose in Part 2 of the study, received 1mg of tarlatamab for the first dose (Cycle 1, Day 1) before receiving step-up doses of 10mg tarlatamab. This was to mitigate the risk of cytokine release syndrome.
[b] Median and quantiles were estimated using Kaplan-Meier method and 95% CI of median were estimated using log-log transformation of KM survival estimate by Brookmeyer and Crowley (1982) method[34]
[c] 95% Cis were estimated using Kalbfleisch and Prentice (1980) method[35]

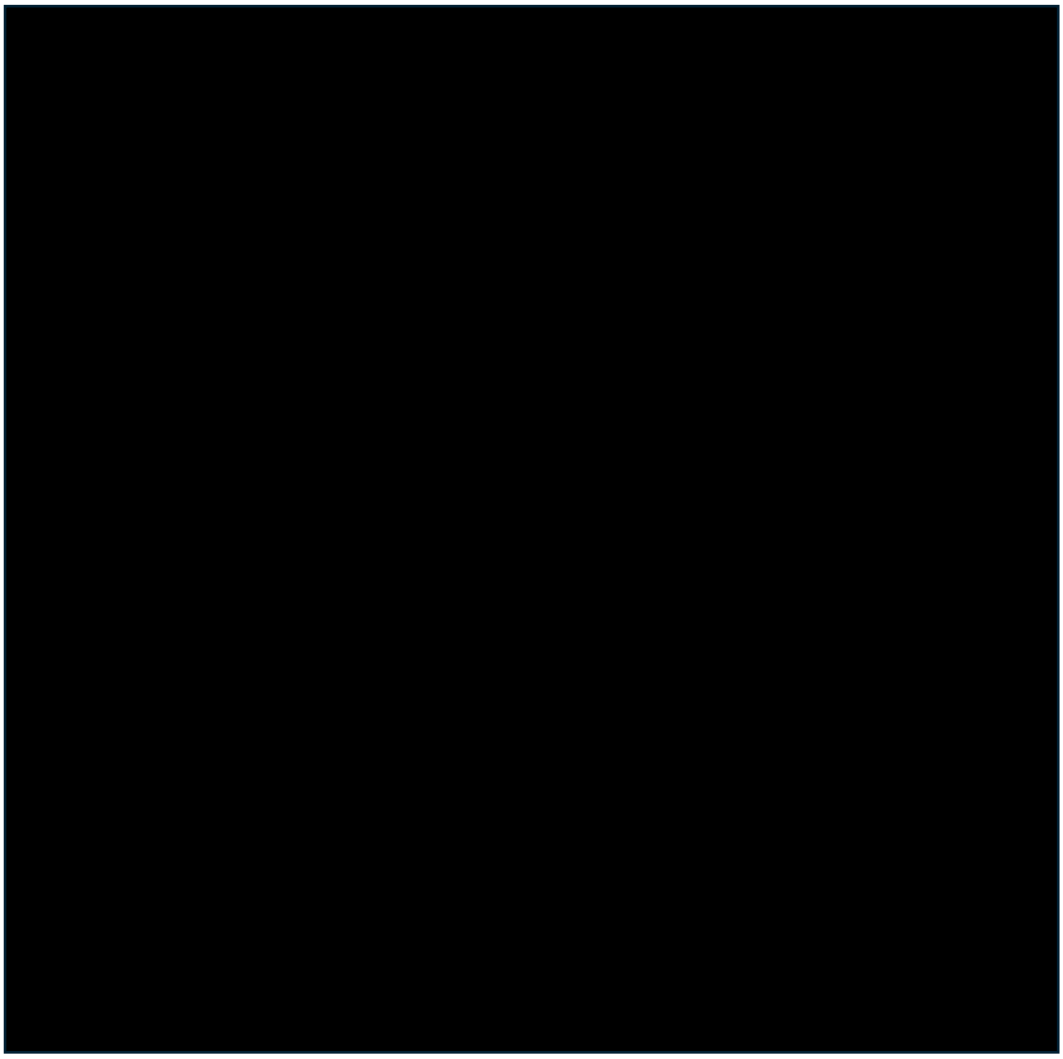

Figure 3 KM plot for time to discontinuation (unadjusted and adjusted)

Source: Reproduced from company response to clarification question A11, Figure 3.
CI, confidence interval; TTD, time to treatment discontinuation

3.2.5.5 HRQoL outcomes

CS section B.2.6.3 presents results of patient-reported outcomes which were exploratory outcomes in DeLLphi-301 and in response to clarification question A13 the company provided a summary of the VAS scores. These results are summarised in Table 11.

The company state in clarification response A13 that, reflecting the trial entry criteria that patients should have an ECOG performance status of 0 or 1, patients had generally moderate to good health at baseline by the EQ-5D-5L visual analogue scale (where 0 is the worst and 100 is the best health imaginable). The global health status/QoL and physical function baseline scores from the EORTC QLQ-C30 were below those for the normative sample of patients with SCLC (i.e. they were worse).

EORTC QLQ-LC13 least squares mean (LSM) score changes from baseline for cough, chest pain and the dyspnoea composite [redacted] indicating a [redacted] in these symptoms up to cycle 12. The EORTC-QLQ-C30 global health status/QoL LSM score [redacted] the baseline value indicating [redacted] QoL at cycle 12. The MMRM analysis of the EQ-5D-5L VAS scores showed that this [redacted] with respect to baseline in cycle 12 and during treatment the change from baseline ranged from [redacted] to [redacted].

Health state utility values in the economic model were based upon EQ-5D-5L data collected in the DeLLphi-301 trial which is discussed further in section 4.2.8.2 of this report.

Table 11 Patient reported outcomes results in DeLLphi-301

	Mean baseline score	LSM change from baseline
EORTC QLQ-LC13[a]		Up to cycle 12
- cough	[redacted]	[redacted] (95% CI: [redacted])
- chest pain	[redacted]	[redacted] (95% CI: [redacted])
- dyspnoea composite	[redacted]	[redacted] (95% CI: [redacted])
EORTC QLQ-C30[b]		At cycle 12
Global health status/QoL	[redacted][c]	[redacted] (95% CI: [redacted])
Physical function	[redacted][c]	Not reported
EQ-5D-5L		MMRM analysis LSM change from baseline in cycle 12
EQ-5D-5L VAS[d]	[redacted]	[redacted][e] (95% CI: [redacted])
		During treatment (up to cycle 15) Mean change from baseline
EQ-5D-5L VAS[d]	[redacted]	Range [redacted] to [redacted]

Source: EAG table based on information presented in CS section B.2.6.3 and the company response to clarification question A13.
EORTC, European Organization for Research and Treatment of Cancer; LSM, least squares mean; QLQ-C30, quality of life questionnaire core 30; QLQ-LC13, quality of life questionnaire lung cancer module
[a] Score range 0-100. A high score represents a high level of symptoms.[36]
[b] Score range 0-100. A high score for global health status/QoL represents high QoL. A high score for physical function represents a high level of physical functioning.[29,37]
[c] These baseline scores were below those for the normative sample of patients with SCLC ([redacted])
[d] Score range 0-100, where 100 is the best possible health state.

[e] In the response to clarification question A13 this is given as [redacted] which appears to be an error. The value in CSR Table 14-4.2.806 is [redacted].

3.2.5.6 Subgroup analyses

CS Table 4 lists the 10 patient characteristics for which pre-planned subgroup analyses were conducted. Subgroup analysis results are presented in CS section B.2.7 and CS Figure 8. An objective response was seen across [redacted] subgroups (with the exception of the [redacted] which included [redacted] and [redacted] patients, respectively).

3.2.5.7 Safety outcomes

The CS presents safety outcomes for the 99 participants in the Safety analysis set who were in the 10mg target dose group because tarlatamab is [redacted] (CS section B.2.10). Participants who were enrolled in Part 3 of the study (n=34) were not included in CS section B.2.10 so the EAG asked the company to provide a summary of the adverse event data for these patients in clarification question A14.

Adverse events with an onset that occurred after the first dose of tarlatamab was administered and up to and including 47 days after the last dose of tarlatamab or the end of study date (whichever was earlier) were defined as treatment-emergent. The CS summarises the incidence of treatment-emergent adverse events (TEAEs) in section B.2.10.1 and CS Table 21. Greater detail on TEAEs is then provided in CS sections B.2.10.3 (most frequent TEAEs), B.2.10.4 (treatment-related TEAEs), B.2.10.5 (serious TEAEs), B.2.10.6 (fatal TEAEs) and B.2.10.7 (adverse events of special interest).

3.2.5.7.1 Summary of TEAEs for participants in Part 1 and Part 2 10mg target dose group in the DeLLphi-301 study

In the tarlatamab 10mg target dose group for Part 1 and Part 2 of the study 97% of the 99 participants had at least one adverse event. Over half the participants (57.6%) experienced a TEAE of Grade ≥3 (Table 12) and TEAEs were considered treatment-related for 29.3% of participants. For serious adverse events, which were reported for 58.6% of the 10mg target dose participants, 37.4% were considered treatment-related. Seven participants (7.1%) discontinued tarlatamab because of an adverse event and there were three fatal adverse events.[23] None of the fatal adverse events were considered to be related to tarlatamab.

Table 12 Summary of incidence of treatment-emergent adverse events (Safety Analysis Set; 10 mg Parts 1 and 2)

Treatment emergent adverse events, n (%)	Tarlatamab 10mg target dose (N=99)
All treatment-emergent adverse events	96 (97.0)
- Grade ≥2	86 (86.9)
- Grade ≥3	57 (57.6)
- Grade ≥4	16 (16.2)
- Serious adverse events	58 (58.6)
Leading to dose interruption and/or reduction of tarlatamab	31 (31.3)
- Leading to discontinuation of tarlatamab	7 (7.1)
- Serious	[redacted]
- Nonserious	[redacted]
- Fatal adverse events	3 (3.0)

Source: Reproduction of CS Table 21 with some minor editing by the EAG.
N or n, number

3.2.5.7.2 Most frequent treatment-emergent adverse events

CS Table 23 presents the incidence of TEAEs that occurred in more than 10% of patients. The most frequent event, experienced by almost half of the participants (n=49, 49.5%) was CRS and pyrexia (fever) occurred in over a third of participants (n=38, 38.4%). Constipation (28.3%), anaemia (26.3%), decreased appetite (25.3%), dysgeusia (24.2%), fatigue (21.2%) and asthenia (20.2%) were the next most frequently experienced events, affecting 20% or more of the participants. The remaining nine TEAEs listed in CS Table 23 were experienced by between 10% and 15% of participants.

3.2.5.7.3 Treatment-related treatment-emergent adverse events

Treatment-related TEAEs that occurred in more than 5% of participants overall are summarised in CS Table 24. The eight most common treatment-related TEAEs are the same as the eight most common TEAEs (CS Table 23). We have combined elements of CS Table 23 and CS Table 24 to provide an overview of the eight most common TEAEs and show what proportion of these events were considered to be treatment-related (Table 13). All of the CRS TEAEs were considered to be treatment related and 89.5% of the pyrexia TEAEs. Decreased appetite and dysgeusia (altered taste) were also commonly considered treatment related. Of the top eight TEAEs, constipation was the event that was considered to be treatment related in the smallest proportion of participants (42.9%).

Table 13 Overview of the eight most common TEAEs and the proportion of these that were considered treatment-related

Preferred Term, n (%)	Tarlatamab 10mg target dose (N=99)		
	Patients with TEAEs N (%)	Patients with treatment-related TEAEs N (%)	Proportion of TEAEs considered treatment-related
Overall	96 (97.0)	89 (89.9)	92.7%
The eight most common TEAEs			
Cytokine release syndrome	49 (49.5)	49 (49.5)	100%
Pyrexia	38 (38.4)	34 (34.3)	89.5%
Constipation	28 (28.3)	12 (12.1)	42.9%
Anaemia	26 (26.3)	16 (16.2)	61.5%
Decreased appetite	25 (25.3)	21 (21.2)	84.0%
Dysgeusia	24 (24.2)	20 (20.2)	83.3%
Fatigue	21 (21.2)	14 (14.1)	66.7%
Asthenia	20 (20.2)	14 (14.1)	70.0%

Source: Table compiled by EAG by combining elements of CS Table 23 and CS Table 24 with the proportion of TEAEs considered treatment-related calculated by the EAG.
TEAE, treatment-emergent adverse event

3.2.5.7.4 Serious adverse events

Serious adverse events (TEAEs of Grade 3 or higher) were reported for 57 participants in the tarlatamab 10mg target dose group in Part 1 and Part 2 of the DeLLphi-301 study. These are summarised in CS Table 25. The Grade ≥3 adverse events by preferred term that were the most frequently reported in the tarlatamab 10mg target dose group were anaemia ([redacted] %), lymphocyte count decreased ([redacted]%), lymphopenia ([redacted]%), fatigue ([redacted]%) and hyponatremia ([redacted]%). There were [redacted] serious adverse events among those with CRS.

3.2.5.7.5 Fatal treatment-emergent adverse events.

Three fatal adverse events occurred in the tarlatamab 10mg target dose group in Part 1 and Part 2 of the DeLLphi-301 study (CS Table 26). None of the fatal adverse events were considered by the investigator to be related to tarlatamab treatment (CS section B.2.10.1).

3.2.5.7.6 Adverse events of special interest

CS Table 27 reports data for four adverse events of special interest:

- CRS
- ICANS and associated neurological events

- Neurological events and psychiatric disorders
- Neutropenia

Neurological events ([redacted]%) and CRS (49.5%) were the most common adverse events of special interest among participants in the 10mg target dose group in Part 1 and Part 2 of the DeLLphi-301 study. There were [redacted]%) Grade ≥3 neurological events and no Grade ≥3 CRS events. For neutropenia, which was seen in [redacted]% of participants there were 6 (6.1%) Grade ≥3 events and ICANS was reported in 7 (7.1%) of participants with [redacted] Grade ≥3 ICANS events.

3.2.5.7.7 Adverse events for participants in Part 3 (10mg target dose) of the DeLLphi-301 study

In response to clarification question A14 the company provided a summary of TEAEs for the 34 participants in Part 3 of the DeLLphi-301 study and we have reproduced the company's table below as Table 14. The proportion of participants experiencing a TEAE in Part 3 of the study (100%) was similar to that of participants in Part 1 and Part 2 of the study (97%). The distribution of Grade ≥2, Grade ≥3 and Grade ≥4 events was also similar. The proportion of serious adverse events (41.2%) was a little lower than in Part 1 and Part 2 of the study (58.6%). A lower proportion of participants had an event leading to dose interruption and/or reduction of tarlatamab (14.7% versus 31.3% in Part 1 and Part 2 of the study). Most TEAEs were considered treatment-related (85.3%) which is similar to Part 1 and Part 2 of the study (89.9%). There were [redacted] fatal TEAEs, [redacted] of which was considered treatment related. The majority of other treatment-related TEAEs were Grade ≥2 (67.6%).

Table 14 Summary of incidence of TEAEs (modified safety monitoring dose group; Part 3)

Treatment-emergent adverse events, n (%)	Tarlatamab 10mg target dose (N=34)
All treatment-emergent adverse events	34 (100.0)
Grade ≥2	33 (97.1)
Grade ≥3	22 (64.7)
Grade ≥4	7 (20.6)
Serious adverse events	14 (41.2)
Leading to dose interruption and/or reduction of tarlatamab	5 (14.7)
Leading to discontinuation of tarlatamab	3 (8.8)
Serious	[redacted]
Nonserious	[redacted]
Fatal adverse events	[redacted]

Treatment-emergent adverse events, n (%)	Tarlatamab 10mg target dose (N=34)
Treatment-related treatment-emergent adverse events	29 (85.3)
Grade ≥2	23 (67.6)
Grade ≥3	5 (14.7)
Grade ≥4	2 (5.9)
Serious adverse events	7 (20.6)
Leading to dose interruption and/or reduction of tarlatamab	3 (8.8)
Leading to discontinuation of tarlatamab	0 (0.0)
Fatal adverse events	[redacted]

Source: Reproduction of Company's response to clarification question A14, Table 4 with minor amendment to heading of column 2.
N or n, number

3.2.6 Pairwise meta-analysis of intervention studies

The company state that meta-analysis was not conducted as DeLLphi-301 is the only trial of tarlatamab that they identified in decision problem indication (CS section B.2.8).

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

3.3.1 Rationale for ITC

The company carried out an ITC because the DeLLphi-301 trial was a single-arm study and the company's systematic literature review did not identify any relevant head-to-head trials of tarlatamab for which data were available (CS section B.2.9). As discussed in section 3.1, the company did not include search terms for randomised controlled trials in their search strategies, and therefore RCTs would not necessarily have been retrieved. Despite this limitation, the EAG do not expect there to be any RCTs available of the use of tarlatamab as a third-line therapy+ (see section 3.1). We agree it is appropriate that the company has conducted an ITC to compare the efficacy of tarlatamab to that of standard of care.

The company used unanchored MAIC methodology to compare the clinical efficacy of tarlatamab, in terms of OS and PFS, versus the single basket standard of care comparator defined in the company decision problem, which consisted of CAV, topotecan and carboplatin + etoposide (see section 2.3). The OS and PFS estimates from the unanchored MAIC informed the company's economic model (CS section B.3.3). An unanchored MAIC approach is suitable when only single-arm trials are available or there is a disconnected network, as in the present case.[1] The company describe the unanchored MAIC methodology

used in CS section B.2.9 and CS Appendix D, sections D.1.4 to D.1.9. The company additionally supplied a data-on-file, confidential report of the MAIC with the CS.[2]

3.3.2 Identification, selection and feasibility assessment of studies for ITC

The DeLLphi-301 trial, identified from the company's systematic literature review, provided individual patient data for the MAIC for tarlatamab. As described in section 3.2.1.1, [redacted]

Of these, two patients had not received two prior lines of therapy and one did not receive tarlatamab, which resulted in an available sample of 97 participants for the ITC of patients who had received the target 10 mg dose in both study parts (CS section B.2.9.5). The trial is described in more detail in section 3.2.1.

[redacted] was conducted to [redacted] for the standard of care comparators, as the company's original systematic literature review did not identify any studies of these comparators in a third-line setting (section 2.4) [redacted] In the EAG's opinion, the methodology used to [redacted] lacks transparency. [redacted] a retrospective, UK real-world evidence study the company had conducted – the UK CAS study – was selected to be the external control arm for the unanchored MAIC (CS Appendix D, section D.1.4). The confidential report of the MAIC[2] describes this as [redacted] (pages 7, 14-15, and 17).

