REAL-WORLD EVIDENCE (RWE) AT NICE: DEMONSTRATING METHODS FOR USING THE SYSTEMIC ANTI CANCER THERAPY (SACT) DATASET IN TECHNOLOGY APPRAISALS

REPORT BY THE DECISION SUPPORT UNIT

14 March 2025

Authors: Andrew Metry[1], Nicholas Latimer[1], Allan Wailoo[1], Paul Tappenden[1]

[1]Sheffield Centre for Health and Related Research, University of Sheffield

Decision Support Unit,
SCHARR,
University of Sheffield,
Regent Court,
30 Regent Street,
Sheffield, S1 4DA

Tel (+44) (0)114 222 0734
E-mail dsuadmin@sheffield.ac.uk
Website www.nicedsu.org.uk
Twitter @NICE_DSU

ABOUT THE DECISION SUPPORT UNIT

The Decision Support Unit (DSU) External Assessment Centre is based at the University of Sheffield with members at York, Bristol, Leicester and the London School of Hygiene and Tropical Medicine. The DSU is commissioned by The National Institute for Health and Care Excellence (NICE) to provide a research and training resource to support the Institute's Centre for Health Technology Evaluation Programmes. Please see our website for further information www.nicedsu.org.uk.

The production of this document was funded by the National Institute for Health and Care Excellence (NICE) through its Decision Support Unit. The views, and any errors or omissions, expressed in this document are of the authors only. NICE may take account of part or all of this document if it considers it appropriate, but it is not bound to do so.

Acknowledgement

We would like to thank the team from NICE and the HTA Lab (Jamie Elvidge, Hannah Patrick, Jane Adam, Jeremy Braybrooke, Adam Brooke, Ravinder Claire, and Stephen O'Brien) for their regular contributions to this project through the steering group meetings, commenting on the draft report, and helping shape the Discussion Section. We would also like to thank Donna Davis, Programme Administrator, SCHARR, for providing administrative support and in preparing and formatting the report.

Many thanks to NICE staff who provided source documents; Kieron Smith, Magda Bugajska, and Celia Mayers.

This report should be referenced as follows:

Metry A, Latimer N, Wailoo A, Tappenden P. (2025) Real-world evidence (RWE) at NICE: Demonstrating methods for using the systemic anti-cancer therapy (SACT) dataset in technology appraisals. Report by the Decision Support Unit.

CONTENTS

1. EXECUTIVE SUMMARY 7
2. INTRODUCTION 9
3. METHODS 12
3.1. METHODS FOR TOPIC SELECTION 12
Identification of case studies 12
3.2. DATA EXTRACTION FROM SACT REPORTS 14
3.3. METHODS FOR THE COST-EFFECTIVENESS ANALYSIS 15
3.3.1. Methods for modelling OS and ToT 16
3.3.2. Methods for modelling PFS 17
3.3.3. Other assumptions applied in the DSU analysis 18
4. COST-EFFECTIVENESS RESULTS 19
4.1. CASE STUDY 1 – TA766: PEMBROLIZUMAB FOR ADJUVANT TREATMENT OF RESECTED MELANOMA WITH HIGH RISK OF RECURRENCE[11] 19
4.2. CASE STUDY 2 – TA725: ABEMACICLIB WITH FULVESTRANT FOR HR+, HER2-NEGATIVE ADVANCED BREAST CANCER AFTER ENDOCRINE THERAPY[12] 26
4.3. CASE STUDY 3 – TA629: OBINUTUZUMAB IN COMBINATION WITH BENDAMUSTINE FOR TREATING RITUXIMAB-REFRACTORY FOLLICULAR LYMPHOMA[13] 32
4.4. CASE STUDY 4 – TA653: OSIMERTINIB FOR TREATING LOCALLY ADVANCED OR METASTATIC EGFR T790M MUTATION-POSITIVE NON-SMALL-CELL LUNG CANCER[14] 40
4.5. CASE STUDY 5 – TA736: NIVOLUMAB FOR TREATING RECURRENT OR METASTATIC SQUAMOUS CELL CARCINOMA OF THE HEAD AND NECK AFTER PLATINUM-BASED CHEMOTHERAPY[16] 47
4.6. CASE STUDY 6 – TA975: TISAGENLECLEUCEL FOR TREATING RELAPSED OR REFRACTORY B-CELL ACUTE LYMPHOBLASTIC LEUKAEMIA IN PEOPLE AGED UP TO 25 YEARS[17] 50
4.7. CASE STUDY 7 – TA962: OLAPARIB FOR MAINTENANCE TREATMENT OF BRCA MUTATION-POSITIVE ADVANCED OVARIAN, FALLOPIAN TUBE OR PERITONEAL CANCER AFTER RESPONSE TO FIRST-LINE PLATINUM-BASED CHEMOTHERAPY[18] 55
4.8. CASE STUDY 8 – TA655: NIVOLUMAB FOR PREVIOUSLY TREATED SQUAMOUS NON-SMALL-CELL LUNG CANCER[19] 62
5. SUMMARY AND DISCUSSION 67
5.1. SUMMARY OF FINDINGS 67
5.2. DISCUSSION 72
5.2.1. Implications for cost-effectiveness analyses 72
5.2.2. Data limitations 74
5.2.3. Real-world survival 76
5.2.4. Limitations of our analyses 76
5.2.5. Further research 78
5.2.6. Conclusions 78
REFERENCES 80

TABLES

Table 1: Included and excluded case studies and rationale 13
Table 2: Cost-effectiveness results – TA766 23
Table 3: Disaggregated outcomes for TA766 24
Table 4: Cost-effectiveness results – TA725 30
Table 5: Disaggregated outcomes for TA725 31
Table 6: Cost-effectiveness results – TA629 38
Table 7: Disaggregated outcomes for TA629 39
Table 8: Cost-effectiveness results – TA653 45
Table 9: Disaggregated outcomes for TA653 46
Table 10: Cost-effectiveness results – TA975 53
Table 11: Disaggregated outcomes for TA975 54
Table 12: Cost-effectiveness results – TA962 60
Table 13: Disaggregated outcomes for TA962 60
Table 14: Cost-effectiveness results – TA655 65
Table 15: Disaggregated outcomes for TA655 66
Table 16: The change in cost-effectiveness results based on SACT data vs trial data 69

FIGURES

Figure 1: KM plots for OS and ToT outcomes in TA766 20
Figure 2: OS predictions used in TA766 and the DSU's SACT-based analysis 22
Figure 3: KM plots for OS and ToT outcomes in TA725 27
Figure 4: Modelled OS used in TA725 and the DSU's SACT analysis 28
Figure 5: Modelled ToT used in TA725 and the DSU's SACT analysis 29
Figure 6: Parametric model fits used in TA725 and the DSU's SACT analysis 30
Figure 7: KM plots for OS and ToT outcomes in TA629 33
Figure 8: Modelled OS in TA629 and the DSU's SACT analysis 34
Figure 9: Different OS curves fitted to the SACT data for TA629 34
Figure 10: KM plots from the trial data in TA629 35
Figure 11: Modelled OS and PFS used in TA629 and the DSU's SACT analysis 36
Figure 12: ToT data used in TA629 and the DSU's SACT analysis 37
Figure 13: KM plots for OS and ToT outcomes in TA653 41
Figure 14: Modelled OS in TA653 and the DSU's SACT analysis 41
Figure 15: Modelled PFS and ToT in TA653 42
Figure 16: Modelled ToT in TA653 and the DSU's SACT analysis 43
Figure 17: Modelled OS for the comparator arm using SACT data versus the DSU approach for TA653 44
Figure 18: OS KM plots from SACT compared to trial – TA736 48
Figure 19: The different OS extrapolations using week 20 piecewise models fitted to SACT compared to trial observed data and extrapolation – TA736 48
Figure 20: ToT KM plots from SACT compared to trial – TA736 49
Figure 21: OS KM plots from SACT compared to trial – TA975 51
Figure 22: Parametric fits for OS in TA975 and the DSU's SACT analysis 52
Figure 23: Parametric fits for EFS in TA975 and the DSU's SACT analysis 53
Figure 24: OS and ToT KM plots from SACT compared to trial – TA962 56
Figure 25: Parametric fits for OS in TA962 and the DSU's SACT analysis 57
Figure 26: Parametric fits for PFS2 in TA962 and the DSU's SACT analysis .. 58
Figure 27: Parametric fits for PFS in TA962 and the DSU's SACT analysis 58
Figure 28: OS and ToT KM plots from SACT compared to trial – TA655 63
Figure 29: Parametric fits for OS in TA655 and the DSU's SACT analysis 64
Figure 30: Parametric fits for PFS and ToT in TA655 and the DSU's SACT analysis 65

ABBREVIATIONS

Abbreviation	Description
AC	Appraisal Committee
CDF	Cancer Drugs Fund
DM	Distant metastases
DSU	Decision Support Unit
EFS	Event-free survival
ERG	Evidence Review Group
FAD	Final appraisal determination
FP-NMA	Fractional polynomial meta-analysis
HR	Hazard ratio
ICER	Incremental cost-effectiveness ratio
IPD	Individual-patient data
KM	Kaplan-Meier
LR	Locoregional recurrence
LY	Life year
MAA	Managed Access Agreement
NICE	National Institute for Health and Care Excellence
NuTH	Newcastle upon Tyne Hospitals NHS Foundation Trust
OS	Overall survival
PFS	Progression-free survival
QALY	Quality-adjusted life year
RF	Recurrence-free
RWD	Real-world data
RWE	Real-world evidence
SACT	Systemic Anti-Cancer Therapy
TA	Technology Appraisal
ToT	Time on treatment

1. EXECUTIVE SUMMARY

The National Institute for Health and Care Excellence (NICE) is committed to using the best available evidence, including real-world evidence (RWE), to inform decision making and make recommendations as relevant and appropriate as possible. In oncology-related appraisals, some interventions enter the Cancer Drugs Fund (CDF), where real-world data (RWD) are collected for a specific period of time. The intention is for RWD to create RWE to aid Appraisal Committees (ACs) when treatments placed in the CDF are reviewed. However, recent reviews have suggested that this RWE does not often make a significant contribution to decision-making. This may be due to committee preferences, or the limitations associated with the data made available for treatments placed in the CDF, such as the absence of data on comparator treatments, the limited availability of data on long term outcomes, and the aggregate nature of the data.

NICE's Health Technology Assessment Innovation Laboratory (HTA Lab) has collaborated with the Decision Support Unit (DSU) to illustrate how the data provided for treatments placed in the CDF could be used in their current form, and what future recommendations regarding the best use of such data might include.

The DSU studied eight cases to demonstrate how the RWE provided for treatments placed in the CDF could be used to produce cost-effectiveness estimates. The analyses used were based on aggregate data provided for the intervention placed in the CDF, with corresponding data for relevant comparators predicted using treatment effect estimates obtained from the pivotal trial. This approach required important assumptions around the modelling of overall survival (OS), progression-free survival (PFS), and time on treatment (ToT), which are described in the report. We acknowledge that other modelling techniques could be used, and other assumptions could be made, but this does not detract from the primary objectives of this report, which are to demonstrate what could be done with the data provided for treatments placed in the CDF to inform cost-effectiveness estimates, and what impact the use of these data could have on NICE decision-making.

Results show that, in general, survival observed for the interventions appraised is worse in the RWD compared to the trial data. When time on treatment is also reduced in the real world, estimates of cost-effectiveness may not be substantially impacted. However, in instances where survival was substantially reduced in the real world but time on treatment was not, estimates of cost-effectiveness could change dramatically. In addition, we demonstrate that in some instances substantial reductions in survival in the RWD can lead to considerable increases in incremental cost-effectiveness ratios, even when treatment durations are also reduced and relative treatment effects are assumed to be retained.

According to our analyses, we consider that using the RWD to inform estimates of cost-effectiveness could potentially have led to altered NICE recommendations in four of the eight case studies. However, we note important limitations associated with the RWD provided, and the strong assumptions associated with our analyses, which could impact the validity of the results and how these analyses would be interpreted by NICE ACs. Related to this, we consider circumstances in which the RWD-based analyses may be more likely to impact NICE decision-making. For instance, we believe that RWD-based analyses are more likely to impact decision-making when SACT data indicates reduced survival in the real-world population but similar time-on-treatment, which is likely to lead to worsened cost-effectiveness estimates even if it is assumed that the relative treatment effect observed in pivotal trials is retained.

We also consider enhancements in the data that may result in more reliable and impactful analyses. In particular, we note that if data were made available for comparators, and if access to patient-level data were permitted (allowing adjustments to be made for potential confounding factors), RWD-based analyses are likely to be more reliable and more acceptable to decision-makers.

2. INTRODUCTION

Managed Access Agreements (MAAs) between NHS England and pharmaceutical companies enable patient access to a drug for a limited period, during which time further evidence is gathered to allow National Institute for Health and Care Excellence (NICE) Appraisal Committees (ACs) to determine whether the treatment should be recommended for routine commissioning. The Cancer Drugs Fund (CDF) was introduced in 2011 for cancer treatments that had been rejected for routine commissioning by NICE[1, 2].

In 2016, the CDF was judged to be exceeding its allocated budget and needed urgent review. NHS England, NICE, Public Health England, and the Department of Health partnered to introduce the framework for the new CDF which offered a mechanism for conditional approval. Under the reformed CDF, drugs could be recommended for use within the CDF if the NICE AC decided that the treatment had the potential to be cost-effective at the manufacturer's proposed price, but evidence was too uncertain to permit a routine commissioning decision to be made, and if it was decided that, alongside longer follow up in the pivotal clinical trial, further data collection could address the uncertainty in the evidence. This new approach sought to provide:

- Access to promising new treatments, via MAAs, while further evidence is collected to address uncertainty.
- Interim conditional funding for all newly recommended cancer drugs, giving patients access to these treatments many months earlier than before the new CDF[1, 2].

When treatments are placed in the CDF, data collection arrangements are put in place to ensure that the data collected will address key areas of uncertainty identified by the AC. There are two main data collection sources for the new CDF: additional trial data with longer follow-up, and the national mandatory Systemic Anti-cancer Therapy (SACT) data collection submitted by all NHS England oncology service providers. A time frame is set up to allow meaningful data to be collected - normally up to two years[2] - and both data sources are presented to the AC at a subsequent NICE committee meeting, during which the AC makes final recommendations on whether the drug should be available for routine use in the English NHS[1, 3]. The SACT data presented

to the AC primarily consists of overall survival (OS) and time-on-treatment (ToT) curves for the CDF drug, for the period since it was placed in the CDF.

Compared to clinical trials, RWD collected in the SACT dataset are more reflective of UK clinical practice and could be expected to play a major role in decision making when treatments placed in the CDF come back to NICE for review. However, a recent study that looked into the first 24 drugs that exited the reformed CDF suggests that limited use was made of SACT data when these treatments were reappraised, and that economic modelling of subsequent treatments or treatment duration mainly used clinical trial data to overcome initial uncertainties[4]. The main reasons for the greater use of trial data were the inherent limitations in the way data are collected within SACT (lack of co-morbidity data and limited granularity of outcome data), the absence of comparator data, and the immaturity of survival data collected given the period during which treatments are available via the CDF[4]. This is in addition to the absence of coherent analytical plans to assess comparative effectiveness of SACT data to support the reduction of uncertainties[5].

In 2024, the Decision Support Unit (DSU) reported pilot cost-effectiveness analyses that used aggregate SACT data for four treatments placed in the CDF. These analyses demonstrated a tendency for clinical outcomes to be worse in the real world than in clinical trials, which, in turn, impacted on cost-effectiveness outcomes.

In this report we extend the analyses to eight further case studies of CDF reviews where trial and aggregate SACT data were available. For each case study, we assess whether or how OS data from the trial and SACT data differ, and whether or how ToT differed between the two data sources. We incorporate evidence from the SACT data into the economic model, demonstrating how this can be done, including an exploration of how comparative effectiveness could be estimated given that evidence on comparators is not routinely included in the data provided to NICE for treatments placed in the CDF. We report whether the incremental cost-effectiveness ratio (ICER) using SACT data differs from that using the trial data.

In Section 3 of this report we describe how case studies were selected, and present the methods used for incorporating evidence from the aggregate SACT data provided to NICE into cost-effectiveness models. Section 4 presents results for each of the eight case studies. Section 5 begins with a summary of the results observed across the case studies, identifying potential patterns before providing a discussion of the lessons learned from this investigation.

