CONFIDENTIAL UNTIL PUBLISHED

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

Cemiplimab for treating cutaneous squamous cell carcinoma (CDF review TA592)

Produced by	**Southampton Health Technology Assessments Centre (SHTAC)**
Authors	Neelam Kalita, Senior Research Fellow, Health Economics Lois Woods, Senior Research Assistant, Evidence synthesis and Information Specialist Joanne Lord, Professor, Health Economics Geoff Frampton, Senior Research Fellow, Evidence Synthesis
Correspondence to	Dr Geoff Frampton Southampton Health Technology Assessments Centre (SHTAC) School of Healthcare Enterprise and Innovation Alpha House Enterprise Road, University of Southampton Science Park Southampton SO16 7NS www.southampton.ac.uk/shtac
Date completed	16th February 2022 (updated 3rd March 2022 following factual accuracy check)

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 135279.

Acknowledgements

We are grateful for clinical advice provided by Eugene Healy, Professor of Dermatology, University of Southampton and Honorary Consultant Dermatologist, University Hospital Southampton NHS Foundation Trust; Dr Kate Fife, Consultant Clinical Oncologist, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust; and Dr Agustin Martin Clavijo, Consultant Dermatologist, University Hospitals Birmingham NHS Foundation Trust. We also thank Jo Picot, SHTAC, for commenting on a draft of this report.

Declared competing interests of the authors

None

Declared competing interests of the clinical experts

Professor Healy is currently a collaborator in a study looking at genomic and transcriptomic factors that might associate with development of cutaneous squamous cell cancer metastases, funded by Sanofi (manufacturer of cemiplimab). Professor Healy declares that the only funds paid to the University of Southampton were costs of cutting tissue sections of Southampton samples. Dr Martin-Clavijo has received honoraria from Almirall (manufacturer of fluorouracil) for consultancy, chairing meetings and contributing to an educational steering committee in dermatology (primarily relating to inflammatory dermatosis) but has not been involved in any funded work on cutaneous squamous cell carcinoma or on fluorouracil. Dr Kate Fife received honoraria from Sanofi to present cases and experiences of patients who had received cemiplimab under the Cancer Drugs Fund, at a virtual meeting (February 2021); for training Sanofi employees in new national guidelines for squamous cell carcinoma (February 2021); to develop and present case studies for a virtual platform (April 2021); and to present patient cases and a summary of published trial data at an East of England Plastic Surgery Educational Meeting (September 2021). Dr Fife confirms that she did not participate in the company's cemiplimab trials, did not receive funding from the company specifically to work on cemiplimab, and did not contribute to the company's submission. Dr Fife also received honoraria from Pfizer (manufacturer of carboplatin, cisplatin and fluorouracil) for participating in a global advisory board (April 2021) which discussed renal cancer therapy.

Rider on responsibility for the report

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

This report should be referenced as follows:

Kalita N, Woods L, Lord J, Frampton G. Cemiplimab for treating cutaneous squamous cell carcinoma. A NICE Technology Appraisal. Southampton Health Technology Assessments Centre*, 2022.*

Confidential information

Please note that: content highlighted in yellow and underlined is 'academic in confidence' (AIC). Content underlined in blue is 'commercial in confidence' (CIC).

Word count: 26,675

1 EXECUTIVE SUMMARY

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

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

1.1 Critique of the company's adherence to the committee's preferred assumptions from the Terms of Engagement

The company have largely adhered to the committee's preferred assumptions but there are some major limitations in the evidence base which limit how well the committee's assumptions can be addressed, as summarised below.

1.2 Summary of the key issues in the clinical effectiveness evidence

The company have submitted several new sources of evidence for cemiplimab, chemotherapy and best supportive care (summarised in section 2.3) and conducted indirect treatment comparisons (ITCs) on these (discussed in sections 3.1.6 and 3.1.7). Key issues relating to these evidence sources and comparisons are:

SACT dataset

- The SACT (Systemic Anticancer Therapy) dataset which the company provided as new evidence is reflective of current clinical practice and indicates that an older, frailer population than that of the company cemiplimab trials can be treated with cemiplimab in practice, albeit with lower overall survival. However, relatively limited population characteristics are available for the SACT dataset, and generalisability of the findings may be influenced by the COVID-19 pandemic. The company use the SACT data only in a limited way to validate outcomes of the cemiplimab trials, as they did not digitise the SACT overall survival KM data for comparison against results from the company's trials and modelled extrapolations.

Comparator studies

- Three retrospective comparator studies are included in the submission representing chemotherapy (company UK chart review – new evidence, Jarkowski study – existing evidence) and best supportive care (Sun study – new evidence). These all have major limitations, and the company themselves regard the chart review as having "poor face-

validity", meaning that despite the study being conducted in the UK the population characteristics and results are highly uncertain. The Jarkowski and Sun studies are both very small (N≤20) and were conducted in the USA, therefore of questionable reliability and relevance to UK clinical practice. The population characteristics and results of these studies are therefore also highly uncertain.

- These limitations mean that none of the included studies provide a reliable estimate of the effects of chemotherapy or of BSC in a UK setting. Thus, uncertainties in these comparators have not been reduced relative to the information that was available prior to the CDF appraisal in TA592.

Indirect treatment comparisons

- In the absence of any trials directly comparing cemiplimab against chemotherapy or best supportive care the company conducted several ITC analyses to enable these comparisons. However, the ITCs are all limited by the high uncertainty in the population characteristics and results of the comparator studies that they included.
- Additionally, the ITC methods are all subject to uncertainty due to the inability of ITC models to balance measured prognostic covariates and/or a lack of sufficient data to enable all measured prognostic covariates to be modelled.
- One of the ITC approaches comparing cemiplimab against chemotherapy estimated the average treatment effect in the treated group (ATT) and the average treatment effect in the comparator group (ATC) which provides an opportunity to select which study represents the target population of interest (i.e. the company trials according to the ATT approach, or the chart review according to the ATC approach). Poor face validity of the chart review study makes it unclear whether the chart review reflects a UK clinical practice population and hence whether an ATC model would be more appropriate than the company's preferred approach which uses an ATT model. However, all the tested ATT and ATC models failed to adequately balance the prognostic covariates so results of all the models are at high risk of confounding.
- Hazard ratios obtained from the ITCs require the proportional hazards assumption to be satisfied. This assumption does not hold for the ITCs comparing cemiplimab against chemotherapy and appears unlikely to hold for the comparison of cemiplimab against BSC. Hazard ratios therefore cannot be relied upon to assist interpretation of relative treatment effects from the ITC results.

Summary

Whilst the company have largely adhered to the Terms of Engagement, the new evidence from comparator studies provided for this CDF review has not reduced the uncertainty in the effectiveness of cemiplimab as used in the UK compared to platinum-based chemotherapy and BSC. The longer-term data available from the EMPOWER-CSCC 1 trial have limited value in establishing relative effectiveness of cemiplimab since comparable long-term data do not exist for the comparator studies.

The areas where uncertainty has been reduced are:

- Improved confidence in the stopping rule and improved follow up of survival outcomes in the trial setting as a result of longer-term data being available in the EMPOWER-CSCC 1 trial.
- The SACT dataset suggests that the company cemiplimab trials lack generalisability to UK clinical practice. However, the SACT dataset has limitations due to relatively few population characteristics collected, whilst the overlap between the SACT dataset and COVID-19 pandemic could influence generalisability of the SACT data.

1.3 Summary of the key issues in the cost effectiveness evidence

- Extended trial follow up has provided better evidence for the cemiplimab PFS and OS extrapolations than was available for the TA592 analysis. It also supports the assumption of longer maintenance of survival with cemiplimab, with data now available for a maximum follow up of 5 years. The company assume loss of the relative treatment effect (hazards equal to those of comparators) at 5 years, extended from 3 years in the TA592 analysis. The ERG consider this to be reasonable.
- However, the SACT dataset has demonstrated that the patients treated with cemiplimab in UK practice were on average older and less fit than those in the company's trials. This suggests that the OS and PFS extrapolations based on the trial data that are used in the company's base case are likely to be more favourable than one would expect in routine NHS practice.
- The indirect treatment comparisons between the cemiplimab trials and comparator studies (company chart review, Jarkowski study and Sun study) inform the OS and PFS extrapolations for chemotherapy and best supportive care. However, as noted above, the results of these ITCs are all highly uncertain, meaning that there is significant uncertainty over the comparability of the survival extrapolations for cemiplimab and the two comparators. In particular, the evidence used in the model

for best supportive care was very sparse (20 immunocompetent patients from the US, Sun study cohort).

- For progression free survival, the company rely on different sources to model OS and PFS for chemotherapy, and they assume that all patients on best supportive care start in the 'post-progression' health state.
- The company's approach to selecting distributions for the survival extrapolations appears reasonable. However, the rationale for the choice of the base case distribution for OS with chemotherapy (fitted to the chart review data) is not clearly explained. The company noted that distributions with a better fit to the chart review data had a plateau in long-term survival, which the company's advisory board did not consider plausible. Similarly for best supportive care, the company chose a log-logistic distribution for OS, fitted to the Sun study data. This was not the best-fitting distribution but was selected based on clinical opinion.

1.4 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG conducted additional scenario analyses on the company's CDF model. The results of the company and ERG scenarios indicate that:

- The results are most sensitive to assumptions about: i) patient characteristics (age at treatment initiation); ii) waning of treatment effects; iii) OS & PFS extrapolations.
- Assumptions relating to efficacy and PFS adjustment also influenced the cost-effectiveness results, but to a lesser extent.
- The ICERs for the comparison with chemotherapy ranged between £33,195 (Gompertz distribution for PFS with chemotherapy) and £43,233 (Treatment waning at 42 months). For the comparison with best supportive care, the ICER ranged between £32,646 (mean age at baseline of 81 years, 80% male, based on the population in an Italian cemiplimab cohort[7]) and £28,859 (no population adjustment of indirect comparison).

Note that these scenarios do not capture the more fundamental uncertainties arising from the limitations of the comparative evidence described above.

1.5 Summary of ERG's preferred assumptions and resulting ICERs

We prefer the scenario based on the demographics of the SACT cohort, as this reflect the population treated with cemiplimab in UK NHS practice, see Table 1 below. This increases the ICER of cemiplimab versus chemotherapy by £1,612 (from the company's base case

ICER of £36,163 to £37,775) and that of cemiplimab versus BSC by £1,514 (from £29,438 in the company's base case to £30,952).

Table 1 ERG preferred assumption (PAS price)

Assumption	ICER vs PBC	ICER vs BSC
Company base case	£36,163	£29,438
+ Population characteristics from SACT database (age: 77 years; 74% male)	£37,775	£30,952
ERG preferred assumption	£37,775	£30,952
Abbreviations: BSC best supportive care; PBC platinum-based chemotherapy; O; ICER incremental cost-effectiveness ratio; PAS patient access scheme		

A range of scenarios was conducted on the ERG preferred assumption, which included varying assumptions on:

- Time to waning of treatment effects (e.g. at 4 years, and between 5 and 8 years)
- Using different models for population adjustment in the ITCs
- Extrapolating survival for cemiplimab and the comparators using different distributions (e.g. Weibull, Second order P(0,-1), Loglogistic, Gompertz, Lognormal)
- Adjusting the method for estimating PFS for the comparators (based on the relationship between PFS and OS in the Jarkowski cohort).

ERG scenarios around these assumptions showed that:

- The ICER for cemiplimab versus chemotherapy ranged between £44,379 and £33,942 and between £33,246 and £30,793 comparing cemiplimab versus BSC.
- The cost-effectiveness results are most sensitive to assumptions about OS extrapolations, treatment waning, and adjusting the PFS for the comparator arms.
- However, these analyses do not capture the underlying uncertainties related to generalisability of the trial data and weaknesses in the evidence base for the indirect comparisons.

2 INTRODUCTION AND BACKGROUND

2.1 Introduction

This report is a critique of the company's submission (CS) to the NICE Cancer Drugs Fund (CDF) review of TA592 on the clinical effectiveness and cost effectiveness of cemiplimab for treating cutaneous squamous cell carcinoma. Clarification on some aspects of the CS was requested on 24th January 2022. The company's response was received by the ERG on 4th February 2022 and a corrected version of the company's economic model was received on 8th February 2022.

The CS accurately reports the recommended use of cemiplimab within the CDF (CS section A1) and the licensed indication (CS section A4).

2.2 Background

Cemiplimab (Libtayo®) is a fully human immunoglobulin G4 (IgG4) monoclonal antibody that binds to the programmed death 1 (PD-1) receptor. It was granted a conditional marketing authorisation in July 2019 and a full marketing authorisation is anticipated in [redacted]. Cemiplimab is indicated as a monotherapy for adults with locally advanced or metastatic cutaneous squamous cell carcinoma (laCSCC or mCSCC) who are not candidates for curative surgery or curative radiation. The company acknowledge that the licensed dose according to the Summary of Product Characteristics is 350mg every three weeks via intravenous infusion over 30 minutes. Treatment with cemiplimab may be continued until disease progression or unacceptable toxicity. We note that the pivotal studies applied a stopping rule of 22 months, disease progression or death, whichever is sooner (CS section A.6.1.2).

In the original appraisal (TA592), NICE recommended cemiplimab for use within the Cancer Drugs Fund (CDF) according to the licensed indication and conditions in the Managed Access Agreement. The recommendation includes a 24-month stopping rule. TA592 concluded that the cemiplimab trial data were promising but uncertain. The data were considered immature as median overall survival had not been reached. The cost-effectiveness estimates were above what is normally considered a cost-effective use of NHS resources. Additionally, there was little evidence available on life expectancy with current treatments for advanced CSCC making it uncertain how long cemiplimab might prolong life and whether the end-of-life criteria apply. More mature data and more data on life

expectancy with current treatments may confirm the expectation that end-of-life criteria apply and that cemiplimab may be a cost-effective treatment.

In the previous appraisal TA592 the company's evidence for cemiplimab was obtained from the phase II EMPOWER CSCC-1 study (N=193) and the phase I study 1423 (N=26) which were pooled to give a single cohort (N=219). The TA592 NICE Committee agreed with the combining of these studies into a single cohort. In the current CDF review the company have also combined the two cemiplimab studies into one cohort. For brevity we refer to this pooled cohort throughout this report as the "company trials".

The relevant comparators in the current CDF review are chemotherapy and best supportive care (BSC). The company have limited their interpretation of chemotherapy specifically to mean platinum-based chemotherapy which is consistent with their submission in TA592. BSC is not clearly defined in oncological literature generally[1 2] and the CS states it is "where treatment options are palliative" (CS section A.6). For the purposes of this review we understand BSC to mean any treatment or care that is for managing symptoms and is without curative intent.

2.3 New evidence

The company's CDF review submission provides new evidence for clinical effectiveness from several sources. New evidence for cemiplimab is summarised in Table 2 below and discussed further in sections 3.1.2 and 3.1.3. New evidence for the comparators is listed below and discussed further in sections 3.1.5.1 and 3.1.5.3.

New evidence for the comparators:

1. **Company chart review:** a retrospective study of patient records conducted on behalf of the company by a third party provides OS evidence for platinum-based chemotherapy (PBC) for advanced CSCC in a UK population. Discussed in section 3.1.5.1 of this report.
2. **Sun et al. 2019 study [3]** a retrospective study of patient records identified by the ERG during the technical engagement phase of TA592 provides OS evidence for best supportive care (BSC). Discussed in section 3.1.5.3 of this report.

Table 2 New evidence for cemiplimab

Study		Phase I Study (Study 1423) N=26	Phase II EMPOWER-CSCC 1 (Study 1540) N=193	Systemic Anti-Cancer Therapy (SACT) dataset N=352
Overview		Additional data from a more recent data cut of the company phase I multicentre safety study	Additional data from a more recent data cut of the company phase II non-randomised, non-comparative, three-group multicentre study	Real-world data collected by Public Health England (PHE)/NHS Digital on UK patients treated with cemiplimab within the CDF during the two-year Managed Access Agreement period
Median follow-up, months (range)	**CDF review**	31.7 (1.1 to 47.0)	[redacted]	10.2 (6.3 to 26.3)
	TA592	11.1 (1.1 to 17.0)	8.6 (0.8 to 15.9)	Not applicable
Dosing [a]	**CDF review**	(n=26): weight-based dose (laCSCC & mCSCC)	Group 1 (n=59): weight-based dose (mCSCC) Group 2 (n=78): weight-based dose (laCSCC) Group 3 (n=56): flat dose (mCSCC) New evidence includes data from 56 new patients (23 in Group 2 and 33 in Group 3).	Not explicitly reported. The SACT Report refers to a 3-week treatment cycle which corresponds with the flat dose regimen.
	TA592	(n=26): weight-based dose (laCSCC & mCSCC)	Group 1 (n=59): weight-based dose (mCSCC) Group 2 (n=55): weight-based dose (laCSCC) Group 3 (n=23): flat dose (mCSCC)	Not applicable

Outcomes and data cuts (data in bold inform the economic analysis)	**CDF review**	**OS (2019)** [b] **PFS (2019)** [b] **Safety** Overall response rate Duration of response	**OS (July 2021)** **PFS (IRC assessed, July 2021)** **Safety (October 2020)** [c] **HRQoL (October 2020)** Overall response rate Duration of response Treatment duration	OS Treatment duration **Patient characteristics** (scenario analysis only)
	TA592	OS (October 2017) PFS (October 2017)	OS (October 2017) HRQoL (October 2017)	Not applicable
Pooled data provided by the company	**CDF review**	Combined results from both studies form a single cohort (N=219) referred to in this report as 'the company trials'.		Not applicable
	TA592	Combined results from both studies formed a single cohort (N=163)		

[a] Weight based dose: 3mg/kg IV every 2 weeks; flat dose: 350mg IV every 3 weeks.
[b] The company confirmed at the clarification meeting that the 2019 data cut was used for the analysis in this review (CS Table 3 erroneously reports a 2021 data cut)
[c] CS Table 3 says July 2021 but clarification question A28 says October 2020.
CDF: Cancer Drugs Fund; HRQoL: Health Related Quality of Life; IRC: independent review committee; IV: intravenous infusion; OS: overall survival; PFS: progression-free survival.
Sources: CS Table 3, CS section A.6.5, SACT Report, Clarification questions response, TA592 company submission.

2.4 Critique of the company's adherence to the Terms of Engagement

The ERG's critique of the company's adherence to the Terms of Engagement (ToE) is shown in Appendix 1 and a summary is provided in Table 3 below. Overall, the company have addressed the NICE Committee's preferred assumptions as stated in the ToE. However, the evidence for relative effectiveness of cemiplimab compared against chemotherapy and best supportive care remains highly uncertain, primarily because of methodological limitations with the comparator studies (discussed in sections 3.1.5.1 to 3.1.5.3 below) which produce uncertainty in the results of the indirect treatment comparisons (section 3.1.7 below). The company have considered SACT data in their submission but do not explicitly use the SACT results to validate survival extrapolations. Whilst the SACT data inform an economic scenario analysis, this only reflects the impact of SACT cohort demographics (age and gender) on general population mortality rates and utilities (section 4.2 below). SACT data reflect patients treated with cemiplimab during the COVID-19 pandemic which could influence generalisability of the SACT results (section 3.1.3 below).