The company selected the UK CAS study as the external control for the MAIC because they argue it is the most up-to-date source of evidence and it is expected to be representative of the patients seen and treatments used in UK clinical practice (CS Appendix D, section D.1.4). The supplied MAIC report shows that unanchored MAICs were [redacted] for the standard of care comparator [redacted] We note that these were [redacted] and thus are less likely to be representative of UK practice.

[redacted]

[redacted]In the UK CAS study, [redacted][2] The databases used and the data that were sourced from them are shown in Table 15. [redacted][2] The CS states that patient-level data were *"not readily available"* (CS section B.2.9.5).

Table 15 Data sources used in the UK CAS study

Source	Data used	Data availability
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]		[redacted]
[redacted]		[redacted]
[redacted]	[redacted]	[redacted]

Source: EAG created table, using information sourced from the company's Data on File MAIC report.[2]
[a] In their response to clarification question A22, the company stated that the diagnosis identification period of the UK CAS study was 1st January 2013 to 31st December 2020.
[b] CS Appendix Table 15 states that the time period of data used from the SACT was 1st January 2013 to 31st December 2020. CS Appendix D.1.4 states that data *"between 1st January 2013 and 31st May 2021"* were used. The company clarified in their response to clarification question A22 that the whole UK CAS study period was 1st January 2013 to 31st May 2022. The company stated the reference to May 2021 was an error and that this should read "*2022*". The company stated in clarification response A23 that 31st May 2022 was the latest date of data availability in the UK CAS study at the time of the analysis.

Participants from the CAS database were included in the UK CAS dataset for this study if they had:

[redacted]Table 15 [redacted]

In total, [redacted] patients in England were included in the UK CAS study (CS Table 18). Outcomes were available for participants up to [redacted]

Table 15 [redacted]

A MAIC approach to an ITC essentially adjusts the intervention's treatment effect (in this case tarlatamab), using selected covariates, to reflect the effect that would be observed in the population of patients receiving the intervention who have the characteristics of the patients included in the comparator study (in this case, of standard of care).[1] Therefore a

fundamental assumption of a MAIC is that the population included in the comparator study is more representative of the target patient population who will be treated with the intervention in clinical practice than the intervention's trial's population.[1] A NICE Decision Support Unit (DSU) Technical Support Document (TSD18) suggests that, when conducting a MAIC, the target population for a decision is likely to be appropriately reflected through the use of a UK registry or cohort study.[1] The UK CAS study included adult patients in England treated within the NHS who had had two previous lines of therapy and who had an ECOG performance status of 0 or 1 at initiation of third-line therapy, as per the population of patients expected to be treated with tarlatamab in clinical practice (see section 2.2.3). One of the EAG's experts commented that the characteristics of the patients included in the UK CAS study and the treatments used were a reasonable representation of the patients treated and treatments used in clinical practice. The other two experts did not comment on this. Overall, in our opinion, the UK CAS study provides an appropriate comparator to which to adjust the DeLLphi-301 trial clinical efficacy estimates in an unanchored MAIC as it is likely to be a representative UK study (with the exception of prior PD-L1 inhibitor treatment; please see bullet point three below and section 3.3.3.3).

We do, though, note the following uncertainties about the use of the UK CAS study, the first of which was, as detailed below, addressed through the company's response to one of our clarification questions:

- From the information available in the CS and the MAIC report,[2] it was unclear how the [redacted] patients identified for inclusion in the UK CAS study were selected from the patients available in the [redacted].[2] The company did not provide a flowchart or a description of the number of patients available and numbers excluded from the final sample due to not meeting the study inclusion criteria or other reasons, such as data availability. Thus, there was a lack of transparency about how patients were selected for inclusion in the final sample. The EAG asked the company to provide a flowchart and/or to describe how the final sample was selected in clarification question A24. The company responded with both a flowchart and a description. Based on this information, the selection process appears reasonable. We note that as part of the process, the company applied inclusion criteria to ensure the selected sample aligned with DeLLphi-301 trial sample in terms of having previously received platinum-based chemotherapy (clarification response A24).
- Clarification response A22 states that the diagnosis identification period was 1st January 2013 to 31st December 2020, whereby any patients meeting diagnostic criteria during this time were included in the UK CAS study, while it is reported in the

MAIC report [redacted] were included in the study. There is a lack of clarity about whether the start date of the diagnostic period and therefore the study was [redacted] or 2013. This may be due to an error in reporting but if it is not and the start date was 2013 then this raises [redacted] of the patients included in the UK CAS study had had exposure to a prior PD-L1 inhibitor. The reason for this is unclear, but it might be because the dataset included patients who had been [redacted] and NICE guidance (TA638) on atezolizumab with carboplatin and etoposide for untreated extensive-stage small-cell lung cancer was not published until the 1st July 2020.[38] Clinical expert advice to the EAG is that most patients in practice receive atezolizumab with carboplatin and etoposide as a first-line therapy. In this respect, the UK CAS study cohort does not represent the patients treated in clinical practice. Based on our experts' advice, we understand that it is unknown if this would affect outcomes from the MAIC. Furthermore, the company have not provided subgroup analyses of survival in the CS for the DeLLphi-301 study by prior PD-1 or PD-L1 inhibitor therapy (the only subgroup analyses by prior PD-1 or PD-L1 inhibitor therapy (Yes/No) presented in the CS are of the objective response rate (ORR) outcome; CS section B.2.7).

3.3.3 Clinical heterogeneity assessment

A key assumption of an unanchored MAIC is that all treatment effect modifiers and prognostic factors that are imbalanced between the included studies have been accounted for in the analysis.[1] As is noted in DSU TSD18, this assumption is considered nearly impossible to meet, which means that there will likely be residual bias present in a treatment effect estimate from an unanchored MAIC.[1] When conducting a MAIC, efforts should be undertaken to include relevant prognostic factors, selected *a priori* using clinical expert opinion or from undertaking a thorough review of the topic area, and to match these patient characteristics across the studies of the treatments being compared in the analyses.[1]

3.3.3.1 Company's approach to identifying prognostic factors

Prior to conducting the CS MAIC, the following steps were undertaken to identify prognostic factors that could potentially be adjusted for in the analyses:

- A **targeted literature review** was identified from the systematic literature review that included 39 studies that reported factors prognostic of OS. We asked the company in clarification question A26 if they performed a risk of bias assessment of the review. The EAG found the company's response ambiguous, with the company stating that they did not conduct an additional assessment to that which may or may not have been carried out in the publication. From the targeted literature review, 35 potential prognostic factors were identified. The 35 factors were validated in a **survey of [redacted] clinical experts** and the experts considered the following to be 'very important': ECOG performance status, disease stage, and response to previous treatment (as captured by platinum sensitivity).
- **Analyses were undertaken of the DeLLphi-301 trial data** and this process identified ECOG performance status and brain metastases as significant predictors of OS, which clinical experts confirmed to be important.
- A **meta-regression** was carried out of two systematic literature reviews of second-line and third-line therapy+ studies, respectively,[40] in patients with relapsed or refractory SCLC to identify statistically significant factors that moderated OS, PFS and ORR outcomes. The company provided the references for these systematic literature reviews in response to clarification question A25. The following factors were identified as *"possibly"* (CS section B.2.9.2) predictive: sex, class of treatment, ECOG performance status, treatment line, and extensive stage disease. It was noted that age was not found to be a prognostic factor.

Using a combination of expert opinion about the important prognostic factors and the factors that were found to be statistically significant in the meta-regression, the following were selected for use in the MAIC base case as covariates (CS section B.2.9.2):

- Age at diagnosis
- Sex
- Smoking
- ECOG performance status (0 vs 1) (at initiation of third-line+ treatment, as clarified in the company's response to clarification question A28)
- Presence of brain metastases
- Presence of liver metastases
- Chemotherapy-free interval (≥ 180 days) (this captures response to previous treatment. Patients with this interval between the end of first- and second-line treatment are considered to be platinum-sensitive; clarification response A27)

- Chemotherapy-free interval (≥ 90 and < 180 days) (this captures response to previous treatment. Patients with this interval between the end of first- and second-line treatment are considered to be platinum-sensitive. Those with a chemotherapy-free interval of <90 days are considered to be refractory or resistant to platinum; clarification response A27)
- Extensive-stage disease at diagnosis (as there was a lack of real-world data to obtain stage at initiation of third-line treatment)
- Time from diagnosis to line of therapy

The company notes that age and sex were not universally agreed by the clinicians they consulted to be prognostic factors, but they were included as covariates in the MAIC analyses as they had previously been included in other population-adjusted analyses in SCLC.

We note that smoking is stated in CS section B.2.9.2 to have been selected as a prognostic factor for inclusion in the base case, but it was not actually used as a covariate in the end (CS Table 16). We asked the company the reason for this in clarification question A29. In the company's written response, they stated that following discussions with the EAG it was agreed that an answer was not required (clarification response A29). The company provided their reason during the clarification questions meeting with NICE and the EAG noted it was due to data availability. We therefore agree with the company's approach of not including smoking as a covariate, particularly as one of our experts did not believe that it was a prognostic factor.

3.3.3.2 EAG's critique of the company's approach to selecting prognostic factors

In the EAG's opinion, broad work was undertaken to identify potential prognostic factors on which to match patients in the MAIC. Both clinical expert opinion and reviews of the topic area were used, thus reflecting both the recommended approaches in DSU TSD18.[1] A minor criticism is that there is a lack of transparency about how the targeted literature review and the [redacted] systematic literature reviews used for the meta-regression were selected, leading to some uncertainty about whether or not these were the most relevant data sources available. A further minor criticism is that the clinical expert survey was limited to [redacted] respondents. This is probably sufficient for eliciting opinion, but a larger sample would have been ideal.

All three of our clinical experts considered that the company's selected list of prognostic factors was reasonable. One noted, though, that sex and, as stated above, smoking are not convincing factors (the latter because most patients will have smoked). As is also noted

above, smoking was not included in the company's unanchored MAIC base case. Our experts all agreed that previous response to treatment (i.e. platinum resistance or sensitivity as measured by chemotherapy-free interval) is one of the most important prognostic factors. The other factors considered to be most important by at least one of our three experts were: time since diagnosis, ECOG performance status and maintaining performance status, burden of disease, presence of brain/liver metastases, and progression-free interval.

Two of the three clinical experts we consulted pointed out that prognostic scoring systems are available for SCLC (e.g. the Manchester Scoring Index) and it was noted that paraneoplastic syndromes (e.g. low sodium, neuropathy), which affect a small number of patients, can have a negative prognostic effect. This was not a prognostic factor identified for inclusion in the company's analyses, however, as noted by one of our experts, if this information is not prospectively recorded, then it will not be accessible for use in analyses.

An uncertainty about the company's selected prognostic factors is that age and sex were included in their analyses as covariates, yet there was not a consensus between the clinicians they consulted that these were prognostic and the meta-regression did not find age (when measured as the median or the mean) to be a prognostic factor either. We therefore asked the company to carry out a MAIC sensitivity analysis excluding these covariates to explore the impact of this on the results (clarification question A37), which the company provided in response along with a cost-effectiveness scenario analysis incorporating these results, as we requested (clarification response A37).

To explore the impact of the covariate selection further on the results of the unanchored MAIC analyses, we also asked the company to provide sensitivity analyses of OS and PFS in which only the prognostic factors considered to be 'very important' by the company's clinical experts were controlled for (i.e. ECOG performance status, disease stage and response to previous treatment) (clarification question A38). We also requested the company provided a cost-effectiveness scenario analysis using the results. The company provided the requested analyses (clarification response A38) and the results are presented in section 3.5.

3.3.3.3 Heterogeneity between the DeLLphi-301 trial participants and the CAS Control Cohort patients

We note there were baseline characteristic differences between the DeLLpi-301 trial participants and the CAS Control Cohort on the following prognostic factors controlled for in the analyses: presence of brain or liver metastases (the company state in CS section B.2.9.4 that these characteristics were generally similar, but the EAG disagree), platinum-sensitivity

(chemotherapy-free interval of ≥180 days), extensive disease at 'diagnosis' (measured by a proxy of the presence of extensive disease at initiation of third-line therapy+) and gender (CS Table 18). Clinical expert advice to the EAG is that the difference between the studies in the presence of extensive disease would favour tarlatamab, while the difference between the studies in the presence of liver metastasis would favour standard of care. A clinical expert commented that it is unclear whether the gender differences are important. Additionally, there was a substantial difference in the proportions of participants who had previously received a PD-L1 inhibitor (DeLLphi-301: 72.7%; CAS Control Cohort: [redacted]) but this was not a covariate included in the MAIC analyses. It is unknown what impact this difference might have on the results (please see discussion in section 3.3.2). As noted in section 3.3.2, clinical expert advice to the EAG is that most patients in practice receive atezolizumab with carboplatin and etoposide as a first-line therapy.

It is unclear if there were any other baseline differences between the patients in the study that could potentially affect the outcomes, as baseline characteristics were reported for only a limited number of variables in CS Table 18 that were not selected prognostic factors. We therefore asked the company to confirm if the list of characteristics the DeLLphi-301 and CAS Control Participants were compared on in CS Table 18 was comprehensive and that there were no missing characteristics (clarification question A31). In response to this, the company provided a comparison on some additional characteristics (clarification response A31). This showed that the two samples differed in terms of the proportions of participants who were Asian or White. As discussed in section 3.2.1.2, our experts do not expect that race or ethnicity would impact on treatment response. There were differences in the presence of the following comorbidities too: hypertension (DeLLphi-301: 45.4%, CAS Control Cohort: [redacted]), chronic obstructive pulmonary disease (10.3% versus [redacted]) and diabetes mellitus (20.6% versus [redacted]). Our clinical experts were of the opinion that these differences would likely not impact or have limited impact on the OS and PFS outcomes.

3.3.4 Similarity of treatment effects

As outlined in section 3.3.1, the company included unanchored MAICs of OS and PFS in the CS. Table 16 shows how these outcomes were defined in DeLLphi-301 and UK CAS study. The OS outcome definitions are comparable. In terms of PFS, real-world evidence data was not available for this outcome from the UK CAS study and thus TTD was used as a proxy for PFS for the standard of care comparator. We note that in the literature, in non-small-cell lung cancer (i.e. a different indication to this appraisal), TTD and PFS have been found to have a high but not perfect correlation (r = 0.87).[41] In the absence of PFS data, the EAG consider using TTD as a proxy is acceptable, but it may potentially disadvantage standard of care as

TTD may be shorter than PFS, because patients may discontinue treatment for reasons other than disease progression, such as unacceptable toxicity or to switch to a new therapy prior to progression. To address this uncertainty, we asked our clinical experts how well the median PFS found in the company's MAIC analyses for standard of care, which was [redacted] months (95% CI: [redacted], [redacted]) (see section 3.5), reflects that expected in clinical practice. All three felt it was a reasonable reflection, which provides some confidence that TTD may adequately represent PFS in the present analyses. We also asked the company to provide an unanchored MAIC sensitivity analysis using TTD from DeLLphi-301, instead of PFS, versus TTD from the UK CAS study and a cost-effectiveness scenario analysis using the result (clarification question A39). In response, the company provided information on how the analysis can be performed in the economic model, using the option of TTD instead of PFS for tarlatamab, and provided the result of the requested cost-effectiveness scenario analysis (clarification response A39). This indicated that when TTD was used for tarlatamab, comparable cost-effectiveness results were obtained to the company's updated base case (based on deterministic analyses).

Table 16 Definitions of OS and PFS outcomes in the DeLLphi-301 and UK CAS studies

Outcome	Tarlatamab (DeLLphi-301)	Standard of care (UK CAS study)
OS	Time from index treatment initiation to death from any cause.	
PFS	Time from third-line+ treatment initiation to progression or death (whichever happened first). Assessed by BICR using the RECIST v 1.1 criteria.	TTD used as a proxy for PFS, as PFS not available. TTD was defined as time from the start of third-line therapy to the end of that treatment or death (whichever happened first).

Source: CS Appendix D, section D.1.5
BICR, blinded independent central review; CAS, Cancer Analysis System; OS, overall survival; PFS, progression-free survival; RECIST, response evaluation criteria in solid tumours; TTD, time to treatment discontinuation.

3.3.5 Risk of bias assessment for studies included in the ITC

The company critically appraised the DeLLphi-301 trial using the Downs and Black checklist[42] (CS Appendix D, section D.1.9). Please see section 3.2.2 for a discussion of the results of this and the EAG's critical appraisal of the study. The company did not, however, carry out a risk of bias assessment of the UK CAS study. We therefore carried out an assessment using the criteria recommended by NICE for critically appraising non-randomised and non-controlled studies.[28] Our full appraisal of the study is available in Appendix 4. As this shows, in our opinion, there is an unclear risk of selection bias because

it is unclear whether the start of the diagnosis identification period of the study for patient inclusion was in [redacted] or 2013 due to inconsistency in reporting of these dates. Additionally, the CAS Control Cohort is not fully representative of the patients treated in practice, as [redacted] had received a prior PD-L1 inhibitor and the impact of this on the results is unknown (see section 3.3.2). We also have concerns about potential detection bias due to TTD being used as a proxy for PFS, which could potentially lead to an underestimation of PFS for standard of care. However, as discussed in section 3.3.4, our experts were satisfied that the median PFS found in the MAIC for standard of care was reflective of that seen in clinical practice, so this may not be an issue.

EAG comment on the studies included in the ITC

The company's selection of the UK CAS study to represent the standard of care comparator in the unanchored MAIC is appropriate, as it is a large UK study that is likely generally representative of the characteristics of the patients treated in clinical practice and the treatments they receive at third-line therapy+, except in terms of receipt of a prior PD-L1 inhibitor. The potential impact the latter could have on the results of the MAIC is unknown.

There is an unclear risk of selection bias in how participants were selected for inclusion in the UK CAS study, as the dates during which patients meeting the diagnostic criteria for eligibility are inconsistently reported. The EAG suspects this is likely due to an error in reporting. Additionally, arguably, a more recent data cut of the UK CAS study than 31st May 2022 could have been used, but the company indicate that this was the end of data availability at the time of the analyses (there is no indication that the analyses were specifically carried out for the NICE submission).

Our experts confirmed that the company's selection of covariates to use in their unanchored MAIC base case analyses was generally reasonable. The use of TTD as a proxy for PFS for standard of care also appears reasonable in the absence of PFS data.