3. METHODS

3.1. METHODS FOR TOPIC SELECTION

Decisions regarding topic selection were made via a Steering Group comprising members from the NICE HTA Lab, NICE consultant clinical advisors, former and current NICE AC Chairs, and the authors of this report. The 38 technology appraisals (TAs) that had exited the CDF as of November 2024 were reviewed, of which three with terminated or withdrawn guidance were not further considered. When the previous pilot DSU study on this topic was conducted, the four case studies selected were from a group of 11 TAs that fell within the timeframe of that review. Aside from the four TAs selected, the other seven considered did not adequately satisfy the inclusion criteria set for the pilot study (criteria then included: high number of OS and ToT events; SACT data that could have been used but was not explored in the appraisal; coverage of various cancer types). TA766 was excluded from the pilot study due to the structure of the economic model – a state transition model with 4 states was used, and it is difficult to incorporate evidence from SACT in such a model. In the present study, it was decided to re-consider this TA, as an example that considered how SACT data could be used in cases where a standard partitioned survival model was not used. In addition, at the time of our pilot study, TA872 was excluded because committee papers were not available – these papers were available at the time of the present study, and so this TA was re-considered for inclusion. Hence, 26 TAs were considered for case study selection.

Identification of case studies

Case studies for inclusion in the present study were identified based on the following criteria:

i) A relatively large sample size of patients in the SACT data (roughly 100 people or more) except for one case study to explore an example with a low sample size, demonstrating potential additional uncertainties associated with this;
ii) SACT data were available, there appeared to be differences between the SACT data and the trial data, but the SACT data were seemingly not used as a key input to AC decision-making during the review appraisal;

iii) The economic model was available and structured in such a way that allowed SACT data to be relatively easily incorporated into the cost-effectiveness analysis;
iv) Case studies covering various cancer types and complex features including one-time and maintenance therapies, treatment stopping rules, and non-standard modelling approaches.

The second criterion was selected to prioritise appraisals where SACT data for OS and ToT were available and showed differences compared to the clinical trial, but these had not been explored in the CDF review.

Table 1 shows the 26 TAs considered and the rationale for inclusion or exclusion in this study. The eight selected case studies for analysis in this report included TA766 as an example of an economic model that did not use a partitioned survival structure, TA629 as a case where the SACT data had a sample size less than 100, and TA975 where the technology cost is accrued at a single time-point at the start of treatment. Based on an assessment of the 26 TAs, it was considered that eight case studies offered a sufficiently large and varied sample to allow the potential for general conclusions to be drawn.

Table 1: Included and excluded case studies and rationale

TA	Indication	Included?	Rationale
1. TA524	Hodgkin lymphoma	No	No SACT collected
2. TA531	Lung cancer	No	Follow-up period too short to present useful SACT data
3. TA629	Non-Hodgkin lymphoma	Yes	Satisfied all criteria, chosen as a case of low sample size in SACT (N=92)
4. TA653	Lung cancer	Yes	Satisfied all criteria
5. TA655	Lung cancer	Yes	Satisfied all criteria
6. TA683	Lung cancer	No	No mention of SACT data collected in the appraisal papers
7. TA684	Lung cancer	No	No SACT KM data for OS or ToT were reported
8. TA687	Breast cancer	No	No SACT KM data for OS were reported

9. TA691	Merkel cell carcinoma	No	Low SACT sample size
10. TA692	Urothelial cancer	No	No SACT KM data for OS or ToT were reported
11. TA713	Lung cancer	No	Low SACT sample size
12. TA725	Breast cancer	Yes	Satisfied all criteria
13. TA736	Squamous cell carcinoma of head and neck	Yes	Satisfied all criteria
14. TA739	Urothelial cancer	No	Low sample size
15. TA766	Melanoma	Yes	Selected as example of non-partitioned survival model
16. TA770	Lung cancer	No	No mention of SACT data collected
17. TA872	Lymphoma	No	Unclear level of SACT data accrued
18. TA897	Multiple myeloma	No	SACT data used in decision making
19. TA908	Fallopian tube or peritoneal cancer	No	Low SACT sample size
20. TA939	Cervical cancer	No	No mention of SACT data collected in the papers
21. TA946	Fallopian tube or peritoneal cancer	No	Low SACT sample size
22. TA962	Fallopian tube or peritoneal cancer	Yes	Satisfied all criteria (also chosen as an example of maintenance treatments)
23. TA967	Hodgkin lymphoma	No	SACT data used in decision making
24. TA975	Lymphoblastic leukaemia	Yes	Satisfied all criteria (also chosen as an example of one-off treatments)
25. TA1007	Fallopian tube or peritoneal cancer	No	SACT data used in decision making
26. TA1018	Myelofibrosis	No	Low sample size and SACT data used in decision making

Abbreviations: KM - Kaplan Meier; OS - overall survival; SACT - Systemic anti-cancer therapy; TA - technology appraisal; ToT - time on treatment

3.2. DATA EXTRACTION FROM SACT REPORTS

In CDF reviews, SACT data on OS and ToT are provided in the form of summary statistics and Kaplan-Meier (KM) survival functions. We used Guyot *et al*'s commonly used survival data reconstruction method to recreate pseudo individual-level patient data (IPD), in order that survival models could be fitted to the SACT data and

incorporated in the economic models[6]. Estimated survival probabilities were first derived from each KM curve using plot digitisation software, WebPlotDigitizer (https://apps.automeris.io/wpd/). Then, the pseudo IPD were created using the R package IPDfromKM Shiny application[7].

Whenever possible, other parameters including age, gender, and subsequent treatments were extracted from the SACT reports and used to inform parameter values in the economic model, when appraisal documents indicated that the AC deemed these values to be reflective of the target population.

3.3. METHODS FOR THE COST-EFFECTIVENESS ANALYSIS

For the selected case studies, NICE shared with the DSU the unredacted committee papers and the economic models used for decision making (seven were partitioned survival models and one state-transition model; all built in Microsoft Excel).

To estimate lifetime OS and ToT for the CDF treatment, based on SACT data, we fitted the parametric distribution favoured by the AC in the final guidance document to the pseudo IPD for OS and ToT from SACT. This was done using the 'streg' package in Stata version 18[8]. This assumes that the preferred parametric distributions for OS and ToT would stay the same, irrespective of whether models are being fit to trial or SACT data. This is a simplification, and in practice a full model selection process may be preferred[9]. However, for the purposes of the illustrative case studies included in this report we deemed this simplification to be reasonable, especially given that whilst the absolute survival curves may differ according to the SACT and trial data sources, the nature of the curves may be similar. That is, due to the underlying nature of the disease, the shape of the hazard function may reasonably be considered to be similar irrespective of whether models are fitted to SACT or trial data.

Given that SACT data for the comparator are not provided in CDF reviews, assumptions had to be made to derive survival curves for the comparator. A variety of approaches could be taken, with each having advantages and disadvantages. For instance, hazard ratios (HRs) reported from the pivotal trial under consideration could be applied to the intervention group parametric OS and ToT curves fitted to the pseudo

IPD from SACT, to derive curves for the comparator group. This would enforce a constant treatment effect (i.e., proportional hazards) assumption, which could be problematic if the AC did not believe that such an assumption was appropriate. In particular, we wished to avoid assuming a constant treatment effect, because this is often deemed inappropriate in TAs due to violation of the proportional hazards assumption and uncertainty around long-term relative treatment effects[10].

3.3.1. Methods for modelling OS and ToT

For our exploratory analyses, for each case study, we derived OS and ToT curves for the comparator using the following approach:

i) In each case, the company had fitted the AC-preferred parametric models to the trial data, and from the models had derived cumulative survival probabilities (or the probability of remaining on treatment for ToT) over time (for each cycle of the model), for each treatment arm.

ii) From the survival (and ToT) probabilities over time, we calculated the hazard function over time.

iii) For each discrete cycle in the economic model, we then calculated the HR, by dividing the hazard in the intervention group by the hazard in the comparator group – resulting in an estimate of a time-dependent HR.

iv) We then applied this time-dependent HR to the OS (and ToT) survival curves fitted to the pseudo IPD from SACT for the intervention group, in order to derive curves for the comparator group (in the survival curves presented in Section 4, these are referred to as 'if SACT' curves, reflecting outcomes that we predict would have been observed if we had comparator data for the SACT population).

This approach makes the strong assumption that the relative treatment effect estimated from the preferred parametric models fitted to the trial data (i.e., the treatment effect estimated from a head-to-head comparison) would be replicated in the NHS population treated with the new treatment while it was in the CDF.

In addition, from a technical perspective our derivation approach is somewhat problematic, because, where the AC preferred accelerated failure time (AFT) parametric models, we apply time-dependent HRs to AFT models, whilst treatment effects in AFT models should be measured on the time scale rather than the hazard scale. However, the situation here is complex. Consider a situation where the AC preferred parametric models independently fitted to the intervention and control groups. Because the models are independently fitted (and indeed may have different distributions, and could be either proportional hazards or AFT models), there is no model-based 'true' HR or time ratio that summarises the treatment effect, because the survival probabilities for each treatment arm were estimated from different models.

Importantly, our approach assumes that the relative difference in the hazards derived from curves fitted to the trial data and preferred by the AC would be replicated in the SACT-based analyses. This preserves the relative differences estimated by the trial-based models, and allows the relative difference (i.e., the treatment effect) to change over time. We believe that this is a valid approach, since trial-based analyses represent the best estimates of relative treatment effects. As previously mentioned, an alternative would be to use a single trial-based HR to derive the comparator curves, but this would enforce a constant treatment effect assumption which is frequently rejected in NICE appraisals[10].

3.3.2. Methods for modelling PFS

SACT does not provide data on progression-free survival (PFS), but this is frequently a key component of economic models used to inform decisions for CDF treatments. Therefore, PFS curves estimated for the SACT population had to be derived, based on the SACT data on OS and ToT that were provided. In each case study we considered two approaches to derive PFS 'if SACT' curves for each treatment arm:

i) Calculating time-dependent HRs between the AC-preferred trial-based OS and PFS parametric models used in the CDF review economic model. These HRs were then applied to the OS curves estimated for the SACT population, to derive PFS outcomes for the SACT-based analysis, assuming the relationship between OS and PFS from the trial would be replicated in the SACT population.

ii) Calculating time-dependent HRs between the AC-preferred trial-based PFS and ToT parametric models used in the CDF review economic model. These HRs were then applied to the ToT curves estimated for the SACT population, to derive PFS outcomes for the SACT-based analysis, assuming the relationship between PFS and ToT from the trial would be replicated in the SACT population.

3.3.3. Other assumptions applied in the DSU analysis

Constraints were applied whenever required to ensure that ToT estimates never exceeded PFS estimates where a drug was licensed to be taken until progression, or where the extrapolations from trial data showed this to be the case and the views of clinical experts expressed in appraisal documents did not mention treatment use after disease progression. Other parameters such as costs (excluding treatment costs associated with ToT), utilities and various other modelling assumptions were kept the same as those used in the NICE CDF re-appraisal.

Given the time constraints for this project, results are only reported in the form of deterministic ICERs with disaggregated outcomes for life years (LYs), quality-adjusted life years (QALYs), and costs.

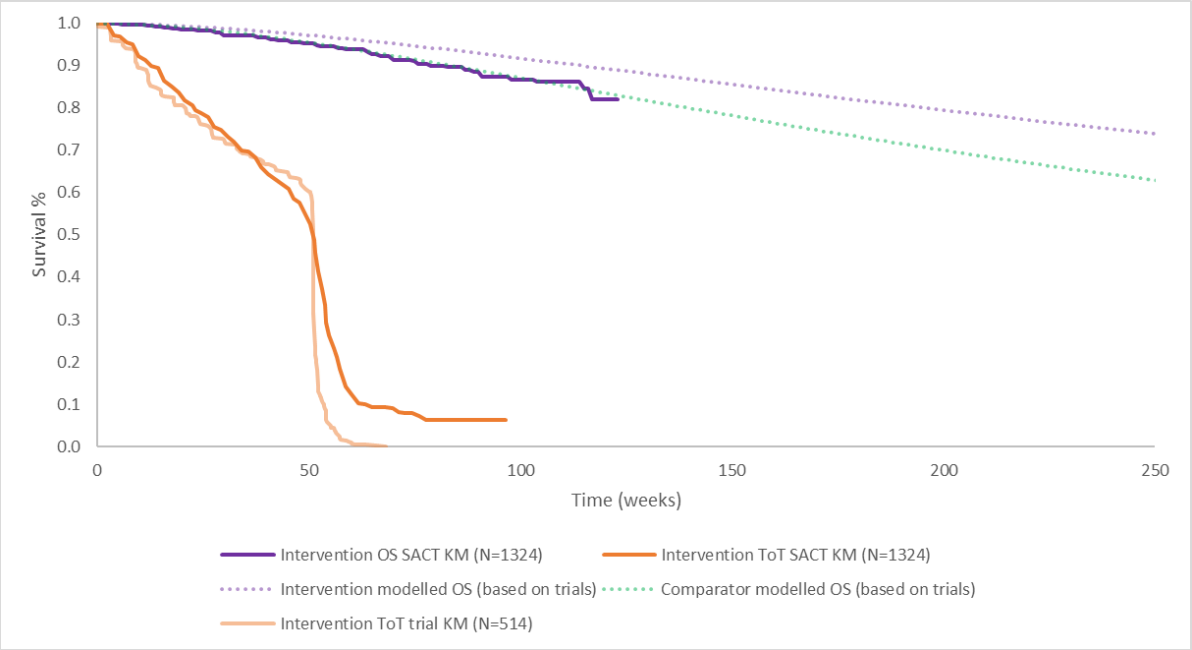
4. COST-EFFECTIVENESS RESULTS

4.1. CASE STUDY 1 – TA766: PEMBROLIZUMAB FOR ADJUVANT TREATMENT OF RESECTED MELANOMA WITH HIGH RISK OF RECURRENCE[11]

TA766 compared pembrolizumab with placebo for the adjuvant treatment of stage III melanoma in adults with lymph node involvement who have had complete surgical resection. The median age of patients in the SACT cohort was higher than those in the pivotal KEYNOTE-054 trial (64 years vs. 54 years). A higher number of patients were assessed to have an ECOG performance status (PS) of 0 in the trial compared with SACT (94.4% vs. 69%), and the proportion of patients with a BRAF V600 positive mutation was lower in the SACT dataset (19% versus 47.5% in the trial).

Figure 1 shows KM plots for OS and ToT from SACT and the pivotal clinical trial. In this appraisal a state-transition model was used, and therefore OS was indirectly modelled based on transition probabilities between health states, estimated using trial data. OS outcomes for patients treated with pembrolizumab in the SACT dataset appear worse than those estimated using the trial data, and in fact were similar to the trial-based OS curves estimated for the comparator. ToT for pembrolizumab was similar between the trial and SACT populations, although approximately 10% of SACT patients appear to have continued treatment beyond 12 months, whereas a 12-month stopping rule was used in the clinical trial, was enforced in the economic model, and is referred to in the marketing authorisation for pembrolizumab. This is surprising and the reasons for this are unclear. Notably, in the CDF re-appraisal the AC-preferred base case analysis used SACT data as the source for assumptions about subsequent treatments.

Figure 1: KM plots for OS and ToT outcomes in TA766



The methodology for deriving economic model inputs for PFS and OS described in Section 3.3 could not be applied in this case study because the economic model was not a partitioned survival model; so all cost-effectiveness outcomes are driven by the transition probabilities between the health states constituting the model. These were recurrence-free (RF), locoregional recurrence (LR), distant metastases (DM), and death. There were six allowed transitions; three allowing transitions to death (RF to death, LR to death, DM to death), and three to states other than death (RF to LR, RF to DM, LR to DM).

In keeping with the illustrative nature of the analyses contained within this report, we used a simplistic but practical approach to perform a cost-effectiveness analysis consistent with the OS data provided by SACT for this case study, by applying multipliers to the transition probabilities included in the economic model such that the resulting OS predicted by the model resembled the SACT OS KM curve. For each time cycle included in the economic model, the squared difference between the modelled OS probability (based on the trial data) and the corresponding estimate from the SACT OS KM data was calculated. This was done for all model cycles up to the end of the follow-up period of the SACT OS KM data (123 weeks). Then all these differences were averaged to give a mean of squared differences.

The Excel Solver functionality was then used to generate a multiplier that could be applied either to the three death transition probabilities in one analysis (multiplier = 1.86), or to all the six transition probabilities in a second analysis (multiplier = 1.34), to achieve the least mean of squared differences between the OS predicted by the economic model and the SACT OS KM data. This means that the model was calibrated to ensure similar OS predictions to those observed in the SACT population. We also applied the multiplier to the transition probabilities for the comparator, such that the relative treatment effect observed in the trial was approximately retained.