Table 3 Summary of the company's adherence to the Terms of Engagement (ToE)

Terms of Engagement item	Addressed by company	ERG comments (for details see Appendix 1)
Population	Yes	The company have addressed the ToE, but there are uncertainties regarding data validity in the company's chart review.
Comparators	Yes	Note that chemotherapy is limited to platinum-based, which is consistent with TA592.
Generalisability of trial evidence	Yes, with limitations	The company have considered the SACT data. This indicated differences between the population treated with cemiplimab in NHS practice and the trial populations, although these data were collected during the COVID-19 pandemic which would likely influence generalisability.
Survival outcomes	Partly	Survival extrapolations based on updated trial data were explored but were not informed by data from the SACT dataset.
Comparator data	Yes, with limitations	The company have addressed the ToE, but the results of all comparator studies are uncertain due to methodological limitations.

Relative effectiveness	Yes, with limitations	The company have addressed the ToE, but results of all indirect treatment comparisons are uncertain because of the uncertainty in the comparator studies and the indirect comparison methods.
Treatment effect duration	Yes	The company have addressed the ToE: they have used the updated survival data from EMPOWER-CSCC 1 and explored the impact of a 24-month stopping rule on long-term outcomes.
End of life	Yes	The company argue that cemiplimab meets end-of-life criteria compared to both chemotherapy and BSC. However, their base case model indicates that the criteria are met for the comparison with BSC, but not for the comparison with chemotherapy (as the life expectancy exceeds 2 years). The ERG preferred scenario reiterates this conclusion. Overall, it remains unclear if cemiplimab meets end-of-life criteria due to high uncertainty in the comparator data. This issue warrants further discussion with clinical experts.

3 CLINICAL EFFECTIVENESS

3.1 Critique of new clinical evidence

3.1.1 Updated systematic review of clinical effectiveness evidence

The company performed an updated systematic literature review (SLR) on 17th July 2021 (CS section A.15.16), but the CS does not provide details. The company identified one potential additional study on cemiplimab[6] but excluded it for several reasons, including it being a non-UK study and having relatively high proportions of immunocompromised patients (24%) and those with ECOG PS ≥2 (27%). The ERG agree with the exclusion of this study.

Given the lack of details about the company's updated SLR, the ERG conducted a search to check whether any evidence might have been missed. We searched MEDLINE and Embase using the company's search strategies from the original TA592 submission without the study design filters and using date limits to cover the period since the latest search in November 2018. We found nine retrospective real-world studies for an advanced CSCC population including the French study identified by the company (Appendix 2). All were non-UK studies and we concluded that none of these would be eligible for inclusion in this CDF review, primarily because most studies did not report either OS or PFS, or because outcomes were not reported for population subgroups relevant to this review.

Details of the company's updated SLR were subsequently provided in a separate Systematic Review Technical Report (clarification response A1). The company carried out thorough searches for studies that assessed the efficacy of cemiplimab and all alternative interventions (not only PBC or BSC) for treatment of patients with advanced CSCC. They identified 42 new citations bringing the total number in their review to 66 citations representing 50 studies. Appendices B-E in the company SLR Report clearly detail study characteristics, patient characteristics and outcomes of the included studies and the ERG did not identify further studies relevant to this CDF review from them. We are satisfied that all relevant available published evidence is included in the review.

Following the literature searches described above, one of the ERG's clinical experts identified a recent conference abstract reporting a retrospective study of UK patients in the

UK Named Patient Scheme who had received cemiplimab before it was funded via the CDF.[9] We understand that the full paper has been submitted for publication. Information available in the study abstract is summarised in section 3.1.4 below.

3.1.2 Company trials

Overview

The strengths and limitations of the company trials were discussed in the original appraisal (see TA592 ERG report sections 3.1.3 to 3.1.6). The main concerns were that they were non-comparative and had relatively short follow-up with immature data. The new evidence included in this review is summarised in Table 2 above. It shows that the median follow-up has increased by 35.1 months (from [redacted] to [redacted] months) with a maximum follow-up of [redacted] months. Fifty-six additional patients were recruited to the phase II trial, increasing the pooled cohort to 219 patients.

Generalisability of the dosing regimens

The cemiplimab phase II trial contains three subgroups of patients (CS Table 5):
Groups 1 and 2 received the weight-based dose (3mg/kg IV every 2 weeks, total 137 patients) and group 3 received the flat dose which is the licensed indication (350/mg IV every 3 weeks, total 56 patients). CS section A.6.1.1 states that the results remain consistent between the three groups despite the different dosing regimens, but no evidence was presented for the ERG to verify this. In clarification response A2 the company provided objective response rates and KM curves for OS and PFS for each group from an October 2020 data cut because formal analyses of these subgroups were not yet available for the July 2021 data cut (the company have requested these but it is unclear whether they will be available within the timeframe of this CDF review). These data are relatively limited, without hazard ratios and confidence intervals for the survival data provided. However, we agree the results appear broadly consistent across the weight-based and flat dose groups.

Results

OS data are still immature (median survival has still not been reached) with the survival rate reported as [redacted] at a median follow up of [redacted] (maximum [redacted]) months (CS section A.6.1.2). The increased follow up is sufficient to analyse a 24-month stopping rule as specified in the Terms of Engagement. Median PFS was [redacted] (CS section A.6.1.3).

ERG conclusion

The company trials are the main source of cemiplimab treatment data used to inform the economic model in this CDF review. The additional data provide longer follow-up with a modest increase in the sample size. The new data, although limited, suggest that weight-based and flat cemiplimab doses have similar effects on OS, although information from the latest data cut would be desirable to confirm this.

3.1.3 SACT dataset

Overview

The SACT (Systematic Anti-Cancer Therapy) dataset is a cohort of 352 patients who received cemiplimab under the CDF from July 2019 to April 2021. It has a median follow-up of 10.2 months (range 6.3 to 26.3 months) for OS and 5.5 months (maximum 21.9 months) for treatment duration both of which are much shorter than the median follow-up for the company trials. The SACT dataset is reported in CS section A.6.5 and in a Public Health England report [4] (hereafter referred to as the SACT Report) that was provided in clarification response A3.

Eligibility criteria

Eligibility criteria in the SACT Report are consistent with those in the Managed Access Agreement.[10] There were 393 CDF applications during the review period and 41 were excluded (8 duplicate applications; 16 who received cemiplimab prior to the CDF; 12 died before treatment; 3 did not receive treatment; 2 were missing). The company did not apply any additional eligibility criteria to the SACT dataset. No reasons are given for why three people did not receive cemiplimab; one of the ERG's clinical experts suggested this may have been due to clinical deterioration. The SACT Report does not say whether any patients received cemiplimab other than as intended.

The Managed Access Agreement specifies an ECOG performance status of 0-1 in order to receive treatment with cemiplimab. However, 4% of patients in the SACT dataset had a status of ECOG PS 2 and 14% had no status recorded meaning up to 18% of patients might have had a performance status greater than 1 (we note that other recent real-world studies reported that between 20% and 27% of patients with an ECOG performance status greater than 1 had received cemiplimab, although none of these were UK studies[5-8]). Patient eligibility in the Managed Access Agreement allows for cemiplimab to be used with caution in immunosuppressed patients, and only 4% of patients in the SACT dataset were immunocompromised (SACT Report Table 5). Reasons for missing data for ECOG

performance status (14%) and (7%) are not provided so it is unclear whether these may have been related to the therapy or outcome.

Population characteristics

A limited set of population characteristics were collected for the SACT dataset: disease severity (laCSCC or mCSCC), median age, gender, and ECOG performance status. Median age was 77 years, compared to 72 years in the company trials. SACT also represents an older population than those of the comparator studies (Appendix 3). Based on the limited data available, the ERG's clinical experts considered the SACT dataset to be a good reflection of UK clinical practice and noted that the population could be considered frailer than that of company trials (NB the experts referred to frailty in a general sense, mainly reflecting the older population age; instruments that specifically assess frailty were not reported in the studies).

Influence of the COVID-19 pandemic

The company acknowledge (CS section A.6.5) that the COVID-19 pandemic, which started eight months after cemiplimab entry into the CDF, may have affected treatment with cemiplimab, and hence the SACT dataset. The ERG's clinical experts suggested the pandemic would have caused service disruption for several reasons, including delayed referrals when patients were unable to access GPs or other clinicians and the cancellation of all surgery, precluding patients who were candidates for surgery from receiving it. The pandemic would likely have impacted on clinical assessments, treatment options and outcomes, and there may have been extended dosing intervals or missed doses of cemiplimab. Contributory factors include staffing shortages in infusion centres, lack of transport to hospital if relatives/drivers were infected, fear of catching COVID-19 at the hospital or in transport, and clinician uncertainty about the effect of COVID-19 on CSCC patients, such as the risk of autoimmune side effects. The ERG's clinical experts suggested that during the pandemic patients presented with more advanced disease and progressed more, with one expert observing that the proportion of patients with laCSCC increased.

Generalisability of SACT

The company argue in CS section A.6.5 that patients in the SACT dataset may have had poorer PS than recorded, which we agree is plausible, albeit speculative. The company's and ERG's clinical experts concur that cemiplimab is used to treat an older population than that included in the company trials, although the ERG's clinical experts noted that cemiplimab may be less effective in older patients. The company consider that as the SACT

cohort has shorter follow up than the company trials, longer trial data would be preferable (CS section A.6.5). However, we note that median OS was reached in the SACT cohort and the value of the cohort (and its purpose as stated in the Terms of Engagement) is to reflect UK clinical practice rather than a trial setting. The company explored the impact of the SACT population characteristics (age and gender) as a scenario analysis in their economic model (CS section A.12).

Results

OS was defined as starting from treatment initiation in both the SACT dataset and the company trials. Median follow up for OS was 10.2 months (range 6.3 to 26.3 months) compared to [redacted] in the company trials. OS was 63% at 12 months, 56% at 18 months and 46% at 24 months which is lower than in the company trials. Median OS was 21 months (Table 4) whereas median OS was not reached in either of the company trials. The company argue that an older and frailer population in the SACT dataset may be an explanation for reaching median OS sooner. The ERG also note that the COVID-19 pandemic may have impacted negatively on OS, and some patients could have died of other causes, e.g. COVID-19.

Treatment duration in the SACT dataset was defined as the patients' median observed time from the start of their treatment to their last treatment date in SACT + prescription length. Median treatment duration for all patients was 8 months (95% CI 6.2 to 9.3 months).

Table 4 Survival estimates in the cemiplimab and comparator studies

Study	Median OS	Median PFS
Company trials (cemiplimab) (July 2021 data cut)	Median OS not reached. [redacted] OS at a median of [redacted] months (maximum [redacted] months) follow up. (CS section A.6.1.2)	[redacted] in the phase II trial (CS section A.6.1.3) [a]
SACT dataset (cemiplimab)	21 months [4] [b]	Not reported
Company chart review (chemotherapy)	[redacted] months" (CS section A.6.2.2) [b]	Not reported
Jarkowski study (chemotherapy)	15.1 months [11] [b]	9.8 months [11] [b]
Sun study (best supportive care)	5.0 months [3] (95% CI 2.6 to 14.4 months)	Not reported
CI: confidence interval		

[a] CS section A.6.1.3 reports median PFS as 18.4 months for the pooled studies whereas CS Table 7 reports median PFS as 22 months for each individual study. The reason for this discrepancy is unclear.
[b] Range and confidence interval not reported

ERG conclusion

The SACT dataset is representative of a UK population receiving treatment for advanced CSCC with cemiplimab, confirmed by expert clinical opinion. Patients in the SACT dataset are older, possibly frailer, than those of the company trials, reflecting that an older population can receive cemiplimab in clinical practice. Follow up in SACT was shorter than in the company trials, but median OS was reached (21 months). OS in the SACT population is lower than that of the company trials, likely reflecting the older (perhaps frailer) population and impact of the COVID-19 pandemic, although the extent to which these factors influenced OS is uncertain, since only four population characteristics are reported, limiting detailed interpretation.

3.1.4 Named Patient Scheme study

This study[9] was identified by the ERG (section 3.1.1) but not included in the CS. A summary of the information available in the abstract is provided here as the study is relevant to the scope of this CDF review. The full paper has been submitted for publication.

This was a retrospective study[9] of UK patients in the Named Patient Scheme who received cemiplimab for laCSCC or mCSCC prior to CDF funding. Forty-seven patients were enrolled across 17 centres. Nine patients progressed and were deemed unfit for treatment prior to starting cemiplimab, leaving a total study population of 38. Patients enrolled from November 2018 to July 2019 and the data cut is May 2020, with a median of 8 (range 1-24) treatment cycles and 8.7 (range 0.3 to 16.1) months of follow up. Patients were younger (median age 74 years) than those in the SACT dataset (median 77 years), a greater proportion of patients had metastatic or nodal disease, or both (Appendix 3) and 3/38 (8%) were immunocompromised. Median OS was 12.6 months (compared to 21 months in the SACT dataset); 60.5% of patients were alive at one year; median PFS was 7.7 months; and 34.2% of patients were continuing on cemiplimab at data cut off. The abstract reports that survival outcomes were significantly affected by disease stage and not by age, performance status or line of treatment.

ERG conclusion

The Named Patient Scheme study provides relevant context for the use of cemiplimab in a

UK clinical setting. However, limitations are that it is a small, retrospective study with a very short follow-up period, and only conference abstract details are currently available.

3.1.5 Comparator studies

Three studies were identified by the company as providing relevant comparators (i.e. chemotherapy or best supportive care for people with advanced CSCC. These are all retrospective chart reviews. The ERG and our clinical experts agree that these studies represent the most relevant available comparator data for the decision problem. Other related advanced CSCC cohort studies exist but are either not generalisable to UK practice or have other limitations (Appendix 2).

3.1.5.1 Company chart review: chemotherapy (OS)

Eligibility criteria

The company provided a protocol for their retrospective chart review (clarification response A9[a]). Eligibility criteria for the chart review are reported in section 3.1.2.1 of the Chart Review Report[12] and are consistent with those stated in the protocol. The eligibility criteria aimed to obtain a population of patient records with characteristics comparable to those of people enrolled in the company trials. We note two differences between the eligibility criteria of the chart review and those of the company trials:

- The proportions of laCSCC and mCSCC patients was not specified as an eligibility criterion for the company trials but a 60:40 balance of mCSCC to laCSCC patients was specified for the chart review (protocol page 9). The final ratio of mCSCC to laCSCC patients reported for the chart review (CS Table 23) therefore may not reflect the ratio of these groups seen in clinical practice.
- The chart review eligibility criteria do not specify any limits on the ECOG performance status of patients whereas those enrolled in the company trials had ECOG PS ≤1.

Data collection

To allow for potential (retrospectively observed) follow up of at least 24 months the chart review included patients whose diagnosis of laCSCC or mCSCC fell between 1st January 2011 and 31st December 2015. For the purposes of this appraisal the analysis was restricted to UK patients (N=106, from 25 centres). The data collection was contracted to a third-party

vendor, Medical Data Analytics (MDA) and the company had no direct contact with the study sites or de-identified patient data (clarification response A11).

After initial data collection the company raised concerns that several aspects of the data "did not align with clinical expectation" and were not fully comparable to the company trials (CS section A.15.6; section 3.1.2.2 and Appendix A in the Chart Review Report). The ERG critiqued the company's concerns, and our clinical experts commented that incomplete and ambiguous reporting of key information limits the usefulness of the chart review and impedes interpretation of the results (see Appendix 4).

The company provided the original MDA data collection forms in response to clarification question A7. However, the forms do not confirm whether investigators were expected to complete all fields in the forms, nor which IPD were finally collected. The IPD were not provided to the ERG, so we are unable to check validity of the summary results presented in the CS and Chart Review Report.[12]

Chart review audit

The company state that "*on the advice of clinical experts an audit of the UK cohort was conducted to investigate data quality*". Detailed methods of the audit (e.g. whether it was based on a protocol, and how many investigators collected and checked data) are not provided in the CS. The Chart Review Report[12] states that additional data elements were collected which included the reason why a patient was unresectable, confirmation of the date of metastatic disease, location of metastases, baseline biopsy date and location, confirmation of additional biopsies, and reasons for radiation at baseline. The audit aimed to clarify why laCSCC patients were deemed not to be candidates for surgery or radiation and why these patients had extended survival compared to other published estimates in the same population (Chart Review Report Appendix A1).

Site participation in the audit (which occurred during the COVID-19 pandemic) was voluntary. Of the original population, [redacted]/106 patients from [redacted]/25 of the centres were audited. The company integrated the data from these [redacted] audited patients into the original data set (i.e. N=106) to create "*enriched profiles*" of patients whose records were audited (CS section A.6.2.1). The company explain in clarification response A8 that "enriched profiles" were those in which the data points changed between the original chart review, and the original chart review data were replaced with the audit data as these were anticipated to be more accurate. The ERG requested clarification on which data points had changed

following the audit but the company's explanation is not specific: "*Additional data points around excisional biopsies, surgeries, radiotherapy, and systemic treatments both before and after their advanced CSCC diagnosis*" (clarification response A7).

Post-hoc eligibility criteria applied after audit

Following the audit the company observed residual differences between the chart review population and that of the company trials. They applied three further post hoc decision rules to identify those patients who received treatment in real-world clinical practice who would have been eligible for treatment with cemiplimab in the company trials. These decision rules are shown in CS Figure 3 and summarised in Table 5 below. These rules led to the exclusion of [redacted]/106 patients, leaving [redacted] patients (of which [redacted] were audited) in the final chart review analysis population.

Table 5 Post hoc decision rules for excluding patients from the chart review

Rule	Company exclusion rationale	ERG comments
1	[redacted]	The company's reasons for excluding these [redacted] patients appear broadly reasonable, but the ERG are unable to check consistency of the cited data since only summary statements of the IPD are provided. The ERG are unclear whether the [redacted] patients described would all have been eligible for BSC, since Chart Review Report section 3.1.2.2 states that for two of these patients the intent of radiotherapy was recorded as curative, whilst for some other patients the intent was not recorded.
2	[redacted]	The ERG and our clinical experts agree that these exclusions are appropriate as they are consistent with the eligibility criteria for the company trials.
3	[redacted]	

Baseline characteristics of the final analysis population ([redacted]) and the set of audited patient records ([redacted]) (Chart Review Report Table A1) suggest that the audited subset, despite

not being a random sample, is broadly reflective of the characteristics of the final analysis population. OS KM curves are reported for the [redacted] chart review patients, [redacted] chart review patients "enriched" with audit data for [redacted] of them, and for the [redacted] audited patients alone in Chart Review Figure A5. These curves mostly overlap, suggesting that these groups had similar OS.