3.4 Critique of the indirect treatment comparison

As described in sections 3.3.1 and 3.3.4, the company conducted unanchored MAICs comparing tarlatamab to standard of care on the outcomes of OS and PFS (using TTD as a

proxy for PFS for standard of care). The single-arm trial design of the DeLLphi-301 study necessitated an ITC.

3.4.1 Data inputs to the ITC

As noted in section 3.3.2, the company used the DeLLphi-301 trial to represent tarlatamab in the unanchored MAIC and the UK CAS study as the external control arm to represent standard of care. Individual patient data were used from 97 participants in the DeLLphi-301 trial, but CS section B.2.9.5 highlights that that because individual patient data were not *"readily"* (CS section B.2.9.5) available for the CAS Control cohort, pseudo-patient level data were created using a published method[43] and used for the selected [redacted] patients. This is a limitation, as the DeLLphi-301 trial participants were matched to the aggregate values from the CAS control cohort rather than the variance, which creates some uncertainty around the reliability of the data.

As noted in section 3.2.1.1, participants in DeLLphi-301 could receive tarlatamab past disease progression, [redacted].[17] CS section B.2.9.3 states that in the unanchored MAIC analyses, participants from DeLLphi-301 who received tarlatamab after progression were censored when disease progression occurred in the OS analyses. The EAG agrees that this is appropriate. CS section B.2.9.3 also states that these participants were censored from the PFS analyses, but information in CS section B.3.3.1 and clarification response A39 suggest that they were not (there may be an error in reporting in CS section B.2.9.3). The EAG believes it would be appropriate to not censor these participants from the PFS analyses and to include their PFS data points. Participants receiving tarlatamab post-progression were also censored from the EAG-requested sensitivity analysis provided in response to clarification question A39, in which TTD from DeLLphi-301 was used instead of PFS and compared with TTD from the UK CAS study (clarification response A39). The censoring rules are not clearly outlined in the clarification response. It is unclear if the same approach, as set out in CS Table 38 for TTD and in clarification response A11 for the censored TTD analyses from the DeLLphi-301 trial, was used. If it was, then we agree that the company's approach is appropriate. We note that there is a lack of clarity in the CS about how many participants received tarlatamab post-progression (see Table 7 in section 3.2.4).

3.4.2 Statistical methods for the ITC

The individual patient data from DeLLphi-301 and the pseudo-individual patient data from the CAS Control cohort were used in the unanchored MAIC matching process to adjust

imbalances in the covariates and to derive statistical weights (CS section B.2.9.5). A form of propensity score weighting was used to create the weights (CS Appendix D.1.6). The weights were used to adjust the DeLLphi-301 trial outcomes so that the tarlatamab patient population aligned to that of the CAS Control cohort (CS section B.2.9.5). Hazard ratios were generated from weighted Cox proportional hazards models, with treatment used as a covariate. Standard errors were also generated and effective sample sizes provided. The effective sample size is the number of non-weighted participants that would be needed to obtain the same estimate as in the weighted sample (CS Appendix D.1.6). When the effective sample size is considerably reduced compared to the original sample size, estimates may be unreliable.[1] The company provided the results of the matching process showing the standardised mean differences on each of the prognostic factors before and after matching (CS section B.2.9.6 Figure 9). However, they did not report the raw data pre- and post-match, and so we requested these in clarification question A33, which the company provided in response (clarification response A33). We also asked for the standardised mean difference results before and after matching for any unmatched baseline characteristic variables, which the company also provided (clarification response A34).

The company carried out two MAIC scenario analyses in which they removed selected prognostic factors and reported results of these in the CS:

- Scenario 1: chemotherapy-free interval data were missing for 30% of participants in the DeLLphi-301 trial and an assumption had been accordingly used in the base case that these participants had a chemotherapy-free interval of ≥ 180 days (CS section B.2.9.2). To explore whether this was a potential source of bias, the company excluded chemotherapy-free intervals as prognostic factors from the MAIC.
- Scenario 2: Stage of disease at diagnosis had to be used as a covariate in the base case in the absence of the preferred prognostic factor of presence of extensive disease at initiation of treatment. Therefore, this factor was excluded in a scenario analysis.

As detailed above, the EAG also requested three sensitivity analyses in clarification questions A37, A38 and A39, to explore the robustness of the results to the removal of selected covariates and to explore the impact of using TTD from DeLLphi-301 instead of PFS in the comparison against TTD for standard of care:

- EAG clarification question A37 requested sensitivity analysis: analyses of OS and PFS omitting sex and age as prognostic factors (requested for the reasons outlined in section3.3.3).

- EAG clarification question A38 requested sensitivity analysis: analyses of OS and PFS that only include the three prognostic factors considered to be 'very important' by the company's clinical experts (i.e. ECOG performance status, disease stage, and response to previous treatment). This analysis was requested as the company's MAIC base case effective sample size was markedly reduced and we wished to explore the robustness of the results when fewer prognostic factors were included in the analyses.
- EAG clarification question A39 requested sensitivity analysis: an analysis using TTD from DeLLphi-301, instead of PFS, versus TTD from the UK CAS study (requested to explore the impact on the results when a more comparable outcome from the DeLLphi-301 trial was used to that from the UK CAS study to examine the potential robustness of using TTD as a proxy for PFS for standard of care). The company did not clearly state the prognostic factors used in this analysis but stated that the MAIC weights are the same for the TTD and PFS outcomes, which suggests the company used the base case matching factors.

We also asked the company to provide cost-effectiveness scenario analyses using the results of each of these sensitivity analyses. We summarise the prognostic factors included in the company's unanchored MAIC base case, two scenario analyses and the EAG's two requested sensitivity analyses in which we requested alterations to the included covariates in Table 17.

Table 17 Prognostic factors adjusted for in the unanchored MAIC

Prognostic variable	Company base case ESS = [redacted]	Company scenario 1 ESS = [redacted]	Company scenario 2 ESS = [redacted]	EAG clarification A37 ESS = [redacted]	EAG clarification A38 ESS = [redacted]
Age at diagnosis (prior to first-line treatment)	✓	✓	✓		
Sex (male vs female)	✓	✓	✓		
ECOG PS (0 vs 1) [a]	✓	✓	✓	✓	✓
Brain metastases [a]	✓	✓	✓	✓	
Liver metastases [a]	✓	✓	✓	✓	

Chemotherapy-free interval (≥ 180 days) [b]	✓		✓	✓	✓
Chemotherapy-free interval (≥ 90 and < 180 days) [b]	✓		✓	✓	✓
Extensive-stage disease at diagnosis [c]	✓	✓		✓	✓
Time from diagnosis to line of therapy	✓	✓	✓	✓	

Source: Partly reproduced from CS Table 16, with additional information sourced from clarification questions A37 and A38
CS, Company Submission; EAG, External Assessment Group; ECOG, Eastern Cooperative Oncology Group; ESS, effective sample size; PS, performance status; TNM, tumour, node, metastases.
[a] At initiation of index line of treatment.
[b] At completion of first-line treatment.
[c] Only TNM stage was reported in the DeLLphi-301 trial; stage IV was assumed to be extensive-stage.

The company provided the results of the requested EAG sensitivity analyses and cost-effectiveness scenarios in their clarification responses A37 and A38. In response to clarification question A39, the company provided information on how the analysis can be performed in the economic model, using the option for TTD instead of PFS for tarlatamab, and provided just the result of a requested cost-effectiveness scenario analysis (clarification response A39) rather than the MAIC results for TTD from the DeLLphi-301 trial versus TTD from the UK CAS study.

We note that the company did not provide the distribution of weights for the company's unanchored MAIC scenario analyses 1 and 2, so we requested these in clarification question A32, which the company accordingly provided (clarification response A32). The company also did not provide the median OS and PFS for either the tarlatamab or standard of care group before and after weighting for the two company scenario analyses, so we requested these results in clarification question A35, and the company responded with the requested information (clarification response A35).

The EAG considers that the unanchored MAIC methodology has been appropriately chosen, with appropriate prognostic factors and treatment effect modifiers adjusted for in the base case analyses. As stated in section 3.3.3.2, the EAG's clinical experts confirmed that the

selection of prognostic factors and treatment effect modifiers was reasonable. The MAIC methodology appears to have been correctly implemented, although no individual patient data was available to the EAG to confirm this. The matching process was successful, and the prognostic variables appeared to be well-balanced post-match. The distribution of weights also appears reasonable, but there was a high reduction in the base case effective sample size ([redacted]), which resulted in an effective sample size of just [redacted] patients. The company's two scenario analyses provided similar OS and PFS results to the base case results (see section 3.5). The EAG's two requested sensitivity analyses, requesting the removal of selected covariates from the analyses, resulted in smaller reductions in the effective sample size ([redacted] and [redacted], respectively). A limitation of these analyses, however, is that stronger assumptions were made about what factors were prognostic and there were differences between the samples in brain and liver metastases (which would not have been controlled for in the sensitivity analysis conducted in response to clarification question A38) and gender between the two studies (see section 3.3.3.3) and we do not know what the differences post-match for these unmatched variables were. Both the EAG requested sensitivity analyses produced results that were [redacted] to tarlatamab for PFS or [redacted] to tarlatamab for OS than the base case (see section 3.5).

3.4.3 Summary of EAG critique of the ITC

The unanchored MAIC methodology has been appropriately chosen and implemented. We are satisfied with the use of the MAIC base case OS and PFS results in the cost-effectiveness analyses. However, a limitation of the base case analyses of OS and PFS is that they resulted in an effective sample size of just [redacted] participants. When the effective sample size is greatly reduced, MAIC estimates may become unreliable.[1] In addition, a limitation of an unanchored MAIC approach is that systematic error is likely to still be present due to the likelihood that there are unaccounted for covariates in such analyses.[1]

3.5 Results from the indirect comparison

3.5.1 OS

OS with tarlatamab treatment was significantly longer in comparison to the available treatment options, represented by the UK CAS patients treated with systemic anti-cancer therapies, in both the unadjusted and MAIC-adjusted comparisons (Table 18). In the unadjusted analysis there was a [redacted]% lower mortality risk with tarlatamab and in the MAIC-adjusted analysis the mortality risk with tarlatamab was [redacted]% lower. Figure 4 (which is a reproduction of CS Figure 11) shows the Kaplan-Meier plots for both the unadjusted (unweighted) and MAIC-adjusted (weighted) OS data. As stated in section 3.2.5.2 median

OS with tarlatamab (unadjusted) was 14.3 months (95% CI: 10.8, NE). In the MAIC-adjusted analysis it was [redacted] months (95% CI: [redacted]), as reported in CS section B.2.9.6. We note, however, that median OS after MAIC weighting and with censoring of patients who received at least one dose of tarlatamab post-progression is reported to be [redacted] months in CS Table 36. In comparison, median OS in the UK CAS study was [redacted] months. Based on the information presented in CS Table 36, the EAG suggests that in CS section B.2.9.6 the company has reported the median OS for tarlatamab after weighting in the MAIC, but without censoring for patients who received tarlatamab post-progression.

Table 18 OS for tarlatamab versus available treatment options (unadjusted and MAIC adjusted results)

	ESS	HR (95% CI)	SE	p-value
Unadjusted		0.278 (0.194, 0.399)	[redacted]	<0.0001
MAIC-adjusted	[redacted]	0.367 (0.202, 0.667)	[redacted]	0.001

Source: Table reproduced from CS Table 19
CI, confidence interval; ESS, effective sample size; HR, hazard ratio; MAIC, matching-adjusted indirect comparison; OS, overall survival; SE, standard error.

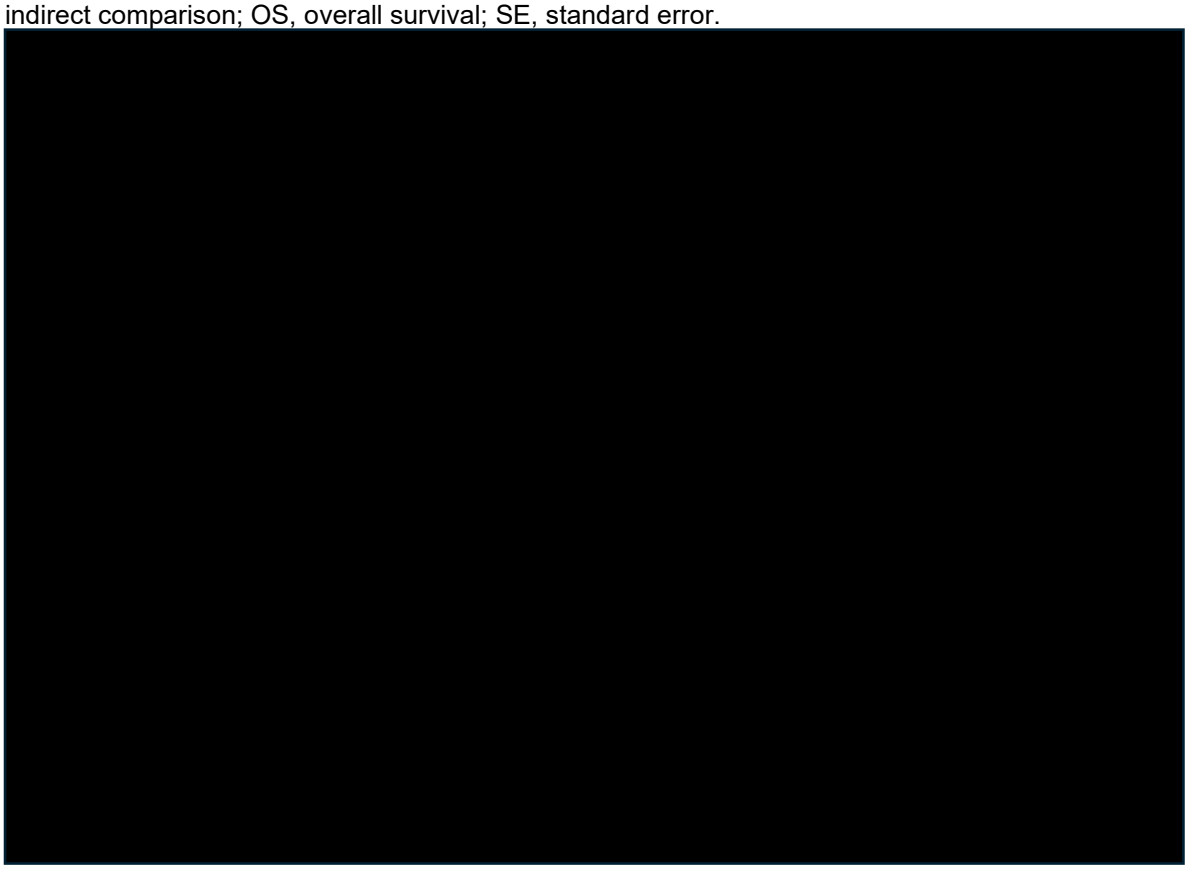

Figure 4 Kaplan-Meier plot of OS for tarlatamab (unadjusted and MAIC adjusted) versus available treatment options

Source: Reproduction of CS Figure 11
MAIC, matching-adjusted indirect comparison; OS, overall survival

3.5.1.1 ITC scenario analyses for OS

The company's two scenarios and the EAG requested scenarios have been discussed in section 3.4.2 and shown in Table 17. The company scenario results were reported in CS Appendix D.1.7 and the results for two of the EAG requested scenarios were provided in response to clarification questions A37 and A38. As can be observed from the hazard ratios reported in Table 19, under company scenario 1 the MAIC-adjusted hazard ratio [redacted] in comparison to the base case MAIC-adjusted hazard ratio thus [redacted] the relative efficacy of tarlatamab from [redacted]% in the base case to [redacted]%. Under company scenario two the MAIC-adjusted hazard ratio was [redacted] the base case MAIC-adjusted value. The results from both company scenario analyses still demonstrate significantly improved survival with tarlatamab in comparison to the available treatment options. For the two scenario analyses that the EAG requested, both led to [redacted] in the MAIC-adjusted hazard ratio in comparison to the base case (i.e. there was [redacted] in the relative efficacy of tarlatamab).

Table 19 OS for tarlatamab versus available treatment options in the company and EAG scenarios compared to the base case (MAIC adjusted results)

		ESS	HR (95% CI)	SE	p-value
Company base case[a]	Unadjusted	-	0.278 (0.194, 0.399)	[redacted]	<0.0001
	MAIC-adjusted	[redacted]	0.367 (0.202, 0.667)	[redacted]	0.001
Company scenario 1[b]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	<0.001
Company scenario 2[c]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	<0.001
EAG scenario clarification A37[d]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario clarification A38[e]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	[redacted]

Source: EAG compiled table combining CS Table 19, CS Appendix D.1.7 Table 11 and Table 12, Clarification response to A37 Table 10 and Clarification response to A38 Table 14.
CI, confidence interval; EAG, External Assessment Group; ESS, effective sample size; HR, hazard ratio; MAIC, matching-adjusted indirect comparison; OS, overall survival; SE, standard error.
[a] The prognostic factors adjusted for in the Company base case unanchored MAIC were: Age at diagnosis (prior to first-line treatment); Sex (male vs female); ECOG performance status (0 vs 1) at initiation of index line of treatment; Brain metastases at initiation of index line of treatment; Liver metastases at initiation of index line of treatment; Chemotherapy-free interval (≥ 180 days) at completion of first-line treatment; Chemotherapy-free interval (≥ 90 and < 180 days) at completion of first-line treatment; Extensive-stage disease at diagnosis (Only TNM stage was reported in the DeLLphi-301 trial; stage IV was assumed to be extensive-stage); Time from diagnosis to line of therapy
[b] Chemotherapy-free interval (≥ 180 days) and Chemotherapy-free interval (≥ 90 and < 180 days) were not adjusted for in this scenario.
[c] Extensive-stage disease at diagnosis was not adjusted for in this scenario.

[d] Age at diagnosis (prior to first-line treatment) and sex (male vs female) were not adjusted for in this scenario
[e] Four prognostic factors were adjusted for in this scenario: ECOG performance status (0 vs 1) at initiation of index line of treatment; Chemotherapy-free interval (≥ 180 days) at completion of first-line treatment; Chemotherapy-free interval (≥ 90 and < 180 days) at completion of first-line treatment; Extensive-stage disease at diagnosis

3.5.2 PFS

PFS was not available in the CAS Control cohort so TTD was used as a proxy for PFS.