As acknowledged in Section 3.3 for the methods used for all the other case studies included in this report (i.e., those that used partitioned survival models), alternative approaches could have been used to derive SACT-based model inputs for this case study. The approach we used represents a pragmatic and achievable method for using SACT data in a state-transition economic model in the context of a time-limited technology appraisal. However, the calibration technique is simplistic and involves important limitations – for example, the multipliers were applied equally to all the transition probabilities, and the same weight was given to all the KM estimates regardless of the sample size involved at each time point.

ToT KM data from SACT were used in place of the trial data in our SACT-based analyses.

Figure 2 shows the OS predictions used in the original CDF review and the base case of our SACT-based analyses.

Figure 2: OS predictions used in TA766 and the DSU's SACT-based analysis

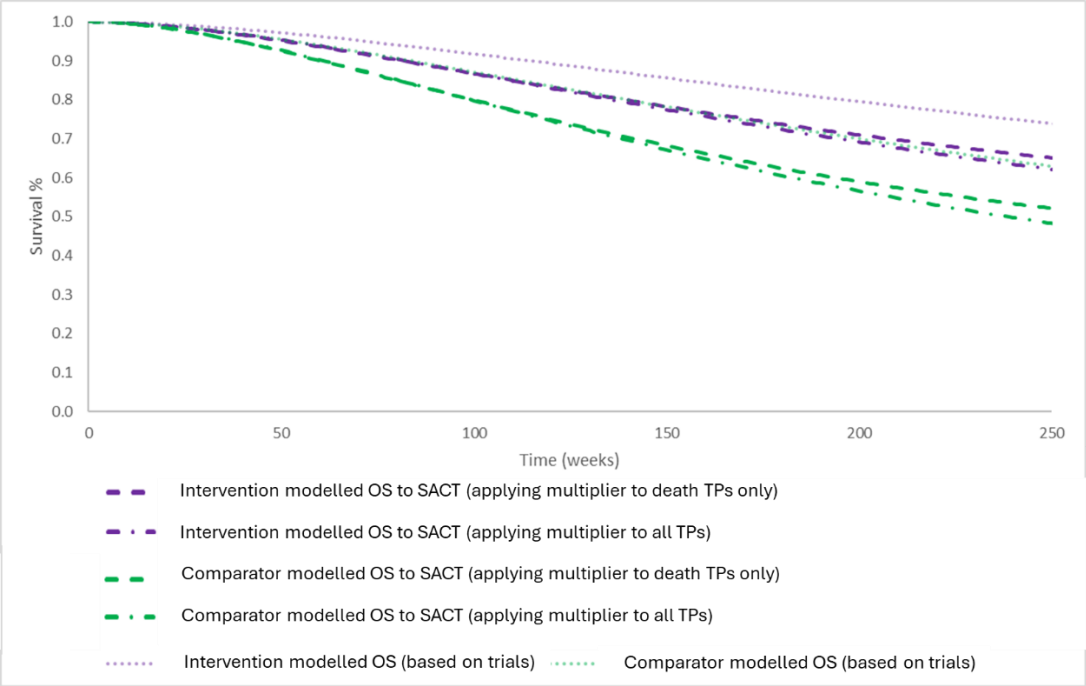


Table 2 presents the cost-effectiveness results using the AC preferences across various model inputs and the discounted price for pembrolizumab, based on the trial data (this is the ICER value reported in the final appraisal determination (FAD) document), and using the SACT data as described above in two alternative base cases.

Table 2: Cost-effectiveness results – TA766

Technologies	Total costs (£)	Total LY*	Total QALYs	Incr. costs (£)	Incr. LY*	Incr. QALYs	ICER (£/QALY)
Trial data used in the CDF review							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	14.38	[redacted]	[redacted]	[redacted]	[redacted]	£26,493
SACT data applying a multiplier of 1.86 to death transition probabilities							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	13.48	[redacted]	[redacted]	[redacted]	[redacted]	£27,667
SACT data applying a multiplier of 1.34 to all transition probabilities							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	11.06	[redacted]	[redacted]	[redacted]	[redacted]	£27,087

Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYs: life years; QALYs: quality-adjusted life years
*Undiscounted

Disaggregated results are shown in Table 3. Analyses based on the SACT data provide broadly similar ICERs to those based on the trial data. Applying the survival multiplier only to death probabilities resulted in less time in the DM (distant metastases) state owing to the higher DM to death transition probabilities compared to death from other states – this was also reflected by less treatment and management costs attributable to the DM state. Applying the survival multiplier to all transition probabilities had a considerable impact on time spent in the RF (recurrence-free) state, as this state has three outgoing transitions to the other health states which all increased in this analysis.

Table 3: Disaggregated outcomes for TA766

Outcome	Trial			SACT applying multiplier of 1.86 to death probabilities			SACT applying multiplier of 1.34 to all probabilities		
	Intervention	Comparator	Difference	Intervention	Comparator	Difference	Intervention	Comparator	Difference
LYs RFS	[redacted]	12.07	[redacted]	[redacted]	12.07	[redacted]	[redacted]	9.02	[redacted]
LYs LR	[redacted]	0.44	[redacted]	[redacted]	0.42	[redacted]	[redacted]	0.39	[redacted]
LYs DM	[redacted]	1.87	[redacted]	[redacted]	0.99	[redacted]	[redacted]	1.65	[redacted]
Total LYs	[redacted]	14.38	[redacted]	[redacted]	13.48	[redacted]	[redacted]	11.06	[redacted]
QALYs RFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs LR	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs DM	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALY lost due to ageing and AEs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs (RFS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs (LR)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Treatment costs (DM)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Disease management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE - adverse event; DM - distant metastasis; LR - locoregional; LY - life year; QALY - quality adjusted life year; RFS – recurrence-free survival

4.2. CASE STUDY 2 – TA725: ABEMACICLIB WITH FULVESTRANT FOR HR+, HER2-NEGATIVE ADVANCED BREAST CANCER AFTER ENDOCRINE THERAPY[12]

TA725 compared abemaciclib with fulvestrant (intervention) versus exemestane and everolimus (comparator) for treating hormone receptor-positive, human epidermal growth factor receptor 2 (HER2)-negative, locally advanced or metastatic breast cancer after prior endocrine therapy. The pivotal trial compared the intervention to fulvestrant alone which was not considered as a relevant comparator for the decision problem, so the OS extrapolations for both the intervention and the relevant comparator were estimated using a fractional polynomial network meta-analysis (FP NMA) built on an indirect treatment comparison conducted by the company.

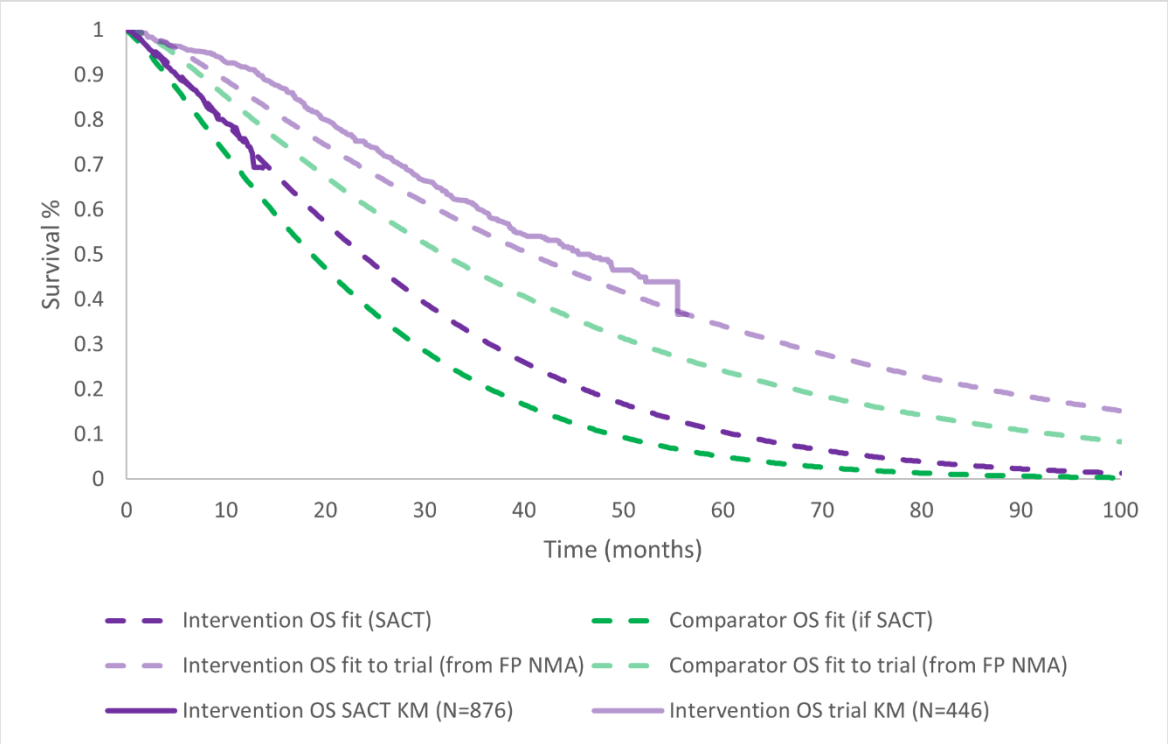
The median age of patients in the SACT cohort was higher than those in trial (65 years vs. 59 years), and a higher number of patients had an ECOG PS of 0 in the trial compared with SACT (59.2% vs. 31%). The AC concluded that the SACT data was immature and that clinical-effectiveness data from the trial was more appropriate for decision making.

Figure 3 shows KM plots for OS and ToT from both SACT and the pivotal trial. OS outcomes from the SACT dataset appear to be noticeably worse compared to those observed in the trial, whereas ToT appears to be more similar.

Figure 3: KM plots for OS and ToT outcomes in TA725

For the trial-based economic model used in the appraisal, a Weibull distribution was used to model OS for the intervention. We used the same approach, fitting a Weibull model to the pseudo IPD constructed from the SACT data, and the OS curve for the comparator was derived as described in Section 3.3.1. Figure 4 shows the trial-based and SACT-based OS extrapolations. It is notable that the trial-based OS curve for the intervention data does not provide a good visual fit to the trial KM estimates, because the curve was derived from the FP-NMA rather than by directly fitting a parametric survival model to the intervention arm of the trial.

Figure 4: Modelled OS used in TA725 and the DSU's SACT analysis



In the CDF review, ToT for both the intervention and the comparator was estimated by applying an HR to PFS curves. However, the HR value was highly uncertain and a significant point of contention in the AC discussions. The ToT extrapolation preferred by the AC for the intervention was actually lower than ToT estimated using either the SACT or trial data. For simplicity, and to be consistent with the preferences of the AC, we decided to retain the ToT extrapolation used in the appraisal. This implied that the PFS extrapolations used in the appraisal should also be retained, given that ToT was estimated to be a function of PFS. The ToT curves are illustrated in Figure 5.

Figure 5: Modelled ToT used in TA725 and the DSU's SACT analysis

Figure 6 presents the extrapolations used for OS, PFS and ToT in the trial-based analyses used in the appraisal, and in our SACT-based analyses. We note that although we intended to use the same PFS and ToT fits that were used in the appraisal, these crossed the SACT-based OS curves and therefore had to be capped. Hence, the analysis based on SACT resulted in a shorter time spent in OS, PFS and on treatment, for both the intervention and the comparator. In addition, the median age (65) of patients reported from SACT differed from the average age in the trial (59). We incorporated this in the model, but it had minimal effect because life tables were not used to cap OS probabilities and age was only used to estimate age-adjusted utility multipliers.

Figure 6: Parametric model fits used in TA725 and the DSU's SACT analysis

Table 4 presents the cost-effectiveness results using a set of the AC preferences (the FAD illustrates that some combinations of plausible assumptions gave ICERs over £30,000 per QALY gained, and a single set of AC preferences was not reported) based on the trial data, and using the SACT data.

Table 4: Cost-effectiveness results – TA725

Technologies	Total costs (£)	Total LY (und*)	Total QALYs	Incr. costs (£)	Incr. LY (und*)	Incr. QALYs	ICER (£/QALY)
Trial data used in the CDF review							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£32,586
SACT data							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	£34,289

Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYs: life years; QALYs: quality-adjusted life years
*Undiscounted

Disaggregated results are shown in Table 5. Analyses based on the SACT data provide broadly similar ICERs to those based on the trial data because the decrease in the QALY differential was offset by a decrease in the treatment costs for the intervention arm, as well as a decrease in subsequent treatment costs associated with the reduced time spent in the progressive disease (PD) health state.

Table 5: Disaggregated outcomes for TA725

Outcome	**Trial**			**SACT**		
	Intervention	**Comparator**	**Difference**	**Intervention**	**Comparator**	**Difference**
LYs PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LYs PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Subsequent therapy costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PFS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Pain management costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; LYs: life years; QALYs: quality-adjusted life years; PFS: progression-free survival; PD: progressed disease

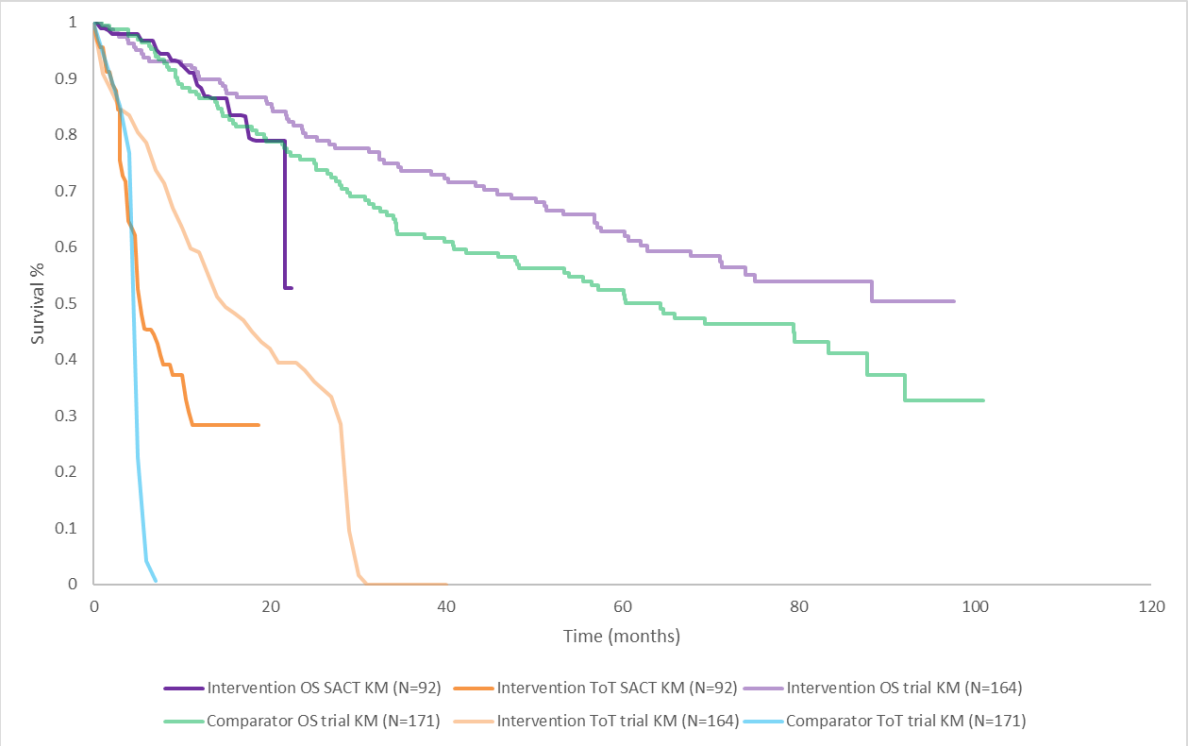
4.3. CASE STUDY 3 – TA629: OBINUTUZUMAB IN COMBINATION WITH BENDAMUSTINE FOR TREATING RITUXIMAB-REFRACTORY FOLLICULAR LYMPHOMA[13]

In the CDF review of TA472, TA629 compared obinutuzumab plus bendamustine to bendamustine alone for treating rituximab-refractory follicular lymphoma. The main clinical uncertainty identified by the committee in the original appraisal was the survival benefit for obinutuzumab with bendamustine.