Chart Review Report Appendix A1 states that the audit provided some clarification on the significant treatment data gaps for some patients but "*did not provide any additional clarification on the absence of PFS events or absence of prior treatments, whilst also raising questions about validity of reporting on treatment duration in the population*." The ERG note that higher than expected survival estimates observed for patients receiving systemic therapy (CS section A.15.6) were resolved following the audit process; however, the other face validity concerns were not resolved by the audit process (Appendix 4). Results reported for the chart review (see below) therefore remain uncertain.

Company and ERG assessments of study validity

The company and ERG did not conduct separate risk of bias assessments on the chart review using any published risk of bias tools. Limitations to the validity of the chart review study have been extensively discussed by the company and ERG as stated above and we deemed it unlikely that an additional assessment of the risk of bias would add new information, given that several key threats to validity were identified. These include: a risk of selection bias due to the retrospective data collection and application of post-hoc decision criteria; unavailability of the source IPD precluding validation of data by the ERG; and considerable uncertainty in the study methods due to lack of clarity in reporting.

The company acknowledge that following initial data collection there may have been patient selection bias (Chart Review Report Appendix A1). The company provided an extract of the study protocol, stating the intended process for checking data (clarification response A11), but did not provide information on the degree of protocol adherence or deviations, or on the number of reviewers who checked data or the rate of errors identified, either in the initial data collection or in the chart review audit.

Results of the chart review

Baseline characteristics of the chart review cohort post-audit [redacted]) are provided in Appendix 3 (this is an ERG- corrected version of CS Table 23). As the chart review eligibility criteria aimed to provide a cohort with baseline characteristics similar to those of the

company trials it would be expected that the chart review characteristics would be closer to those of the company trials than to those of the SACT dataset. This is true for median age (70 years compared to 72 years in the company trials and 77 years in SACT). But the chart review included a smaller proportion of patients with ECOG PS=0 ([redacted]) than the SACT dataset (18%) and company trials (45%). The chart review also included fewer patients than the company trials who had prior systemic therapy ([redacted] versus 34%) or prior radiotherapy ([redacted] versus 68%) but a greater proportion with undifferentiated tumours ([redacted] versus 66%) and T3-T4 stage tumours ([redacted] versus 32%) (SACT data are unavailable for these characteristics) (Appendix 3). We conclude that there is heterogeneity in baseline characteristics between the chart review and the company trials although it is unclear whether this is clinically meaningful given the issues with data reliability in the chart review discussed above and in Appendix 4.

CS Figure 68 compares the OS KM curve from the chart review (post-audit, [redacted]) against KM curves from the other studies that were included in the CDF review. Corresponding median OS estimates are summarised in Table 4 above. The median OS in the chart review is reported imprecisely as "~15 months", without confidence intervals. The company do not report median PFS since PFS events were not reliably recorded (Appendix 4). OS in the chart review is similar to that seen in the Jarkowski chemotherapy study (CS Figure 68 and Table 4 above). However, as with the baseline characteristics these chart review results are uncertain because of the key issues with face validity of the chart review.

ERG conclusion

The company have excluded BSC patients from the chart review and were unable to estimate PFS, meaning that the chart review serves as a comparator study for OS on platinum-based chemotherapy only. The exclusion of BSC patients and of the PFS outcome appear broadly appropriate but IPD were not provided and so the ERG have not been able to verify the summary data reported in the CS and Chart Review Report.[12] The company's chart review suffers from multiple issues of face validity, as acknowledged by the company and ERG's clinical experts, and results are at high risk of selection bias due to the retrospective data collection and post hoc application of eligibility criteria.

3.1.5.2 Jarkowski study: chemotherapy (OS and PFS)

Methods

This was a retrospective chart review of patients diagnosed with CSCC from January 2001 to January 2011 in the United States.[11] The study included 25 patients, of whom only 18 had

received relevant platinum-based chemotherapy and are relevant to the current appraisal. In their original CS for TA592 the company noted that treatment characteristics, such as dose and schedule, were not reported in the study so the company assume that the doses and schedules of systemic treatments in Jarkowski et al. 2016 were similar to those of other CSCC studies or clinical trials conducted in head and neck squamous cell carcinoma populations (TA592 CS Appendix D.1.3.2).

Company and ERG assessments of study validity

The company and ERG assessments of study validity are provided in section 3.1.4 of the TA592 ERG report. The ERG noted that although this study had an adequate duration of follow-up, it is very small (N=18) and at high risk of bias due to the retrospective selection of cases. In addition, the generalisability is unclear due to having a non-UK population, younger age, and higher proportion of trunk lesions than would be expected in NHS clinical practice, as well as limited reporting of baseline characteristics.

Results

Results of the study have been reported and discussed in the CS and ERG report for TA592 and remain unchanged. Patients in the Jarkowski study had a lower median age (66 years) than those in the company chemotherapy chart review (70 years) and company cemiplimab trials (72 years) but it is difficult to compare the studies on other characteristics due to the limited information reported and because locoregional and metastatic are not defined in the Jarkowski study and do not appear to equate to laCSCC and mCSCC as defined by the company (Appendix 3). Results of the ITCs including the Jarkowski study are reported in section 3.1.7 below.

ERG conclusion

Despite having major limitations, the Jarkowski study is the most relevant source of platinum-based chemotherapy data to inform the PFS outcome for this appraisal (company base case). The company also include OS data from the Jarkowski study in a scenario analysis.

3.1.5.3 Sun study: best supportive care (OS)

Methods

The Sun study[3] is summarised in CS section A.6.4. In summary, this was a retrospective chart review of patients who underwent surgical resection and postoperative radiotherapy for primary or recurrent CSCC of the head and neck between 1st January 1995 and 31st

December 2014 in the United States. The study included 72 patients, of whom 32 were immunocompetent and 40 were immunosuppressed. Among the immunocompetent patients, 20 had unresectable lesions that would be reflective of a BSC population relevant to the current appraisal. The remaining 52 patients (40 immunocompromised and 12 immunocompetent with resectable lesions) were excluded from analysis by the company which the ERG agree is appropriate. Baseline characteristics are only reported in the study publication for the 32 immunocompetent patients. The company assume that the baseline characteristics would be similar between immunocompetent patients irrespective of whether their lesions were resectable or unresectable; the ERG's clinical experts agreed this assumption is reasonable.

The Sun study excluded patients with distant metastatic disease at diagnosis, squamous cell carcinoma in situ alone, patients who had trunk or extremity tumours, and palliative doses of radiotherapy. As such, the study excludes some people who would be classified as receiving best supportive care, and thus provides a relatively narrow BSC population. We note that locoregional recurrence was defined in the Sun study as recurrence at the primary site, resection margin, or regional lymph nodes. Spread to regional lymph nodes would be classified as mCSCC according to the company's definitions of disease severity, so locoregional and distant as employed in the Sun study do not correspond to laCSCC and mCSCC as employed by the company. As noted in Appendix 3 below the ERG were unable to identify the source of laCSCC and mCSCC baseline characteristics reported in the CS for this study.

Company and ERG assessments of study validity

The CS does not provide an assessment of study validity. The Sun study has a similar design to the Jarkowski study, with similar limitations: small sample size (N=20), high risk of bias due to the retrospective selection of cases, and unclear generalisability due to having a non-UK population. Few baseline characteristics are reported in the Sun study that could be compared with the company trials (primarily age, gender, tumour location and T-stage) but these limited population characteristics are similar to those of the company trials (Appendix 3). As the key threats to validity are readily discernible the ERG did not consult published risk of bias tools to explore threats to validity in further detail.

Results

Aside from the median age (73 years) in the Sun study, which is comparable to that of the company cemiplimab trials (72 years), it is difficult to compare the baseline characteristics of

the studies because limited details were reported (Appendix 3). Results from the Sun study are only available for OS. The study paper reports that median survival of the 20 immunocompetent patients who had unresectable lesions was 5.0 months, whilst that of the 16 immunosuppressed patients with unresectable lesions was 3.9 months.[3] As would be expected, OS in the Sun study was lower than in the chemotherapy studies (Jarkowski study and the company chart review) (CS Table 68 and Table 4 above). However, results of these studies are uncertain due to the numerous limitations discussed above.

ERG conclusion

Despite having major limitations, the Sun study is the only source of BSC data for this appraisal (company base case).

3.1.6 Indirect treatment comparison methods

No studies directly comparing cemiplimab against chemotherapy or best supportive care are currently available. The company therefore employed indirect treatment comparison methods to compare the company cemiplimab trials against the three single-cohort comparator studies described above.

3.1.6.1 Identification of covariates for adjustment

Clarification responses A14 and A15 confirm that the targeted search for prognostic factors referred to in CS section A.7.1 is the same search that was reported in the previous appraisal TA592. The ERG identified a recent systematic review of prognostic factors for head and neck CSCC[13] and we also sought clinical expert opinion. We conclude that the company have identified the relevant prognostic factors as covariates for inclusion in their indirect treatment comparisons. However, as noted in the CS and Chart Review Report,[12] these variables were not consistently reported in the comparator studies, limiting the number of covariates that could be adjusted for in the analyses to a maximum of eight: median age; sex (% male); disease severity (% with laCSCC or mCSCC); tumour differentiation (% in each class); tumour location (% head/neck, trunk, or extremities); ECOG Performance Score (0, 1, 2); % who received prior systemic therapy; and tumour T-stage (% in each class) (Appendix 5).

For unanchored ITCs both prognostic factors and effect modifiers should be accounted for, [14] although the company considered that incorporating effect modifiers was not feasible due to the limited sample size and they acknowledge this as a limitation of the analyses (Chart Review Report section 5). We note that several of the included covariates are probable

effect modifiers (e.g. age, gender and performance status) and due to limitations of the available data it would not be feasible to include further covariates.

3.1.6.2 Summary of the indirect treatment comparison methods

The company employed four statistical approaches for adjusted indirect comparisons: inverse probability of treatment weighting (IPW), simulated treatment comparison (STC), matched-adjusted indirect comparison (MAIC), and multivariable regression (MVR) (described in more detail in sections 3.1.6.3 to 3.1.6.5 below). These are all forms of "unanchored" indirect comparison which can be used to estimate relative treatment effects by comparing single-arm studies. These statistical approaches aim, as far as is possible, to minimise bias in the measured outcomes (confounding) that results from imbalances in covariates between the studies under comparison.

The company also present naïve (unadjusted) comparisons of cemiplimab against chemotherapy and best supportive care alongside each of these analyses. Naïve comparisons are highly likely to produce biased outcomes, because imbalance in confounding covariates is not accounted for.[14] However, DSU Technical Support Document 18 recommends that results of naïve comparisons should be presented as supporting information alongside those of adjusted indirect comparisons.[14]

An overview of the ITC approaches employed by the company is provided in Table 6, showing which analyses inform the economic analysis. IPW and multivariable regression are appropriate where individual participant data (IPD) are available for both of the studies under comparison.[15] The STC and MAIC analyses are appropriate when IPD are available for one of the studies being compared and aggregate data are available for the comparator study.[14]

Table 6 Overview of the comparators and analysis approaches employed in the company's ITC analyses

Comparator study	Included in TA592?	Full IPD available? [a]	Analysis employed in the current CS	Outcomes analysed
Chemotherapy – company chart review [12]	No	Yes	Main analysis: IPW to estimate the average treatment effect in the treated (ATT). **This analysis informs the economic model** (CS Table 12). Scenario analysis: IPW to estimate the average treatment effect in the comparator (ATC). Scenario analysis: Multivariable regression (summarised in the Chart Review Report[12] but not discussed in the CS). Scenario analysis: Naïve comparison	Overall survival
Chemotherapy – Jarkowski study[11]	Yes	No	Main analysis: STC. [a] **This analysis informs the economic model** (CS Table 12). Scenario analysis: MAIC [a] Scenario analysis: Naïve comparison [a]	Overall survival Progression-free survival
Best supportive care – Sun study[3]	No	No	Main analysis: STC. **This analysis informs the economic model** (CS Table 12). Scenario analysis: MAIC Scenario analysis: Naïve comparison	Overall survival

IPD: individual participant data; IPW: inverse probability of treatment weighting; MAIC: matched-adjusted indirect comparison; STC: simulated treatment comparison
[a] This refers to participant-level data for the covariates, i.e. not including the limited IPD that were reconstructed by the company by digitising published KM curves in the Jarkowski and Sun studies to determine survival and censoring times (CS section A.15.9).
[b] This is the same as the approach employed in TA592.

The ERG agree with the company's overall strategy for indirect comparisons:

- The selection of the method (i.e. the choice of IPW, MVR, STC or MAIC) is broadly consistent with the approaches recommended in the DSU Technical Support Documents.[14 15]
- The company have attempted to include as many prognostic covariates as possible and have explored the statistical models that provide the best balance of covariates, acknowledging that analyses are limited by the availability of data on the covariates.

However, each of the ITC analyses has substantial limitations, which we discuss below for each analysis approach (sections 3.1.6.3 to 3.1.6.5).

The CS does not discuss the multivariable regression analysis when reporting results of the ITCs (CS Table 27). The best-fitting multivariable regression model (a marginal model based on backward selection of covariates) provided hazard ratios which according to their wide confidence intervals were not significantly different from 1.0 (Chart Review Report Table 5), although the company regard this as an exploratory analysis only due to model instability. The ERG believe that, due to limitations of the data, difficulty in selecting models, and the relatively wide range of hazard ratios produced by the models (Chart Review Report section 4.3) the MVR approach is less suitable than the IPW approach for indirect treatment comparisons in the present CDF review. We therefore do not discuss the MVR results further in this report.

3.1.6.3 Cemiplimab (company trials) versus chemotherapy (company chart review): IPW method

Summary of the statistical method

The company employed IPW based on propensity scores to improve the balance of covariates between the company trials for cemiplimab and the company chart review for chemotherapy) (CS section A.7.1). Propensity scores are defined as the predicted probability of treatment based on relevant covariates and were derived by the company using a logistic regression of treatment assignment (membership of the company trials or chart review) against a set of measured baseline covariates. These covariates were chosen because they were considered prognostically important (see section 3.1.6.1).

The IPW approach uses a patient's propensity score to generate a weight for each patient as summarised in section 3.2.1.2 of the company's Chart Review Report.[12] Using the IPW approach, patients from one study can be reweighted to match the baseline covariates of

those in the other study, thereby balancing the covariates between the studies to reduce the risk of baseline characteristics being confounded with the outcome (i.e. reduce the risk that effects on OS or PFS are explained by prognostic covariates rather than by cemiplimab or chemotherapy treatment). The reweighting approach was used by the company to estimate two different treatment effects:

- The **average treatment effect in the treated (ATT)**: Patients who received chemotherapy in the company chart review (i.e. the comparator population) were reweighted to match the baseline characteristics of those who received cemiplimab in the company trials (i.e. the treatment population). The ATT is an estimate of the treatment effect that would have been observed if the chart review patients had the same baseline characteristics as those enrolled in the company trials. This is the relevant analysis if the target population of interest is patients enrolled in the cemiplimab trials.
- The **average treatment effect in the comparator (ATC)**: Patients who received cemiplimab in the company trials (i.e. the treatment population) were reweighted to match the baseline characteristics of those who received chemotherapy in the company chart review (i.e. the comparator population). The ATC is an estimate of the treatment effect that would have been observed if the patients who were enrolled in the company trials had the same baseline characteristics as those in the company chart review. This is the relevant analysis if the target population of interest is patients in the real-world UK clinical practice (assuming that the company chart review population reflects that seen in UK clinical practice).

The company explored ATT and ATC models with varying inclusion of covariates (Appendix 5). The “full model” was designated the ATT or ATC model that incorporated the full set of available covariates. Ten further ATT and ATC models incorporating different combinations of the covariates were run, numbered sequentially 1-10 according to their statistical fit, with ATT model 1 having a better fit than ATT model 2 and so on (see ‘Assessment of model fit’ below). In the CS and Chart Review Report[12] the company focus mainly on the full ATT model and ATT model 1, and the full ATC model and ATC model 1 which the ERG agree is appropriate since these optimise both the incorporation of covariates and the statistical fit. To reduce the influence of extreme weights which lead to poor covariate balance, the company ran each analysis again, with trimmed weights capped at the 95th percentile (“trimmed analysis”), and explored whether this improved the model fit.

Assessment of model fit

The aim was to select the propensity score model with weights that resulted in the best balance of the key relevant covariates. The main parameters used for this judgement are the balance statistic (higher values indicating greater homogeneity of the covariates after reweighting); effective sample size (ESS; higher values are preferable); the proportion of covariates which had a low (<10%) or high (≥10%) absolute standardised difference (ASD) between the studies after reweighting (Appendix 6); and histograms of the distributions of patient weights (Chart Review Report Figures 2-3, C10-C18, D1-D2, D14-D22).

Appropriateness of the target population

The company preferred to use the ATT approach for their "base case" IPW analysis and the ATC approach in sensitivity analyses (CS section A.7.1). The ERG believe that the target population should be patients in a real-world clinical setting, so the ATC analysis would be logical as the base case analysis (provided that the company chart review population can be assumed to reflect that seen in UK clinical practice).

Appropriateness of the statistical models and assumptions

- The ATT and ATC analyses are standard statistical analysis approaches derived from causality theory and the company's overall approach to the IPW analyses is consistent with TSD 17 guidance.[15]
- Calculation of hazard ratios requires that the assumption of proportional hazards (PH) is satisfied. The company present tests of the PH assumption in Chart Review Report Appendix E. Both the company (clarification responses A22-A25) and ERG agree that these tests suggest the PH assumption is violated for most if not all ITC comparisons (some subjectivity of interpretation is inevitable). The hazard ratios presented in the CS and ERG report are therefore uncertain and should not be used to infer relative treatment effects.
- The company did not model time-varying hazard ratios, for three reasons stated in clarification response A22 which the ERG agree are reasonable.
- Instead of using hazard ratios to provide relative treatment effects for the economic model, the model is informed by the separate IPW-adjusted KM curves which do not assume proportional hazards (clarification response A24). This is consistent with the approach employed in TA592.

Uncertainties in the ITC methods

Results of the IPW analyses comparing cemiplimab to chemotherapy are uncertain because:

- The company chart review population had a retrospective design with, post-hoc data selection, and has poor face-validity (section 3.1.5.1).
- IPW adjustment was not fully successful at balancing all the covariates (all analysis models had at least two covariates with a standardised absolute difference > 10% after reweighting: see Appendix 6).
- A maximum of eight prognostic factors could be included as covariates due to limited details being reported in the studies; prior radiation therapy was not included as a covariate in any analysis models (Appendix 5).
- HRs should be interpreted with caution due to lack of support for the proportional hazards assumption.
- The ERG do not have access to IPD for the company trials and company chart review and therefore cannot verify that the analyses were conducted as stated.