PFS with tarlatamab treatment was significantly better than in the comparator therapy group in both the unadjusted and MAIC-adjusted comparisons (Table 20). In the unadjusted analysis the risk of progression with tarlatamab was [redacted]% lower than the available treatment options, in the MAIC-adjusted analysis it was [redacted]% lower. Figure 5 shows the Kaplan-Meier plots for the unadjusted (unweighted) and MAIC-adjusted (weighted) PFS data. Median PFS with tarlatamab (unadjusted) was [redacted] months (95% CI: [redacted]) and in the MAIC-adjusted analysis it was [redacted] months (95% CI: [redacted]). In the comparator CAS cohort median PFS (for which TTD or death was used as a proxy) was [redacted] months (95% CI: [redacted]).

Table 20 PFS for tarlatamab versus available treatment options (unadjusted and MAIC adjusted results)

	ESS	HR (95% CI)	SE	p-value
Unadjusted		0.206 (0.150, 0.282)	[redacted]	<0.0001
MAIC-adjusted	[redacted]	0.184 (0.100, 0.340)	[redacted]	<0.0001

Source: Reproduction of CS Table 20
CI, confidence interval; ESS, effective sample size; HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; SE, standard error.

Figure 5 Kaplan-Meier plot of PFS for tarlatamab (unadjusted and MAIC adjusted) versus available treatment options

Source: Reproduction of CS Figure 12
MAIC, matching-adjusted indirect comparison; PFS, progression-free survival.

3.5.2.1 ITC scenario analyses for PFS

Table 21 shows the MAIC-adjusted results for the two company scenarios and the two EAG requested scenarios in comparison to the company MAIC base case for PFS of tarlatamab compared with available treatment options. Across all four scenarios the MAIC-adjusted hazard ratio for PFS was [redacted] giving relative efficacies for tarlatamab that lay between [redacted]% and [redacted]% in comparison to the company base case MAIC-adjusted value of [redacted]%. The results from both company scenario analyses [redacted] indicate significantly improved PFS with tarlatamab in comparison to the available treatment options.

Table 21 PFS for tarlatamab versus available treatment options in the company and EAG scenarios compared to the base case (MAIC adjusted results)

		ESS	HR (95% CI)	SE	p-value
Company base case[a]	Unadjusted		0.206 (0.150, 0.282)	[redacted]	<0.0001
	MAIC-adjusted	[redacted]	0.184 (0.100, 0.340)	[redacted]	<0.0001
Company scenario 1[b]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	<0.001

		ESS	HR (95% CI)	SE	p-value
Company scenario 2[c]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	<0.001
EAG scenario clarification A37[d]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	[redacted]
EAG scenario clarification A38[e]	MAIC-adjusted	[redacted]	[redacted]	[redacted]	[redacted]

Source: EAG compiled table combining CS Table 20, CS Appendix D.1.7 Table 13 and Table 14, Clarification response to A37 Table 10 and Clarification response to A38 Table 14.
CI, confidence interval; EAG, External Assessment Group; ESS, effective sample size; HR, hazard ratio; MAIC, matching-adjusted indirect comparison; PFS, progression-free survival; SE, standard error.
[a] The prognostic factors adjusted for in the Company base case unanchored MAIC were: Age at diagnosis (prior to first-line treatment); Sex (male vs female); ECOG performance status (0 vs 1) at initiation of index line of treatment; Brain metastases at initiation of index line of treatment; Liver metastases at initiation of index line of treatment; Chemotherapy-free interval (≥ 180 days) at completion of first-line treatment; Chemotherapy-free interval (≥ 90 and < 180 days) at completion of first-line treatment; Extensive-stage disease at diagnosis (Only TNM stage was reported in the DeLLphi-301 trial; stage IV was assumed to be extensive-stage); Time from diagnosis to line of therapy
[b] Chemotherapy-free interval (≥ 180 days) and Chemotherapy-free interval (≥ 90 and < 180 days) were not adjusted for in this scenario.
[c] Extensive-stage disease at diagnosis was not adjusted for in this scenario.
[d] Age at diagnosis (prior to first-line treatment) and sex (male vs female) were not adjusted for in this scenario.
[e] Four prognostic factors were adjusted for in this scenario: ECOG performance status (0 vs 1) at initiation of index line of treatment; Chemotherapy-free interval (≥ 180 days) at completion of first-line treatment; Chemotherapy-free interval (≥ 90 and < 180 days) at completion of first-line treatment; Extensive-stage disease at diagnosis.

3.5.3 TTD

The EAG asked the company to conduct a MAIC sensitivity analysis using TTD from the DeLLphi-301 trial (instead of PFS) versus TTD from the UK CAS study, to provide the results of this analysis and then to conduct a cost-effectiveness scenario using these data (clarification question A39). In response, the company stated that the MAIC weights are the same regardless of whether the outcome is TTD or PFS. The company provided the requested cost-effectiveness scenario analysis, and we summarise the results in section 5.2.2.

3.6 Conclusions on the clinical effectiveness evidence

The company's decision problem adequately reflects the NICE scope but focuses on a narrower population. The NICE scope specified a population of patients with advanced SCLC who had had disease progression on or after prior therapy, while the company has focused on patients who have advanced SCLC and who have received platinum-based chemotherapy and at least one other treatment, reflecting the company's proposed

positioning of tarlatamab as a third-line therapy+. The EAG's clinical experts agreed with the company's proposed positioning of tarlatamab and their decision problem population. Our experts confirmed that the company's selected comparators, carboplatin + etoposide, CAV and topotecan, reflect those used in clinical practice at third-line therapy+. The EAG agree with the use of these as a single basket standard of care comparator. The company have not included best supportive care as a comparator, even though this was specified in the NICE scope. One of the EAG's three clinical experts felt it was reasonable to not include best supportive care as a comparator, with the other two not commenting on this. Obtaining further clinical expert opinion about this during the appraisal may be beneficial.

The company identified and included one single-arm trial of tarlatamab in the CS – the DeLLphi-301 trial. The included participants are generally representative of the patients seen in clinical practice and those expected to be treated with tarlatamab (i.e. third-line therapy+ patients with an ECOG performance status of 0 or 1). In an unanchored MAIC analysis, the company compared tarlatamab, using data from DeLLphi-301, to standard of care using a generally representative UK real-world evidence study – the UK CAS study – as the source of data for standard of care. The EAG was satisfied with the selection of these studies as the data sources. The EAG were also satisfied with the covariates selected for inclusion in the analysis.

The unanchored MAIC base case, adjusted-analyses resulted in an OS hazard ratio of 0.367 (95% CIs: 0.202, 0.667), favouring tarlatamab. This represented a [redacted] lower mortality risk for patients treated with tarlatamab versus standard of care. Median OS for tarlatamab in the adjusted-analyses was [redacted] months (95% CIs: [redacted]). Median OS for standard of care was [redacted] months (95% CIs not reported). Additionally, the MAIC base case adjusted-analyses resulted in a PFS hazard ratio of 0.184 (95% CIs: 0.100, 0.340), favouring tarlatamab. Median PFS for tarlatamab in the adjusted-analyses was [redacted] months (95% CIs: [redacted]). Median PFS for standard of care population, using TTD as a proxy for PFS, was [redacted] months (95% CIs: [redacted]). However, in the base case analyses the effective sample size was reduced to [redacted] participants. The OS and PFS estimates from the MAIC were used in the company's economic model. Neurological events and CRS were the most common adverse events of special interest experienced by the participants receiving the 10mg target dose of tarlatamab in Parts 1 and 2 of the DeLLphi-301 trial (experienced by [redacted]% and 49.5% of these 99 participants, respectively).

The EAG have identified the following concerns and uncertainties associated with the OS and PFS estimates (both of which inform the economic model) in the CS:

- No RCT, head-to-head trials of tarlatamab versus the standard of care comparators, carboplatin + etoposide, CAV and topotecan, were identified, so comparative efficacy data are only available from an indirect treatment comparison (i.e. the unanchored MAIC), which is inferior to direct comparative data. Residual systematic error is likely in an unanchored MAIC due to unobserved prognostic variables and effect modifiers.[1]
- Interim data from the tarlatamab DeLLphi-301 trial informed the OS and PFS estimates presented in the CS both from the trial and from the unanchored MAIC. For OS, 57.6% of participants were still alive at the time of analysis and for PFS 25.3% remained on the study without disease progression or death. The results could potentially change when the trial is complete or updated analyses are undertaken.
- There was potential for blinding of the DeLLphi-301 trial independent central review body who assessed disease progression to have been undermined if they were [redacted] which might affect the accuracy of the assessment of the PFS outcome.
- The UK CAS study participants were not representative of the patients currently seen in practice in terms of prior PD-L1 inhibitor treatment. The potential impact of this on the CS MAIC analyses is unknown.
- The unanchored MAIC base case effective sample size was just [redacted] participants, reduced from 97 participants. A greatly reduced effective sample size may make MAIC estimates unreliable.[1]

An additional likely minor point is that there is an unclear risk of selection bias in the patient selection for the UK CAS study due to inconsistency in the reporting of the dates of the diagnosis identification period (the EAG suspects this is an error).

4 COST EFFECTIVENESS

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

The company conducted a targeted literature search for economic models for interventions used for relapsed SCLC. The inclusion and exclusion criteria are shown in CS Table 28. The searches were conducted in Medline and Embase on April 19 2023 and the search strategy is outlined in CS Appendix G. Six studies were identified and are shown in CS Table 29. One study was from the UK comparing topotecan vs best supportive care and is the Evidence Assessment Group report and economic model for the NICE appraisal of topotecan (TA184).[44] None of the other studies were for tarlatamab.

A search was also undertaken for submissions for SCLC to Health Technology Assessment (HTA) agencies, including NICE, Canada's Drug and Health Technology Agency (CADTH), Pharmaceutical Benefit Advisory Committee (PBAC) and the Institute for Clinical and Economic Review. For relapsed SCLC, one submission was identified for NICE (TA184) for topotecan,[44] one for CADTH (lurbinectedin)[45] and one for PBAC (topotecan).[46] Details of the studies are shown in CS Tables 30-32.

EAG comment

We consider the cost-effectiveness search strategy and review to be reasonable, however the searches are a year out of date so it is unclear if any studies have been missed. The company responded to clarification question B8 that they were not aware of any further studies published in the last year.

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

4.2.1 NICE reference case checklist

The EAG assessment of the company's economic analysis in relation to the NICE reference case is shown in Table 22.

Table 22 NICE reference case checklist

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Yes (patient only) Carer outcomes are not included.

Element of health technology assessment	Reference case	EAG comment on company's submission
Perspective on costs	NHS and PSS	Yes, NHS and PSS costs.
Type of economic evaluation	Cost–utility analysis with fully incremental analysis	Yes
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	Yes, effectively lifetime (10 years from initial age of [redacted] in base case).
Synthesis of evidence on health effects	Based on systematic review	Yes.
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.	Yes, EQ-5D data used.
Source of data for measurement of health-related quality of life	Reported directly by patients and/or carers	Yes, from DeLLphi-301 trial.
Source of preference data for valuation of changes in health-related quality of life	Representative sample of the UK population	Yes. Utilities mapped from EQ-5D-5L to UK 3L values using the Hernández-Alava algorithm.
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	Yes.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Yes.
Discounting	The same annual rate for both costs and health effects (currently 3.5%)	Yes.

Source: Table produced by EAG
EAG, External Assessment Group; EQ-5D, European Quality of Life Working Group Health Status Measure 5 Dimensions; NHS, National Health Service; PSS, personal social services; QALY, quality-adjusted life year.

4.2.2 Model structure

4.2.2.1 Overview of the model structure

The company's model structure is described in CS section B.3.2.2. A partitioned survival model was developed in Microsoft Excel with a time horizon of 10 years and a cycle length of one week. The model structure comprises three health states: progression-free, progressed disease, and death. The structure is illustrated in CS Figure 14 (reproduced in Figure 6 below).

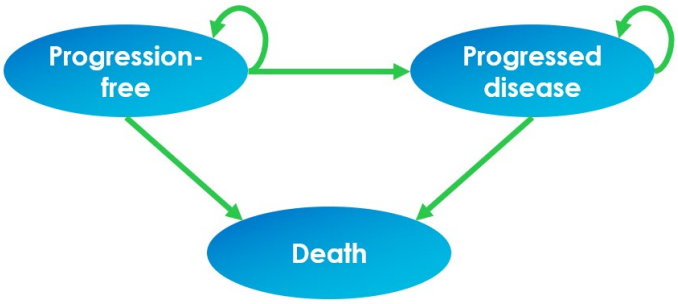

Figure 6 Company economic model structure

Source: Reproduced from CS Figure 14

Patients enter the model in the progression-free health state and are able to transition to the progressed disease or death health states. Patients in the progressed disease health state are only able to remain in the progressed disease state or transition to the death state. The proportion of patients in the progression-free health state is estimated directly from the PFS curves modelled for each comparator, whilst the proportion of patients in the death state at a given time point is calculated using the inverse probability of the OS curve at that time point. The proportion of patients in neither the progression-free nor death health states make up the progressed disease health state. A proportion of patients in the progressed disease state are provided with subsequent treatment; see section 4.2.9.3.3. Costs and QALYs were estimated using the proportion of patients in each of the health states over time. The model ensures that there is consistency between PFS, TTD and OS, such that PFS and TTD remain below OS.

EAG comment on model structure

The company clearly justify their use for a partitioned-survival model using OS and PFS data from the DeLLphi-301 study with parametric survival curves. The EAG

notes that partitioned-survival models were used in previous appraisals for SCLC (TA184 for Topotecan[44] and TA638 for atezolizumab with carboplatin and etoposide[20]). The EAG have no concerns with the model structure used.

4.2.3 Population

The modelled population is described in CS section B.3.2.1. It comprises adult patients with resistant or relapsed SCLC after two or more prior lines of treatment, at least one of which is a platinum-based chemotherapy. The baseline mean age of patients in the model is [redacted] years, and [redacted] of the patients are female; these figures are obtained from the CAS historical control study. The population reflects those patients who are fit enough for treatment with an ECOG performance status of 0 or 1. See section 3.2.1.2 for the EAG's critique of the representativeness of the patient characteristics from the DeLLphi-301 trial of the patients treated in clinical practice.

EAG comment on model population

The population modelled differs from the original NICE scope, as it comprises people who have received two prior treatments, one of which must have been a platinum-based chemotherapy. The EAG consider the focus on this population to be acceptable as it appears to match the clinical evidence provided, and experts we consulted were also in agreement with the population chosen.

4.2.4 Interventions and comparators

The economic model compares the incremental cost-effectiveness of tarlatamab to a basket of treatments as a comparator. Tarlatamab is administered as a 60-minute intravenous (IV) infusion with a step-up dose of 1mg on Cycle 1 Day 1, followed by a 10mg dose on Cycle 1, Day 8, and Day 15, and every two weeks thereafter (Q2W) in a 28-day treatment cycle. Hospitalisation is required for 24 hours post-infusion on Cycle 1 Day 1 and Cycle 1 Day 8. The EAG notes that this differs to the draft SmPC recommendation, which states that [redacted]. In the model, it is assumed that patients stop tarlatamab treatment once their disease has progressed.

The comparator treatment is a basket of treatments comprising the following:

- Cyclophosphamide, doxorubicin and vincristine (CAV) (38%)
- Platinum and etoposide chemotherapy (20%)

- Topotecan (42%)

These treatments and proportions are based upon those in the CAS historical control study (see section 3.3.2 of this report for more information on the CAS study). Seven percent of patients in the CAS study had treatments different from those above and so this proportion was re-allocated among the three treatments (clarification response question B11). CS Table 34 reports the dosing and administration schedule of the treatments in the comparator arm.

EAG comment on intervention and comparators

Clinical experts consulted for this appraisal by the EAG agree that the basket of treatments suggested by the company is an appropriate comparator. The percentage split of treatments is in line with that seen in UK clinical practice, as agreed by our clinical experts. Two of the three experts were also in agreement that tarlatamab administration would cease following disease progression, [redacted]; see section 3.2.1.1 for further information.[17] The EAG do not suggest any changes to the economic base case regarding the intervention or comparators. However, as is stated in section 3.6, obtaining further clinical expert opinion about the relevance of best supportive care as a comparator for tarlatamab may be beneficial (we raise this as an 'other key issue' in section 1.6).

4.2.5 Perspective, time horizon and discounting

The model follows the NHS reference case with respect to the perspective for costing (NHS and Personal Social Services), the time horizon (10 years, effectively lifetime) and discounting (3.5% for costs and health effects).

4.2.6 Treatment effectiveness and extrapolation

The economic model uses parametric curves fit to the observed data for OS, PFS and TTD. The observed survival data are from the DeLLphi-301 trial for tarlatamab and the UK CAS study for standard of care. More details about these studies are given in section 3.

In order to compare tarlatamab with standard of care (SOC), the company conducts a MAIC analysis by applying propensity score weighting to the DeLLphi-301 trial data to balance characteristics of the trial population with those from the UK CAS study. The EAG's description and critique of the MAIC is given in sections 3.3 and 3.4. The Kaplan-Meier curves of the OS and PFS for tarlatamab and standard of care before MAIC weighting are shown in CS Figures 16 and 17 respectively.

In the DeLLPhi-301 trial, [redacted] of 97 patients were allowed to continue treatment with tarlatamab beyond radiologic disease progression. However, the intended marketing authorisation for tarlatamab is for treatment to stop at disease progression. To adjust for this, patients who received tarlatamab post-progression were censored at the time of progression for OS and TTD. The median OS and PFS for tarlatamab before and after adjustment and standard of care are shown in Table 23 (CS Table 36). The Kaplan-Meier curves of OS and PFS for tarlatamab and standard of care after MAIC weighting with adjustment for post-tarlatamab use are shown in CS Figure 18. Clinical advice to the EAG confirmed that the survival estimates for the standard of care arm were reasonable for third-line SCLC patients.

Table 23 Median OS, PFS and OS of tarlatamab and standard of care

Outcomes (months)	Tarlatamab			SOC
	Before weighting	After weighting (without adjustment)	After weighting (with adjustment)	
OS	[redacted]	[redacted]	[redacted]	[redacted]
PFS	[redacted]	[redacted]	[redacted]	[redacted]
TTD	[redacted]	[redacted]	[redacted]	[redacted]

Source: CS Table 36
OS, overall survival; PFS, progression-free survival; SOC, standard of care; TTD, time to treatment discontinuation

The survival data described in the following sections for tarlatamab refers to the data after MAIC weighting and adjustment for post-progression tarlatamab treatment.