The length of follow-up in SACT was limited, in part because the company provided further trial data earlier than intended, triggering a quicker CDF re-appraisal. The AC accepted that the SACT population included a higher proportion of people with multiple-relapsed disease, who therefore had a poorer prognosis compared with the trial population; and concluded that despite being insightful, SACT data were not robust enough to be used in economic modelling.

Figure 7 shows KM plots for OS and ToT, from the SACT data and from the pivotal trial. SACT data was less mature and indicated poorer survival. However, this was informed by low numbers of patients and high levels of censoring (out of 92 patients, 13 had died and 79 were censored). Observed ToT also appeared lower in the SACT data.

Figure 7: KM plots for OS and ToT outcomes in TA629



In the TA, Weibull distributions were used to model OS for the intervention based on the trial data. Hence, we fitted Weibull models to the SACT data, and the OS curve for the comparator was derived as described in Section 3.3.1. Figure 8 shows the trial-based and SACT-based extrapolations for OS. We note that the large dip in the OS KM estimate from the SACT data between 18 and 21 months had an impact on the Weibull model fitted to the SACT data, whilst being informed by only five patients. Hence a scenario analysis was performed fitting the parametric curve to only the first 18 months of the SACT data. Figure 9 illustrates the impact of taking this approach on the parametric curve fit.

Figure 8: Modelled OS in TA629 and the DSU's SACT analysis

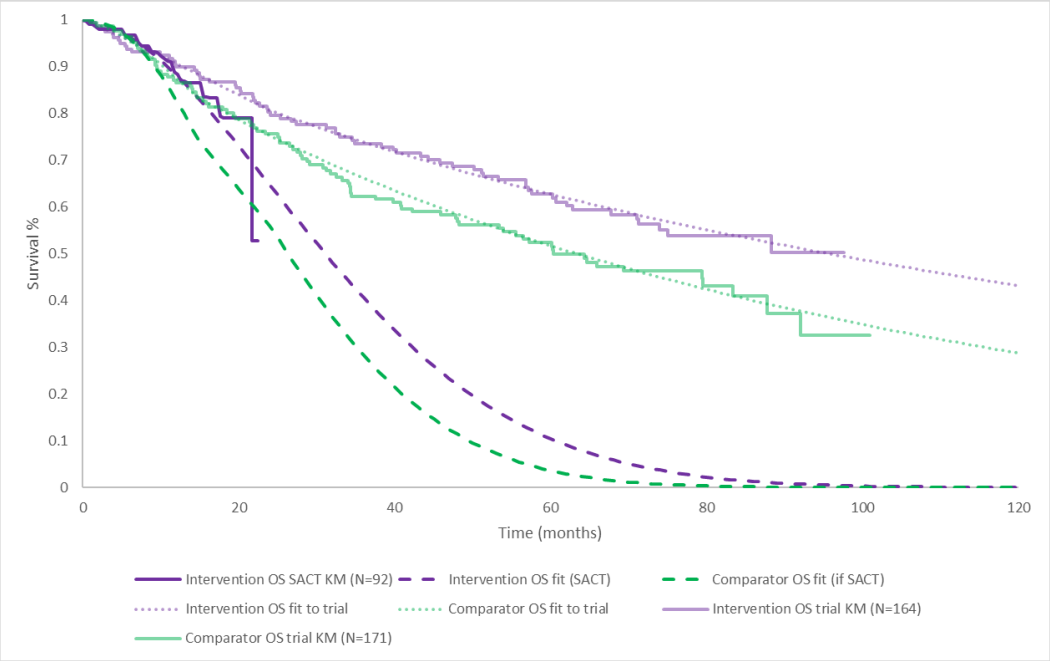
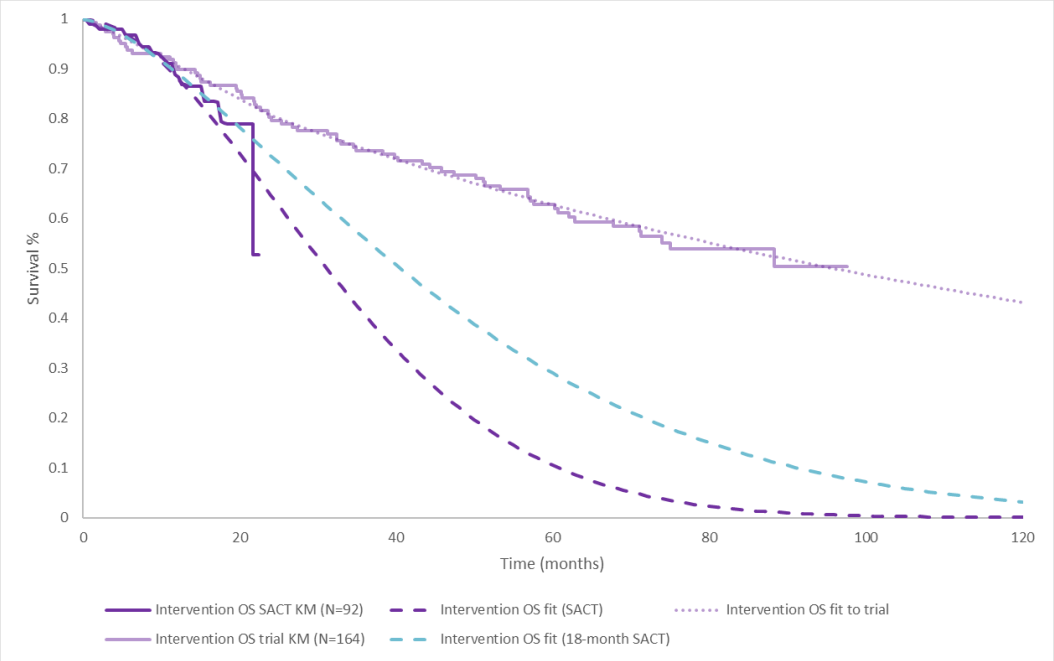


Figure 9: Different OS curves fitted to the SACT data for TA629



For both treatment arms there were stopping rules used in the trial and in the associated economic model. These were 2.5 years for the intervention and 6 months for the comparator (as shown in Figure 10). Because of the stopping rules and because, visually, PFS outcomes seem more related to OS than ToT, we simulated PFS outcomes for the SACT population using approach 1 described in Section 3.3.2,

rather than approach 2 (that is, assuming a HR between OS and PFS, rather than between ToT and PFS).

Figure 10: KM plots from the trial data in TA629

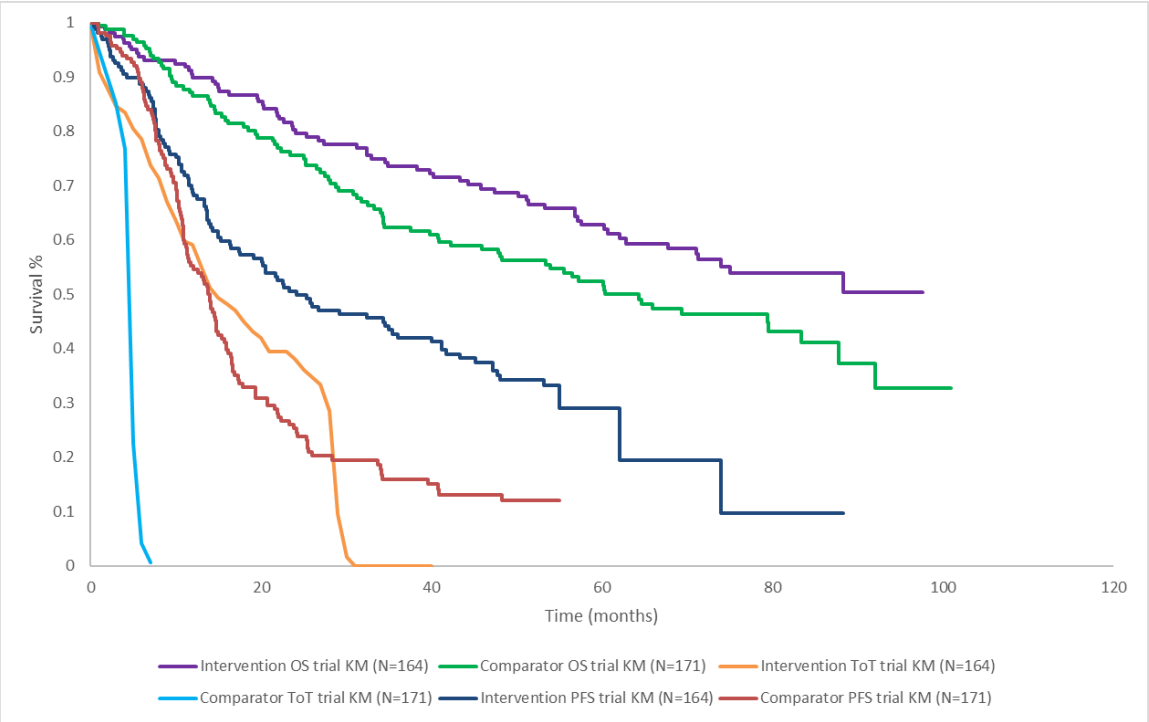
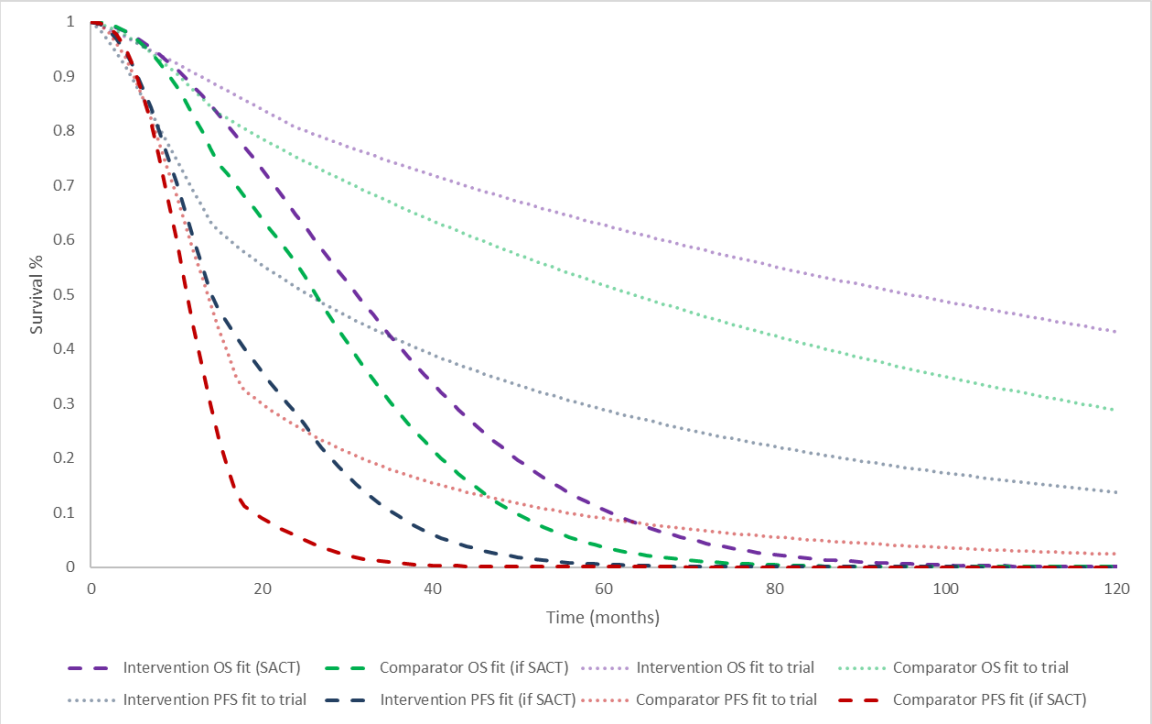


Figure 11 compares the OS and PFS extrapolations used in the AC's preferred base case (based on trial data), and using our SACT-based analysis.

Figure 11: Modelled OS and PFS used in TA629 and the DSU's SACT analysis



In the appraisal, KM plots for ToT were mature and were used directly to model treatment costs. The SACT ToT KM plot was less mature compared to the trial data, and we conducted two analyses in the presence of this uncertainty: (i) in a 'base case', we assumed that treatment did not continue beyond the SACT follow-up period; (ii) as a scenario analysis we assumed a gradual decrease in the ToT curve up to the stopping rule time-point of 2.5 years. These scenarios are shown in Figure 12. For the comparator, we used the same ToT curves that were used in the CDF review, as it was deemed that the 6-month stopping rule would limit any impact alternative data could have. Finally, the median age (65) of patients reported from SACT was used in place of the average age from the trial (62).

Figure 12: ToT data used in TA629 and the DSU's SACT analysis

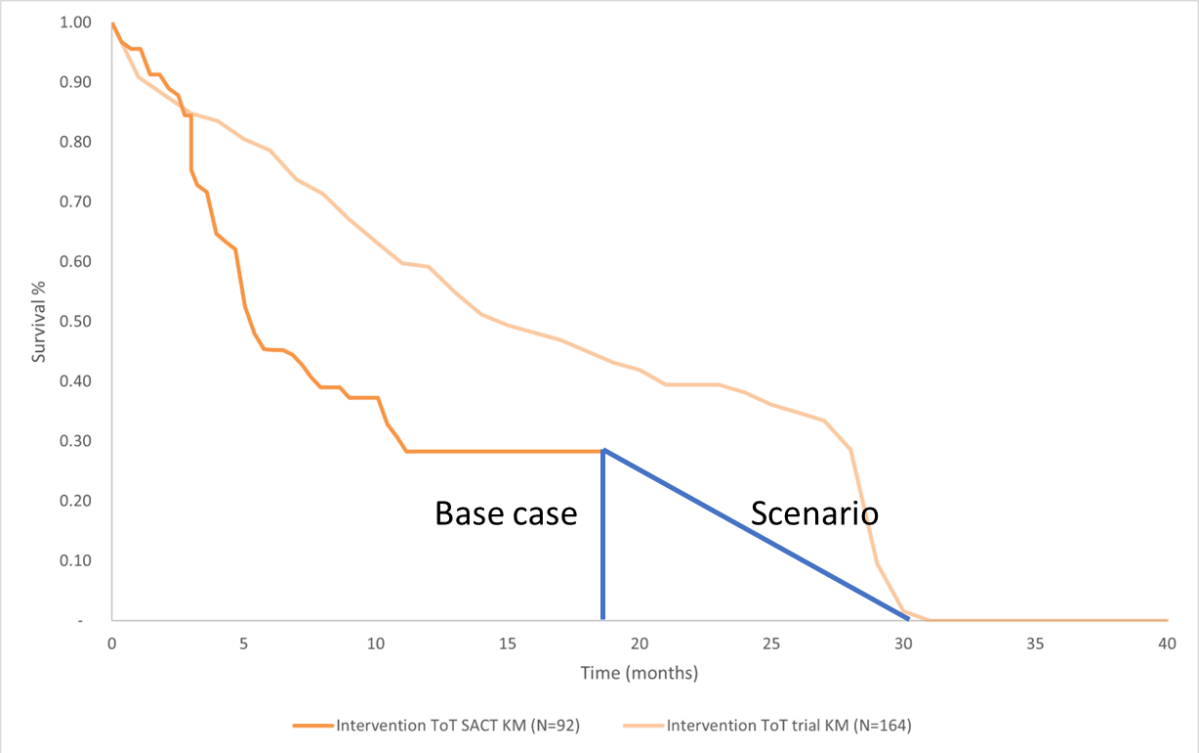


Table 6 presents the cost-effectiveness results using the AC preferences based on the trial data (this is the ICER reported in the final guidance), and using the SACT data. Comparing the SACT base case to that based on the trial, there was a 79% decrease in incremental QALYs gained, and a 36% decrease in incremental costs. This resulted in a tripling of the ICER. Assuming the more optimistic fit for SACT OS (by fitting Weibull models to 18-month SACT data) still results in a doubling of the ICER. Assuming that a proportion of the SACT population would stay on treatment until the stopping rule timepoint of 2.5 years increases the SACT-based ICER by approximately £3,700.

Table 6: Cost-effectiveness results – TA629

Technologies	Total costs (£)	Total LY (und*)	Total QALYs	Incr. costs (£)	Incr. LY (und*)	Incr. QALYs	ICER (£/QALY)
Trial data used in the CDF review							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	7.69	[redacted]	[redacted]	[redacted]	[redacted]	£15,045
SACT data base case							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	2.41	[redacted]	[redacted]	[redacted]	[redacted]	£45,043
SACT scenario assuming Weibull fit to the first 18 month of SACT OS data							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	3.15	[redacted]	[redacted]	[redacted]	[redacted]	£28,240
SACT scenario assuming gradual decrease to the stopping rule timepoint for the intervention							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	2.41	[redacted]	[redacted]	[redacted]	[redacted]	£48,765

Disaggregated results are shown in

Table 7. Time spent in both the PFS and PD health states were impacted substantially by the use of SACT data, resulting in vastly reduced QALY estimates.