3.1.6.4 Cemiplimab versus chemotherapy (Jarkowski study): STC & MAIC

Summary of the statistical method

The company confirmed (clarification response A19) that the statistical methods applied for STC and MAIC to compare the company trials to the Jarkowski study were identical to those employed in TA592. The ERG agree that the company's rationale for selecting STC as the main analysis, with the MAIC and naïve analyses as scenarios is appropriate (TA592 ERG Report section 3.1.7.4). The company explored two models: a core model which incorporated two covariates (disease stage and tumour location) and an extended model which incorporated four covariates (disease stage, tumour location, gender and prior systemic therapy) (Tables 5 & 6 in clarification response A18). The company selected the core model as it had the better fit, but this was based on only a marginally lower Akaike Information Criterion (AIC) value for the STC analysis ([redacted] versus [redacted]) (Clarification response A18). The company report the OS and PFS curves produced by the STC and MAIC analyses (CS Figures 9 and 10) and corresponding hazard ratios (CS section A.7.2) but these are only for the core model. Given the closeness of the AIC values, and the larger number of covariates included in the extended model, the ERG suggest that STC and MAIC results for the extended model should also be provided.

Appropriateness of the target population

STC analysis simulates adding a "missing" trial arm such that outcome predictions are made for the company trial population using the mean characteristics of the Jarkowski study population. In the MAIC analysis the Jarkowski study is modelled as the target population

(i.e. company trial IPD are reweighted to match those of the Jarkowski study). Appropriateness of the target population is therefore contingent on the Jarkowski study population being reflective of that seen UK clinical practice which (as noted in TA592) is questionable given that the study was conducted in the USA. (NB it is difficult to compare the Jarkowski study with the UK SACT dataset to clarify its UK relevance due to the limited SACT population characteristics reported; Appendix 3).

Uncertainties in the ITC methods

Results of the STC and MAIC analyses comparing cemiplimab to chemotherapy using the Jarkowski study are uncertain because:

- The Jarkowski study has several limitations including retrospective design, small sample size and being a non-UK study (section 3.1.5.2).
- Only two covariates could be included in the core STC and MAIC model due to limited details being reported in the studies. Results of the extended model, which included four covariates, are not provided despite a similar model fit.
- As noted in the IPW analyses (section 3.1.6.3) the proportional hazards assumption is not supported for comparisons of cemiplimab against platinum-based chemotherapy and therefore hazard ratios would be unreliable for estimating relative treatment effects.
- The ERG do not have access to IPD for the company trials and therefore cannot verify that the analyses were conducted as stated.

3.1.6.5 Cemiplimab versus best supportive care (Sun study): STC & MAIC

Summary of the statistical method

The comparison of the company trials against the Sun study followed the same approach using STC and MAIC as for the comparison against the Jarkowski study described above (section 3.1.6.4). A core model incorporated four covariates (age, disease stage, tumour location and tumour stage) and an extended model incorporated a further three covariates (gender, ECOG performance score and prior radiation therapy) (Table 7 in clarification response A18). The CS does not discuss whether any intermediate models incorporating other combinations of covariates could have been developed. The AIC values favoured the extended model over the core model ([redacted] and [redacted] respectively) (clarification response A18). The company report the OS curve produced by the STC and MAIC analyses (CS Figure 11) and corresponding hazard ratios (CS section A.7.3) for the extended model

only. The ERG believe this is acceptable given that the core model included only four covariates, without improved model fit.

Appropriateness of the target population

As discussed above in section 3.1.6.4, appropriateness of the target population is contingent on the comparator study population, i.e. in this case the Sun study, being reflective of that seen in UK clinical practice. This is questionable given that the study was conducted in the USA. (NB it is difficult to compare the Sun study with the UK SACT dataset to clarify its UK relevance due to the limited SACT population characteristics reported; Appendix 3).

Uncertainties in the ITC methods

Results of the STC and MAIC analyses comparing cemiplimab to BSC are uncertain because:

- The Sun study population has several limitations including retrospective design, small sample size and being a non-UK study (section 3.1.5.3).
- As noted in Appendix 3 the ERG were unable to identify the source of disease severity (laCSCC and mCSCC) baseline characteristics provided by the company for this study.
- The proportional hazards assumption does not appear to be supported for the comparison of cemiplimab against BSC (CS Figure 25), although the company do not discuss this explicitly. Hazard ratios for this comparison therefore may be unreliable.
- The ERG do not have access to IPD for the company trials and therefore cannot verify that the analyses were conducted as stated.

3.1.7 Indirect treatment comparison results

As discussed above (sections 3.1.6.3 to 3.1.6.5) the proportional hazards assumption is not satisfied for most if not all ITC analyses, meaning that hazard ratios describing the relative treatment effects from the ITCs will be unreliable and should be interpreted with caution. The primary output from the ITCs which inform the economic model are the ITC-adjusted/weighted KM curves for OS and PFS which do not assume proportional hazards.

Cemiplimab versus chemotherapy: IPW and naïve analyses (OS)

The company report a range of KM curves, hazard ratios and model fit parameters for the eleven ATT models and the eleven ATC models, for trimmed and untrimmed analyses as summarised in Appendix 6. The main models likely to be of interest for the economic analysis are the full ATT model (Chart Review Report Figure 4), ATT model 1 (Chart Review

Report Figure 5), full ATC model (Chart Review Report Figure D3) and ATC model 1 (Chart Review Report Figure D4) since these optimise model fit and inclusion of covariates. The KM curves for all models indicate OS is higher with cemiplimab than with platinum-based chemotherapy; the models primarily differ in the degree of overlap of the confidence intervals for the curves.

For their base case economic analyses the company preferred ATT model 1 (see Table 7). This model included 5 covariates and appears to have a slightly better fit than the full ATT model which included 7 covariates (ATT model 1 has higher ESS and fewer reweighted covariates with ASD >10%, but a lower balance statistic – illustrating that decisions on model fit can be somewhat subjective) (Appendix 6). CS Figure 8 shows the weighted KM curve for chemotherapy (company chart review) compared against the KM curve for cemiplimab for ATT model 1. Both the trimmed and untrimmed analyses demonstrate that patients receiving chemotherapy in the company chart review had lower OS than those receiving cemiplimab in the company trials, as expected.

If the company chart review reflects a real-world UK clinical practice cohort it may be more appropriate to treat this as the target population of interest, i.e. using the ATC analysis approach. As shown in Appendix 6, it is difficult to separate the full ATC model and ATC model 1 based on the balance statistic, ESS, number of unbalanced covariates remaining after reweighting (trimmed analysis) or on the distribution of weights (Chart Review Figures D1 and D2) which were broadly similar for both models. The full ATC model has the advantage that it incorporates eight covariates whereas ATC model 1 incorporates seven. We note that the full ATC model incorporates one more covariate than the company's preferred ATT model 1 model and the histograms of weights are suggestive of a marginally better balance for the full ATC model than ATT model 1 (compare Chart Review Figures 1-2 versus D1-D2). However, it is important to stress that all ATT and ATC models had at least 2 covariates with absolute standardised exceeding 10% in trimmed analyses, indicating that none of the models was fully successful at balancing the covariates, meaning that for all models there is a residual risk of confounding.

The figures in the Chart Review Report which present the results of the ATT and ATC analyses also include naïve comparisons (i.e. unadjusted curves are included within the figures). Comparisons of the unadjusted cemiplimab and chemotherapy curves provide a similar interpretation to those of the adjusted curves, in all cases clearly showing OS to be higher with cemiplimab than with platinum-based chemotherapy. Results of naïve

comparisons are also presented as hazard ratios alongside those of the ATT and ATC model results in Chart Review Report Table 5. There is general overlap of the hazard ratios across the models, and the hazard ratios tend to be slightly higher for the trimmed analyses (Appendix 6) but these results should be interpreted with caution as the proportional hazards assumption is not supported.

ERG conclusion

Results of the IPW-adjusted ITC analyses are highly uncertain because the comparator study lacks face-validity (section 3.1.5.1) and none of the IPW models were fully successful at balancing all covariates (section 3.1.6.3). Hazard ratios do not assist interpretation since the proportional hazards assumption is not supported. It is unclear conceptually whether an ATT or ATC model would be most appropriate, since the extent to which the company chart review is reflective of UK clinical practice (external validity) is uncertain (the chart review aimed to collect data relevant to clinical practice, but the eligibility criteria for the chart review also aimed to match the population characteristics of the company trials). A critical consideration to enable causal inference is that the ITC outcomes should be free from confounding (high internal validity) and therefore IPW models should be selected which successfully balance the covariates. Given that none of the ATT and ATC models achieved this (Appendix 6) it is inadvisable to apply causal inference to these ITC results.

Cemiplimab versus chemotherapy: STC, MAIC and naïve analyses (OS and PFS)

The company preferred the core STC model to inform their economic analysis, but this incorporates only two covariates (disease stage [laCSCC or mCSCC] and tumour location [head and neck versus other]) (Clarification response Table 5). Results are only reported for the core model.

CS Figure 9 (OS) and CS Figure 10 (PFS) show the predicted cemiplimab KM curves from the STC and MAIC analyses for the core model, compared to the chemotherapy curve from the Jarkowski study. Both OS and PFS are higher for cemiplimab than for platinum-based chemotherapy, although for PFS the tails of the cemiplimab and chemotherapy curves overlap after 36 months where numbers at risk are small. Confidence intervals are missing from the KM curves so there is no indication of the uncertainty.

As noted above (section 3.1.6.4), the ERG believe OS and PFS results of the extended model which incorporates four covariates should also be presented for the comparison, given the similar model fit

The figures which present the results of the MAIC and STC analyses (CS Figure 9 for OS and CS Figure 10 for PFS) also include naïve comparisons (i.e. unadjusted curves are included within the figures). Comparisons of the unadjusted cemiplimab and chemotherapy curves provide a similar interpretation to those of the adjusted curves, in all cases clearly showing OS to be higher with cemiplimab than with chemotherapy. Hazard ratios from a naïve (unadjusted) comparison of the company trials against the Jarkowski study are compared against those from the STC and MAIC in CS section A.7.2. There is general overlap of the hazard ratios although these results should be interpreted with caution as the proportional hazards assumption is not supported.

ERG conclusion

Results of this ITC analysis are highly uncertain because of limitations in the comparator study (section 3.1.5.2), use of a suboptimal model that incorporates only two covariates (section 3.1.6.4) and absence of confidence intervals for the KM curves so the uncertainty is not displayed. The ERG suggest that the extended model results should have been provided alongside those of the core model, with confidence intervals provided for all KM curves.

Cemiplimab versus BSC: STC, MAIC and naïve analyses (OS)

The company preferred the extended model for STC to inform their economic analysis. The extended model incorporates 7 covariates and appears to have a better fit (lower AIC) than the core model (which incorporates 4 covariates) (clarification response Table 7).

CS Figure 11 shows the predicted cemiplimab KM curves from the STC and MAIC analyses for the extended model, compared to the best supportive care curve from the Sun study. There is some disagreement between the STC and MAIC curves, with the STC more closely matching the observed cemiplimab data. Both curves clearly demonstrate higher OS with cemiplimab than with best supportive care. However, the company note that the extended model resulted in a substantially reduced ESS of [redacted], indicating an overall poor model fit (clarification response A21).

A visual naïve comparison of the unadjusted KM curves for cemiplimab and chemotherapy is provided in CS Figure 11, which is consistent with the results of the STC and MAIC analyses. Hazard ratios from a naïve (unadjusted) comparison of the company trials against the Sun study are compared against those from the STC and MAIC in CS section A.7.3. There is general overlap of the hazard ratios but these results should be interpreted with

caution as it appears unlikely that the proportional hazards assumption is supported for comparison of cemiplimab against BSC.

ERG conclusion

Results of this ITC analysis are highly uncertain because of limitations in the comparator study (section 3.1.5.3), poor model fit, and absence of confidence intervals for the KM curves so the uncertainty is not displayed. The ERG suggest that confidence intervals should be provided for all KM curves.

3.2 Safety

The Terms of Engagement do not specify safety monitoring. The ERG requested an update on adverse events given that longer follow-up is now available in the company trials. A summary of adverse events up to an October 2020 data cut are provided in Table 9 of the clarification response document since formal statistical analysis of safety data has not been performed for the July 2021 data cut (clarification response A28). The October 2020 data do not identify any new safety concerns and demonstrate comparable safety between the weight-based and flat dose groups within the EMPOWER CSCC-1 study (as stated in clarification response A27). The company confirmed that safety data were not collected in the SACT dataset (clarification response A29).

3.3 Additional work on clinical effectiveness undertaken by the ERG

The company provided details of their updated systematic literature review on 4th February 2022 (clarification response A1). Meanwhile, the ERG carried out brief searches to identify whether any new evidence published since the original appraisal was missed. See section 3.1.1 of this report for details.

3.4 Conclusions of the clinical effectiveness section

New evidence

- The company have identified all relevant studies. New evidence is available from 4 sources: an updated data cut in the company trials (cemiplimab), SACT dataset (cemiplimab), company chart review (chemotherapy), and Sun study (best supportive care). Existing data from a previous study included in the TA592 appraisal (Jarkowski, chemotherapy) were also used in a scenario analysis.

SACT dataset

- The ERG's clinical experts confirmed that the SACT dataset is reflective of current clinical practice and is therefore suitable as a benchmark against which to assess the external validity of the company clinical trials. However, patient behaviour and clinical practice represented in the SACT dataset are likely to reflect the impact of the covid-19 pandemic.
- There are differences between the company trials and SACT dataset. The SACT dataset reflects that an older, frailer population with comorbidities such as autoimmunity can be treated with cemiplimab in practice. Overall survival in the SACT dataset is lower than in the company trials, likely reflecting the younger, fitter population enrolled in the company trials.

Comparator studies

- The three comparator studies (company chart review, Jarkowski study, Sun study) all have major limitations. The company chart review contains data and assumptions which the ERG's three clinical experts considered clinically implausible, as well as missing data, and the company themselves regard the chart review as having "poor face-validity". The company, ERG and ERG's clinical experts concur that the population characteristics and results of the chart review are highly uncertain.
- The Jarkowski and Sun studies are both small (N≤20), retrospective, and conducted in the USA therefore of questionable reliability and relevance to UK clinical practice. The population characteristics and results of these studies are therefore also highly uncertain.
- These limitations mean that none of the included studies provide a reliable estimate of the effects of chemotherapy or of BSC in a UK setting. Thus, uncertainties in these comparators have not been reduced relative to the pre-CDF appraisal TA592.

Indirect treatment comparisons

- The company used indirect treatment comparisons to compare the company trials (cemiplimab) against the comparator studies, i.e. the company chart review (chemotherapy, OS), the Jarkowski study (chemotherapy, OS and PFS), and the Sun study (best supportive care, OS). Three methods of indirect treatment comparison were employed (IPW approach, STC and MAIC) which are appropriate for the types of data available.

- The ITC analyses are all limited by the high uncertainty in the population characteristics and results of the comparator studies that they included, rendering the results of the ITC analyses themselves highly uncertain.
- Additionally, the ITC methods are subject to uncertainty, primarily due to the inability of ITC models to balance all measured prognostic covariates (IPW approach), and lack of sufficient data to enable sufficient prognostic covariates to be modelled (STC and MAIC approaches).
- The IPW approach estimated the average treatment effect in the treated (ATT) and the average treatment effect in the comparator (ATC) which provides an opportunity to select which study represents the target population of interest (i.e. the company trials according to the ATT approach, or the chart review according to the ATC approach). Unfortunately, the poor face validity of the chart review study makes it unclear whether the chart review reflects a UK clinical practice population and hence whether an ATC model would be more appropriate than the company's preferred approach which uses an ATT model. In practice, however, all models failed to adequately balance the prognostic covariates so their results are at high risk of confounding.
- Hazard ratios obtained from the ITCs require the proportional hazards assumption to be satisfied. This assumption does not hold for the ITCs comparing cemiplimab against chemotherapy and appears unlikely to hold for the comparison of cemiplimab against BSC. Hazard ratios therefore cannot be relied upon to assist interpretation of the ITC results, which is primarily limited to the visual inspection of KM curves.

Summary

Whilst the company have largely adhered to the Terms of Engagement, the new evidence from comparator studies provided for this CDF review has not reduced the uncertainty in the effectiveness of cemiplimab as used in the UK compared to platinum-based chemotherapy and BSC. The longer-term data available from the EMPOWER-CSCC 1 trial have limited value in establishing relative effectiveness of cemiplimab since comparable long-term data do not exist for the comparator studies.

The areas where uncertainty has been reduced are:

- Improved confidence in the stopping rule and improved follow up of survival outcomes in the trial setting as a result of longer-term data being available in the EMPOWER-CSCC 1 trial.

- The SACT dataset has also helped to establish that the company cemiplimab trials lack generalisability to UK clinical practice. However, the SACT dataset has limitations due to relatively few population characteristics collected, whilst the overlap between the SACT dataset and COVID-19 pandemic could influence generalisability of the SACT data.

4 COST EFFECTIVENESS

4.1 Model structure

In response to clarification question B1, the company submitted a revised version of their CDF review model capable of replicating the ICERs used in the committee's decision making at the point of CDF entry. All discussion and results reported below relates to this revised CDF review model (version 8 submitted 8 February 2022).

The model has a partitioned survival structure with 3 health states: pre-progression, post-progression and death, which the TA592 committee considered acceptable. This structure has not changed for the CDF review. The company have made some minor corrections and changes to model assumptions and parameters, listed in Table 7 below. We critique these changes the in the following sections of this report.