4.2.6.1 Overall survival

The company assess whether the proportional hazard assumption holds for OS. Based on the assessment, the proportional hazard assumption did not hold and therefore standard parametric curves were fit to the OS curves for tarlatamab and standard of care separately. The EAG agrees with the company's judgement that the proportional hazard assumption is not supported.

The fitted parametric curves (exponential, Weibull, log-logistic, lognormal, Gompertz, generalised gamma and gamma) compared to the tarlatamab OS trial data are shown in CS Figure 19. The goodness of fit was calculated for Akaike information criterion (AIC) and Bayesian Information criterion (BIC) and are shown in CS Table 39. The extrapolated 1- and

2-year OS outcomes for each of the parametric functions are shown in CS Table 40. Based on the AIC/BIC data, the exponential curve is chosen (lowest AIC/BIC values).

The fitted parametric curves compared to the standard of care OS data are shown in CS Figure 20. The goodness of fit using AIC/BIC was calculated and are shown in CS Table 41. The extrapolated 1- and 2-year OS outcomes for each of the parametric functions are shown in CS Table 42. The CS states that the exponential is chosen for the standard of care arm as the same distribution should be chosen for both arms (NICE DSU 14 guidance [47]). The CS states that the exponential produces higher OS estimates at 1 and 2 years than the observed data and is therefore conservative against tarlatamab. The company chooses the exponential distribution for its base case and provides scenarios for the alternative distributions in CS table 68.

The EAG notes that the company fits the parametric curve to the tarlatamab Kaplan-Meier data and then uses the same curve for standard of care. The EAG prefers to fit the parametric curve to the standard of care data and then use the same curve for the tarlatamab data because the standard of care data are mature and the population is much larger for the standard of care arm (n=540) than the tarlatamab arm (n=[redacted]. Based on visual fit and AIC/BIC statistical fit to the standard of care Kaplan-Meier data, the EAG considers the most appropriate model is the gamma and this model also provides a good fit for the tarlatamab arm. An alternative would be the Weibull distribution. The EAG uses the gamma distribution in our base case in section 6 and the Weibull in a scenario analysis (Table 38). The company and EAG's preferred curves for OS compared to the observed data are shown Figure 7.

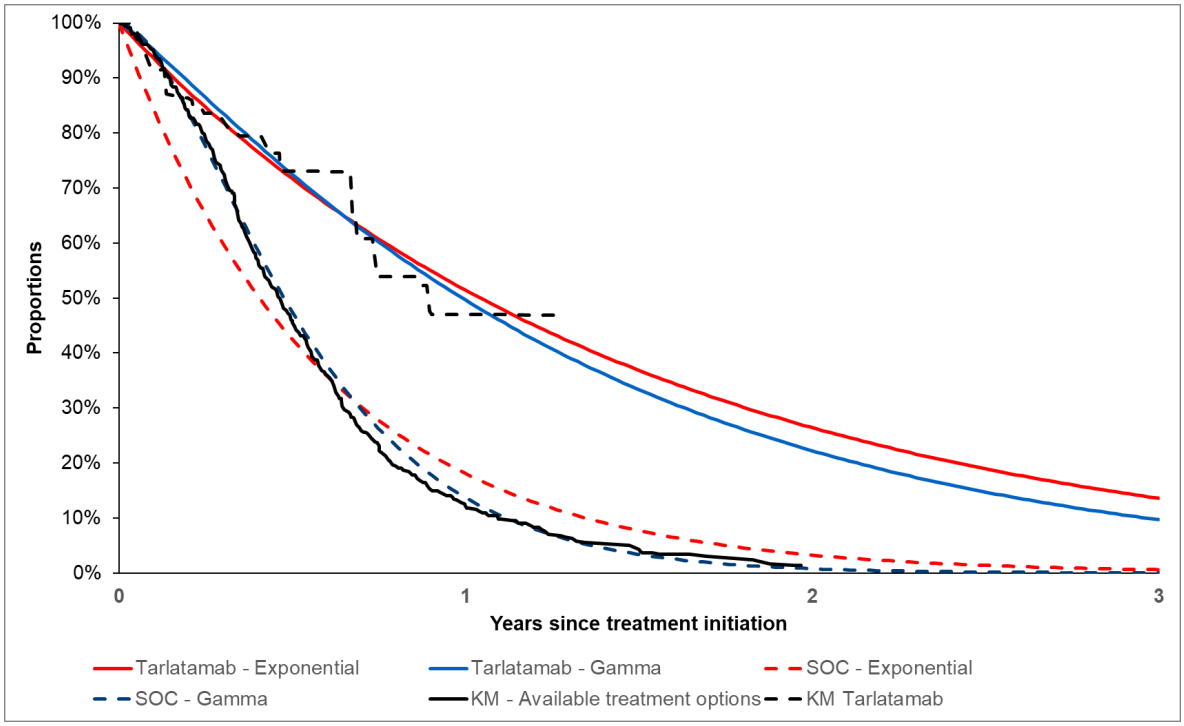

Figure 7 Selected parametric curves versus observed data for tarlatamab and standard of care for OS

Source: EAG created figure

4.2.6.2 Time to treatment discontinuation

The fitted parametric curves (exponential, Weibull, log-logistic, lognormal, Gompertz, generalised gamma and gamma) compared to the tarlatamab TTD trial data are shown in CS Figure 22. The goodness of fit was calculated for AIC and BIC and are shown in CS Table 45. The extrapolated 1 and 2 year TTD outcomes for each of the parametric functions are shown in CS Table 46. The exponential curve is chosen by the company based on this being the most clinically plausible.

The fitted parametric curves compared to the standard of care TTD data are shown in CS Figure 23. The goodness of fit using AIC/BIC was calculated and are shown in CS Table 47. The extrapolated 1 and 2 year TTD outcomes for each of the parametric functions are shown in CS Table 48. The CS states that the exponential is chosen as the same distribution should be chosen for both arms (NICE DSU 14 guidance [47]). The company chooses the exponential distribution for its base case.

The EAG notes that for the tarlatamab arm, all fitted curves for TTD cross the OS or PFS curves except for the exponential. Therefore, we prefer to also use the exponential curve in

our base case. The company and EAG's preferred curves for TTD compared to the observed data are shown in Figure 8.

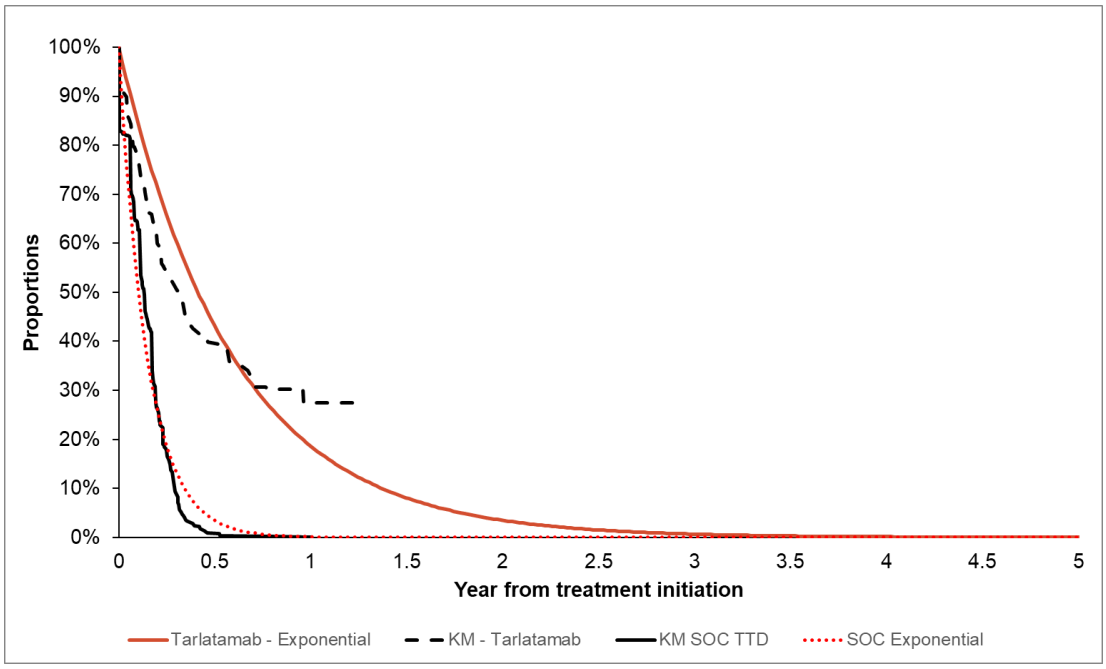

Figure 8 Selected parametric curves versus observed data for tarlatamab and standard of care for TTD

Source: EAG created figure

4.2.6.3 Progression-free survival

PFS was not available from the UK CAS study for the standard of care arm, and therefore the PFS is assumed to be the same as TTD. The company assessed whether the proportional hazard assumption holds for PFS. Based on the assessment, the proportional hazard assumption did not hold and therefore standard parametric curves were fit to the PFS curves for tarlatamab and standard of care separately. The EAG agrees with the company's judgement that the PH assumption is not supported.

The fitted parametric curves compared to the tarlatamab PFS trial data are shown in CS Figure 21. The goodness of fit was calculated for AIC and BIC and are shown in CS Table 43. The extrapolated 1- and 2-year PFS outcomes for each of the parametric functions are shown in CS Table 44. Based on the AIC/BIC data, the log-normal was chosen (lowest AIC/BIC values).

For the standard of care arm, as the PFS is assumed to be the same as the TTD, the exponential distribution is used.

We note that the distributions used for PFS are different between arms, and this is contrary to the advice provided in NICE DSU 14.[47] We therefore prefer to use the exponential distribution for both treatment arms.

EAG comment on treatment effectiveness and extrapolation

For OS, the company uses the exponential curve fitted to the observed data for both tarlatamab and standard of care. The EAG considers the gamma curve provides a better fit to the standard of care arm and prefers to use this curve for both treatments. We agree with the company's choice for modelling TTD. For PFS, the company uses the log-normal for the tarlatamab arm. For consistency between arms, we prefer to use the exponential curve for both the tarlatamab and standard of care arms.

4.2.7 Adverse events

Adverse events with grade ≥ 3 are included in the economic model for both arms. The frequency of serious adverse events is reported in CS Table 50 and included events such as anaemia, diarrhoea and febrile neutropenia. In addition, grade 1/2 CRS and grade 1/2 ICANS are included as these are specific to tarlatamab and its mode of action. [redacted] of patients had a CRS and [redacted] had an ICANS. Disutilities and costs were applied for each adverse event, reported below in sections 4.2.8.2 and 4.2.9. The duration of the adverse event is assumed to be 28 days for the serious adverse events and four days for grade 1/2 CRS and ICANS events.

We note that the adverse events for CAV were 0% for several AEs (e.g. diarrhoea, febrile neutropenia, lymphopenia, non-sepsis infection). Clinical advice to the EAG suggest that some of these AEs may have been underestimated, such as febrile neutropenia and diarrhoea. We conducted a scenario assuming the frequency of AEs for diarrhoea and febrile neutropenia was the average of that for topotecan and platinum-based chemotherapy (Table 38).

4.2.8 Health-related quality of life

4.2.8.1 Systematic literature review for utilities

The company conducted a systematic literature review of existing HRQoL studies, detailing the search and findings in CS Appendix H. The original search was conducted on 9th May

2022, with an updated search on 19th December 2023. The inclusion criteria are shown in CS Appendix H Table 37, with the main element being patients with small-cell lung cancer with disease progression on or after two lines of therapy. Two studies were identified and reported in CS Appendix H Table 38, however no references were provided for these texts and the EAG were unable to locate them. The EAG looked for further studies and found informative non-small-cell lung cancer papers. Clinical experts consulted on this appraisal agreed that non-small-cell lung cancer was a suitable proxy for small-cell lung cancer for HRQoL. A study performed by Chouaid et al.[48] comprised 263 patients in 25 hospitals internationally, with the EQ-5D questionnaire used to capture HRQoL, and a publication by Nafees et al. 2008[49] elicited UK utility values from 100 participants using a standard gamble approach.

4.2.8.2 Health state utility values used in the economic model

Health state utility values in the economic model were based upon EQ-5D-5L data collected in the DeLLphi-301 trial. These EQ-5D-5L utilities were mapped to European Quality of Life Working Group Health Status Measure 5 Dimension, 3 level (EQ-5D-3L) utility scores using the mapping algorithm by Hernandez-Alava,[50] and an age and sex-related utility adjustment was also applied to the health state utilities using general population utility values. The company developed a mixed-effects model to account for the repeated and longitudinal data collected in the DeLLphi-301 trial. The same utilities are assumed for both tarlatamab and standard of care arms in each health state; the death state is assumed to have a utility of zero. Table 24 below reports the health state utilities implemented in the model. The utilities implemented in the company model are derived from the entire DeLLphi-301 population, rather than the MAIC population (n=[redacted]); the utilities appear to be more favourable given the status of the patient population receiving third-line treatment compared with similar populations with NSCLC. Therefore, the EAG have used non-small-cell lung cancer as a proxy and have used utilities from Chouaid et al.[48] in the EAG base case. A scenario analysis using utilities from Nafees et al. 2008[49] is also performed in section 6.3.1.

Table 24 Model estimates on EQ-5D-3L utility scores in the DeLLphi-301 trial

Health states	N patients	N assessments	Mixed effects model		EAG base case (Chouaid et al. [48])
			Mean	SE	
Before treatment	[redacted]	[redacted]	[redacted]	[redacted]	-
Progression-free	[redacted]	[redacted]	[redacted]	[redacted]	0.62

Health states	N patients	N assessments	Mixed effects model		EAG base case (Chouaid et al. [48])
			Mean	SE	
Post-progression	[redacted]	[redacted]	[redacted]	[redacted]	0.47

Source: Reproduced from CS Table 49
EQ-5D-3L, European Quality of Life Working Group Health Status Measure 5 Dimensions, 3 levels; SE, standard error.

4.2.8.2.1 Adverse event disutilities

The company's economic model includes disutilities for grade 3+ for adverse events occurring in greater than 3% of patients in either treatment arm. These disutilities, based on the literature and assumptions, were assumed to last 28 days, and are reported in CS Table 51. The model also included disutilities for grade 1 and 2 CRS and ICANS, which were assumed to be equivalent to the maximum disutility of the adverse events and last 4 days. The EAG notes that the disutility for neutropenia from Sullivan et al[51] is reported incorrectly in CS Table 51; the correct disutility is -0.09. Table 25 below presents the utilities and disutilities used in the company model.

Table 25 Adverse event disutilities used in the revised company model

Adverse event disutilities	Value	Reference
Anaemia	-0.02	Sullivan et al. 2011[51]
Diarrhoea	-0.05	Nafees et al. 2008[49]
Fatigue	-0.07	
Febrile neutropenia	-0.09	
Leukopenia	-0.04	Sullivan et al. 2011[51]
Lymphocyte count decreased	-0.01	
Lymphopenia	-0.07	
Nausea	-0.05	Nafees et al. 2008[49]
Neutropenia	-0.09	Sullivan et al. 2011[51]
Neutrophil count decreased	-0.07	
Non-sepsis infection	-0.22	Stein et al. 2018[52]
Thrombocytopenia	-0.07	Sullivan et al. 2011[51]
CRS and ICANS disutilities		
CRS	-0.22	Assumed the same as maximum adverse event disutility
ICANS	-0.22	

Source: Reproduced from CS Table 51

CRS, cytokine release syndrome; ICANS, immune effector cell associated neurotoxicity syndrome.

EAG comment on HRQoL

Published literature for patient populations with NSCLC suggests that the baseline utility for patients in the progression free and progressed disease health states should be lower than those reported by the company. The company have used the entire DeLLphi-301 population (n=97) to inform baseline utilities, rather than the population used in the MAIC (n=[redacted]) which would better match the standard of care population. The EAG were unable to find well conducted studies for health state utilities for SCLC, and have therefore opted to use non-small-cell lung cancer as a proxy to inform baseline health state utilities in the EAG base case, which was deemed suitable by clinical experts consulted on this appraisal.

4.2.9 Resources and costs

4.2.9.1 Drug acquisition

CS Table 52 reports the dosing schedule and costs of drugs used in the model. The cost for tarlatamab used in the model is [redacted] per 10mg vial and [redacted] per 1mg vial. These prices are using the current Patient Access Scheme (PAS) discount. The DeLLphi-301 study informed the dosing schedule of tarlatamab, and the model incorporates pre-treatment with dexamethasone on day 1 and day 8 of the 1st cycle only. For standard of care, costs were obtained from the British National Formulary (BNF) and the drugs and pharmaceutical electronic market information tool (eMIT).[53,54] The EAG notes that the costs reported in the company submission for dexamethasone, carboplatin, etoposide, cyclophosphamide, doxorubicin and vincristine, all obtained from eMIT, were incorrect. Following clarification question B12, the company updated the economic model with the current prices. Table 26 below reports the prices used in the revised company model.

The company selected a cost of £7.50 per 0.25mg capsule of topotecan from the BNF. However, as the dose is 2.3mg/m^2 and the average body surface area of patients is 1.78m^2, the average dose would be 4.094mg. Therefore, it is more suitable to use 4 x 1mg capsules, rather than 16 x 0.25mg. The EAG have included this change in cost and capsule size as an EAG correction in section 5.3.3.

Table 26 Drug costs and dosing schedule used in the revised company model

Treatment	Admin Route	Dosing	Price	Reference
Tarlatamab				
Dexamethasone	IV	8mg on days 1 and 8 of cycle 1	£0.38 /6.60mg vial	DeLLphi-301 trial[24], eMIT[54]
Tarlatamab	IV	1mg on cycle 1 day 1, 10mg target dose on cycle 1 day 8 and day 15, Q2W thereafter in a 28-day cycle	[redacted]/10mg vial, [redacted]/1mg vial	DeLLphi-301 trial[24]
Topotecan				
Topotecan	Oral	2.30mg/m^2/d, 5 consecutive days every 21 days	£7.50/0.25mg capsule £36 /1 mg capsule	TA184[44], BNF[53]
Carboplatin + etoposide chemotherapy				
Carboplatin	IV	5mg/mL/min, on day 1 of each 21-day cycle	£48.09/450mg vial	Mansfield et al. 2020[55], eMIT[54]
Etoposide	IV	100mg/m^2, on days 1, 2, and 3 of each 21-day cycle	£13.40/500mg vial	
CAV				
Cyclophosphamide	IV	1000mg/m^2, on day 1 of 21-day cycle	£13.14/1000mg vial	Aix et al. 2023[56], eMIT[54]
Doxorubicin	IV	45mg/m^2, on day 1 of 21-day cycle	£15.98/200mg vial	

Treatment	Admin Route	Dosing	Price	Reference
Vincristine	IV	2mg, on day 1 of 21-day cycle	£6.64/2mg vial	

Source: Reproduced from CS Table 52
BNF, British National Formulary; IV, intravenous; CAV, cyclophosphamide, doxorubicin and vincristine; eMIT, electronic market information tool; min, minute; Q2W, once every 2 weeks.