Table 7: Disaggregated outcomes for TA629

Outcome	Trial			SACT		
	Intervention	Comparator	Difference	Intervention	Comparator	Difference
LYs PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LYs PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Obinutuzumab costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Bendamustine costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PFS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

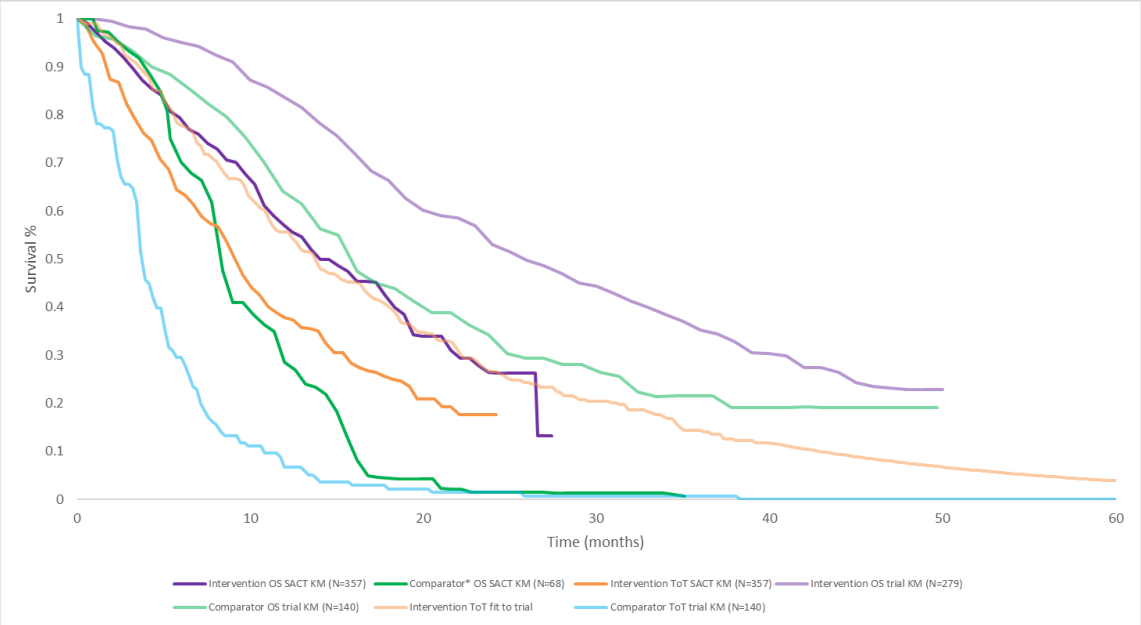
Abbreviations: AE: adverse event; LYs: life years; QALYs: quality-adjusted life years; PFS: progression-free survival; PD: progressed disease

4.4. CASE STUDY 4 – TA653: OSIMERTINIB FOR TREATING LOCALLY ADVANCED OR METASTATIC EGFR T790M MUTATION-POSITIVE NON-SMALL-CELL LUNG CANCER[14]

TA653 compared osimertinib (intervention) versus platinum doublet chemotherapy (PDC) (comparator) for treating locally advanced or metastatic EGFR T790M mutation-positive non-small-cell lung cancer. This case study represented an unusual case where SACT data for OS were available both for the intervention and, to some extent, for the comparator – with the data for the comparator taken from a previous CDF review in which the manufacturer had gained access to additional SACT data relevant to that appraisal.[15] The data for the comparator were by no means perfect – the data were not from exactly the same population (e.g., T790M status was unknown) and the make-up of the comparator treatment was not identical (e.g., the comparator group was not exclusively made up of patients who received PDC).

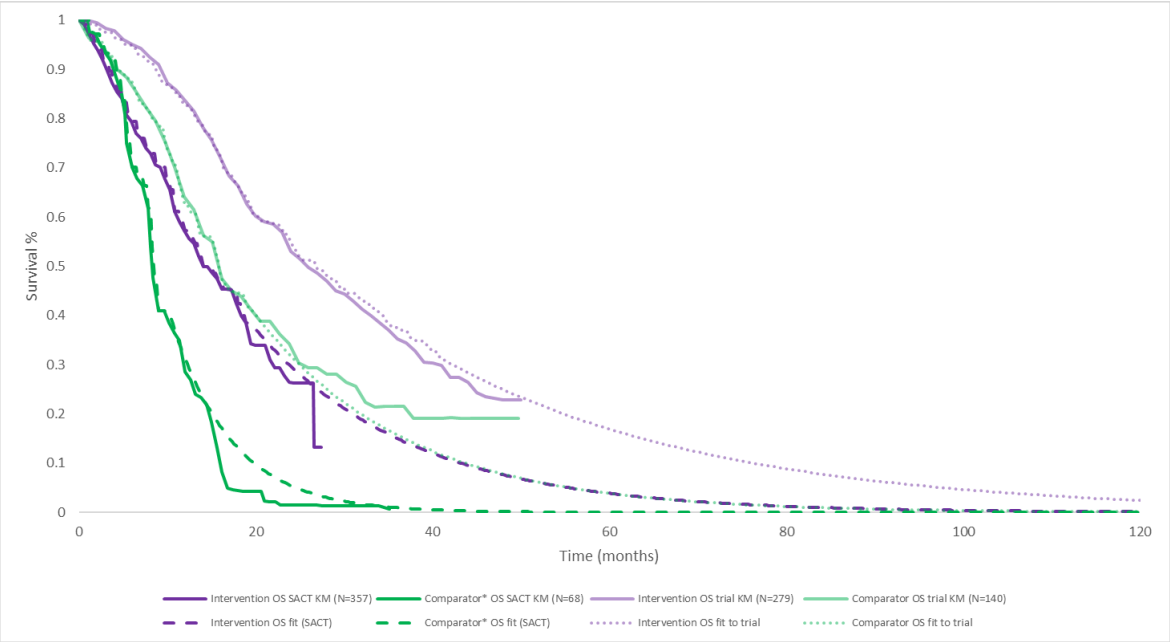
Figure 13 shows KM plots for OS and ToT from SACT and from the pivotal trial. It is apparent that OS outcomes for patients treated with the intervention in the SACT population appear to be considerably worse than those observed in the intervention arm from the trial. The FAD discusses potential reasons for this, including differences in age, ethnicity, ECOG PS, frequency of cerebral metastases, previous intake of tyrosine kinase inhibitors, and variation in time to receiving biopsy results. However, the AC determined that the key factors could not be conclusively determined, and preferred to use data from the trial to inform decision-making.

Figure 13: KM plots for OS and ToT outcomes in TA653



In the CDF review, mature KM estimates for OS, PFS, and ToT were available from the trial. It was determined that a constant hazard trend became evident before the end of the KM data for each of these outcomes, and so it was considered appropriate to extrapolate the available KM data using exponential functions. In our SACT-based analysis we used the same approach, using exponential models to extrapolate beyond the SACT KM data for OS for both treatment arms and for ToT for the intervention. Figure 14 shows the trial-based and SACT-based OS extrapolations.

Figure 14: Modelled OS in TA653 and the DSU's SACT analysis



The trial-based PFS and ToT KM curves and exponential fits used in the appraisal are shown in Figure 15. Based on these plots, we decided that ToT from SACT could be used as a proxy for SACT-based PFS estimates in this case study, noting however that the intervention was allowed to be given beyond progression in the trial. If this post-progression treatment also occurred within the SACT dataset, our approach would slightly over-estimate SACT-based PFS for the intervention (because PFS would actually be slightly shorter than ToT). The ToT 'if SACT' fit for the comparator was modelled using the methods described in Section 3.3.1. ToT extrapolations are shown in Figure 16.

Figure 15: Modelled PFS and ToT in TA653

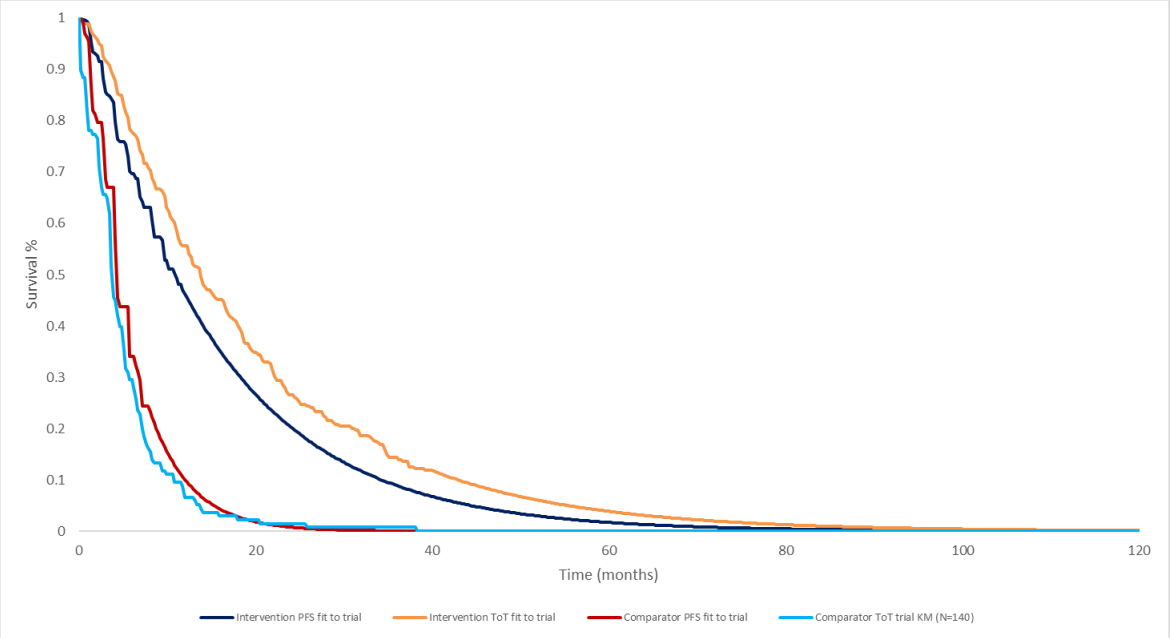
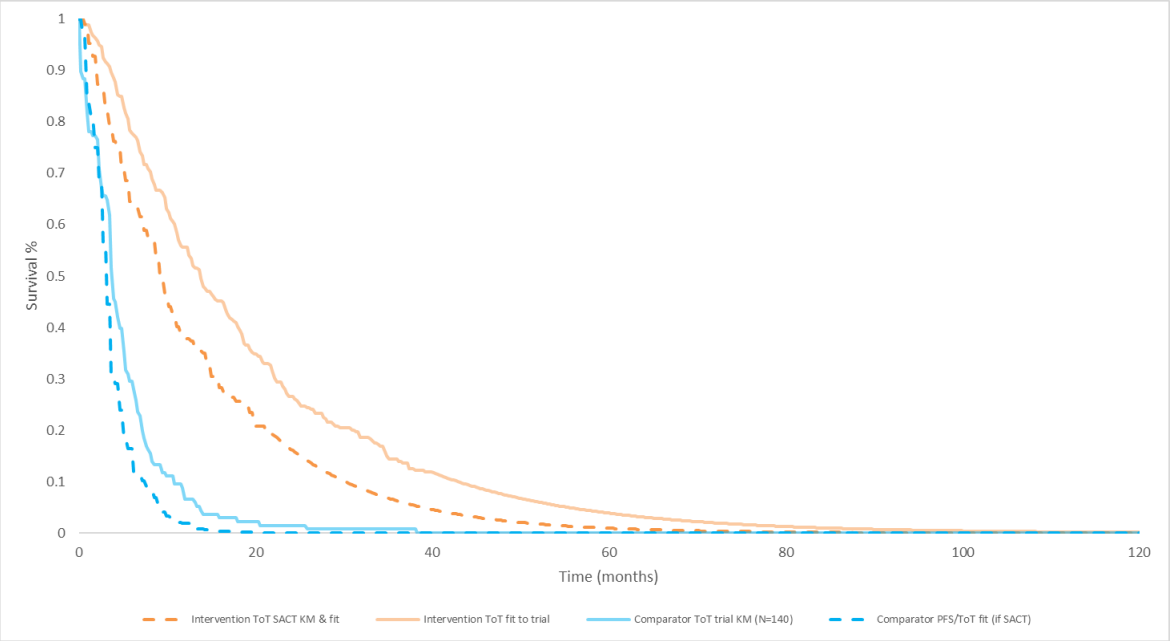


Figure 16: Modelled ToT in TA653 and the DSU's SACT analysis



It is important to reiterate that in this case study, SACT data for the comparator were available. This presented an opportunity to test the approach described in Section 3.3.1 to derive OS curves for the comparator, used in all the other case studies presented in this report. To do this, we conducted a scenario analysis where we estimated 'if SACT' OS for the comparator using the approach described in Section 3.3.1, and compared this OS curve to the actual comparator SACT OS curve that was available. Figure 17 compares the observed SACT OS KM for the comparator (and the model based on this), versus the OS curve derived for the comparator using our usual 'if SACT' approach. The curves are relatively similar and over-lapping, indicating that in this case, our approach of deriving SACT-based OS for the comparator by using time-dependent estimates of treatment effects from the pivotal trial and applying these to SACT OS curves for the intervention, appears to result in adequate estimates of the OS that would be expected to be observed for the comparator in the SACT population.

Figure 17: Modelled OS for the comparator arm using SACT data versus the DSU approach for TA653

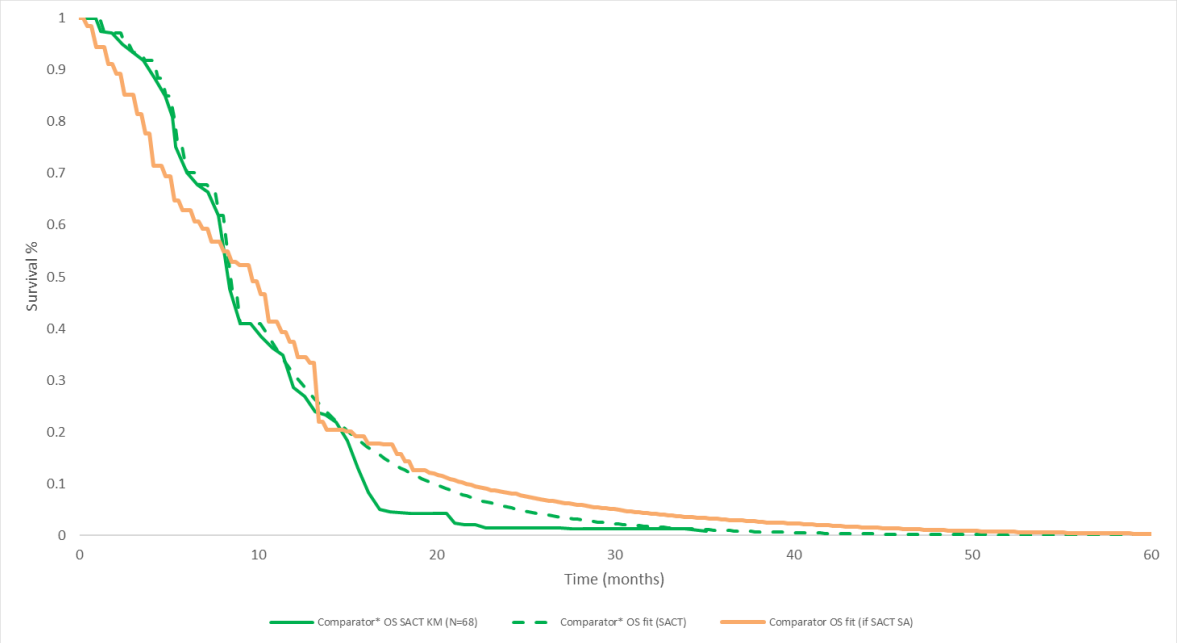


Table 8 presents the cost-effectiveness results using the trial-based AC preferences (this is the highest estimate for the ICER value reported in the final guidance), and based on the SACT data.