Table 7 List of changes to the company model for the CDF review

Change to model	Location in submission	ERG discussion
Population baseline characteristics		
Mean age (71.2 years) and gender (83.1% male) from 2021 trial data cut (includes flat dose group)	CS Table 35	4.2 below
Overall survival extrapolations		
Cemiplimab: company trial data updated to July 21 (no change to log-normal survival function)	CS A.8.3.1 and Table 15	4.5.2 below
Chemotherapy: UK Chart Review [12], ATT model 1 trimmed, log-logistic survival function	CS A.8.3.2, Table 15 and Clarification Response B2	
BSC: Sun et al. 2019 [3], STC analysis with log-logistic survival function	CS A.8.3.3 and Table 15	
General population mortality cap: updated to 2018-2020 life tables, with gender-specific population	CS A.15.13 and model	
Progression free survival extrapolations		
Cemiplimab: company trial data updated to July 2021, fractional polynomial (p1 = 0 p2 = -1) survival function	CS A.8.4.1 and Table 15	4.5.3 below
Chemotherapy: no change to data source (Jarkowski et al. 2016)[11] or survival function (Weibull).	CS A.8.4.2 and Table 15	
BSC: patients start in post-progression state	CS A.8.4.3 and Table 15	
Waning of treatment effect on OS and PFS		
Duration of cemiplimab relative effects extended to 60 months. (No change to 2-year stopping rule).	CS A.8.5	4.5.4 below
Adverse event rates		

Cemiplimab rates updated to July 2021 trial data.	CS Table 34	4.5.4 below
Exclusion of adverse events with <5% incidence	CS A.15.13	
Utilities		
Updated EORTC QLQ-C30 from company trials with October 2020 data cut (no change to mapping)[16]	CS Table 33	4.7 below
Correction to cap for age-related utility decrement for PFS health state, and inclusion of multiplicative option.	CS A.15.13	
Resource use and costs		
Cemiplimab PAS price discount [redacted] per 350 mg vial).	CS Table 2	4.8 below
Unit costs updated: 2021 eMIT, 2019/20 NHS Reference Costs, 2020 PSSRU and inflation (NHSCII index)[17-19]	CS A.15.12	

4.2 Population

The modelled cohort is based on the population in the cemiplimab trials. The company revised the baseline patient characteristics in their base case to reflect the dataset in the CDF review, which includes an additional patient group allocated to a flat dose of cemiplimab in EMPOWER (CS section A.6.1.1). This increased the mean age of the modelled cohort from 70.44 years in TA592 to 71.16 years in the CDF review (CS Table 35), which causes a small increase in the ICERs.

The model uses separate sources for survival outcomes with cemiplimab (company trials), chemotherapy (chart review and Jarkowski study) and BSC (Sun study), which is a potential source of bias. The company attempt to adjust for population differences in their ITC analyses but results of the ITCs are highly uncertain due to limitations of the comparator studies and residual imbalances in prognostic factors (see discussion in sections 3.1.6 and 3.1.7 above).

The committee noted that the modelled cohort in TA592 (based on the cemiplimab trials) did not completely represent patients expected to have cemiplimab in UK clinical practice. The company state that baseline characteristics in the CDF review model are thought to be generalisable to the UK patient population "*as demonstrated in the SACT dataset and the chart review study*" (CS section A.14.3). But they go on to note differences between the trial and real world populations: the latter being generally older and frailer, with more prior systemic therapy and autoimmune comorbidities.

For the base case comparison with chemotherapy, the company use an ATT model: adjusting survival with chemotherapy from the chart review to reflect the population in the cemiplimab trials. They also present a scenario with an ATC model: adjusting the cemiplimab trial data to reflect the population in the chart review (CS Table 20). This raises the question of which approach best reflects outcomes in UK practice (see section 3.1.7.1).

There are particularly notable differences between the patients in the cemiplimab trials and those treated with cemiplimab in the SACT dataset (CS Table 23).[4] The company report a scenario with SACT demographics (median age 77 years and 74% male), which increases the ICERs for cemiplimab (CS Table 20). The company have not included SACT survival data in the model, arguing that the trial provides 'more robust longer term trial data' (CS Table 9).

Results from the SACT dataset so far indicate that survival has been worse under the CDF (median OS 21 months) than in the trial [redacted] alive at [redacted] months). The company present various explanations for these differences in survival, including the patient populations and the impact of COVID-19 on clinical presentation and treatment (CS section A.6.5).

ERG conclusions

- The SACT dataset comprises patients treated with cemiplimab in UK practice. This indicates that clinicians will offer cemiplimab to patients who are on average older and less fit than those in the trials, and also that patients with some degree of immunocompromise may be offered cemiplimab. This view is supported by expert advice to the ERG. We therefore prefer the company's scenario with baseline patient characteristics derived from the SACT dataset (median 77 years of age, 74% male).
- There is uncertainty over the comparability of the populations in the cemiplimab trials and the chart review, and which source is more generalisable to UK practice. The 'real world' chart review should better reflect UK practice, but it is subject to bias due to problems of face-validity including missing and ambiguous data and post hoc exclusion of patients from the analysis. This uncertainty translates to uncertainty over which IPW method (ATT or ATC) should be used to adjust for prognostic factors. Although there is a more fundamental uncertainty, as the ERG does not have confidence that any of the ATT or ATC models successfully balanced all covariates (see Section 3.4 above.

4.3 Interventions and comparators

4.3.1 Cemiplimab

The base case uses survival curves for cemiplimab estimated from the company's trials, including groups 1 and 2 treated with a weight-based dose and group 3 treated with a flat dose in the EMPOWER-CSCC 1 trial (CS Table 5). In response to clarification question A2, the company provided a provisional comparison of outcomes between the three groups, see section 3.1.2 above. The model uses costs for cemiplimab based on the flat dose of 350 mg IV every week recommended in the marketing authorisation.

The model assumes that all patients continue treatment until progression or a maximum of 24 months, as recommended in TA592. The analysis does not account for patients who may stop treatment before progression, for example because of adverse effects. The maximum duration of treatment in the company's trials was 22 months (shorter for the fixed dose group). The CDF submission does not report the duration of treatment in the cemiplimab trials, but median PFS was [redacted] (CS section A.6.1.3).

4.3.2 Platinum-based chemotherapy

The NICE committee concluded that platinum based chemotherapy and best supportive care are both relevant comparators for cemiplimab (TA592 section 3).

The company agree, but state that UK clinical opinion is that BSC may be considered a more relevant comparator, as cemiplimab can be used for patients who cannot tolerate chemotherapy (CS section A.1). This view is supported in the submission by the British Association of Dermatologists, who state that in the UK very few patients will be offered EGFR inhibitors or chemotherapy (BAD submission p5).

A clinical expert advising the ERG reported that although patients were occasionally treated with platinum based chemotherapy prior to the availability of cemiplimab, many more patients are suitable for treatment with cemiplimab. Another expert noted that the views of dermatologists and/or oncologists in different centres in the UK may differ regarding which patients would be suitable for chemotherapy or which individual patients in the dermatology clinic should be offered the option of chemotherapy.

4.3.3 Best supportive care

The company excluded patients on BSC from the Chart review analysis and instead relied on data for the 20 immunocompetent patients in the Sun study for their base case analysis. They also report a scenario with survival outcomes for BSC based on data for chemotherapy from the Jarkowski study (as in the analysis for TA592).

ERG conclusions

- The company do not report a full incremental analysis between cemiplimab, chemotherapy and BSC. The ERG consider that this is reasonable because, although clinical advice suggests that cemiplimab is likely to provide an alternative for patients who would otherwise have chemotherapy and for those who would have BSC, these groups of patients may be considered as largely distinct.
- The model reflects the TA592 recommendation for a maximum 24-month stopping rule for cemiplimab, but with the assumption that no patients stop treatment prior to disease progression. This latter assumption does not reflect experience from the SACT dataset and is likely to overestimate the costs of cemiplimab.
- There is very sparse data on outcomes with BSC, as patients treated with BSC were excluded from the UK chart review and the Sun study cohort is limited. There is therefore high uncertainty over the cost-effectiveness estimates for the comparison with BSC.

4.4 Perspective, time horizon and discounting

The model uses a lifetime horizon (30 years from an initial mean age of 71 years in the base case). In accordance with the original submission and the NICE reference case, costs are estimated from the perspective of the NHS and personal social services and a discount rate of 3.5% per year is applied to both costs and QALYs. The model uses a monthly cycle, with a half-cycle correction.

4.5 Treatment effectiveness and extrapolation

4.5.1 Overview of methods for survival extrapolations

The company outline their approach to estimating PFS and OS in CS section A.8.2. As in the original submission, they fit independent survival curves from separate single-arm data sources for cemiplimab, chemotherapy and BSC.

Evidence regarding the proportional hazards assumption for OS and PFS comparisons is

presented in CS A.15.2 and A.15.3, including log-log plots and hazard plots. Additional information, including Schoenfeld residual plots, is provided in the technical report on the UK chart review (Sanofi 2021).[12] See discussion in sections 3.1.6.3 and 3.1.7 above.

The comparisons are adjusted for population differences, using methods described in CS section A.7 and Appendices A.15.8 and A.15.9, see discussion in section 3.1.6 above. The economic model uses the IPW and STC approaches:

- For the comparison with chemotherapy, OS is estimated from IPD from the company trials and the chart review, with IPW-based indirect comparisons used to weight the data to achieve similar population characteristics for the two data sources (CS sections A.7.1 and A.15.8), as explained in section 3.1.6.3 above. The base case uses an ATT approach (chart review results adjusted to reflect the trial population), with ATT model 1 as the preferred model (CS Table 10). The company also report scenarios with the full ATT model and ATC model 1 (CS Table 20). All economic analyses use 'trimmed' weights, capped at the 95% percentile.
- The Jarkowski study cohort provides another data source for chemotherapy. This is used in the base case for PFS, which is not available from the chart review, and as a scenario for OS. The company only include the STC method of population adjustment in the economic model. For this analysis, results for cemiplimab are adjusted to reflect characteristics of the Jarkowski cohort (analogous to an ATC approach). As noted in section 3.1.5.2 above, the ERG has concerns over the robustness of this analysis due to limitations in the face validity of the comparators and lack of covariates to adequately match the populations.
- For the comparison with BSC, OS is estimated from the Sun study cohort, using the STC method to adjust the cemiplimab results to reflect characteristics of the Sun cohort (CS section A.7.3). This source does not report PFS and the company make an assumption that the BSC population start in a progressed health state (CS A.8.4.3).

Finally, the company fitted survival distributions to the (population adjusted) data (CS section A.8.2). For each survival outcome, four parametric distributions (Weibull, Gompertz, log-normal and log-logistic) and ten fractional polynomial (FP) distributions were fitted. The company reported following the steps recommended in NICE DSU guidance (TSD 14 and 21) to select preferred distributions for OS and PFS: assessment of statistical (AIC/BIC) and visual fit to KM data; assessment of the shape of the hazard over time, and consideration of the plausibility of the extrapolations (clinical expert opinion from an advisory board).

However, they did not explore uncertainty over the choice of survival distributions in scenario analysis.

4.5.2 Overall survival extrapolations

Cemiplimab OS (CS section A.8.3.1)

- Fitted to unadjusted integrated trial data (CS Figure 1).
- Log-normal distribution (as in TA592): best AIC/BIC statistics (CS Table 24), good visual fit (CS Figure 32) and decreasing hazards (CS Figure 45).

Platinum based chemotherapy OS (CS section A.8.3.2)

- Chart review data, adjusted for trial population (IPW ATT model 1, trimmed analysis).
- Log-logistic distribution: revised from Gompertz fitted to Jarkowski data in TA592.
- Note that there is a reporting error in CS section A.8.3.2 and CS Table 15, as confirmed in the company's response to clarification question B2. However, without a correction to the text it is difficult to understand the rationale for the company's choice of log-logistic distribution for their base case. The model fit statistics for the base case model are also missing from CS Table 24 and it is very difficult to assess the visual fit to the KM data (CS Figure 37) or the trends in hazards (CS Figure 50), given the scale and numbers of series shown on these graphs.
- From visual inspection in the model, it does appear difficult to reconcile fit to the chart review KM curve with the 3-5 year life expectancy estimated by the company's advisory board. The distributions with a better fit to the chart review data have a plateau in long-term survival.

Best supportive care OS (CS A.8.3.3)

- Fitted to the Sun study data for immunocompetent patients (n=20), with STC adjustment of the cemiplimab curve to reflect Sun study population characteristics.
- Log-logistic distribution chosen, based on clinical opinion (survival landmarks for BSC). This is not the best-fitting distribution.
- In TA592, the same OS curve was used for BSC as for chemotherapy (Jarkowski study, Gompertz distribution due to the lack of other data.

General population mortality rates

- Updated for 2018-2020 National Life Tables, England and Wales (ONS).[20]

- Applied as a lower limit to the modelled mortality rates, as in the TA592 model.

4.5.3 Progression free survival extrapolations

Cemiplimab PFS (CS section A.8.4.1)

- Fitted to updated integrated trial data (CS section A.6.1.3; CS Figure 2)
- Second order fractional polynomial (p1 = 0, p2 = -1) chosen based on the statistical fit and advice on clinical plausibility of the extrapolations from the company's clinical advisory group.
- The company note that the Weibull distribution used in TA592 had the poorest statistical fit to the updated cemiplimab trial data.

Platinum based chemotherapy (CS section A.8.4.2)

- Base case: Jarkowski STC analysis (CS section A.6.3; CS Figure 5), Weibull distribution (as in TA592)
- PFS data from the chart review were not considered reliable (44/47 PFS events were deaths).
- Company argues similarity of chart review and Jarkowski populations (CS section A.6.2.1) and OS results (CS section A.6.2.2; CS Figure 4).

Best supportive care (CS section A.8.4.3)

- BSC is assumed to be palliative; all patients start in a post-progression state.

4.5.4 Waning of treatment effects

The company describe their approach to modelling the waning of treatment effects in CS section A.8.5. The analysis at CDF entry (TA592) had assumed waning of the relative treatment effects of cemiplimab (equal hazards for progression and mortality) at 36 months.

The revised CDF review company base case assumes loss of relative benefit at 60 months, based on the maximum follow up of EMPOWER-CSCC 1 trial data for cemiplimab. The CS presents two scenarios to test less conservative assumptions:

- No waning, with continuation of fitted OS and PFS extrapolations for cemiplimab.
- Gradual waning between 60 to 96 months.

ERG conclusions

- A high degree of uncertainty remains over the survival extrapolations in the CDF review model due to limitations in data for the comparators, and reliance on data from different populations and healthcare contexts.
- The company fit independent survival curves to separate data sources for each comparator in the economic model, rather than using relative treatment effects (hazard ratios) estimated from the population adjusted indirect comparisons. This is reasonable, as proportional hazards are not supported, but the ERG has serious concerns over whether the any of the methods of population adjustment for the indirect comparisons (IPW and STC) provide an adequate balance of prognostic covariates (see section 3.4 above).
- The base case OS extrapolation for chemotherapy is adjusted to align with the population in the cemiplimab trials (ATT analysis), which does not reflect the population treated with cemiplimab in practice under the CDF (SACT dataset). However, the generalisability of the alternative ATC scenario is also questionable.
- The reported methods for fitting survival extrapolations are consistent with guidance.[21 22] However, the rationale for the company's choice of the log-logistic distribution for chemotherapy OS is not clearly explained and they do not explore the impact of using alternative survival distributions. We consider a range of alternatives in ERG scenario analysis, see section 6.1 below.
- Company assumptions used to estimate PFS for the comparators are also uncertain. For chemotherapy, they use different sources for OS (chart review) and PFS (Jarkowski study). And for BSC they assume that all patients start in the post-progression state. We explore another approach in an ERG scenario analysis, using the relationship between PFS and OS in the Jarkowski study (see section 6.1 below).
- The company's approach to modelling waning of the relative treatment benefit for cemiplimab is consistent with that in TA592. They have assumed a longer persistence of the advantage in their updated base case (5 rather than 3 years), based on extended data from EMPOWER. Alongside company scenarios with no waning and gradual waning between 5 and 8 years, we report additional ERG scenarios to test the impact of earlier loss of relative effects (see section 6.1 below).

4.6 Adverse effects

The company have updated adverse event rates for cemiplimab from the July 2021 EMPOWER-CSCC 1 trial data cut (CS A.15.9 Table 34). Incidence of 'failure to thrive' and fatigue were set to zero, as the observed rates did not reach the 5% threshold for inclusion

(CS A.15.13). The change in adverse event rates has a minimal impact on the cost-effectiveness results.

4.7 Health related quality of life

Utility estimates were also updated, based on EORTC quality of life data from the EMPOWER-CSCC 1 trial October 2020 data cut (CS section A.15.10). As in the TA592 analysis, data were mapped to EQ-5D-3L UK tariff values using the algorithm reported by Longworth et al. (2014).[16] [redacted] (CS Table 33).

The model includes a cap on utility that prevents utilities exceeding general population values (adjusted for age and the gender split). The company made a correction to the way in which this utility cap was applied. The ERG agree with this correction.

4.8 Resources and costs

The model includes a revised price discount for cemiplimab ([redacted] at CDF entry to [redacted] in the present analysis). A list of resource use and unit cost parameters is provided in CS Table 35. Resource use assumptions have not been changed from those in the analysis at CDF entry. Unit costs have been updated for all drugs in the model, drug administration, monitoring, adverse events and other resource use.[17-19]

5 COST EFFECTIVENESS RESULTS

5.1 Company's cost effectiveness results

5.1.1 Deterministic base case

The company's deterministic base case results are reported in CS Section A.10, Tables 16 and 17. Revised versions of these tables provided in response to ERG Clarification Question B1 show the correct ICERs for cemiplimab at CDF entry, as specified in the terms of engagement for the CDF review: £45,693 per QALY compared with chemotherapy; and £47,463 per QALY compared with BSC (see Table 8).

Table 8 Cost effectiveness results at CDF entry (deterministic, PAS price)

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER vs comparator
Comparison with platinum-based chemotherapy					
PBC	[redacted]	[redacted]			

Cemiplimab	[redacted]	[redacted]	[redacted]	[redacted]	£45,693
Comparison with best supportive care					
BSC	[redacted]	[redacted]			
Cemiplimab	[redacted]	[redacted]	[redacted]	[redacted]	£47,463
Source: Clarification response B3 and ERG replication from company model submitted 8/2/22					

The deterministic ICERs for the company's new base case are £36,163 per QALY gained for the comparison with chemotherapy and £29,438 for the comparison with BSC. These results include all of the revisions listed in Table 7 above and the PAS price discount of [redacted] for cemiplimab. The ERG replicated the reported ICERs using the revised version of the company's model submitted with clarification questions on 8 Feb 2022. We found very small (£1) discrepancies with the reported incremental costs in the company's revised base case analyses that we could not explain, see Table 9 below for the ERG results.

Table 9 Company's revised base case results at CDF review (deterministic, PAS price)

	Total costs	Total QALYs	Incremental costs	Incremental QALYs	ICER vs comparator
Comparison with platinum-based chemotherapy					
PBC	[redacted]	[redacted]			
Cemiplimab	[redacted]	[redacted]	[redacted]	[redacted]	£36,163
Comparison with best supportive care					
BSC	[redacted]	[redacted]			
Cemiplimab	[redacted]	[redacted]	[redacted]	[redacted]	£29,438
Source: ERG replication from company model submitted 8/2/22					

5.1.2 Probabilistic sensitivity analysis

The company report probabilistic sensitivity analysis (PSA) results in CS section A.11. For the comparison with chemotherapy (CS Table 18), the reported probabilistic ICER (£35,995) is similar to the deterministic result (£36,163). For the comparison with BSC (CS Table 19), the probabilistic ICER (£26,211) is somewhat lower than the deterministic ICER (£29,438).