4.2.9.2 Drug administration

The CS reports the drug administration costs for oral chemotherapy (topotecan) and intravenous chemotherapy (carboplatin + etoposide and CAV) in CS Table 53. Costs were obtained from the NHS England Payment Scheme 23/24.[57] The company assumed that the administration cost of dexamethasone and tarlatamab were equivalent to the cost of a simple IV chemotherapy. In the DeLLphi-301 trial, patients on tarlatamab were hospitalised for 24 hours following the tarlatamab infusion on days 1 and 8 of cycle 1. This incurred an additional cost of £488 which was included as an administration cost in the model. The EAG notes that the cost for oral chemotherapy provided in the CS was incorrect; the reported cost was £27.40, whilst the correct cost is £137. The company rectified the model in response to clarification question B3. Topotecan oral medication is provided to patients to take at home. The EAG notes that the topotecan administration in the model is incorrect, as it is applied five times per cycle rather than once. The EAG have made this change in the model as an EAG correction; see section 5.3.3. Table 27 reports the administration costs included in the company base case model.

Table 27 Drug administration costs used in the revised company model

Administration	Unit cost	Reference
Oral chemotherapy	£137	NHS England Payment Scheme 23/24 prices: Unbundled chemotherapy delivery SB11Z- Deliver exclusively oral chemotherapy[57]
IV chemotherapy (first attendance)	£172	NHS England Payment Scheme 23/24 prices: Unbundled chemotherapy delivery SB12Z- Deliver simple Parenteral Chemotherapy at first attendance[57]
IV chemotherapy (first administration of cisplatin etoposide	£515	NHS England Payment Scheme 23/24 prices: Unbundled chemotherapy delivery SB14Z- Deliver Complex Chemotherapy, including Prolonged infusional Treatment, at First Attendance[57]

Administration	Unit cost	Reference
IV chemotherapy (subsequent attendance)	£343	NHS England Payment Scheme 23/24 prices: Unbundled chemotherapy delivery SB15Z- Deliver subsequent elements of a chemotherapy cycle[57]
Required hospitalisation for 24 hours post-tarlatamab	£488	NHS England Payment Scheme 23/24 prices: Healthcare resource group (HRG) WH16A observation or counselling with CC score1+[57]

Source: Reproduced from CS Table 53 with oral chemotherapy cost corrected by the company
IV, intravenous.

4.2.9.3 Resource use

4.2.9.3.1 Medical resource use and monitoring costs

The medical resource use (MRU) and monitoring frequency and costs for patients receiving tarlatamab or standard of care was taken from the NICE submission for atezolizumab (TA638).[20] The costs and frequencies were assumed to be the same for both treatment arms and are applied until progression. CS Table 54 reports the MRU costs and frequencies, whilst CS Table 55 presents the monitoring and test costs and frequencies. Unit costs were acquired from the NHS England Payment Scheme 23/24,[57] PSSRU 2022,[58] and the NHS Schedule of Reference Costs 2011/22.[59] All frequencies were obtained from the NICE submission for atezolizumab (TA638).[20] Costs were inflated to 2023 values where required.[58] The CS originally applied MRU and monitoring costs to patients in the progression-free health state only. However, a proportion of patients in the progressed disease health state are on subsequent treatment (CS Table 58) and would incur additional resource use and monitoring costs. The EAG raised this issue in clarification question B7, and the company updated the economic model to incorporate MRU and monitoring costs for those patients in the progressed disease health state. Table 28 below presents the MRU and monitoring costs and frequencies used in the company's base case model.

The EAG notes that there are no additional costs for patients in the progressed disease state who are not on subsequent treatment. Clinical experts agree that these patients would still receive basic monitoring and therefore incur health care costs. However, the EAG consider that excluding these costs are not likely to have a significant impact on the model results. Further some of these costs may be included in the terminal care costs. Clinical advice to the EAG suggested that patients not on treatment would receive 0.1 hospital visits /week, 0.25 GP visits / week and 0.5 community nurse visits / week. We have used these values in a scenario in section 6.3.1.

Clinical experts consulted by the EAG suggest that the frequency of blood tests should be higher than those used by the company (0.13/week). As a result, the EAG have amended the frequency as an EAG correction to match that of outpatient visits, at 0.29 blood tests per week; see section 6.3.

Table 28 Medical resource use and monitoring used in the revised company model

Resource	Unit cost	Frequency/week	
Outpatient visit	£141	0.29	NICE submission for atezolizumab (TA638)[20]
GP visit – surgery	£42	0.06	
GP visit – home	£90	0.05	
Cancer nurse visit	£56	0.07	
Community nurse visit	£85	0.04	
Electrocardiogram	£134	0.01	
Chest X-ray	£29	0.11	
CT scan unit cost	£99	0.09	
Brain MRI unit cost	£209	0.01	
Blood test unit cost	£3	0.13	

Source: Reproduced from CS Table 54 and CS Table 55
GP, general practitioner; CT, computed tomography; MRI, magnetic resonance imaging

4.2.9.3.2 Adverse event costs

The company economic model includes costs for grade 3+ adverse events occurring in greater than 3% of patients in either treatment arm, as well as those for grades 1 or 2 CRS and ICANS. Unit costs per adverse event were acquired from NHS England Payment Scheme 23/24,[57] and costs for CRS and ICANS were obtained from the DeLLphi-301 trial,[24] the BNF,[53] and the NHS Reference Costs 2021/22,[59] inflated to 2023 prices.[58] For almost all adverse events where multiple potential HRG codes could apply, the company selected the most expensive HRG code. However, as the same set of HRG codes applied for febrile neutropenia and leukopenia, the company chose the most expensive cost for the former (SA08G) and the least expensive for the latter (SA08H). The cost per cycle was calculated based on the trial duration and the proportion of patients with adverse events.

The management cost per case of CRS comprised one dose of tocilizumab for 10% of patients and hospitalisation for 53% of patients (see CS Table 57). The EAG noted that the costs provided in CS Table 56 for adverse events are incorrect; the company acknowledged that the costs in the CS are incorrect, but that the costs in the company model are correct

with the exception of nausea, following clarification question B1. The EAG preferred choice of adverse event costs is to use a weighted average across all severity of complication levels available. We also consider the choice of HRG code for febrile neutropenia chosen by company to be unsuitable; instead, we have matched the cost for febrile neutropenia to non-sepsis infection. The EAG base case costs for adverse events are presented in Table 29 below.

Table 29 Adverse event costs used in the EAG base case model

Parameter	HRG code	EAG costs	Description	Reference
Anaemia	SA01G-SA01K	£2,055	Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia, CC score 0-8+	NHS Reference Costs 2021/22[59], NHS Cost Inflation Index[58]
Diarrhoea	FD01E-FD01H	£3,689	Non-Malignant Gastrointestinal Tract Disorders with Single Intervention, CC score 0-9+	
Fatigue	SA01G-SA01K	£2,055	Acquired Pure Red Cell Aplasia or Other Aplastic Anaemia, CC score 0-8+	
Febrile neutropenia	WH07C-WH07D	£5,071	Infections or Other Complications of Procedures, with Single Intervention, CC score 0-2+	
Leukopenia	SA08G-SA08J	£1,392	Other Haematological or Splenic Disorders, CC score 0-6+	
Lymphocyte count decreased	800	£303	Clinical Oncology Service	
Lymphopenia	SA08G-SA08J	£1,392	Other Haematological or Splenic Disorders, CC score 0-6+	
Nausea	800	£303	Clinical Oncology Service	

Parameter	HRG code	EAG costs	Description	Reference
Neutropenia	SA08G-SA08J	£1,392	Other Haematological or Splenic Disorders, CC score 0-6+	
Neutrophil count decreased	800	£303	Clinical Oncology Service	
Non-sepsis infection	WH07C-WH07D	£5,071	Infections or Other Complications of Procedures, with Single Intervention, CC score 0-2+	
Thrombocytopenia	SA12G-SA12K	£1013	Thrombocytopenia, CC score 0-8+	

Source: Partly reproduced from CS Table 56 and company economic model
CRS, cytokine release syndrome; EAG, External Assessment Group; HRG, Healthcare Resource Group; ICANS, immune effector cell-associated neurotoxicity syndrome; NHS, National Health Service

4.2.9.3.3 Subsequent treatment costs

A proportion of patients in the progressed disease health state were assumed to receive subsequent treatments and the remainder received no further treatment. The proportion of patients and the distribution of treatments were sourced from the DeLLphi-301 trial and the UK CAS natural history study for tarlatamab and standard of care, respectively. CS Table 58 presents the subsequent treatment distributions implemented in the company's economic model; this has been reproduced in Table 30 below. Clinical advice to the EAG confirmed that most patients would have no further treatment and that the subsequent treatment distribution for those patients who do receive further treatment was reasonable. Patients not on subsequent treatment may also incur monitoring costs in the progressed disease state, which have not been explicitly modelled by the company. However, these costs may be included in end-of-life costs; therefore, the EAG have included this as a scenario in section 6.3.1.

Table 30 Subsequent treatment distribution in the revised company model

Treatment	Duration (weeks)	Proportion post-tarlatamab	Proportion post-standard of care	Source
Topotecan	[redacted]	[redacted]	[redacted]	

Treatment	Duration (weeks)	Proportion post-tarlatamab	Proportion post-standard of care	Source
Clinical trials	[redacted]	[redacted]	[redacted]	UK CAS natural history study[60], DeLLphi-301 trial[24]
Platinum-based regimen	[redacted]	[redacted]	[redacted]	
CAV	[redacted]	[redacted]	[redacted]	
No treatment	[redacted]	[redacted]	[redacted]	

Source: Reproduced from CS Table 58
CAV, cyclophosphamide, doxorubicin and vincristine; SOC, standard of care

4.2.9.3.4 End of life costs

The company implemented an end-of-life cost of £8,408, applied when a patient transitions to the death state.[61]

EAG comment on resources and costs

The EAG has made changes to the economic model for dosing and administration frequency of topotecan.

For adverse events, the company selected costs corresponding to the highest CC score from the NHS England Payment scheme. The EAG prefer to have a weighted average approach across all CC scores; this has been implemented in the EAG base case model.

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

The company's base case results are shown in CS Table 64 with an incremental cost-effectiveness ratio (ICER) of £35,012 per QALY for tarlatamab versus standard of care for patients with previously treated advanced SCLC. This and all other cost effectiveness results in this report are conducted with a confidential PAS price discount for tarlatamab of [redacted]. The results are also shown with a severity multiplier of 1.7 which is applied to the incremental QALYs (see section 7). In response to the EAG's clarification questions, the company corrected the following input parameters in a revised base case:

- Adverse event costs of nausea of £303 instead of £313 (clarification question B1),
- Administration cost for oral chemotherapy of £137 instead of £27.40 (clarification question B3),
- Medical resource use and monitoring costs for patients receiving subsequent therapy (clarification question B7),
- Comparator drug costs taken from eMIT (clarification question B12), see table below.
- The EAG notes that the company also reduced the frequency of CRS / ICANS adverse events to zero in the model. We have corrected this is in the model, see section 5.3.3.

Table 31 Corrected comparator drug costs from eMIT

Drug/ dose/ unit	Price in submission appendix K (per vial)	Price in EMIT accessed May 2024 (per vial)
Carboplatin 450 mg vial	£14.69	£48.09
Etoposide 500mg vial	£10.69	£13.40
Cyclophosphamide 1g vial	£12.96	£13.14
Doxorubicin 200mg vial	£17.18	£15.98
Vincristine 2mg vial	£6.78	£6.64
Dexamethasone 6.60mg vial	£4.84	£0.375

Source: Clarification question B12

The revised base case results are shown in Table 32. The ICER for tarlatamab vs standard of care is £33,785 per QALY.

Table 32 Company revised base case after clarification response

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER with severity modifier (£/QALY)	ICER without severity modifier (£/QALY)
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		
Tarlatamab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£33,785	£57,434

ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year

5.2 Company's sensitivity analyses

5.2.1 Deterministic sensitivity analyses

The CS reports deterministic sensitivity analyses in section B.3.9.2 using the company's original model. Parameters were varied by +/- 20% of the deterministic point estimate for each parameter. The top 10 most influential variables were then plotted in a tornado diagram (CS Figure 29). The three most influential parameters were the utility value for tarlatamab patients who are progression-free, the relative dose intensity for tarlatamab and the unit cost for tarlatamab. The EAG notes that the results are shown without the severity multiplier and the results vary between £47,158 and £76,556 per QALY. Most parameters have been included in the deterministic sensitivity analyses. Adverse event frequency for patients on topotecan, adverse event duration and proportions on standard of care treatments were not included.

5.2.2 Scenario analysis

The list of the company's 18 scenarios and the rationale for including them are shown in CS Table 67. These include scenarios for the following:

- Discount rate,
- Time horizon,
- OS distribution,
- Utility,
- standard of care composition,
- Post-infusion costs,
- Impacts of CRS and ICANS,
- Post-progression tarlatamab use
- Vial wastage.

The results of the scenario analyses are shown in CS Table 68 using the company's original model. All scenarios were conducted probabilistically using 5000 iterations. The scenario ICERs ranged from £13,810 (OS distribution selection: generalised gamma) to £44,433 per QALY (No adjustment for post-progression tarlatamab use). The EAG has conducted these scenarios using the company's revised model and the results are shown in Table 36.

The company also conducted scenario analyses using their revised model following clarification questions A37-39. These scenarios are omitting sex and age at diagnosis from the MAIC (A37), including only the three prognostic factors considered to be the most important by the clinicians consulted by the company (A38) and using TTD instead of PFS from the DeLLphi-301 trial (A39). The cost effectiveness of tarlatamab versus standard of

care reduced to £23,290 for clarification question A37, £21,328 for clarification question A38, £33,168 per QALY for clarification question A39 respectively.

5.2.3 Probabilistic sensitivity analysis

The company conducted a probabilistic sensitivity analysis (PSA) with input parameters and distributions detailed in CS Table 62. The PSA was run for 5000 iterations. The standard errors for the parameters were taken, where possible, from the parameters' data source or else the standard error of the parameter was assumed to equal 20% of the mean value. Most parameters have been included in the PSA and the EAG considers that the distributions used are reasonable. Adverse event frequency for patients on topotecan, comparator unit costs were not included. The cost-effectiveness plane and cost-effectiveness acceptability curve using the company original model are shown in CS Figure 27 and Figure 28 respectively. The probabilistic results, shown in Table 33 below (CS Table 64), were in line with the deterministic results. Tarlatamab was associated with 36.4% probability of being cost-effective versus standard of care assuming a willingness-to-pay of £30,000 and including the severity modifier of 1.7.

Table 33 Base case results (probabilistic)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY)
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	-
Tarlatamab	[redacted]	[redacted]	[redacted]	[redacted]	£34,507

Source: EAG created table
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year..
Results include 1.7 severity multiplier applied to QALYs.

5.3 Model validation and face validity check

5.3.1 Company validation

The CS does not give any details of the technical checks conducted by the company to ensure the accuracy of the model.

The company discussed validation of the model by considering the survival extrapolation of the model compared to the observed data for tarlatamab and standard of care in CS section 3.12.1. The comparisons between the fitted curves against the observed data are shown in

CS Figure 30 and 31. The CS states that the predicted curves fitted the observed data well. The EAG provides a detailed critique of the fitted survival curves in section 4.2.6 above.

5.3.2 EAG validation

We conducted a range of checks on the company's model using an EAG checklist:

- Input checks: comparison of all parameter values in the model against the values stated in the CS and cited sources.
- Output checks: replication of results reported in the CS using the company model.
- 'White box' checks: manual checking of formulae which includes reviewing the calculations across each cycle and working backwards to trace links to input parameters and forwards to the results.
- 'Black box' checks: working through a list of tests to assess whether changes to key model inputs or assumptions have the expected effects on the model results.

5.3.3 Corrections to the company model

The EAG corrected the costs for topotecan and the administration costs for topotecan, as described in section 4.2.9.1. In addition, we increased the frequency of blood tests so this would be the same as the frequency of outpatient visits. The EAG notes that the company also reduced the frequency of CRS / ICANS adverse events to zero in the revised company model. We have corrected this is in the model.

The corrected model results are shown in Table 34. The corrected revised company model gives an ICER of £34,958 per QALY for tarlatamab vs standard of care.

Table 34 EAG corrections to the company revised base case model

Technologies	Total costs (£)	Total LYG	Total QALYs	Incremental costs (£)	Incremental LYG	Incremental QALYs	ICER with severity modifier (£/QALY)	ICER without severity modifier (£/QALY)
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]		
Tarlatamab	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£34,958	£59,429

Source: EAG created table
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALY, quality-adjusted life year.

5.3.4 EAG summary of key issues and additional analyses

A full summary of EAG observations on key aspects of the company's economic model is presented in Table 35.

Table 35 EAG observations of the key aspects of the company's economic model

Parameter	Company base case	EAG comment	EAG base case
Model structure			
Model structure	Partitioned survival model with states for PFS, PD and OS	We agree	No change
Population	Section 4.2.3	We agree	No change
Comparators	Standard of care (section 4.2.4)	We agree	No change
Perspective	NHS and PSS	We agree	No change
Time horizon	Lifetime	We agree	No change
Discounting	3.5% for costs and outcomes	We agree	No change
Survival curves			
OS	Exponential distribution for tarlatamab and SOC arms	We consider gamma distribution provides better fit to SOC (section 4.2.6.1)	Gamma distribution for tarlatamab and SOC arms
PFS	Lognormal distribution for the tarlatamab arm and exponential distribution for the SOC arm	We consider the same distribution should be used for both arms (section 4.2.6.3)	Exponential arm for both treatment arms.
TTD	Exponential distribution for tarlatamab and SOC arms	We agree	No change
Adverse events			
Frequency of adverse events	CS Table 50	We agree	No change
Utilities			
Patient utilities	Section 4.2.8.2	We consider the utility values not to be generalisability to patients receiving third line treatment.	We use the utility values from Chouaid et al. in patients with NSCL.