Table 8: Cost-effectiveness results – TA653

Technologies	Total costs (£)	Total LY (und)	Total QALYs	Incr. costs (£)	Incr. LY (und)	Incr. QALYs	ICER (£/QALY)
Trial data used in the CDF review							
Intervention	[redacted]	3.01	1.83	-	-	-	-
Comparator	[redacted]	1.76	1.08	[redacted]	1.25	0.76	£49,649
SACT data using SACT comparator data							
Intervention	[redacted]	1.67	1.06	-	-	-	-
Comparator	[redacted]	0.89	0.54	[redacted]	0.78	0.52	£45,364
SACT data using the 'if SACT' curves for the modelling OS for the comparator							
Intervention	[redacted]	1.67	1.06	-	-	-	-
Comparator	[redacted]	0.91	0.56	[redacted]	0.75	0.50	£46,358

Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYs: life years; QALYs: quality-adjusted life years

Disaggregated results are shown in Table 9. Our SACT-based analyses result in similar ICERs to the trial-based analyses because the decrease in the incremental QALY gain associated with the intervention was counteracted by a decrease in intervention treatment costs. Notably, our approach to derive OS for the comparator using the 'if SACT' approach generates similar cost-effectiveness results in this case.

Table 9: Disaggregated outcomes for TA653

Outcome	Trial			SACT		
	Intervention	Comparator	Difference	Intervention	Comparator	Difference
LYs PFS	1.28	0.49	**0.79**	1.11	0.31	**0.80**
LYs PD	1.73	1.27	**0.45**	0.56	0.58	**-0.02**
Total LYs	3.01	1.76	**1.25**	1.67	0.89	**0.78**
QALYs PFS	0.84	0.33	**0.51**	0.73	0.21	**0.52**
QALYs PD	1.00	0.78	**0.22**	0.34	0.37	**-0.03**
QALY loss due to AEs	-0.00	-0.03	**0.02**	-0.00	-0.03	**0.02**
Total QALYs	1.83	1.07	**0.76**	1.06	0.54	**0.51**
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Administration costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Subsequent therapy costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PFS)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; LYs: life years; QALYs: quality-adjusted life years; PFS: progression-free survival; PD: progressed disease

4.5. CASE STUDY 5 – TA736: NIVOLUMAB FOR TREATING RECURRENT OR METASTATIC SQUAMOUS CELL CARCINOMA OF THE HEAD AND NECK AFTER PLATINUM-BASED CHEMOTHERAPY[16]

TA736 compared nivolumab (intervention) versus docetaxel (comparator) for treating recurrent or metastatic squamous cell carcinoma of the head and neck after platinum-based chemotherapy. Upon investigating the committee papers, we found that the company carried out analyses using SACT data that were similar to the approaches described in this report. This was in response to a clarification question (B6) that was put to the company during the course of the appraisal, in which the company was asked to provide a scenario analysis using the SACT data to estimate OS and ToT for nivolumab. Given this, we have not conducted new analyses for this case study, but instead report what was done and how this was interpreted.

The company generated the pseudo-IPD using the approach described by Guyot *et al.*,[6] and then extrapolated OS using parametric approaches (including piecewise modelling).

The observed OS in the SACT cohort was similar to the OS observed in the trial for the duration of the SACT follow-up (as shown in Figure 18). However, three piecewise models were considered for selection when fitted to data from the pivotal trial (Weibull, log-logistic and log-normal piecewise models), and when these were fitted to the SACT data they produced estimates of OS that were notably dissimilar both from each other and from the outcomes from the longer-term follow-up of the pivotal trial (as shown in Figure 19). Therefore, the AC agreed that the OS data from SACT were potentially immature and would increase the uncertainty if used – the trial-based OS data were preferred.

Figure 18: OS KM plots from SACT compared to trial – TA736

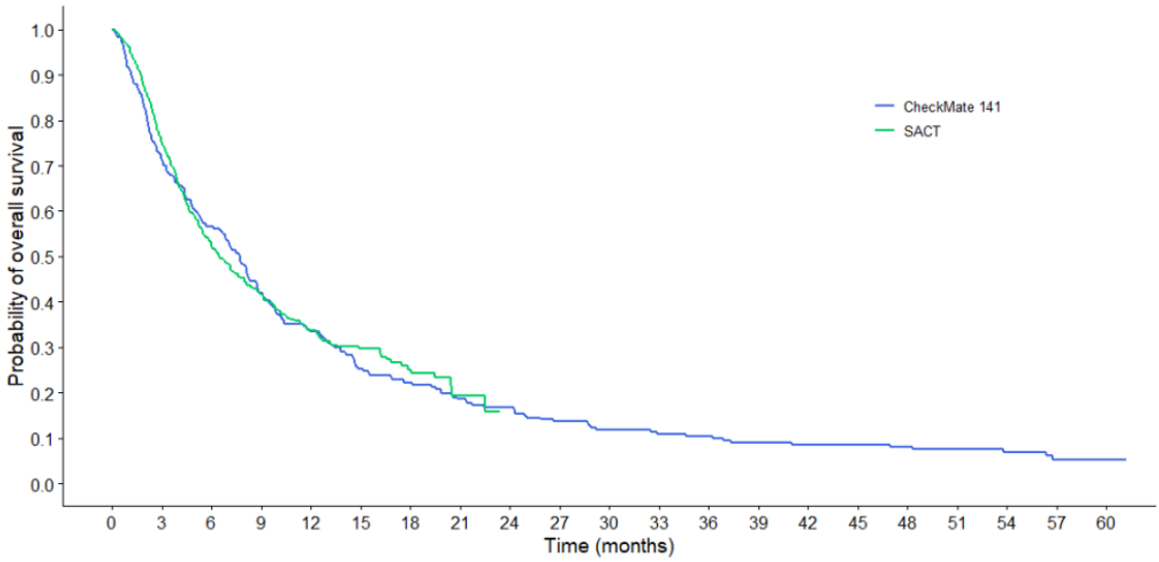


Figure 19: The different OS extrapolations using week 20 piecewise models fitted to SACT compared to trial observed data and extrapolation – TA736

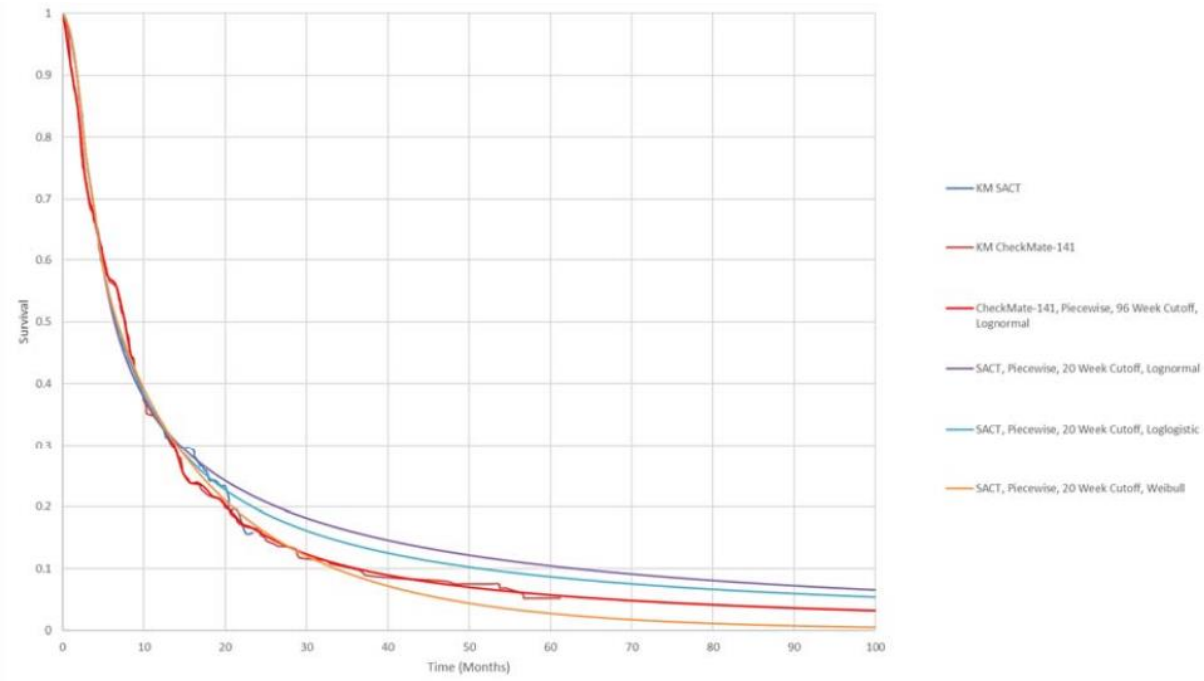
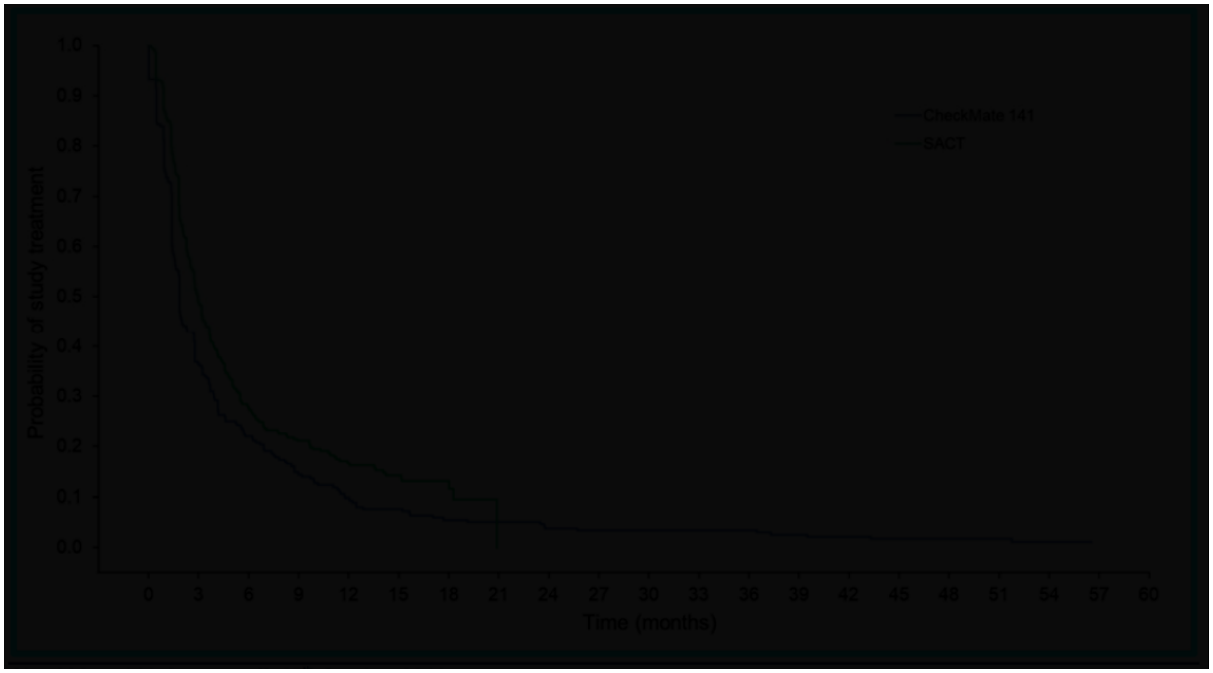


Figure 20 shows KM plots for ToT from the SACT data compared to the trial. As shown in Figure 20, ToT in the SACT population appeared to be slightly higher than ToT observed in the trial, which resulted in a higher estimate of the ICER when ToT data from SACT was used in the economic model. A spline-based parametric model with 1 knot was used to extrapolate the SACT ToT data, and when this was used instead of

the trial-based ToT curve, the company's base case ICER increased from £37,236 to £51,434. However, this analysis is not referred to in the guidance document, in which it is acknowledged that OS was similar in the trial and SACT, ToT was longer in SACT, and follow-up in the trial was substantially longer. The guidance document concludes that the end-of-life criteria were met in this appraisal, but that due to the substantial uncertainty around the cost-effectiveness of nivolumab, a maximum acceptable ICER was substantially below £50,000 per QALY gained. The AC concluded that, based on a commercial arrangement, it was likely that the ICER was indeed substantially below £50,000 per QALY gained and thus nivolumab was recommended. The AC's preferred assumptions did not use SACT data.

Figure 20: ToT KM plots from SACT compared to trial – TA736

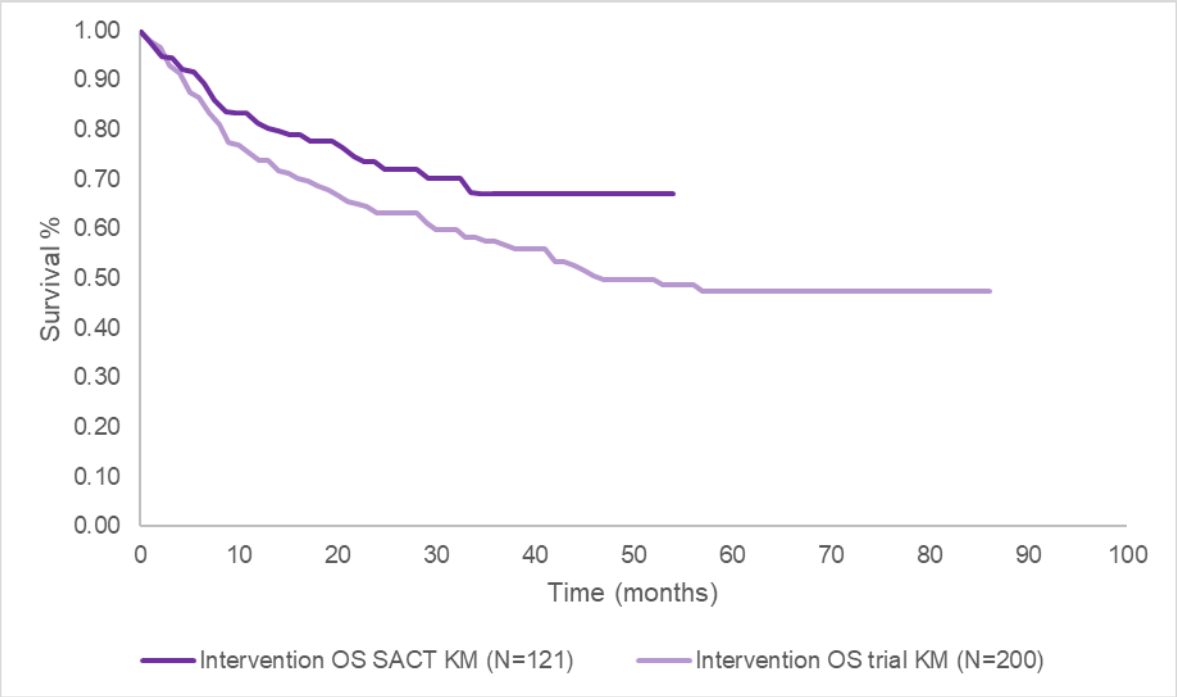


4.6. CASE STUDY 6 – TA975: TISAGENLECLEUCEL FOR TREATING RELAPSED OR REFRACTORY B-CELL ACUTE LYMPHOBLASTIC LEUKAEMIA IN PEOPLE AGED UP TO 25 YEARS[17]

TA975 compared tisagenlecleucel (intervention) versus blinatumomab or salvage chemotherapy for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged 25 years and younger. Comparisons versus salvage chemotherapy produced higher ICER estimates than those versus blinatumomab, hence results reported here are for the intervention versus salvage chemotherapy (comparator) as the resulting ICERs the AC used for decision making would be expected to be closer to the ICER threshold. We note that the AC concluded that a severity weight of 1.7 applied to QALYs was appropriate for this topic.