The ERG re-ran the PSA and obtained ICERs that were very similar to the deterministic ICERs: £36,246 per QALY for the comparison with chemotherapy and £30,688 per QALY compared with best supportive care.

5.1.3 Deterministic sensitivity analysis

One-way deterministic sensitivity analyses are reported in tornado plots (CS Figures 21 and 22). These suggest that the ICERs are most sensitive to changes in monthly costs pre and post progression for cemiplimab, and to OS parameters for cemiplimab and the comparator.

The company's scenario analyses are reported in CS Table 20. They provided a revised version of this table in response to clarification question B3, correcting an error in Scenario analysis 1 (comparator OS based on Jarkowski data). They also provided an Excel file listing the inputs required to re-run their scenarios.

The ERG replicated the company's scenario ICERs, but with some small discrepancies that we could not explain (see scenarios 4, 5 and 6 in Table 10 below, and discussion of ERG model verification procedures in section 5.2.1). We consider that the company have provided limited justification (i) for their choice of scenario analyses and (ii) that other plausible scenarios would not have a more substantial impact on the cost-effectiveness results. See sections 6.1 and 6.2 below for additional ERG analysis.

Table 10 Company scenario analyses (deterministic, PAS price)

Scenario	Company ICERs		ERG ICER replication	
	PBC	BSC	PBC	BSC
Analysis at CDF entry	£45,693	£47,463	£45,693	£47,463
Company base case	£36,163	£29,438	£36,163	£29,438
1) Comparator survival: Jarkowski OS (Gompertz) and PFS (Weibull)	£36,446	£39,340	£36,446	£39,340
2) Population adjustment: ATC model 1	£39,346	NA	£39,346	NA
3) Population adjustment: ATT full model	£36,621	NA	£36,621	NA
4) SACT baseline characteristics: mean age 77 years, 74% male	£37,775	£30,953	£37,775	***£30,952***
5) No waning of treatment benefit	£26,263	£24,663	£26,263	***£24,662***
6) Waning between 60 and 96 months	£32,466	£26,002	***£32,465***	***£26,001***

Source: Clarification response B3 Table 20 and ERG analysis of company's model (dated 8/2/22)

5.2 Model validation and face validity check

5.2.1 Model verification procedures

The ERG conducted a range of manual checks to verify model inputs, calculations, and outputs ('white box' tests) on the company model submitted on 17th January 2022:

- Checking parameter inputs against values in the CS, excel model and cited sources.
- Checking all model outputs against results cited in the CS, including the base case, PSA and DSA and company's scenarios.
- Checking the calculations within the model engines (Arm 1, Arm 2 and Arm 3) sheets
- Running a range of tests by changing the input parameters and checking if results are plausible ('black box' tests)

Due to time constraints, we could not repeat all of the above checks on the revised company model that was received on 8th February 2022 as response to clarification question B1. We did complete the following tests on this model version:

- Reproducing the results from the CDF entry model (with starting base ICERs of £45,693 versus chemotherapy and £47,463 versus BSC) that was used as the basis for this submission.
- Re-running all of the company's scenarios and sensitivity analyses.

We noted a few minor inconsistencies in reporting of adverse events costs: i) cost of infection is reported as £256.62, ERG views this cost should be £251; and ii) cost of thrombocytopenia is reported as £655, we view the NHS reference cost is £618.28). However, these differences are unlikely to affect the cost effectiveness results.

The company submitted an Excel file with their response to clarification question B3 which listed the model settings for their base case and scenario analyses. The ERG re-ran the model with the assumptions cited in the document and found a few minor inconsistencies that we could not explain: the incremental costs for the revised base case (Table 9); and ICERs for scenarios 4-6 (Table 10).

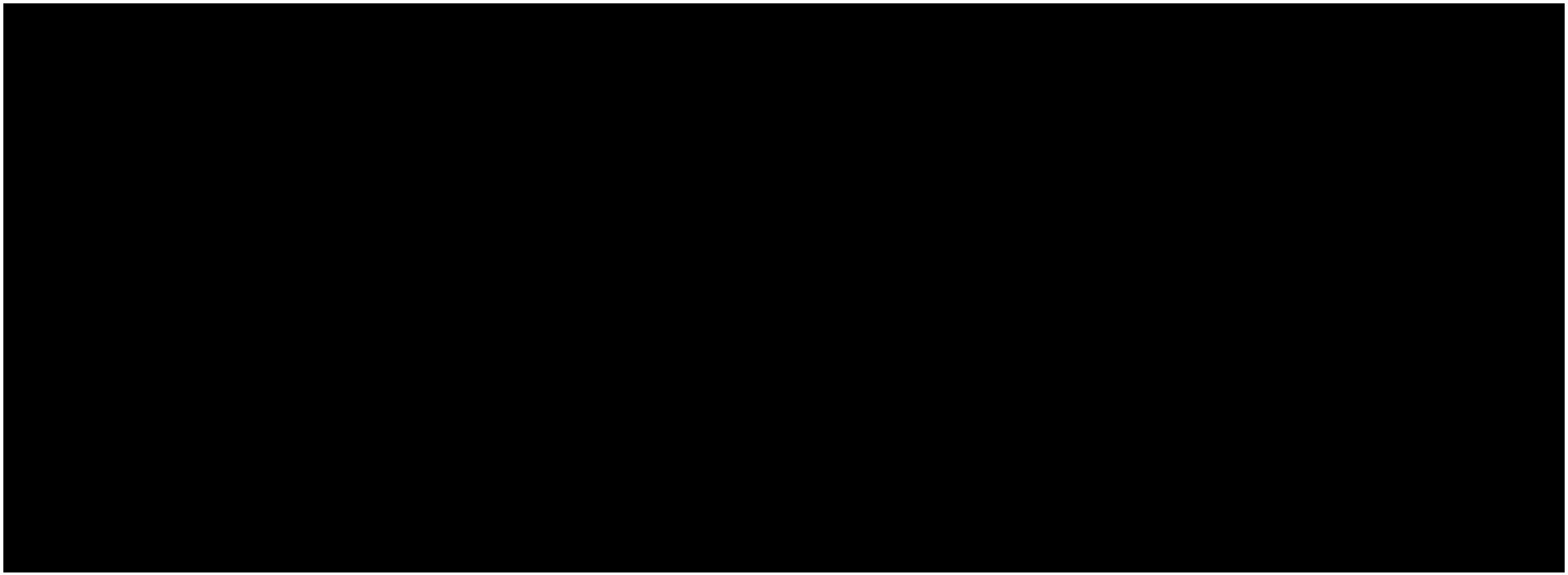
5.2.2 Validation against SACT data

To demonstrate the generalisability of the company trials data, the Terms of Engagement for the CDF review stated that the company should compare the updated results with the data collected through the Systematic Anti-Cancer Therapy (SACT) dataset.[4] See section 3.1.3 above and CS section A.6.5 for discussion of differences between the patient population treated with cemiplimab in the company's trials and the SACT dataset.

The company report a scenario analysis with baseline demographics for the SACT population (see Table 10 above). This scenario adjusts for the older population, but it does not account for other differences between the SACT and trial populations. Clinical opinion is that 'a large proportion' of SACT patients would have received BSC rather than chemotherapy if they had not had access to cemiplimab (as they would not have been able to tolerate side effects of chemotherapy).

The company did not include the SACT data within the economic model or provide any direct validation of trial or modelled survival outcomes against the SACT results. Inspection of the SACT KM survival curve (CS Figure 7) and KM and fitted extrapolation from the company's trials (CS Figure 12) shows that mortality was higher in SACT (see Figure 1 below). This suggests that results from the company's model may not be generalisable to outcomes with cemiplimab in routine NHS use, although we note that the recruitment and outcomes of SACT might have been affected by the onset of COVID 8 months after the entry of cemiplimab into the CDF.

Figure 1: Comparison of the SACT KM curve with the company's KM and fitted OS for cemiplimab



(a) KM curve for cemiplimab from the SACT database

(b) KM curve and fitted OS curve for cemiplimab from the company model

5.2.3 Comparison with survival data from other studies

The company did not provide any comparisons of the extrapolated OS estimates with external data for the population of interest under current treatment. In Table 11 below, we compare the company's survival estimates for chemotherapy with three studies:

- Hillen et al.[23] - a retrospective analysis of 24 German and Austrian patients with median age of 76 years and advanced SCC that comprised metastatic- and locally advanced SCC;
- Amaral et al.[24] - a retrospective study of real world data of 195 German patients with advanced cutaneous squamous cell carcinoma, with a median age of 78 years; and
- Cowey et al.[25] - another a retrospective, observational study of 82 patients in US with unresectable locally advanced CSCC or metastatic CSCC).

We note that the company's OS estimates are within the highest and lowest range of survival estimates as reported in these studies.

Table 11 Comparison of OS estimates for chemotherapy

Study		OS estimates				
		1-year	*2-year*	*3-year*	*5-year*	*10-year*
Company's extrapolations		65%	38%	25%	12%	4%
Hillen et al. (DeCOG study)	Advanced SCC	87%	69%	55%	NR	NR
	Locally advanced SCC	92%	77%	71%	NR	NR
	Metastatic SCC	84%	64%	47%	NR	NR
Amaral et al.		72.9%	58.2%	51.8%	NR	NR
Cowey et al.	Overall	56.1%	30.2%	15.6%	NR	NR
	Locally advanced CSCC	61.1%	32.6%	32.6%	NR	NR
	Metastatic CSCC	54.8%	30.2%	30.2%	NR	NR

Table 12 provides OS estimates for cemiplimab from the company's base case extrapolation, compared with observed survival from the SACT dataset and the study by Strippoli et al.[4 7] The latter is a retrospective cohort of 30 Italian patients with a median age of 81 years, of whom 25 had locally advanced CSCC and the remaining 5 patients had metastatic CSCC. We note that the company's survival estimates in the first two years are significantly higher than those reported from the SACT dataset and Strippoli et al.

Table 12 Comparison of OS estimates for cemiplimab

Study	OS estimates				
	1-year	*2-year*	*3-year*	*5-year*	*10-year*
Company's extrapolations	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SACT database[1]	62.5%	45%	NR	NR	NR
Strippoli et al.[2]	68%	45%	NR	NR	NR

[1]Estimates for the SACT database are approximates based on the KM curve in CS Figure 7 that shows the KM survival plot for patients receiving cemiplimab in the SACT database cohort (N=352)
[2]Estimates from Strippoli et al. are approximates based on the KM curve in Figure 4(B) in the study.

Finally, we compare the company's OS estimates for BSC (extrapolated from the Sun et al. cohort)[3] with outcomes from the study by Amaral et al.[24] This shows large differences in predicted mortality from these sources.

Table 13 Comparison of OS estimates for BSC

Study	OS estimates				
	1-year	*2-year*	*3-year*	*5-year*	*10-year*
Company's extrapolations	65%	38%	25%	13%	5%
Amaral et al.[1]	75%	65%	50%	NR	NR
[1]Estimates are approximates based on the KM curve in CS Figure 1 (f) in the study					

ERG conclusions

- ERG model checks did not identify any errors or inconsistencies that would have a material impact on the cost-effectiveness results.
- The company did not provide validation against SACT outcomes as requested in the terms of engagement for the CDF review. Observed survival with cemiplimab from the SACT dataset was evidently worse than in the company's trials and modelled extrapolations. Whilst the dataset is immature and could have been impacted by the COVID pandemic, we consider that the population is likely to be more relevant to future real-life use of cemiplimab than the population in the clinical trials. This view is supported by the similar survival results from SACT and the Italian cohort reported by Strippoli et al.[7]
- We therefore prefer company's scenario with baseline demographics from the SACT dataset. However, we note that this does not account for other differences that could affect prognosis, such as fitness and prior treatment.

6 EVIDENCE REVIEW GROUP'S ADDITIONAL ANALYSES

6.1 Exploratory and sensitivity analyses undertaken by the ERG

We present a summary of our additional scenario analyses in Table 14.

Table 14 Additional scenarios conducted by the ERG

Issue	***Company analyses***	***ERG analyses***
Patient characteristics- age and gender	Base case: • Age: 71.16 years • Gender: 83.1% Scenario: • Age: 77 years • Gender: 74%	Age: • 75 years (Cowe et al.), • 76 years (Hillen et al.), • 81 years (Strippoli et al.) Gender: • 85.4% (Cowe et al.), • 67% (Hillen et al.), • 80% (Strippoli et al.)
Treatment waning scenario	Base case: • 5 years Scenario: • No waning • Waning between 5 years and 8 years	• Waning at 42 months (3.5 years) • Waning at 48 months (4 years)
Efficacy- IPW analysis for the comparators	Base case: • ATT model 1, STC Scenario: • ATC model 1, ATT full model • None	• ATC full model, No IPW adjustment
Extrapolation of OS	Base case: • Cemiplimab: Lognormal • Chemotherapy: Log-logistic • BSC: Log-logistic Scenario: • Cemiplimab: None • Chemotherapy: Gompertz, log-logistic • BSC: Gompertz, log-logistic	• Cemiplimab: Weibull, Second order P(1, -0.5), log-logistic, Second order P(0, -1), Gompertz • Comparator: Weibull, Second order P(0, -1), lognormal, gompertz

Issue	*Company analyses*	*ERG analyses*
Extrapolation of PFS	Base case: • Cemiplimab: Second order • Chemotherapy: Weibull • BSC: N/A Scenario: • Cemiplimab: None • Chemotherapy: None • BSC: N/A	• Cemiplimab: loglogistic, lognormal, Weibull • Chemotherapy: Gompertz, lognormal, loglogistic, second order P(0, -0.5), second order P(0, -1) • BSC: N/A
PFS for comparators	No adjustment of data from different sources	• Chemotherapy: Adjust PFS by taking the ratio of PFS over OS from Jarkowski et al and apply the ratio to the chart review data. • BSC: Adjust PFS by taking the ratio of PFS over OS from Jarkowski et al and apply the ratio to the Sun et al.

6.2 Impact on the ICER of additional ERG analyses

We present the cost-effectiveness results of the ERG additional scenarios in Table 15 below. The results of the ERG additional scenarios indicated that the ICER for cemiplimab versus chemotherapy ranged between £33,195 (Scenario: PFS for chemotherapy extrapolated using Gompertz) and £43,233 (Scenario: Treatment waning at 42 months). For cemiplimab versus BSC, the ICER ranged between £32,646 (Scenario: Patient demographic with mean age of 81 years and 80% male, based on the population in an Italian cemiplimab cohort reported by Strippoli et al.[7]) and £28,859 (Scenario: without applying population adjusted indirect comparison for efficacy).

Table 15: Additional analyses undertaken by the ERG (PAS price)

Assumption		***ICER vs PBC***	***ICER vs BSC***
Company base case		£36,163	£29,438
Patient characteristics	Age:75 years; Gender ratio: 85.4% male (Cowe et al.)	£36,828	£30,129
	Age:76 years; Gender ratio: 67% male (Hillen et al.)	£37,417	£30,621
	Age: 81 years; Gender ratio: 80% male (Strippoli et al.)	£40,004	£32,646
Treatment waning	Switch to comparator hazard at 48 months (n=[redacted] at risk)	£40,160	£30,775
	Switch to comparator hazard at 42 months (n=[redacted] at risk)	£43,233	£31,676
Efficacy- Population adjustment	ATC full model	£39,191	NA
	None (no IPW or STC)	£36,814	£28,859
OS extrapolation: cemiplimab	Weibull	£36,089	£29,309
	Second order P(1,-0.5)	£35,834	£29,352
	Log-logistic	£36,354	£29,380
	Second order P(0,-1)	£36,132	£29,452
	Gompertz	£35,784	£29,357
OS extrapolation: comparator	Weibull	£43,186	£29,735
	Second order P(0, -1)	£35,652	£29,919
	Lognormal	£38,124	£29,511
	Gompertz	£35,566	£29,871
PFS extrapolation: cemiplimab	Log-logistic	£37,942	£30,574
	Lognormal	£37,998	£30,614
	Weibull	£39,512	£31,609
PFS extrapolation: chemotherapy	Gompertz	£33,195	NA
	Lognormal	£34,560	NA
	Loglogistic	£33,791	NA
	Second order P(0, -0.5)	£33,396	NA
	Second order P(0, -1)	£34,043	NA
PFS adjustment: comparators	Taking the ratio of PFS over OS from Jarkowski et al and applying it to the CHART review data (chemotherapy) and Sun et al (BSC)	£36,852	£31,512

Source: produced by ERG from company's model (dated 08/02/22)
Abbreviations: BSC best supportive care; PBC platinum based chemotherapy; OS overall survival; PFS progression free survival; NA not applicable; ICER incremental cost-effectiveness ratio; PAS patient access scheme

6.3 ERG's preferred assumptions

As discussed in Section 5.2.2, the ERG views the SACT cohort to reflect the patients treated with cemiplimab in UK NHS practice. We present the results of the ERG preferred assumption in Table 16. This increases the ICER of cemiplimab versus chemotherapy to £37,775 (an increase of £1,612 from the company's base case) and that of cemiplimab versus BSC to £30,952 (an increase of £1,514 from the company's base case). We also conduct a range of scenarios on the ERG preferred analysis, presented in Table 17 below

Table 16 ERG preferred analysis (PAS price)

Assumption	***ICER vs PBC***	***ICER vs BSC***
Company base case	£36,163	£29,438
+ Population characteristics from SACT (age: 77 years; 74% male)	£37,775	£30,952
ERG preferred analysis	£37,775	£30,952
Abbreviations: BSC best supportive care; PBC platinum-based chemotherapy; O; ICER incremental cost-effectiveness ratio; PAS patient access scheme		

Table 17 Additional scenarios conducted on the ERG preferred assumption (PAS price)

Assumption	***ICER vs PBC***	***ICER vs BSC***
ERG preferred assumption	£37,775	£30,952
Treatment waning: 48 months	£41,935	£32,380
Treatment waning: Between 60 months and 96 months	£33,942	£27,475
Efficacy- population adjustment: ATC full model	£40,863	-
Efficacy- population adjustment: ATC model 1	£41,021	-
Efficacy- population adjustment: ATT full model	£38,531	-
OS extrapolation for cemiplimab: Weibull	£37,675	£30,793
OS extrapolation for cemiplimab: Second order P(1, -0.5)	£37,503	£30,838
OS extrapolation for cemiplimab: Loglogistic	£37,969	£30,879
OS extrapolation for cemiplimab: Second order P(0, -1)	£37,749	£30,960
OS extrapolation for cemiplimab: Gompertz	£37,412	£30,843
OS extrapolation for comparator: Weibull	£44,379	£31,351
OS extrapolation for comparator: Second order P(0, -1)	£37,627	£31,443
OS extrapolation for comparator: Lognormal	£39,530	£31,069
PFS adjustment for comparators	£38,414	£33,246
Abbreviations: BSC best supportive care; PBC platinum-based chemotherapy; O; ICER incremental cost-effectiveness ratio; PAS patient access scheme		

6.4 Conclusions of the cost effectiveness section

Extended trial data

- The extended follow up data now available from the company's trials has provided better evidence for the survival extrapolations used for cemiplimab in the economic model, and greater confidence that they will be maintained for longer.
- Clinical experts consulted by the ERG reported a good experience of using cemiplimab under the CDF. They expressed enthusiasm over the responses that they had observed and confidence in use of the treatment for a wider group of patients, including some who would not be offered, or who would decline, treatment with platinum base chemotherapy due to concerns over adverse effects.
- The SACT dataset has demonstrated that within the CDF, a wider group of patients have been treated with cemiplimab than in the company's trials, including patients who are older, less fit and with a degree of immune compromise. This is positive, but it adds to uncertainty over the generalisability of the company's trial data to the population who would be treated with cemiplimab in UK practice.
- This suggests that the OS and PFS extrapolations in the company's economic model are likely to be more favourable than one would expect in routine NHS practice.