Parameter	Company base case	EAG comment	EAG base case
AEs disutilities	Section 4.2.8.2	We agree	No change
Severity modifier	Severity modifier of 1.7 used	We agree	No change
Resource use and costs			
Drug acquisition and administration	Section 4.2.9	We agree	No change
Healthcare resource use	Section 4.2.9.3	We agree	No change
Adverse event costs	Section 4.2.9	The adverse event costs were chosen for the most severe complication level.	We use adverse events costs calculated using the weighted average of complication severity.

Source: EAG created table
AE, adverse events; EAG, External Assessment Group; NHS, National Health Service; NICE, The National Institute for Health and Care Excellence; PSS, Personal Social Services; SOC, standard of care.

6 EAG'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the EAG

The EAG repeated the company scenario analyses using the revised version of the model submitted with their clarification response. The results are shown in Table 36. The results are most sensitive to changes in the distributions fitted to the OS curves which vary between £21,665 (lognormal) and £45,030 per QALY (Gompertz).

Table 36 Company scenarios with revised version of the company model, including PAS for tarlatamab

Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Company revised base case	[redacted]	[redacted]	£33,785
Set costs and discount rates to 1.5%	[redacted]	[redacted]	£33,078
Set costs and discount rates to 5%	[redacted]	[redacted]	£34,303
Time horizon 5 years	[redacted]	[redacted]	£35,280
Time horizon 15 years	[redacted]	[redacted]	£33,743
OS distribution selection: Weibull	[redacted]	[redacted]	£40,449
OS distribution selection: lognormal	[redacted]	[redacted]	£21,665
OS distribution selection: log-logistic	[redacted]	[redacted]	£23,778
OS distribution selection: Gompertz	[redacted]	[redacted]	£45,030
OS distribution selection: generalised gamma	[redacted]	[redacted]	£16,664
OS distribution selection: gamma	[redacted]	[redacted]	£39,074
Treatment-specific health state utility values used in the PFS state	[redacted]	[redacted]	£33,011
SOC composition: all patients received topotecan	[redacted]	[redacted]	£32,300
SOC composition: all patients received CAV	[redacted]	[redacted]	£34,983
SOC composition: all patients received platinum-based chemotherapy	[redacted]	[redacted]	£34,631
Do not consider post-infusion hospitalisation costs for tarlatamab	[redacted]	[redacted]	£32,935

Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
Do not consider impacts of CRS and ICANS	[redacted]	[redacted]	£33,785
Do not adjust tarlatamab OS and TTD for post-progression tarlatamab use	[redacted]	[redacted]	£43,548
Do not consider vial wastage	[redacted]	[redacted]	£33,849

Source: From company model; includes severity multiplier of 1.7.
CAV, cyclophosphamide, doxorubicin and vincristine; CRS, cytokine release syndrome; ICANS, immune effector cell-associated neurotoxicity syndrome; ICER, incremental cost-effectiveness ratio; OS, overall survival; SOC, standard of care.

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

6.3 EAG's preferred assumptions

Based on the EAG critique of the company's model discussed in Table 35, we have identified several key aspects of the company base case with which we disagree. Our preferred model assumptions are the following:

- Overall survival: use gamma distribution for tarlatamab and standard of care arms, instead of exponential distribution (section 4.2.6.1)
- Progression-free survival: use exponential distribution for tarlatamab arm, instead of lognormal for the tarlatamab arm (section 4.2.6.3)
- Health-related quality of life: use Chouaid et al. instead of DeLLphi-301 study (section 4.2.8.2),
- Adverse event costs: EAG recalculated values (section 4.2.9.3.2).

The EAG base case results are shown in Table 37 using the EAG's preferred assumptions. When using these assumptions, the ICER increases to £58,847 per QALY for tarlatamab vs standard of care. The model results are most sensitive to using the gamma distribution for OS and changing the source of the HRQoL values.

Table 37 EAG's preferred model assumptions: cumulative impact (deterministic) with PAS for tarlatamab

Preferred assumption	Treatment	Total costs	Total QALYs	Cumulative ICER £/QALY
EAG corrected company base-case model	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£34,958
+ OS: gamma for both arms	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£40,442
+ PFS: exponential for both arms	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£42,045
+ HRQoL: use Chouaid et al. (PFS 0.62; PD 0.47)	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£55,097
+ Adverse event costs: EAG recalculated.	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£58,847
EAG base case	SOC	[redacted]	[redacted]	
	Tarlatamab	[redacted]	[redacted]	£58,847

Source: EAG created table
EAG, evidence assessment group; OS, overall survival; PFS, progression-free survival; HRQoL, health-related quality of life; PD, progressed disease.
Severity multiplier of 1.7 applied to incremental QALYs.

6.3.1 EAG scenario analyses

We performed a range of scenario analyses with the EAG base case to analyse the impact of changing some model assumptions on the final cost-effectiveness results. Table 38 below summarises the results of the scenario analyses on the EAG base case. The following scenarios were conducted:

- Selection of the company's scenario analyses (CS Table 68)
- Alternative QoL values from Nafees et al. (PFS 0.653, PD 0.473)
- Clarification questions A37 (MAIC omitting age and sex) and A38 (MAIC with 3 main prognostic factors).
- Alternative values for AEs for CAV for diarrhoea and febrile neutropenia
- Health care resource costs for patients with progressed disease not on treatment (0.1 hospital visits /week, 0.25 GP home visits / week and 0.5 community nurse visits / week).

The ICERs for the scenarios varied between £26,162 (OS distribution selection: generalised gamma) and £69,309 per QALY (Do not adjust tarlatamab OS and TTD for post-progression tarlatamab use).

Table 38 EAG's scenario analyses with PAS for tarlatamab

Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
EAG base case	[redacted]	[redacted]	£58,847
Company scenarios (CS 67-68)			
Set costs and discount rates to 1.5%	[redacted]	[redacted]	£57,898
Set costs and discount rates to 5%	[redacted]	[redacted]	£59,545
Time horizon 5 years	[redacted]	[redacted]	£59,764
Time horizon 15 years	[redacted]	[redacted]	£58,837
OS distribution selection: Weibull	[redacted]	[redacted]	£60,640
OS distribution selection: lognormal	[redacted]	[redacted]	£33,767
OS distribution selection: log-logistic	[redacted]	[redacted]	£36,959
OS distribution selection: Gompertz	[redacted]	[redacted]	£66,338
OS distribution selection: generalised gamma	[redacted]	[redacted]	£26,162
OS distribution selection: exponential	[redacted]	[redacted]	£51,592
Treatment-specific health state utility values used in the PFS state	[redacted]	[redacted]	£53,056
SOC composition: all patients received topotecan	[redacted]	[redacted]	£54,654
SOC composition: all patients received CAV	[redacted]	[redacted]	£56,036
SOC composition: all patients received platinum-based chemotherapy	[redacted]	[redacted]	£53,171
Do not consider post-infusion hospitalisation costs for tarlatamab	[redacted]	[redacted]	£53,635
Do not consider impacts of CRS and ICANS	[redacted]	[redacted]	£53,857
Do not adjust tarlatamab OS and TTD for post-progression tarlatamab use	[redacted]	[redacted]	£69,309
Do not consider vial wastage	[redacted]	[redacted]	£54,944
Additional EAG scenarios			
QoL values from DeLLphi-301 study	[redacted]	[redacted]	£44,907
QoL values from Nafees et al. (PFS 0.653, PD 0.473)	[redacted]	[redacted]	£56,412

Scenario	Incremental costs	Incremental QALYs	ICER (£/QALY)
CQ A37 MAIC omit age and sex at diagnosis	[redacted]	[redacted]	£45,730
CQ A38 MAIC only include 3 main prognostic factors	[redacted]	[redacted]	£39,720
AEs for CAV, diarrhoea 3.26%, febrile neutropenia 4.56%	[redacted]	[redacted]	£58,639
Health care resource costs for patients with progressed disease not on treatment.	[redacted]	[redacted]	£59,979

Source: From company model; includes severity multiplier of 1.7.
CAV, cyclophosphamide, doxorubicin and vincristine; CRS, cytokine release syndrome; ICANS, immune effector cell-associated neurotoxicity syndrome; ICER, incremental cost-effectiveness ratio; OS, overall survival; SOC, standard of care; CQ clarification question; MAIC matching-adjusted indirect comparison.

6.3.2 Probabilistic sensitivity analysis

The EAG conducted a PSA for the EAG base case analysis with 5000 simulations. The results are shown in Table 39. The ICER is £56,825 per QALY for tarlatamab vs standard of care. Tarlatamab has a 0.6% probability of being cost-effective at a willingness threshold of £30,000 per QALY.

Table 39 Probabilistic results for the EAG base case results (probabilistic)

Technologies	Total costs (£)	Total QALYs	Incremental costs (£)	Incremental QALYs	ICER versus baseline (£/QALY)
Standard of care	[redacted]	[redacted]	[redacted]	[redacted]	
Tarlatamab	[redacted]	[redacted]	[redacted]	[redacted]	£56,825

Source: EAG created table
ICER, incremental cost-effectiveness ratio; LYG, life years gained; QALYs, quality-adjusted life years. Results include 1.7 severity multiplier applied to QALYs.

6.4 Conclusions on the cost effectiveness evidence

The company developed a model to estimate the cost-effectiveness of tarlatamab compared to standard of care for patients with previously treated SCLC. The EAG considers the structure of the model to be reasonable, appropriate and consistent with previous cost-effectiveness models for lung cancer. The company made some minor changes to the model in response to clarification questions. The company's revised base case shows an ICER of

£33,785 per QALY for tarlatamab versus standard of care, including a PAS discount for tarlatamab of [redacted] and including a severity multiplier of 1.7.

The EAG made corrections for the cost and administration of topotecan and the frequency of blood tests.

The EAG disagrees with several of the assumptions in the company's model. Our preferred assumptions include:

- Overall survival: use the gamma distribution for the tarlatamab and standard of care arms, instead of the exponential distribution,
- Progression-free survival: use the exponential distribution for the tarlatamab arm, instead of the lognormal distribution,
- Health-related quality of life: use Chouaid et al. instead of DeLLphi-301 study,
- Adverse event costs: EAG recalculated values.

Incorporating the EAG preferred assumptions, the ICER increase to £58,847 per QALY for tarlatamab vs standard of care. The model results are most sensitive to using the gamma distribution for OS and changing the source of the HRQoL values.

7 SEVERITY

The company calculated the QALY shortfall for tarlatamab by using the online tool published by Schneider et al.[62] The company used the sex distribution ([redacted]) and starting age ([redacted]) from the tarlatamab population after MAIC matching (CS Table 62). The QALYs for patients with third-line SCLC are taken from the standard of care arm. Clinical advice to the EAG confirmed that the OS estimates for the standard of care arm are consistent with those in UK clinical practice. The proportional QALY shortfall for tarlatamab is more than 95% (see Table 40 below). We also calculated the absolute and proportional QALY shortfall using the EAG base case (Table 37) and obtained similar results to the company's revised base case (Table 40) so we agree that there is a case for applying a multiplier for disease severity of 1.7 which the company used in the base case analysis.

Table 40 QALY shortfall analysis

	Expected total QALYs for the general population	Total QALYs that people living with a condition would be expected to have with current treatment	Absolute QALY shortfall	Proportionate QALY shortfall
Company's revised base case	12.03	[redacted]	[redacted]	[redacted]
EAG base case	12.03	[redacted]	[redacted]	[redacted]

Source: Schneider et al. 2021[63]
QALY, quality adjusted life-year.

8 REFERENCES

1. Phillippo DMAAEDSPSAKRWNJ. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submission to NICE. 2016. Available from http://www.nicedsu.org.uk. 2016.
2. Analysis Group Inc (Amgen Data on File). Indirect treatment comparison of tarlatamab versus available treatment options in patients with relapsed or refractory small cell lung cancer after two prior lines of treatment, 2024.
3. Khakwani A, Rich AL, Tata LJ, et al. Small-cell lung cancer in England: trends in survival and chemotherapy using the National Lung Cancer Audit. *PLoS One* 2014;9(2):e89426.
4. National Lung Cancer Audit. *The National Lung Cancer Audit. Data and statistics (for 2022).* https://www.lungcanceraudit.org.uk/data/ (accessed 19/06/2024).
5. Früh M, De Ruysscher D, Popat S, et al. Small-cell lung cancer (SCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24 Suppl 6:vi99-105.
6. Ganti AKP, Loo BW, Bassetti M, et al. Small Cell Lung Cancer, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2021;19(12):1441-64.
7. Cancer Research UK. *Lung cancer incidence statistics (2017-2019)*. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/incidence#heading-Zero (accessed 19/06/2024).
8. Gysling S, Morgan H, Ifesemen OS, et al. The Impact of COVID-19 on Lung Cancer Incidence in England: Analysis of the National Lung Cancer Audit 2019 and 2020 Rapid Cancer Registration Datasets. *Chest* 2023;163(6):1599-607.
9. Royal College of Physicians. National Lung Cancer Audit annual report (for the audit period 2019 England, Wales and Guernsey and 2020 England only). London, 2022.
10. Blackhall F, Girard N, Livartowski A, et al. Treatment patterns and outcomes among patients with small-cell lung cancer (SCLC) in Europe: a retrospective cohort study. *BMJ Open* 2023;13(2):e052556.
11. Li N, Chu Y, Song Q. Brain Metastasis in Patients with Small Cell Lung Cancer. *Int J Gen Med* 2021;14:10131-39.
12. Rudin CM, Brambilla E, Faivre-Finn C, Sage J. Small-cell lung cancer. *Nat Rev Dis Primers* 2021;7(1):3.
13. Bennett BM, Wells JR, Panter C, et al. The Humanistic Burden of Small Cell Lung Cancer (SCLC): A Systematic Review of Health-Related Quality of Life (HRQoL) Literature. *Front Pharmacol* 2017;8:339.
14. Laudicella M, Walsh B, Burns E, Smith PC. Cost of care for cancer patients in England: evidence from population-based patient-level data. *Br J Cancer* 2016;114(11):1286-92.
15. Ben-Ari E. *Tarlatamab shows promise for some people with small cell lung cancer*. https://www.cancer.gov/news-events/cancer-currents-blog/2023/tarlatamab-previously-treated-sclc.
16. Paz-Ares L, Champiat S, Lai WV, et al. Tarlatamab, a First-in-Class DLL3-Targeted Bispecific T-Cell Engager, in Recurrent Small-Cell Lung Cancer: An Open-Label, Phase I Study. *Journal of Clinical Oncology* 2023;41(16):2893-903.
17. Amgen Data on F. Tarlatamab 10 mg Summary of Product Characteristics.
18. National Institute for H, Care E. *Lung cancer: diagnosis and management (NICE Guideline 122). Available at: https://www.nice.org.uk/guidance/ng122. Accessed: 09/11/23*.
19. National Institute of H, Care E. Topotecan for the treatment of relapsed small-cell lung cancer [TA184]. Available at: https://www.nice.org.uk/guidance/ta184/chapter/3-The-technology. [Accessed 01/11/23].

20. National Institute of Health and Care Excellence. Atezolizumab with carboplatin and etoposide for untreated extensive-stage small-cell lung cancer. Available at: https://www.nice.org.uk/guidance/ta638. (Accessed 09/11/23).
21. Azam F, Farooq Latif M, Farooq A, et al. Performance Status Assessment by Using ECOG (Eastern Cooperative Oncology Group) Score for Cancer Patients by Oncology Healthcare Professionals. *Case Rep Oncol* 2019;12(3):728-36.
22. Amgen Data on F. Tarlatamab Draft Summary of Product Characteristics. 2023.
23. Ahn MJ, Cho BC, Felip E, et al. Tarlatamab for Patients with Previously Treated Small-Cell Lung Cancer. *N Engl J Med* 2023.
24. Amgen Data on F. DeLLphi-301 Clinical Study Report (June 2023 DCO).
25. Clinicaltrials.gov. A Phase 2 Study of Tarlatamab in Patients With Small Cell Lung Cancer (SCLC) (DeLLphi-301). Available at: https://classic.clinicaltrials.gov/ct2/show/study/NCT05060016. (Accessed: 14/12/23).
26. Clinicaltrials.gov. *Study Comparing Tarlatamab With Standard of Care Chemotherapy in Relapsed Small Cell Lung Cancer (DeLLphi-304)*. https://clinicaltrials.gov/study/NCT05740566 (accessed 1st July 2024).
27. British Journal of Sports M. Downs and Black's Checklist. 2018. Available at: https://bjsm.bmj.com/content/bjsports/52/6/387/dc3/embed/inline-supplementary-material-3.pdf?download=true (Accessed: 15/05/24).
28. National Institute for Health and Care Excellence. Single technology appraisal and highly specialised technologies evaluation: User guide for company evidence submission template: NICE, 2022.
29. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A Quality-of-Life Instrument for Use in International Clinical Trials in Oncology. *JNCI: Journal of the National Cancer Institute* 1993;85(5):365-76.
30. Griffiths P, Peipert JD, Leith A, et al. Validity of a single-item indicator of treatment side effect bother in a diverse sample of cancer patients. *Support Care Cancer* 2022;30(4):3613-23.
31. Chung HC, Piha-Paul SA, Lopez-Martin J, et al. Pembrolizumab After Two or More Lines of Previous Therapy in Patients With Recurrent or Metastatic SCLC: Results From the KEYNOTE-028 and KEYNOTE-158 Studies. *J Thorac Oncol* 2020;15(4):618-27.
32. Ready N, Farago AF, de Braud F, et al. Third-Line Nivolumab Monotherapy in Recurrent SCLC: CheckMate 032. *J Thorac Oncol* 2019;14(2):237-44.
33. Clopper C, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1934;26(4):404-13.
34. Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. *Biometrics* 1982;38(1):29-41.
35. Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*. New York: John Wiley & Sons; 1980.
36. Bergman B, Aaronson NK, Ahmedzai S, et al. The EORTC QLQ-LC13: a modular supplement to the EORTC Core Quality of Life Questionnaire (QLQ-C30) for use in lung cancer clinical trials. EORTC Study Group on Quality of Life. *Eur J Cancer* 1994;30a(5):635-42.
37. Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Brussels, 2001.
38. National Institute for Health and Care Excellence. Atezolizumab with carboplatin and etoposide for untreated extensive-stage small-cell lung cancer. Available at: https://www.nice.org.uk/guidance/ta638. (Accessed 09/11/23).
39. Amgen Data on F. ITC prognostic variable assessment report.
40. Amgen Data on File. Summary meta-regression in small cell lung cancer.
41. Blumenthal GM, Gong Y, Kehl K, et al. Analysis of time-to-treatment discontinuation of targeted therapy, immunotherapy, and chemotherapy in clinical trials of patients with non-small-cell lung cancer. *Annals of Oncology* 2019;30(5):830-38.

42. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health* 1998;52(6):377-84.
43. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;12:9.
44. Topotecan for the treatment of relapsed small-cell lung cancer. Avaiable at: https://www.nice.org.uk/guidance/ta184/chapter/2-Clinical-need-and-practice. (Accessed: 08/05/24).
45. Canada's Drug and Health Technology Agency, Lurbinectedin. Available at: https://www.cadth.ca/lurbinectedin (Accessed: 13/05/24).
46. The Pharmaceutical Benefits Scheme, TOPOTECAN,capsules, 0.25 mg and 1 mg (as hydrochloride), Hycamtin®, July 2010. 2010. Available at: https://www.pbs.gov.au/info/industry/listing/elements/pbac-meetings/psd/2010-07/pbac-psd-Topotecan-july10 (Accessed: 13/05/24).
47. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data. . 2011.
48. Chouaid C, Agulnik J, Goker E, et al. Health-related quality of life and utility in patients with advanced non-small-cell lung cancer: a prospective cross-sectional patient survey in a real-world setting. *J Thorac Oncol* 2013;8(8):997-1003.
49. Nafees B, Stafford M, Gavriel S, et al. Health state utilities for non small cell lung cancer. *Health Qual Life Outcomes* 2008;6:84.
50. Hernández Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results from a UK Population Study. *PharmacoEconomics* 2023;41(2):199-207.
51. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V. Catalogue of EQ-5D scores for the United Kingdom. *Med Decis Making* 2011;31(6):800-4.
52. Stein EM, Yang M, Guerin A, et al. Assessing utility values for treatment-related health states of acute myeloid leukemia in the United States. *Health Qual Life Outcomes* 2018;16(1):193.
53. British National Formulary (BNF). BNF 83 March - September 2021, date of consultation: 4/6/2023. Available at: https://bnf.nice.org.uk/ (Accessed: 15/05/24).
54. eMIT national database 1 July 2023 to 31 December 2023. Available at: (https://www.gov.uk/government/publications/drugs-and-pharmaceutical-electronic-market-information-emit). (Accessed: 13/05/24).
55. Mansfield AS, Każarnowicz A, Karaseva N, et al. Safety and patient-reported outcomes of atezolizumab, carboplatin, and etoposide in extensive-stage small-cell lung cancer (IMpower133): a randomized phase I/III trial. Ann Oncol. 2020;31(2):310-317.
56. Aix SP, Ciuleanu TE, Navarro A, et al. Combination lurbinectedin and doxorubicin versus physician's choice of chemotherapy in patients with relapsed small-cell lung cancer (ATLANTIS): a multicentre, randomised, open-label, phase 3 trial. The Lancet Respiratory Medicine. 2023;11(1): 74-86.
57. 2023-25 NHS Payment Scheme. 2024. Accessed on March 2024. Available from: https://www.england.nhs.uk/publication/2023-25-nhs-payment-scheme/#heading-6.
58. Personal Social Services Research Unit (PSSRU), Unit Costs of Health & Social Care 2022. Available at: https://www.pssru.ac.uk/project-pages/unit-costs/. (Accessed 10/01/24).
59. 2021-22 NHS Cost Collection. 2023. Accessed on March 2024. Available from: https://www.england.nhs.uk/publication/2021-22-national-cost-collection-data-publication.
60. Amgen Data on F. UK CAS RWE study results.
61. Round J, Jones L, Morris S. Estimating the cost of caring for people with cancer at the end of life: A modelling study. *Palliat Med* 2015;29(10):899-907.

62. Schneider P, McNamara S, Love-Koh J, et al. *QALY Shortfall Calculator. 2022. Available at: https://shiny.york.ac.uk/shortfall. [Last accessed: 17 August 2023]*.
63. Paul Schneider SM, James Love-Koh, Tim Doran, Nils Gutacker. *QALY Shortfall Calculator. 2021. Available at: https://shiny.york.ac.uk/shortfall. [Accessed February 2024]*.
64. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *Bmj* 2019;366.

9 APPENDICES

Appendix 1

In clarification question A2 we asked the company to provide evidence for their statement that their three selected comparators have similar efficacy. The company provided results from two selected studies in response. One study showed no statistically significant differences between topotecan and CAV on ORR, median time to progression and median OS in patients with SCLC who had relapsed after a minimum of 60 days following the end of first-line therapy. The other study showed statistically significant differences between topotecan and CAV on ORR and median PFS, but not OS, among patients with advanced stage IV or locally relapsed SCLC. Additional analyses of UK data used for standard of care that the company submitted in response to clarification question A3 show that TTD and OS are similar for topotecan and CAV, but both outcomes are better for patients treated with platinum. Drawing on a point the company makes in a covering letter to these additional analyses, this may be due to these participants still being platinum-sensitive, but it is possible that the results may also reflect differences in the efficacy of the treatments.

Appendix 2

Table 41 EAG appraisal of systematic review methods

Systematic review components and processes	EAG response	EAG comments
Was the review question clearly defined using the PICOD framework or an alternative?	Yes	The aim of the systematic review was "*to identify relevant clinical evidence on the clinical efficacy and safety outcomes in patients with SCLC whose disease has progressed following two prior treatments*" (CS section B.2.1). CS Appendix D.1.2 Table 9 describes the eligibility criteria in PICOD format.
Were appropriate sources of literature searched?	Yes	Good coverage of sources including MEDLINE, Embase, Cochrane Central Register of Controlled Trials and Cochrane Database of Systematic Reviews, conference abstracts. Did not search clinical trial registries for ongoing studies.
What time period did the searches span and was this appropriate?	The initial search was from 2012 to 19th April 2022. An update search spanned the period from 1st January 2022 to 19th December 2023. This is an appropriate time span.	Conferences were included in the Embase search from 2018 and handsearching of major oncology and lung cancer conferences was from 2022 onwards.
Were appropriate search terms used and combined correctly?	No	Search strategies provided in CS Appendix D.1.1 Tables 1 to 8. Relevant subject headings and relevant free text terms were used in the searches. Terms for RCTs were not included in the search strategies. A reason for this is not given in the CS. However, we do not believe any relevant evidence has been missed.
Were inclusion and exclusion criteria specified? If so, were these criteria appropriate and	Yes, specified inclusion and exclusion criteria were appropriate and relevant to the decision problem.	Eligibility criteria provided in CS Appendix D.1.2 Table 9. For study design, this included RCTs but, as noted in the row above, no search

Systematic review components and processes	EAG response	EAG comments
relevant to the decision problem?		terms for RCTs were included in the search strategies.
Were study selection criteria applied by two or more reviewers independently?	Yes	CS Appendix D.1.2 states two independent reviewers screened both titles and abstracts and then screened the retrieved full papers in a double-blind manner.
Was data extraction performed by two or more reviewers independently?	No	Data extraction was carried out by one reviewer. A second reviewer verified all extracted data.
Was a risk of bias assessment or a quality assessment of the included studies undertaken? If so, which tool was used?	Yes	CS Appendix D.1.2 states the Cochrane risk of bias tool (RoB2) [64] was used to assess each RCT, but as stated previously, RCTs were not searched for, and none are included in the CS. Single-arm trials were assessed using the Downs and Black tool.[42]
Was risk of bias assessment (or other study quality assessment) conducted by two or more reviewers independently?	No	Study quality assessment was carried out by one reviewer. A second reviewer validated the assessment.
Is sufficient detail on the individual studies presented?	Yes	Evidence summarised in CS sections B.2.3 to B.2.7 and B.2.10. The evidence used in the indirect comparison is summarised in CS Appendix D.1.4 to D.1.5 and CS section B.2.9.4.
If statistical evidence synthesis (e.g. pairwise meta-analysis, ITC, NMA) was undertaken, were appropriate methods used?	Yes	The EAG is generally satisfied with the methods the company used to carry out an unanchored MAIC of tarlatamab versus standard of care. Please see sections 3.3 and 3.4.

Source: EAG table.
CS, company submission; EAG, external assessment group; MAIC, matching-adjusted indirect comparison; PICOD, patient, intervention, comparator, outcome, design; RCT, randomised controlled trial

Appendix 3

Table 42 Assessment of Methodological Quality of the DeLLphi-301 study (modified Downs and Black checklist[27]

Outcomes	Company (CS Table 10)		EAG
	Score	Definition	Comments
Reporting			
Is the hypothesis /aim/ objective of the study clearly described?	1	Yes	**Agree**: CS section B.2.2 states the trial is investigation the safety and efficacy of tarlatamab in patients with relapsed or refractory SCLC after two or more lines of treatment.
Are the main outcomes to be measured clearly described in the Introduction or Methods section?	1	Yes	**Agree**: Outcomes are listed in CS Tables 3 and 4 and are defined in CS Table 5
Are the characteristics of the patients included in the study clearly described?	1	Yes	**Agree**: CS Table 6 reports baseline demographics of trial participants in receipt of 10mg tarlatamab dose. CS Table 7 reports the disease characteristics for this group.
Are the interventions of interest clearly described?	1	Yes	**Agree**: Clearly described in CS section B.2.3.3 (Target dose) with additional details elsewhere in the CS.
Are the distributions of principal confounders in each group of patients to be compared clearly described?	0	Unable to determine	**Agree:** No groups compared as this is a single-arm study.
Are the main findings of the study clearly described?	1	Yes	**Agree**: Results relevant to the appraisal (Full analysis set (BICR) and Safety analysis set) presented in CS section B.2.6 with results for other analysis sets available in the CSR. Sufficient information is provided (e.g. numerators and denominators reported, numbers at risk reported).
Does the study provide estimates of the random variability in the data for the	1	Yes	**Agree**: Confidence intervals, 25th and 75th percentiles, minimum and maximum values, quartile values reported.

Outcomes	Company (CS Table 10)		EAG
	Score	Definition	Comments
main outcomes?			
Have all important adverse events that may be a consequence of the intervention been reported?	1	Yes	**Agree**: Adverse events are reported from the DeLLphi-301 study for the 10mg target dose participants in Parts 1 and 2 of the trial. However, we note that reporting of adverse events in the CS is incomplete because data for the 34 participants enrolled in Part 3 of the study were not included. The company provided these data in response to clarification question A14.
Have the characteristics of patients lost to follow-up been described?	1	Yes	**Agree.** The flow of participants through the trial is shown in CS Figure 3, which includes reasons for discontinuing treatment (which were disease progression and death). The CS also reports [redacted] participants were lost to follow-up in the duration of response and duration of disease control analyses. No participants were lost to follow-up in the PFS analyses. [redacted] participant was lost to follow-up in the OS analysis.
Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?	1	Yes	**Disagree**: Single-arm trial so no comparative statistics for the main outcomes for which a p-value would be reported.
External validity			
Were the patients asked to participate in the study representative of the entire population from which they were recruited?	1	Yes	**Agree:** Supplementary table S1 to the published paper Ahn et al. 2023[23] lists the reasons why 124 patients failed entry screening for the DeLLphi-301 trial. There were not many patients in this list (n=21) that our clinical experts would expect to potentially treat in clinical practice (i.e. those excluded due to lack of consent, perceived inability to complete study activities, with another condition that might pose a risk to safety or study activities or with no tumour tissue sample provided). One expert commented that the majority of exclusions were likely to be those whose disease was progressing quickly i.e. poor performance status, deteriorating organ function and untreated or symptomatic brain metastases. Therefore it seems likely that the participants were representative of the entire population from which they were recruited.

Outcomes	Company (CS Table 10)		EAG
	Score	Definition	Comments
Were those patients who were prepared to participate representative of the entire population from which they were recruited?	1	Yes	**Disagree:** No data presented for patients who met the inclusion criteria but who declined to participate in the trial. Ahn et al. 2023[23] supplementary table S1 shows 8 patients did not provide consent prior to initiation of study activities.
Were the staff, places, and facilities where the patients were treated, representative of the treatment most patients receive?	1	Yes	**Disagree**: CS Table 4 shows the study took place across 56 centres in 17 countries. Two centres were in the United Kingdom which recruited five participants in total (clarification response A5). CS Table 6 shows 41.4% of participants were from Asia and 3% from North America where health care systems (staff, places and facilities) may be different from those in Europe and the UK specifically.
Internal validity - bias			
Was an attempt made to blind study patients to the intervention they have received?	0	No	**Agree:** Single-arm study with no blinding.
Was an attempt made to blind those measuring the main outcomes of the intervention?	0	No	**Unclear**: Although this was a single-arm study with no blinding, the primary analysis was based on the disease response assessment by BICR. However, we have concerns that blinding could have been compromised if the BICR body [redacted]. In their response to clarification question A8, the company did not address this point.
If any of the results of the study were based on "data dredging", was this made clear?	1	Yes	**Agree:** CS Table 4 lists the pre-planned subgroups and this matches the subgroups presented in CS Figure 8.
In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in	1	Yes	**Agree:** different lengths of follow-up adjusted for by survival analysis.

Outcomes	Company (CS Table 10)		EAG
	Score	Definition	Comments
case-control studies, is the time period between the intervention and outcome the same for cases and controls?			
Were the statistical tests used to assess the main outcomes appropriate?	1	Yes	**Agree:** As this was a single-arm study there was very limited statistical analysis. The statistical analysis appears appropriate.
Was compliance with the intervention/s reliable?	1	Yes	**Agree**: CS B.2.10.2 reports extent of exposure which shows the mean relative dose intensity was [redacted]% and the median relative dose intensity was [redacted]%.
Were the main outcome measures used accurate (valid and reliable)?	1	Yes	**Agree**: Outcome measures are clearly described and well-known measures.
Internal validity – confounding factors			
Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?	0	Unable to determine	**Agree:** single-arm trial.
Were study patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over	0	Unable to determine	**Agree:** single-arm trial.

Outcomes	Company (CS Table 10)		EAG
	Score	Definition	Comments
the same period of time?			
Were study patients randomised to intervention groups?	1	Yes	**Disagree:** although there was randomisation to two tarlatamab doses, only one of those doses is relevant and there is no comparator for that study arm.
Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was finished?	0	No	**Agree:** single-arm trial.
Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	0	Unable to determine	**Agree:** single-arm trial.
Were losses of patients to follow-up taken into account?	1	Yes	**Disagree**: In response to clarification question A16 the company state that no imputation was performed for efficacy endpoints. However, the extent and type of missing data is unclear as the company did not answer this part of the question. Imputation did occur when dates of adverse events or concomitant medication use were missing or incomplete. As stated above, there were no or minimal loss to follow-up on the outcomes of duration of response, duration of disease control, PFS and OS.
Power			
Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?	1	Yes	**Agree**: Sample size rationale provided in the Statistical Analysis plan with additional detail given in response to clarification question A18.
Final score	20		

Source: CS Table 10 supplemented with EAG comments.

BICR, blinded independent central review; CS company submission; CSR, clinical study report; OS, overall survival; PFS, progression-free survival; SCLC, small-cell lung cancer; TTD, time to treatment discontinuation.

Appendix 4

The EAG's risk of bias assessment of the UK CAS study that formed the external control arm of the MAIC is shown in Table 43.

Table 43 EAG risk of bias assessment of the UK CAS study

Criterion	EAG assessment of UK CAS study (Yes / No / Unclear)
Was the cohort recruited in an acceptable way?	**Unclear.** The company provided information in response to clarification question A24 about how the CAS Control Cohort patients were selected for the MAIC from the wider UK CAS study database. In the EAG's opinion the selection process is transparent and reasonable. As outlined in section 3.3.2 of this report, it is unclear whether the start of the diagnosis identification period of the study for patient inclusion was in [redacted] or 2013 due to inconsistency in reporting of these dates. This may be an error, but if it is not and the start date of the study was 2013 then it raises the possibility that not all patients with data available have been included in the cohort as data appear to have been available from [redacted] [2] Additionally, the CAS Control Cohort is not fully representative of the patients treated in practice, as [redacted] had received a prior PD-L1 inhibitor and the impact of this on the results is unknown (see section 3.3.2).
Was the exposure accurately measured to minimise bias?	**Yes**. CS Table 18 provides information about the proportions of patients in the CAS Control Cohort who were treated with topotecan, CAV and platinum + etoposide. It is unclear though which platinum treatment participants received. One of the EAG's experts commented that the distribution of the standard of care treatments was a reasonable representation of those used in clinical practice.
Was the outcome accurately measured to minimise bias?	**No**. TTD was used as a proxy for PFS due to data availability. There is a risk that this could under-estimate PFS in the standard of care comparator, as patients may discontinue treatment for reasons other than disease progression. The EAG considers the approach to estimating OS, however,

Criterion	EAG assessment of UK CAS study (Yes / No / Unclear)
	acceptable and we do not have concerns about bias in relation to this outcome.
Have the authors identified all important confounding factors?	**Yes**. Confounding factors were not controlled for in the UK CAS study. However, prognostic factors were identified and controlled for in the MAIC analyses. Our experts considered the company's selection of prognostic factors reasonable.
Have the authors taken account of the confounding factors in the design and/or analysis?	**Yes**. As stated above, prognostic factors were identified and controlled for in the company's MAIC.
Was the follow-up of patients complete?	**Yes**. The company allowed a period of at least 1.5 years for follow-up (clarification response A22), which appears sufficient for measuring the OS and PFS (using TTD as a proxy for PFS) outcomes in a SCLC population.
How precise (for example, in terms of confidence interval and p values) are the results?	The confidence intervals for median OS from the UK CAS study are not provided in the CS but were available from the MAIC report[2] and are narrow. Narrow confidence intervals are also reported for median PFS (using TTD as a proxy).

Source: EAG created table, using NICE-recommended criteria for the quality assessment of non-randomised and non-controlled studies.[28]
CAS, Cancer Analysis System; CAV, cyclophosphamide, doxorubicin and vincristine; EAG, External Assessment Group; MAIC, matching-adjusted indirect comparison; OS, overall survival; PD-L1, programmed cell death ligand 1; PFS, progression-free survival; SCLC, small-cell lung cancer; TTD, time-to-treatment discontinuation.