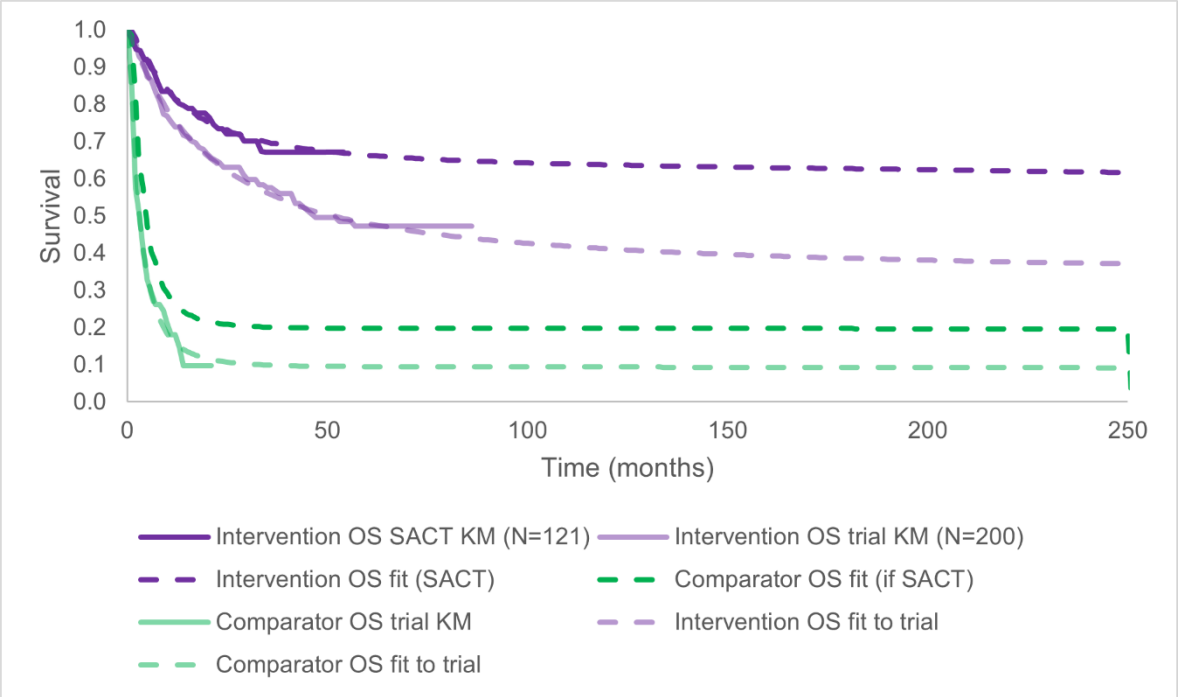
In this case study, the intervention and the comparator are given as one-off treatments at the start of the patient's treatment journey, so modelling ToT was irrelevant. Figure 21 shows KM plots for OS from SACT and from the single-arm trial used for modelling the intervention. In this case, the SACT cohort had better OS outcomes than the trial cohort. Reasons for this could not be pinpointed by the EAG or the AC but the EAG alluded to potential differences in prior and subsequent stem cell transplant rates as a contributing factor.

Figure 21: OS KM plots from SACT compared to trial – TA975



Mixture cure models (MCMs) were used to extrapolate OS in the company's base case and this represented the preference of the AC. In the appraisal, a log-logistic MCM was used for the intervention, and a log normal MCM was used for the comparator. For our analysis, we used the strsmix package in Stata which does not support the log-logistic distribution. Instead, we fitted a log normal MCM for the intervention, which we deemed reasonable because (i) log normal and log-logistic models often result in similar predictions of the survivor function; (ii) a log normal MCM was used for the comparator in the appraisal. In our analysis, after fitting a log normal MCM for the intervention using the pseudo IPD SACT data, we derived OS for the comparator using the approach described in Section 3.3.1. Figure 22 shows the trial-based and the SACT-based OS extrapolations.

Figure 22: Parametric fits for OS in TA975 and the DSU's SACT analysis



For deriving event free survival (EFS), the first approach described in Section 3.3.2 (assuming a relationship between OS and EFS) was used, given the one-off nature of the treatment meaning that it would be inappropriate to assume a relationship between ToT and EFS. The resulting extrapolations for SACT-based EFS are compared to the trial-based EFS estimates in Figure 23. However, it is relevant to note that in the CDF review appraisal, no trial data were available for EFS for the comparator, and instead a HR was applied to OS to estimate EFS – which is similar to the approach taken in this report.

Figure 23: Parametric fits for EFS in TA975 and the DSU's SACT analysis

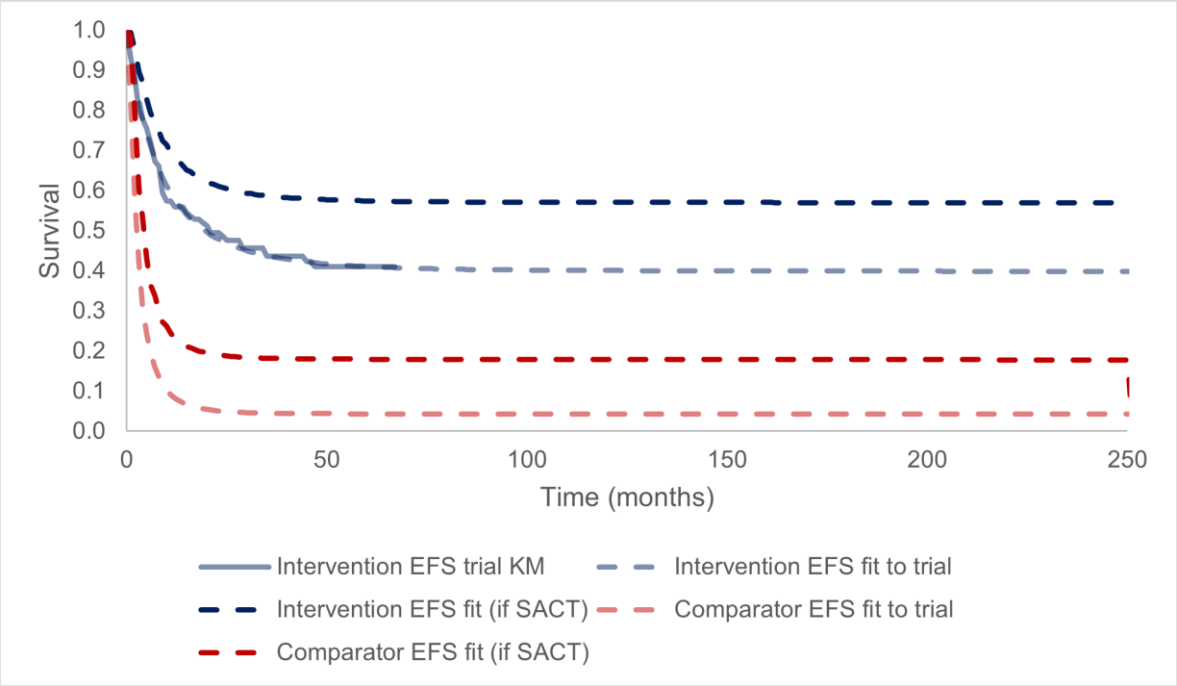


Table 10 presents a comparison of the cost-effectiveness results using the committee preferences in analyses using the trial data, and analyses using the SACT data.

Table 10: Cost-effectiveness results – TA975

Technologies	Total costs (£)	Total LY (und)	Total QALYs	Incr. costs (£)	Incr. LY (und)	Incr. QALYs	ICER (£/QALY)
Trial data used in the CDF review							
Intervention	[redacted]	19.18	[redacted]	-	-	-	-
Comparator	£59,731	5.56	2.22	[redacted]	13.62	[redacted]	£45,124
SACT data							
Intervention	[redacted]	29.10	[redacted]	-	-	-	-
Comparator	£59,817	11.11	4.46	[redacted]	17.99	[redacted]	£35,396

Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYs: life years; QALYs: quality-adjusted life years

Disaggregated results are shown in Table 11. Using SACT data resulted in lower ICERs because the increase in the incremental QALY gain was substantial, combined with no corresponding increase in treatment costs (given the one-off treatment cost).

Table 11: Disaggregated outcomes for TA975

Outcome	Trial			SACT		
	Intervention	Comparator	Difference	Intervention	Comparator	Difference
LYs EFS	18.31	3.92	**14.38**	27.23	8.30	**18.93**
LYs PD	0.88	1.64	**-0.76**	1.87	2.81	**-0.94**
Total LYs	19.18	5.56	**13.62**	29.10	11.11	**17.99**
QALYs EFS	[redacted]	1.61	[redacted]	[redacted]	3.46	[redacted]
QALYs PD	[redacted]	0.72	[redacted]	[redacted]	1.11	[redacted]
QALY loss due to AEs and transplant	[redacted]	-0.10	[redacted]	[redacted]	-0.10	[redacted]
Total QALYs	[redacted]	2.22	[redacted]	[redacted]	4.46	[redacted]
Pre-treatment costs	[redacted]	£0	[redacted]	[redacted]	£0	[redacted]
Treatment costs	[redacted]	£21,409	[redacted]	[redacted]	£21,409	[redacted]
CAR-T NHS tariff	[redacted]	£0	[redacted]	[redacted]	£0	[redacted]
AE-related costs	[redacted]	£1,803	[redacted]	[redacted]	£1,803	[redacted]
Subsequent SCT costs	[redacted]	£22,312	[redacted]	[redacted]	£22,312	[redacted]
Management costs (EFS)	[redacted]	£1,505	[redacted]	[redacted]	£2,620	[redacted]
Management costs (PD)	[redacted]	£800	[redacted]	[redacted]	£1,132	[redacted]
Terminal care costs	[redacted]	£11,901	[redacted]	[redacted]	£10,541	[redacted]
Total costs	[redacted]	£59,731	[redacted]	[redacted]	£59,817	[redacted]

Abbreviations: AE: adverse event; CAR-T: chimeric antigen receptor T-cell therapy; EFS: event-free survival; LYs: life years; QALYs: quality-adjusted life years; PD: progressed disease; SCT: stem-cell transplant

4.7. CASE STUDY 7 – TA962: OLAPARIB FOR MAINTENANCE TREATMENT OF BRCA MUTATION-POSITIVE ADVANCED OVARIAN, FALLOPIAN TUBE OR PERITONEAL CANCER AFTER RESPONSE TO FIRST-LINE PLATINUM-BASED CHEMOTHERAPY[18]

TA962 compared olaparib (intervention) versus routine surveillance (comparator) as an option for maintenance treatment of BRCA mutation-positive, advanced, high-grade epithelial ovarian, fallopian tube or primary peritoneal cancer that has responded to first-line platinum-based chemotherapy. This case study represents an example of a maintenance treatment where the aim of treatment is to achieve long-term remission.

717 NHS patients received the intervention via the CDF. However, there were important differences between the SACT population and the trial population. For example, 100 of the patients in the SACT dataset received the intervention after previously having bevacizumab in combination with first line chemotherapy. Patients previously treated with bevacizumab were not permitted in the trial, and the EAG's clinical advisors stated that patients treated with bevacizumab constitute a worse prognostic group than patients treated with chemotherapy alone. Additionally, the SACT cohort were older than the trial population (61 versus 53 years old), had fewer patients with ECOG PS of 0 (31% versus 77%), had fewer patients with complete response assessment at the end of first-line chemotherapy (64% versus 85%), and had fewer BRCA1 mutations (53% versus 74%).

Figure 24 shows trial-based and SACT-based KM plots for OS and ToT for the intervention. In the appraisal, the guidance document indicated that the EAG and company considered the SACT OS results to be similar to the trial OS results. However, inspection of the SACT and trial OS KMs indicate that the SACT cohort appeared to have worse OS outcomes than the trial cohort, whilst ToT was similar in the two datasets.

Figure 24: OS and ToT KM plots from SACT compared to trial – TA962

Whilst most partitioned survival models partition survival using estimates of PFS (resulting in patients residing in progression-free or post-progression health states), the economic model used in TA962 had an additional partition included for 'PFS2', which is an intermediary state between PFS and progressive disease. MCMs were used to extrapolate OS and PFS in the company's base case and this represented the preference of the AC. For both treatment arms, a log-logistic MCM was used for OS and a generalised gamma MCM was used for PFS. As with our analysis for TA975, we used the strsmix package in Stata to fit MCMs, and this does not support the log-logistic distribution. Therefore, we fitted a log normal MCM to the pseudo IPD from SACT to estimate OS for the intervention. In this case, when fitted to the SACT data, the log normal MCM estimated a zero cure fraction, likely because a plateau in the hazard function was not observed in the SACT data. Owing to the fact that cure models were deemed plausible in the appraisal, we conducted a scenario analysis in which we fitted a flexible parametric non-mixture cure model (NMCM) to the SACT data, setting the boundary knot cure time-point at 15 years, which was approximately the point at which modelled hazards converged to background population hazards in the survival model used in the appraisal.

The SACT-based OS curve for the comparator was simulated using the approach described in Section 3.3.1. Figure 25 shows the trial-based and SACT-based OS extrapolations. It is clear that the MCM models based on the SACT data result in reduced survival compared to the trial-based MCMs, due to the zero cure fraction estimated using the SACT data. When flexible parametric NMCMs are fitted to the SACT data, forcing the models to estimate a cure fraction, survival projections are more similar to those based on the trial data.

Figure 25: Parametric fits for OS in TA962 and the DSU's SACT analysis

For both treatment arms, PFS2 and PFS extrapolations were estimated using the first approach described in Section 3.3.2 (i.e., assuming a relationship between PFS, PFS2, and OS). Trial-based and SACT-based PFS and PFS2 curves are shown in Figure 26 and Figure 27. Note that the SACT-based curves included in these figures are for our base case analysis (i.e., using MCMs fitted to the SACT data, rather than NMCMs).

Figure 26: Parametric fits for PFS2 in TA962 and the DSU's SACT analysis

Figure 27: Parametric fits for PFS in TA962 and the DSU's SACT analysis

In the appraisal, ToT for the intervention was estimated directly from the ToT KM, owing to the completeness of the data (as shown in Figure 24). We took the same approach using the SACT data.

Importantly, in the economic model used in the appraisal, costs associated with subsequent therapies was important, particularly for the comparator arm, because patients who do not receive olaparib as maintenance therapy can instead receive it at a later line of therapy. Because our SACT-based analysis substantially impacted time spent in PFS2 and post-progression survival, we also adjusted the time estimated to be spent receiving subsequent treatments. We did this using the same approach as for PFS and PFS2 – that is, assuming a relationship between time spent receiving subsequent treatments and OS. This resulted in a substantial reduction in subsequent treatment costs in the comparator arm of the model. In addition, age in the economic model was adjusted to reflect the SACT cohort.

Table 12 presents a comparison of the cost-effectiveness results using the committee preferences in analyses using the trial data, and analyses using the SACT data.

Table 12: Cost-effectiveness results – TA962

Technologies	Total costs (£)	Total LY (und)	Total QALYs	Incr. costs (£)	Incr. LY (und)	Incr. QALYs	ICER (£/QALY)
Trial data used in the CDF review							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SACT data base case (using MCMs)							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
SACT data scenario (using non MCMs)							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYs: life years; QALYs: quality-adjusted life years

Using SACT data resulted in higher ICERs because the decrease in the incremental LYs gain was substantial, and this was combined with intervention treatment costs that remained similar and costs associated with the comparator that decreased due to reduced subsequent treatment costs. These impacts were driven by the MCM fitted to the SACT OS data resulting in a zero cure fraction. As would be expected, when we used a flexible parametric NMCM, essentially forcing the model to estimate a cure fraction greater than zero, the cost-effectiveness results moved back towards those based on the trial data – although the ICER still increased in this scenario. Disaggregated results for the trial-based analysis and our base case SACT analysis are shown in Table 13.

Table 13: Disaggregated outcomes for TA962

Outcome	Trial			SACT		
	Intervention	Comparator	Difference	Intervention	Comparator	Difference
LYs PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

LYs PFS2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
LYs PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total LYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs PFS	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs PFS2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
QALYs PD	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total QALYs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Treatment costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
AE-related costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Subsequent therapy costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PFS < 2 years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PFS > 2 years)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PFS2)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Management costs (PD)	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Terminal care costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total costs	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Abbreviations: AE: adverse event; PFS: progression-free survival; LYs: life years; QALYs: quality-adjusted life years; PD: progressed disease

4.8. CASE STUDY 8 – TA655: NIVOLUMAB FOR PREVIOUSLY TREATED SQUAMOUS NON-SMALL-CELL LUNG CANCER[19]

TA655 compared nivolumab (intervention) versus docetaxel (comparator) for treating locally advanced or metastatic squamous non-small-cell lung cancer after chemotherapy. Nivolumab has a 2-year stopping rule, and the decision problem met end-of-life criteria (hence positive recommendations could potentially be based on ICERs up to approximately £50,000 per QALY gained).

348 patients accessed the intervention via the CDF. The ERG noted that these patients were older than patients in the trial (median: 70 years versus 63 years). Comparisons between data sources were limited for ECOG performance status and level of tumour PD-L1 expression as, for 12% and 17% of SACT patients respectively, data on these baseline characteristics were missing. However, the SACT dataset included 10 patients (3%) with ECOG >1, whereas the trial excluded these patients. In addition, 40% of trial patients had PD-L1 expression levels <1, compared to 69% of the SACT cohort.