Comparator data and indirect comparisons

- A high degree of uncertainty remains over survival extrapolations for chemotherapy and best supportive care due to continuing weakness in the evidence base for these comparators, and the lack of data to support adequate population adjustment for the unanchored indirect comparisons with the cemiplimab trial data.
- There is particularly sparse data for best supportive care, as only patients who had received chemotherapy were included in the final UK chart review dataset; and the Sun cohort is very limited.
- Data on progression free survival is also sparse, as the company do not consider the chart review data on progression to be reliable, and this outcome was not reported for the Sun cohort. The company therefore rely on different sources to model the survival parameters for the chemotherapy comparator: the chart review for OS and the Jarkowski study for PFS, and they assume that all patients on best supportive care start in the 'post-progression' health state.

Treatment duration and persistence of effects

- The model reflects the TA592 recommendation that cemiplimab treatment should continue for 24 months or disease progression, whichever is sooner. However, the model assumes that no patients stop treatment prior to disease progression, which does not reflect experience from the SACT cohort. This suggests that the cost of cemiplimab may be overestimated, and ICERs underestimated.
- The company's approach to modelling waning of the relative treatment benefit for cemiplimab is consistent with that in TA592. They have assumed a longer persistence of the advantage in their updated base case (5 rather than 3 years), based on extended data from EMPOWER. Although 5 years is the maximum duration of follow up currently available, the assumption of an instantaneous loss of a relative survival advantage at this time is probably conservative.

Summary of cost-effectiveness results

- The company's revised base case ICER for the comparison with chemotherapy is above £36,163 per QALY gained and remains above £30,000 per QALY in all of their scenarios, except with the assumption of no waning of treatment effects.
- Their base case ICER for the comparison with best supportive care is £29,438 per QALY gained. This rises above £30,000 per QALY when OS and PFS extrapolations are based on data form the Jarkowski cohort (as in the TA592 analysis), or when the initial age of patients at treatment initiation is based on that in the SACT cohort.
- We conducted additional scenario analyses to test a wider range of uncertainties. Our preferred scenario includes the SACT patient demographics (77 years, 74% male). With this assumption all of our scenarios for the comparison with chemotherapy were above £30,000 per QALY. For the comparison with best supportive care, the ICER was below £30,000 with a less conservative waning assumption (gradual loss of the relative benefit between 5 and 8 years), but above this threshold for all other scenarios that we tested.

7 END OF LIFE

The CS argues that cemiplimab meets the NICE end-of-life criteria. They summarise their justification for reaching this conclusion in CS Table 21. Our critique of the company's argument is summarised in Table 18 below.

Whilst the company's analysis confirms that cemiplimab offers an extension of life exceeding 3 months when compared to chemotherapy or BSC, their base case analysis indicates that patients receiving chemotherapy have a longer life expectancy, more than 24 months. Those receiving BSC have a life expectancy shorter than 24 months. This indicates that end of life criteria is only met for patients receiving BSC, and not chemotherapy.

The ERG's preferred analysis (including age and gender based on the SACT dataset) confirms that cemiplimab offers an extension of life which exceeds 3 months when compared to chemotherapy or BSC (gains of 3.61 life years and 5.19 life years respectively). This analysis also indicates that patients receiving chemotherapy have a life expectancy longer than 24 months (2.65 life years or 31.8 months) but that those receiving BSC have a life expectancy shorter than 24 months (1.42 life years or 17.04 months). This suggests that end-of-life criteria are met for patients receiving BSC, and not chemotherapy.

We also note that there is additional uncertainty over whether cemiplimab meets end-of-life criteria because of questions over the generalisability of the cemiplimab and comparator data to the population who would be treated with cemiplimab in routine practice. This is particularly true for BSC, as data for this group is very sparse, as patients treated with BSC were excluded from the UK chart review and the Sun cohort is limited.

Table 18: End-of-life criteria

Criterion	Company's statement	ERG critique
The treatment is indicated for patients with a short life expectancy, normally less than 24 months	The company argue that without cemiplimab patients have a life expectancy less than 24 months. For chemotherapy, they report that median survival was estimated to be ~ 15 months by both the UK chart review and the Jarkowski 2016. Furthermore, they state that	In the company's base case economic model, the mean OS for patients receiving chemotherapy was 2.72 life years (32.64 months) using the retrospective chart review which increases to 2.80 life years (33.6 months) when using the data from Jarkowski et al.

	clinicians they consulted agreed that patients receiving chemotherapy are not expected to survive beyond 2 years. For those receiving BSC, they argue that patients are not expected to survive longer than 6 months and median survival reported by Sun et al was 5 months.[26]	For patients receiving BSC, the mean OS using the study by Sun et al. was 1.46 Life years (17.52 months) which increased to 2.80 Life years (33.6 months) using the data from Jarkowski et al. No discounting was applied to obtain these estimates.
There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	The company argue that cemiplimab offers a substantial extension to life for advanced CSCC patients with survival of [redacted] at [redacted] years. The company state that UK clinical experts support their statement that cemiplimab offers a greater than 3-month extension to life for this patient population. This is further supported by the SACT database that supports an extension to life of greater than 3 months.	In the base case economic modelling, cemiplimab is associated with an incremental gain of 3.0 life-years (36 months) compared to chemotherapy and 4.55 life-years (54.6 months) compared to BSC, when costs and QALYs are discounted at 3.5% pa. With no discounting, cemiplimab is associated with a gain of 3.85 life-years (46.2 months) compared to chemotherapy and 5.89 life-years (70.7 months) compared to BSC respectively.

8 REFERENCES

1. Hui D, De La Cruz M, Mori M, et al. Concepts and definitions for “supportive care,” “best supportive care,” “palliative care,” and “hospice care” in the published literature, dictionaries, and textbooks. *Supportive Care in Cancer* 2013;21(3):659-85. doi: 10.1007/s00520-012-1564-y
2. Zafar SY, Currow D, Abernethy AP. Defining Best Supportive Care. *Journal of Clinical Oncology* 2008;26(31):5139-40. doi: 10.1200/jco.2008.19.7491
3. Sun L, Chin RI, Gastman B, et al. Association of Disease Recurrence With Survival Outcomes in Patients With Cutaneous Squamous Cell Carcinoma of the Head and Neck Treated With Multimodality Therapy. *JAMA Dermatol* 2019;155(4):442-47. doi: 10.1001/jamadermatol.2018.5453
4. Public Health England. Cemiplimab for treating metastatic or locally advanced cutaneous squamous cell carcinoma – data review, 2021.
5. Baggi A, Quaglino P, Rubatto M, et al. Real world data of cemiplimab in locally advanced and metastatic cutaneous squamous cell carcinoma. *European Journal of Cancer* 2021;157:250-58.
6. Hober C, Fredeau L, Pham-Ledard A, et al. Cemiplimab for locally advanced and metastatic cutaneous squamous-cell carcinomas: Real-life experience from the French CAREPI study group. *Cancers* 2021;13(14):3547.
7. Strippoli S, Fanizzi A, Quaresmini D, et al. Cemiplimab in an Elderly Frail Population of Patients With Locally Advanced or Metastatic Cutaneous Squamous Cell Carcinoma: A Single-Center Real-Life Experience From Italy. *Frontiers in Oncology* 2021;11:686308.
8. Valentin J, Gerard E, Ferte T, et al. Real world safety outcomes using cemiplimab for cutaneous squamous cell carcinoma. *Journal of Geriatric Oncology* 2021;12(7):1110-13.
9. Challipalli A, Watkins S, Cogill G, et al. P-251 - Cemiplimab in advanced cutaneous squamous cell carcinoma: UK experience from the Named Patient Scheme. 10th World Congress of Melanoma in conjunction with 17th EADO Congress. Virtual: European Association of Dermato Oncology, 2021.
10. National Institute for Health and Care Excellence. Cancer Drugs Fund Managed Access Agreement: cemiplimab for treating cutaneous squamous cell carcinoma [TA592], 2019.
11. Jarkowski A, 3rd, Hare R, Loud P, et al. Systemic Therapy in Advanced Cutaneous Squamous Cell Carcinoma (CSCC): The Roswell Park Experience and a Review of

the Literature. *Am J Clin Oncol* 2016;39(6):545-48. doi: 10.1097/COC.0000000000000088

12. Sanofi. Indirect treatment comparison of cemiplimab versus other treatments in patients with advanced cutaneous squamous cell carcinoma. Company report, 2021:1-136.
13. Lubov J, Labbé M, Sioufi K, et al. Prognostic factors of head and neck cutaneous squamous cell carcinoma: a systematic review. *Journal of Otolaryngology* 2021;50:54:1-10.
14. Phillippo DM, Ades A, Dias S, et al. NICE DSU technical support document 18: methods for population-adjusted indirect comparisons in submissions to NICE. *Report by the Decision Support Unit* 2016
15. Faria R HAM, Manca A, Wailoo A. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness in technology appraisal: Methods for comparative individual patient data: Decision Support Unit, ScHARR, University of Sheffield, 2015:1-86.
16. Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technology Assessment* 2014
17. Department of Health and Social Care. Drugs and pharmaceutical electronic market information tool (eMIT), 2020.
18. Jones K, Burns A. Unit Costs of Health and Social Care 2021: Personal Social Services Research Unit, University of Kent, Canterbury, 2021.
19. NHS England and NHS Improvement. National Cost Collection: National Schedule of NHS costs - Year 2019-20, 2021.
20. Office for National Statistics (ONS). National life tables: England and Wales, 2021.
21. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. *Report by the Decision Support Unit* 2011
22. Rutherford MJ, Lambert PC, Sweeting MJ, et al. NICE DSU Technical Support Document 21. Flexible Methods for Survival Analysis. *Department of Health Sciences, University of Leicester, Leicester, UK* 2020:1-97.
23. Hillen U, Leiter U, Haase S, et al. Advanced cutaneous squamous cell carcinoma: A retrospective analysis of patient profiles and treatment patterns—Results of a non-interventional study of the DeCOG. *European Journal of Cancer* 2018;96:34-43. doi: https://doi.org/10.1016/j.ejca.2018.01.075

24. Amaral T, Osewold M, Presser D, et al. Advanced cutaneous squamous cell carcinoma: real world data of patient profiles and treatment patterns. *Journal of the European Academy of Dermatology & Venereology* 2019;33 Suppl 8:44-51.
25. Cowey CL, Robert NJ, Espirito JL, et al. Clinical outcomes among unresectable, locally advanced, and metastatic cutaneous squamous cell carcinoma patients treated with systemic therapy. *Cancer Medicine* 2020;9(20):7381-87.
26. Sanofi. HTA advisory Consultation summary report. 2021
27. Chapalain M, Baroudjian B, Dupont A, et al. Stage IV cutaneous squamous cell carcinoma: treatment outcomes in a series of 42 patients. *Journal of the European Academy of Dermatology & Venereology* 2020;34(6):1202-09.
28. Kramb F, Doerfer C, Meiwes A, et al. Real-world Treatment Patterns and Outcomes with Systemic Therapies in Unresectable Locally Advanced and Metastatic Cutaneous Squamous Cell Carcinoma in Germany. *Acta Dermato Venereologica* 2021;29:29.

Appendix 1

Company responses to the NICE Committee's preferred assumptions as stated in the Terms of Engagement

Assumption	Terms of engagement	Addressed by the company submission	Rationale if different	ERG comments
Population	Adults with metastatic or locally advanced cutaneous cell carcinoma that is not appropriate for curative surgery or curative radiotherapy are the relevant population for the CDF review. During technical engagement it was agreed that people with significant autoimmune disease or who have had a solid organ transplant are unlikely to be eligible for treatment.	**Yes** Study populations in all the new evidence match the population stated in the Terms of Engagement. However, there is uncertainty around the rationale for excluding some patients from the company's chart review, i.e. whether some patients relevant to the scope may have been excluded.	Not applicable	The company suggest that people with autoimmune diseases or who have had a solid organ transplant may benefit from cemiplimab (CS Table 1). The SACT dataset,[4] other real-world cohorts,[5-8] and one of the ERG's clinical experts confirmed that some people with autoimmunity or a solid organ transplant have received cemiplimab but numbers were small.
Comparators	The company should present clinical and cost-effective evidence for cemiplimab compared to chemotherapy and best supportive care.	**Yes** Note that chemotherapy is limited to platinum-based chemotherapy (PBC), which is consistent with TA592.	Not applicable	Due to the toxicity of chemotherapy and growing experience of the tolerability of cemiplimab, BSC is becoming the most relevant comparator to cemiplimab. However, BSC evidence is difficult to identify. [1 2]

Generalisability of trial evidence	The company should use data collected through SACT to demonstrate the generalisability of the trial data.	**Yes, with limitations** The company have considered the SACT data. This indicated differences between the population treated with cemiplimab in NHS practice and the trial populations, although these data were collected during the COVID-19 pandemic which would likely influence generalisability.	Not applicable	SACT does not report dosing regimens but according to the stated cycle length the licensed flat dose appears to have been used. Limited baseline characteristics were collected in SACT.
Survival outcomes	The company should use updated survival data from EMPOWER-CSCC 1 and fully explore the most appropriate method to extrapolate survival outcomes. Data collected through SACT should be used to validate the trial outcomes.	**Partly** The latest data cuts for the company trials (July 2021 for EMPOWER-CSCC 1 and 2019 for Study 1423) were pooled and used to inform PFS and OS. Extrapolation methods are explored in the CS based on model fit and clinical opinion. However, SACT data have not been used to validate the company trial survival outcomes.	The rationale for not using SACT data to validate the survival outcomes is not explicitly justified in the CS. The company state that "Data from SACT is short term, in some instances incomplete and contains a number of uncertainties compared to EMPOWER-CSCC 1, therefore use of the longer-term trial data is preferred" (CS section A.6.5).	Although follow up in SACT was shorter than in the company trials, median OS was reached. The CS is not explicit about the "incomplete" and "uncertain" aspects of the SACT data being referred to. Company trial survival extrapolations are reported in CS Table 15 and CS sections A.15.2 to A.15.4. Extrapolations used in the economic model are reported in CS section A.9.
Comparator data	The company should use their UK chart review and any additional data that has become available during the period of managed	**Yes, with limitations** The company use three comparator cohorts:	Not applicable	The company raised concerns about reliability of the data originally collected in the chart review. The ERG's clinical

	access to inform the comparator arms.	• Company chart review: OS for platinum-based chemotherapy. • Jarkowski et al. study (not new): OS & PFS for platinum-based chemotherapy. • Sun et al. study: OS for best supportive care However, reliability of the company's chart review is uncertain.		experts also questioned the appropriateness of eligibility criteria that were applied to the chart review data post-hoc. The ERG and clinical experts did not identify any new studies that were not included by the company.
Relative effectiveness	The company should fully explore the most appropriate treatment comparison method and utilize any updated data that has become available during the period of managed access. The committee concluded that the relative effectiveness estimates for cemiplimab are highly uncertain regardless of ITC method as all used unreliable comparator data.	**Yes, with limitations** The company have used updated cemiplimab and comparator data to compare cemiplimab against chemotherapy and BSC using three adjusted ITC methods (IPW, STC, MAIC). However, high uncertainty remains due to limitations in the comparator data. The proportional hazards assumption is not supported so hazard ratios are illustrative only.	Not applicable	Hazard ratios are not required in the economic model, which is informed by separate ITC-adjusted survival curves (consistent with the approach in TA592). The ERG were not provided with the individual participant data so could not verify that company analyses were conducted as described in the CS and clarification responses.
Treatment effect duration	The company should use updated survival data from EMPOWER-CSCC 1 and fully explore the	**Yes** The company have used the updated survival data from EMPOWER-CSCC 1	Not applicable	Not applicable

	impact of a 24-month stopping rule on long-term outcomes.	and explored the impact of a 24-month stopping rule on long-term outcomes.		
Most plausible ICER	The committee agreed that cemiplimab demonstrated plausible potential to be cost-effective. Due to uncertainty in the evidence base, the committee did not state a preferred ICER	Not applicable	Not applicable	Not applicable
End of life	The company should demonstrate whether cemiplimab meets the end-of-life criteria	**Yes** The company argue that cemiplimab meets end-of-life criteria compared to both chemotherapy and BSC. However, their base case model indicates that the criteria are met for the comparison with BSC, but not for the comparison with chemotherapy (as the life expectancy exceeds 2 years. The ERG preferred scenario reiterates this conclusion.	Not applicable	Overall, it remains unclear if cemiplimab meets end-of-life criteria due to high uncertainty in the comparator data.
BSC: best supportive care; IPW: inverse probability weighting MAIC: matched adjusted indirect comparison; OS: overall survival; PBC: platinum-based chemotherapy; PFS: progression free survival; SACT: Systemic Anti-Cancer Therapy dataset; STC: simulated treatment comparison				

Appendix 2

Nine potentially relevant studies, published in full since November 2018, were identified by the ERG. However, two reviewers excluded them all from this review. The studies and the reason for exclusion are summarised in the table below.

Real-world studies of treatments for advanced CSCC identified and excluded by the ERG

Reference	Setting; design	Population	Intervention	Outcomes	Reason for exclusion
Amaral et al 2019[24]	Germany; retrospective	N=50 (195 total advanced CSCC, 50/195 inoperable); median age 78 years; ECOG PS not reported	Chemotherapy 20/50; BSC 12/50	Overall survival	Outcomes not reported for population subgroups relevant to this review.
Baggi et al 2021[5]	Italy; multicentre (17), retrospective	N=131 (91 laCSCC, 40 mCSCC, 9.2% had autoimmune disease); median age 79 years; ECOG PS 0-1 in 77.9% of 125/131	Cemiplimab	Treatment related adverse events; response rates	Outcomes. OS and PFS not reported.
Chapalain et al 2020[27]	France; single centre, retrospective	N=42 (stage IV CSCC, 31% immunocompromised); median age 75.5 years; ECOG PS 0-1 in 93%	Chemotherapy and/or cetuximab	OS at 4 years; response rate; adverse events	Only 6/42 received chemotherapy alone at 1L, and 20/25 at 2L. Outcomes not reported for population subgroup relevant to this review.

Cowey et al 2020[25]	United States; retrospective, observational	N=82 (17 laCSCC, 65 mCSCC); median age 75 years; ECOG PS 0 in 10%, 1 in 88%, not reported in 2%	Most common 1L regimens: Carboplatin + paclitaxel (27%); Cetuximab monotherapy (24%)	OS	Outcomes not reported for population subgroup relevant to this review (i.e. carboplatin + paclitaxel).
Hillen et al 2018[23]	Germany and Austria; multicentre (24), retrospective	N=190 (76 laCSCC, 114 mCSCC, 24% immunocompromised); median age 78 years; ECOG PS 0-1 in "most"	Chemotherapy including PBC.	Response rates	Outcomes. PFS and OS not reported.
Hober et al 2021[6]	France; multicentre (58), retrospective	N=245 (24% immunocompromised); mean age 77 years; ECOG PS >= 2 in 27%	Cemiplimab	Response rate; OS at 1 year; PFS	Population. Includes immunocompromised and ECOG 2 or greater.
Kramb et al 2021[28]	Germany; single centre, retrospective	N=59 (laCSCC unresectable 20/59, mCSCC unresectable 25/59, immunocompromised were excluded); median age 76 years; ECOG PS not reported	15/45 unresectable patients received systemic treatment	Response rates; PFS; OS	Outcomes not reported for population subgroup relevant to this review (i.e. any of the PBC regimes).

Strippoli et al 2021[7]	Italy; single centre, retrospective	N=30 (25 laCSCC, 5 mCSCC, 5/30 immunocompromised); median age 81 years; ECOG PS 0 in 23%, 1 in 57%, 2 in 20%	Cemiplimab	Response rate; adverse events; OS; PFS	Population. Older, ECOG PS 2.
Valentin et al 2021[8]	France; single centre, retrospective	N=22 (laCSCC and mCSCC); median age 83 years; ECOG PS 0 or 1 in 73%	Cemiplimab	Safety in daily practice – AEs and SAEs	Outcomes. OS and PFS not reported.
ECOG PS: Eastern Cooperative Oncology Group Performance Status scale; laCSCC: local advanced cutaneous squamous cell carcinoma; mCSCC metastatic cutaneous squamous cell carcinoma; OS: overall survival; PBC: platinum-based chemotherapy; PFS: progression free survival					

Appendix 3

Reproduction of CS Table 23 (Summary of baseline patient characteristics for main data sources) with corrections by ERG and including the Named Patient Scheme study for comparison

Red data indicate corrections made by ERG		Cemiplimab					BSC	Chemotherapy	
		EMPOWER-CSCC-1	**Study 1423**	**Pooled company trials**	**SACT dataset (CDF)**	**Named Patient Scheme study**	**Sun et al, 2019**	**Chart review (cohort for analysis),**	**Jarkowksi et al, 2016**
N		193	26	219	352	38	32	[redacted]	25
Disease severity	laCSCC mCSCC	78 (40.4) 115 (59.6)	10 (38.5) 16 (61.5)	88 (40.2) 131 (59.8)	172 (49) 180 (51)	10 (26.4) 28 (73.6) [a]	12 (42.9) [b, c] 16 (57.1) [b, d]	[redacted]	19 (76) [f] 6 (24) [f]
Age median (range)		72 (38-96)	72.5 (52-88)	72 (38-96)	77	74 (28-90)	73 (43-89)	[redacted]	66.4
Gender n (%)	Male	161 (83.4)	21 (80.8)	182 (83.1)	262 (74)	31 (81.6)	26 (81.3)	[redacted]	18 (72)
Differen-tiation n (%)	Well Undiff. [g] Undetermined	[redacted]	2 (7.7) 17 (65.4) 7 (26.9)	[redacted]	- - -	- - -	-	[redacted]	-
Tumour location n (%)	Head and neck Trunk Extremities	[redacted]	19 (73.1) 2 (7.7) 5 (19.2)	[redacted]	- - -	24 (63.2) - -	32 (100) 0 (0) 0 (0)	[redacted]	11(44.0) 7 (28.0) 3 (12.0)
T stage n (%) [h]	T0 Tis Tx T1 T2 T3 T4	[redacted]	- - 8 (30.8) 3 (11.5) 10 (38.5) 2 (7.7) 3 (11.5)	[redacted]	- - - - - - -	- - - - - - -	8 (25.0) T1/T2: 11 (34.4)	[redacted]	- - - - - - -

		[redacted]		[redacted]			T3/T4: 13 (40.6)	[redacted]	
ECOG PS n (%)	0 1 2	86 (44.6) 107 (55.4) -	10 (38.5) 16 (61.5) -	96 (43.8) 123 (56.2)	64 (18) 223 (63) 14 (4) 51 missing	0-1: 33 (86.8)	0-2: 32 (100)	[redacted]	- - -
Prior systemic therapy n (%)		65 (33.7)	15 (57.7)	80 (36.5)	-	9 (23.7)	-	[redacted]	Not reported [i]
Prior radiation n (%)		[redacted]	21 (80.8)	[redacted]	-	-	32 (100)	[redacted]	-

[a] The Named Patient Scheme study distinguishes between nodal disease and distant metastases (26.4% nodal disease, 26.4% distant metastases, 21.1% had both nodal and distant disease).
[b] Data are for 28 of the 32 patients in the study who were immunocompetent; 4 patients recorded as having both locoregional and distant disease are excluded (these four patients would be classified as mCSCC according to the company's definition).
[c] Locoregional disease at 1st recurrence as defined by Sun et al. 2019 which does not correspond to laCSCC as defined by the company (see section 3.1.4.3).
[d] Distant disease at 1st recurrence as defined by Sun et al. 2019 which does not correspond to mCSCC as defined by the company (see section 3.1.4.3).
[e] According to the Chart Review Protocol patient records were included in a 60:40 mCSCC : laCSCC ratio so these proportions may not reflect real world prevalence of laCSCC and mCSCC in this study.
[f] The Jarkowski et al. 2016 paper refers to locoregional and metastatic groups but does not define these. The ERG assume that the locoregional group would overlap with mCSCC as defined by the company, since mCSCC includes regional nodal disease. However, as noted in CS Appendix D.1.3.2 in the TA592 appraisal, Jarkowski et al. use the terms 'locoregional' and 'locally advanced unresectable' interchangeably which might suggest that locoregional could be equivalent to laCSCC. But this is speculative.
[g] Reported as "Moderate/Poor/Undifferentiated" elsewhere (e.g. Chart Review Report Table 3).
[h] The T-stage indicates tumour thickness and local spread into nearby structures.
[i] The proportion who received prior systemic therapy is not reported in the Jarkowski et al. 2016 study publication; the value of 8 (32) reported in CS Table 23 appears to be for therapy received during the study. The company do not explain their interpretation of the data.

Appendix 4

Face validity issues noted by the company for their retrospective chart review

Issue (CS section A.15.6)	Company comments	ERG and clinical experts' comments	Issue resolved by the data audit?
Large data gaps for "many" patients where no events or visits were observed for the laCSCC population.	These gaps appear contrary to standard treatment patterns for patients with advanced cancers	Clarification response A10 indicates 53 patients (50%) had follow up gaps >12 months where no visits were confirmed.	No
Higher uptake of chemotherapy than would be expected.	Once patients have metastasized or progressed following surgery/ radiotherapy, there are no palliative treatments, so patients are cared for in the community (i.e. GPs) rather than in hospitals.	ERG clinical experts commented that patients with metastatic CSCC do not tend to be discharged into the community and doctors may try treatment (e.g. chemotherapy) for patients who they perceive to be fitter. A hospital chart review may be more likely to pick up these cases.	No
Lack of information on prior and palliative treatment (surgery and radiotherapy).	Few patients are recorded having prior radiotherapy, surgery, or systemic therapy. In clinical practice few patients are initially diagnosed with metastatic disease so would usually have received surgery or radiotherapy prior to diagnosis.	ERG clinical experts agreed most patients present with laCSCC before progression to mCSCC. However, the experts questioned the validity and completeness of the chart review since hospital medical records should always include information on prior treatment for CSCC.	No
Longer survival estimates than expected in 90 patients who received ≥1 line of	OS for patients who received at least one line of therapy post advanced diagnosis (n=90) were longer than estimates available in the published literature, particularly for	ERG clinical experts agreed that OS of these 90 patients in the chart review appears longer than would be expected. Some patients may have received radical radiotherapy in combination with	Yes – demonstrated by comparing CS Figure 67 (pre-audit, N=90) against Chart Review

therapy post advanced diagnosis.	laCSCC patients (CS section A.15.6; CS Figure 67). In particular, survival times of patients in laCSCC were considered high compared to clinical expectation that survival in this population would not be expected to be beyond 2 years.	chemotherapy which could lead to better outcomes. The CS does not report OS separately for laCSCC patients.	Report Figure A7 (post-audit, [redacted])
A large proportion of the chart review population were indicated to have received palliative radiation and/or palliative surgery; however, dates of administration were not recorded and may have occurred prior to, during, or following systemic treatment.	It is possible that some of these patients may have received regimens more akin to definitive rather than palliative radiotherapy, though this cannot be determined from the original data collection. Potential under-reporting of palliative treatments could have contributed to the over-estimation of survival.	ERG clinical experts commented that radiotherapy dates should be accessed easily through hospital reporting systems and hospital radiotherapy systems. Lack of these data makes the chart review harder to interpret and raises questions about its usefulness.	No
The reason for excision biopsies was not collected, but the biopsies could have been a form of tumour debulking (CS section A.6.2.1).	The company argue that if patients received tumour debulking they would be incomparable to those in the company trial (i.e. laCSCC, not eligible for surgery or radiotherapy) (CS section A.6.2.1).	ERG clinical experts disagreed with the company assumption that doctors might report an excision biopsy (curative) as tumour debulking (palliative). The experts questioned why patients treated palliatively would be “incomparable” with those in the company trial, since patients receiving BSC would be	No

		eligible for cemiplimab, and the company trials included patients with mCSCC as well as laCSCC.	
Deaths made up a significant portion of the PFS events in the chart review, despite these occurring well after cessation of therapy.	Patients would be expected to progress prior to death, especially in the metastatic cohort, suggesting progression was not formally recorded in many charts. Due to lack of reliable information on progression events PFS was not estimated (CS section A.14.2).	ERG clinical experts commented that these patients may have been discharged to the community then progression would not be recorded. If they had palliative treatment, then progression would be recorded as a reason to discontinue. The ERG agree that the apparent lack of (and possible inconsistency in) reporting progression events precludes reliable estimation of PFS.	No
Relatively few patients experienced any events within the first six months of treatment.	Substantial numbers of patients die within the first six months in the company trials, as well as in all published literature on this patient population. This period of non-events could be due to the data collection process not being sufficiently comprehensive (it may be the case that not all events were recorded) or as patients enrolled in the chart review were generally healthier.	ERG clinical experts commented that patients in the chart review may have been at an earlier stage of laCSCC, whereas substantial numbers of patients dying within the first 6 months in the company trials suggests many had distant metastases. This is where lack of detail in the chart review may explain the outcomes. Poor comprehensiveness of data collection is also plausible.	No

Appendix 5

Covariates included in the company ITC analysis models estimating the average treatment effect in the treatment group (ATT) and average treatment effect in the comparator group (ATC)

Model	Median Age	Disease severity	Differen-tiation	Sex	Tumour location	ECOG PS	Prior systemic therapy	Prior radiation	T-stage
ATT full	Y	Y	Y	Y	Y	Y	N	N	Y
ATT1	Y	Y	Y	Y	N	N	N	N	Y
ATT2	N	Y	Y	Y	N	N	N	N	Y
ATT3	Y	Y	N	Y	Y	N	N	N	Y
ATT4	Y	Y	N	Y	N	N	N	N	Y
ATT5	N	Y	Y	Y	N	N	N	N	Y
ATT6	Y	Y	Y	Y	Y	N	N	N	Y
ATT7	Y	Y	N	Y	N	N	N	N	N
ATT8	Y	Y	N	Y	Y	N	N	N	N
ATT9	N	Y	Y	N	N	N	N	N	Y
ATT10	Y	Y	Y	N	N	N	N	N	Y
ATC full	Y	Y	Y	Y	Y	Y	Y	N	Y
ATC1	Y	Y	Y	Y	Y	Y	Y	N	N
ATC2	N	Y	Y	Y	Y	Y	Y	N	N
ATC3	N	Y	Y	Y	N	Y	Y	N	N
ATC4	Y	Y	Y	N	N	Y	Y	N	N
ATC5	Y	N	Y	Y	Y	Y	N	N	N
ATC6	Y	N	Y	Y	N	Y	N	N	N
ATC7	Y	Y	Y	Y	N	Y	N	N	N
ATC8	Y	Y	Y	Y	Y	Y	N	N	N
ATC9	Y	Y	Y	Y	Y	N	N	N	N
ATC10	Y	Y	Y	Y	N	N	N	N	N

Sources: Chart Review Report Tables 5, 7, C1, D2 and Figures D18, D19, D20, D21, D22

Appendix 6

Overview of ITC model fit, data sources and hazard ratios for ITC comparisons of cemiplimab (company trials) versus chemotherapy (company chart review). All data sources refer to the company Chart Review Report.[12]

Model	Source	Covariates included	Balance [a]	ESS [b]	ESS trimmed	Comparisons of reweighted covariates; KM curves	Number (trimmed analysis) of covariates with ASD>10% after reweighting [c]		HR [d]	HR trimmed [d]
ATT full	Tables 7, C1	[redacted]	[redacted]	[redacted]	[redacted]	Tables 6, B1 KM: Figure 4	[redacted]	Figure 2	[redacted]	[redacted]
ATT1	Tables 5, 7, D3	[redacted]	[redacted]	[redacted]	[redacted]	Table B2 KM: Figure 5	[redacted]	Figure 3	[redacted]	[redacted]
ATT2	Table 7	[redacted]	[redacted]	[redacted]	[redacted]	Table B3 KM: Figure C1	[redacted]	Figure C10	[redacted]	[redacted]
ATT3	Table 7	[redacted]	[redacted]	[redacted]	[redacted]	Table B4 KM: Figure C2	[redacted]	Figure C11	[redacted]	[redacted]
ATT4	Table 7	[redacted]	[redacted]	[redacted]	[redacted]	Table B5 KM: Figure C3	[redacted]	Figure C12	[redacted]	[redacted]
ATT5	Table 7	[redacted]	[redacted]	[redacted]	[redacted]	Table B6 KM: Figure C4	[redacted]	Figure C13	[redacted]	[redacted]
ATT6	Table C1	[redacted]	[redacted]	[redacted]	[redacted]	Table B7 KM: Figure C5	[redacted]	Figure C14	[redacted]	[redacted]
ATT7	Table C1	[redacted]	[redacted]	[redacted]	[redacted]	Table B8 KM: Figure C6	[redacted]	Figure C15	[redacted]	[redacted]
ATT8	Table C1	[redacted]	[redacted]	[redacted]	[redacted]	Table B9 KM: Figure C7	[redacted]	Figure C16	[redacted]	[redacted]

ATT9	Table C1	[redacted]	[redacted]	[redacted]	[redacted]	Table B10 KM: Figure C8	[redacted]	Figure C17	[redacted]	[redacted]
ATT10	Table C1	[redacted]	[redacted]	[redacted]	[redacted]	Table B11 KM: Figure C9	[redacted]	Figure C18	[redacted]	[redacted]
ATC full	Table D2	[redacted]	[redacted]	[redacted]	[redacted]	Tables D1, D4 KM: Figure D3	[redacted]	Figure D1	[redacted]	[redacted]
ATC1	Table D2	[redacted]	[redacted]	[redacted]	[redacted]	Tables 5, D5 [e] KM: Figure D4	[redacted]	Figure D2	[redacted]	[redacted]
ATC2	Table D2	[redacted]	[redacted]	[redacted]	[redacted]	Table D6 KM: Figure D5	[redacted]	Figure D14	[redacted]	[redacted]
ATC3	Table D2	[redacted]	[redacted]	[redacted]	[redacted]	Table D7 KM: Figure D6	[redacted]	Figure D15	[redacted]	[redacted]
ATC4	Table D2	[redacted]	[redacted]	[redacted]	[redacted]	Table D8 KM: Figure D7	[redacted]	Figure D16	[redacted]	[redacted]
ATC5	Table D2	[redacted]	[redacted]	[redacted]	[redacted]	Table D9 KM: Figure D8	[redacted]	Figure D17	[redacted]	[redacted]
ATC6	Figure D18	[redacted]	[redacted]	[redacted]	[redacted]	Table D10 KM: Figure D9	[redacted]	Figure D18	Not reported	Not reported
ATC7	Figure D19	[redacted]	[redacted]	[redacted]	[redacted]	Table D11 KM: Figure D10	[redacted]	Figure D19	Not reported	Not reported
ATC8	Figure D20	[redacted]	[redacted]	[redacted]	[redacted]	Table D12 KM: Figure D11	[redacted]	Figure D20	Not reported	Not reported
ATC9	Figure D21	[redacted]	[redacted]	[redacted]	[redacted]	Table D13 KM: Figure D12	[redacted]	Figure D21	Not reported	Not reported
ATC10	Figure D22	[redacted]	[redacted]	[redacted]	[redacted]	Table D14 KM: Figure D13	[redacted]	Figure D22	Not reported	Not reported
Sensitivity analyses basing ATC model on the best fitting ATT model										

ATC / full ATT model	Table D3	[redacted]	[redacted]	[redacted]	[redacted]	Not reported	Not reported	Not reported	[redacted]	[redacted]
ATC / ATT1 model	Table D3	[redacted]	[redacted]	[redacted]	[redacted]	Not reported	Not reported	Not reported	[redacted]	[redacted]

[a] Higher values indicate greater balance of covariates between the studies

[b] Effective sample size

[c] The number of covariates with an absolute standardised difference (ASD) of >10% between studies after reweighting. Numbers in brackets are for the trimmed analysis. These data were obtained by visually inspecting the source Figures listed. The company report these data numerically only for the ATT full model (the absolute standardized difference was >10% for 5 matching variables) and for ATT model 1 (the standardized mean difference was less than 10% for five of the seven prognostic factors) (Chart Review Report section 4.2.1). The visual observations are more conservative towards detecting the stronger deviations, as they cannot resolve very small differences close to 10%.The data in this column show that for all models at least two covariates had a standardised mean difference >10%, indicative of incomplete balancing of covariates in all models.

[d] Hazard ratios are uncertain and should be interpreted with caution as the assumption of proportional hazards was violated.

[e] Mislabelled Table D4 in the Chart Review Report.