Figure 28 shows trial-based and SACT-based KM plots for OS and ToT. In this case, the plots from both datasets appear to be similar. Based on this, in the appraisal the AC concluded that trial results were generalisable to the NHS in England.

Figure 28: OS and ToT KM plots from SACT compared to trial – TA655

In the appraisal, the preference of the ERG and the AC was to use a flexible parametric model with 2 knots to model and extrapolate OS. We used the stpm2 package in Stata to fit a similar model to the pseudo IPD from SACT, and derived the OS curve for the comparator using the approach described in Section 3.3.1. Figure 29 shows the trial-based and SACT-based OS extrapolations. It can be seen that the OS curves based on the SACT data are more pessimistic than those based on the trial data, which is due to the differences in the shapes of the KM curves for the intervention close to the end of follow-up. In the trial data, the OS KM plot from the trial plateaus to some extent, after approximately 30 months. In contrast, in the SACT data, the KM falls relatively steeply close to the end of SACT follow-up, which is at the earlier time-point of approximately 20 months.

Figure 29: Parametric fits for OS in TA655 and the DSU's SACT analysis

Figure 28 and Figure 30 show that ToT were similar in the trial and in the SACT dataset, and that PFS and ToT were similar in the trial. Hence, if we used the relationship between OS and PFS to estimate SACT-based PFS for the intervention, this would result in PFS estimates that would be lower than the ToT observed in SACT. This is likely to be implausible, as we expect the intervention to be discontinued upon disease progression. Therefore, given that ToT was very similar in the trial and SACT datasets, we decided to use the PFS curves used in the appraisal in our SACT-based analysis. However, PFS was capped by OS which meant that in our analysis, the PFS curve converged with (and would be equal to) the SACT-based OS curve from approximately 25 months onwards.

Figure 30: Parametric fits for PFS and ToT in TA655 and the DSU's SACT analysis

Table 14 presents a comparison of the cost-effectiveness results using the committee preferences in analyses using the trial data, and analyses using the SACT data.

Table 14: Cost-effectiveness results – TA655

Technologies	Total costs (£)	Total LY (und)	Total QALYs	Incr. costs (£)	Incr. LY (und)	Incr. QALYs	ICER (£/QALY)
Trial data used in the CDF review							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	[redacted]	[redacted]	£30,206	1.18	0.75	£40,168
SACT data							
Intervention	[redacted]	[redacted]	[redacted]	-	-	-	-
Comparator	[redacted]	[redacted]	[redacted]	£28,441	0.45	0.35	£82,201

Abbreviations: ICER: incremental cost-effectiveness ratio; Incr.: incremental; LYs: life years; QALYs: quality-adjusted life years

Disaggregated results are shown in Table 15. Using SACT data resulted in higher ICERs due to the decreased incremental QALY gain (caused by reduced survival estimates), whilst treatment costs (and therefore incremental costs) remained similar.

Table 15: Disaggregated outcomes for TA655

Outcome	Trial			SACT		
	Intervention	Comparator	Difference	Intervention	Comparator	Difference
LYs PFS	[redacted]	[redacted]	**1.33**	[redacted]	[redacted]	**0.47**
LYs PD	[redacted]	[redacted]	**-0.15**	[redacted]	[redacted]	**-0.02**
Total LYs	[redacted]	[redacted]	**1.18**	[redacted]	[redacted]	**0.45**
QALYs PFS	[redacted]	[redacted]	**0.76**	[redacted]	[redacted]	**0.31**
QALYs PD	[redacted]	[redacted]	**-0.05**	[redacted]	[redacted]	**-0.01**
QALY loss due to AEs	[redacted]	[redacted]	**0.04**	[redacted]	[redacted]	**0.04**
Total QALYs	[redacted]	[redacted]	**0.75**	[redacted]	[redacted]	**0.35**
Treatment* acquisition costs	[redacted]	[redacted]	**£24,011**	[redacted]	[redacted]	**£24,065**
Treatment* administration costs	[redacted]	[redacted]	**£3,344**	[redacted]	[redacted]	**£3,342**
Treatment* monitoring costs	[redacted]	[redacted]	**£608**	[redacted]	[redacted]	**£546**
AE-related costs	[redacted]	[redacted]	**-£1,076**	[redacted]	[redacted]	**-£1,076**
Management costs (PFS)	[redacted]	[redacted]	**£4,447**	[redacted]	[redacted]	**£1,829**
Management costs (PD)	[redacted]	[redacted]	**-£1,129**	[redacted]	[redacted]	**-£265**
Total costs	[redacted]	[redacted]	**£30,206**	[redacted]	[redacted]	**£28,490**

*Treatment includes subsequent lines of therapy

Abbreviations: AE: adverse event; LYs: life years; QALYs: quality-adjusted life years; PD: progressed disease; PFS: progression-free survival

5. SUMMARY AND DISCUSSION

In this Section we first present a summary of the results of the eight case studies. We then provide a discussion of the findings.

5.1. SUMMARY OF FINDINGS

Error! Reference source not found. summarises the key results of the eight case studies investigated in this report, together with those from the three case studies included in the previous DSU report on this topic. The table includes columns indicating the absolute differences in life years and treatment costs associated with the intervention under appraisal according to the SACT data compared to the trial data. For instance, '↓33%' in life years means that life years estimated based on parametric models fitted to the OS data from SACT were 33% lower than life years estimated based on parametric models fitted to OS data from the pivotal clinical trial. Columns are then included for incremental total costs, incremental QALYs, and the ICER, such that the impact of the SACT data on these estimates can be observed. When changes are 'bad' – that is, they contribute to a worsening of the ICER – they are shaded in red, with deeper tones indicating larger changes. When changes are 'good' – that is, they contribute to improving the ICER – they are shaded in varying tones of green. The table also provides summary details for each appraisal, on whether a stopping rule was incorporated, the line of therapy appraised, the size of the SACT cohort compared to the trial cohort, and, finally, whether we believe that the analysis of the SACT data could potentially have changed the decision made by the NICE AC

Table 16 demonstrates that in most cases (9 out of 11), analyses based on the SACT data resulted in a reduction in overall survival, and a corresponding decrease in incremental estimates of QALY gains associated with the intervention. This was despite the fact that the methods we used retained the treatment effect observed in the pivotal clinical trials and demonstrates that if the absolute level of survival associated with an intervention is reduced in the 'real world' compared to in clinical trials, we can expect the incremental QALY benefits to also reduce, even if relative treatment effects from trials are retained.

Table 16 also demonstrates that in 5 of the 11 case studies, intervention costs reduced appreciably when based on SACT data rather than on data from the clinical trials. In

the remaining 6 case studies, intervention costs either remained approximately the same (i.e., changed by less than 4%), or increased moderately (in one case study). Therefore, while the consistent reductions in survival observed in the SACT data were sometimes accompanied by reduced time-on-treatment and therefore reduced treatment costs, this was by no means always the case. Unsurprisingly, when a reduction in absolute survival (and therefore a reduction in incremental QALYs) was not accompanied by an appreciable reduction in time-on-treatment (and therefore treatment costs), ICERs increased. According to our analyses, ICERs based on SACT data were 20% or more higher than those based on the trial data in 5 of the 11 case studies. SACT-based ICERs were different to trial-based ICERs by 5% or less in 4 of the case studies (typically those where there were similar percentage changes in absolute life years and costs associated with the intervention), with the ICER reducing in 2 case studies – most clearly in TA975 (tisagenlecleucel for treating relapsed or refractory b-cell acute lymphoblastic leukaemia), in which treatment costs were the same in the SACT and trial datasets, whilst survival was improved in SACT.

Table 16: The change in cost-effectiveness results based on SACT data vs trial data

TA	LYs (absolute)	Treatment costs (absolute)	Total costs (incremental)	QALYs (incremental)	ICER	Stopping rule	Line of therapy	Size of SACT vs trial (number of patients)	Decision could potentially change
The case studies detailed in the previous report									
TA870	↓33%	[redacted]	[redacted]	↓17%	[redacted]	No	First	2,460 vs 148	No
TA780 (1)*	↓28%	[redacted]	↑2%	↓23%	↑33%	No	First	814 vs 425	Yes
TA780 (2)*			↓2%	↓18%	↑20%				
TA802	[redacted]	[redacted]	[redacted]	[redacted]	↑4%	Yes	First	352 vs 219	No
The case studies detailed in this report									
TA766	[redacted]	[redacted]	[redacted]	[redacted]	↑4%	12 [redacted]	Adjuvant	1,324 vs 514	No
TA725	[redacted]	[redacted]	[redacted]	[redacted]	↑5%	No	Second	876 vs 443	No
TA629	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	30 [redacted]	Second	92 vs 164	Yes
TA653	↓45%	[redacted]	[redacted]	↓46%	↓9%	No	Second	357 vs 279	No
TA736	Marginal increase	Moderate increase	Likely moderate increase	Likely marginal increase	↑38%	24 months	Second	506 vs 361	Yes
TA975	↑52%	[redacted]	[redacted]	[redacted]	↓22%	No	First	121 vs 200	No
TA962	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	No	Adjuvant	717 vs 260	Yes
TA655	↓50%	↑0%	↓6%	↓53%	↑105%	24 [redacted]	Second	348 vs 135	Yes

*PFS was estimated using two approaches; the first assumed an OS/PFS relationship whereas the second assumed an PFS/ToT relationship.

Key to shading:

	>60% worsening
	40-60% worsening
	20-40% worsening
	0-20% worsening
	0-20% improving
	20-40% improving

We consider that using the SACT data to inform cost-effectiveness estimates could potentially have altered NICE AC decision making in 5 of the 11 case studies – these were the appraisals in which the SACT-based ICER was ≥20% higher than the trial-based ICER. These were: TA780; TA629; TA736; TA962, and TA655. However, whilst we believe that the SACT-based analysis had the *potential* to alter NICE decision-making in these TAs, it is by no means certain that this would have been the case, principally due to substantial uncertainties and limitations associated with the SACT data. We explain this for each TA below.

- TA780 (nivolumab with ipilimumab for untreated metastatic renal cell carcinoma, from our previous study). Whilst the SACT dataset had a larger sample size than the trial, maximum follow-up in the SACT data was approximately 2 years, compared to approximately 6 years in the pivotal clinical trial.
- TA629 (obinutuzumab in combination with bendamustine for treating rituximab-refractory follicular lymphoma). The SACT sample size was smaller than the trial, and maximum follow-up was approximately 2 years, compared to approximately 6 years in the pivotal clinical trial. The AC stated that SACT data were not robust enough to be used in economic modelling.
- TA736 (nivolumab for treating recurrent or metastatic squamous cell carcinoma of the head and neck after platinum-based chemotherapy). The SACT sample size was larger than that in the clinical trial, but maximum follow-up in the SACT data was approximately 2 years, compared to approximately 5 years in the pivotal clinical trial. In this case, the evidence is supportive of survival being similar in the SACT dataset, but also appears reasonably robust in suggesting that time-on-treatment (and therefore treatment costs) was higher in the SACT data. This results in an increased ICER, which was stated in committee papers. However, this was not discussed in the guidance document. Therefore, while we believe that the SACT-based analysis logically could have affected NICE decision making in this appraisal, in practice this did not appear to be the case.
- TA962 (olaparib for maintenance treatment of brca mutation-positive advanced ovarian, fallopian tube or peritoneal cancer after response to first-line platinum-based chemotherapy). In this TA, in committee papers the EAG suggested that

survival in the SACT and trial datasets appeared to be similar, and this was stated in the guidance documents. However, our assessment of the survival data suggests that there are in fact potentially important differences in survival, which could result in a significantly increased ICER. However, whilst the SACT sample size was larger than the trial, the maximum follow-up in SACT was approximately 42 months, compared to approximately 100 months in the trial.

- TA655 (nivolumab for previously treated squamous non-small-cell lung cancer). In this TA, SACT OS did not appear to be substantially different to that observed in the trial, but differences became important when extrapolating, resulting in a substantial increase in the ICER which could have changed NICE decision-making. However, whilst the sample size in the SACT data was larger than in the trial, maximum follow-up in the SACT data was approximately 20 months, compared to approximately 60 months in the trial.

Consideration of these case studies demonstrates that while SACT-based analyses might be indicative of worsened cost-effectiveness for the interventions under appraisal, in each case the follow-up in SACT data was considerably shorter than in the relevant clinical trials. Given that survival extrapolations are frequently key sources of uncertainty in NICE appraisals of cancer treatments, it is unclear whether ACs would be willing to base decision-making on much shorter-term SACT data, when longer-term data from clinical trials are available. In addition to this, the SACT data provided to NICE as part of CDF re-appraisals is limited by the absence of data on comparators, and the absence of information on PFS.

However, whilst it is possible that ACs would be reluctant to fully base decision-making on relatively immature SACT data rather than more mature trial data, we believe that it is possible that cost-effectiveness analyses using SACT data could still be used to *inform* decision-making – for example, if SACT data indicates that ToT is similar in SACT and trial datasets, but that OS may be lower, scenario analyses could be undertaken to explore the impact on the ICER of reduced OS (whilst retaining the relative effect observed in the trial). This could supplement trial-based analyses and allow committees to make decisions based on a wider range of evidence.

5.2. DISCUSSION

Therapies are placed in the CDF because there are key uncertainties around components of cost-effectiveness which are likely to be resolved to some extent by additional follow-up from existing clinical trials, and/or by using RWD from SACT. The analysis contained in this report shows how RWD from SACT can be used to inform estimates of cost-effectiveness. The lack of comparator data provided to NICE in the CDF re-appraisal process, and the lack of information on PFS, means that several important assumptions need to be made in order to derive SACT-based estimates of cost-effectiveness. However, we have demonstrated methods to conduct these analyses, and these could be conducted on a routine basis for CDF reviews. A key issue is whether these analyses provide added value for a decision-making committee, particularly given the limitations in the data available.

5.2.1. Implications for cost-effectiveness analyses

We demonstrate that in several cases, SACT-based cost-effectiveness analyses do not substantially change estimates of the ICER, suggesting that these analyses would not affect decision-making. This is primarily because reductions in survival observed in the RWD are offset by reductions in time-on-treatment and treatment costs. However, our case studies clearly demonstrate that this cannot be assumed to always be the case. Frequently, survival was reduced in the RWD but treatment costs were not, often leading to substantially increased estimates of the ICER. In some circumstances (for instance, when reductions in survival were very substantial), ICERs increased markedly even when treatment costs also reduced. Therefore, we believe that it is likely to be useful to conduct analyses similar to those demonstrated in this report on a routine basis in CDF reviews. We acknowledge the limitations in the SACT data, and the strong assumptions required to derive cost-effectiveness estimates based on these data but believe that these analyses would provide ACs with additional evidence to inform decisions. Whilst base case analyses are likely to still be based on trial-based analyses, scenarios informed by SACT data could helpfully inform estimates of the likely ICER range, which could help ACs make decisions with more confidence (for example, if a SACT-based analysis demonstrated that a treatment was likely to be cost-effective even if survival was reduced in the ‘real world’, and assuming that the relative treatment effect observed in the trial was retained).

We recognise that when results from SACT-based cost-effectiveness analyses do not align with results from trial-based analyses, ACs may be placed in a difficult position. We believe that, in such circumstances, it is unlikely that an AC would choose to use a SACT-based analysis as the base case analysis, due to data limitations (around length of follow-up, as well as comparator data and PFS), and because usually RCT evidence is available, providing unbiased and longer-term data. However, this does not mean that cost-effectiveness analyses using SACT data would not be useful. Below, we make suggestions for how SACT-based analyses could be informatively used in a range of common scenarios:

- Scenario 1. When SACT data is supportive of similar OS as observed in the trial, but increased time-on-treatment. SACT data on ToT is often relatively mature, so reduced follow-up in SACT compared to the trial is less of a concern. In such circumstances, the economic model could simply be re-run using SACT-based ToT instead of trial-based ToT. This scenario was observed in one of our case studies (TA736), but a SACT-based analysis seemingly might not have informed decision-making.

- Scenario 2. When SACT data indicates that ToT is similar in the SACT and trial datasets, but OS appears to be reduced in the RWD. This scenario was observed in four of our case studies. In such cases, the economic model could be re-run using trial-based (or SACT-based) estimates of ToT, but with adjusted estimates of OS (and, where relevant, PFS) – thereby investigating the cost-effectiveness of the treatment if the relative treatment effect is retained, but where survival is simply worse in the real world than in the clinical trial – which is often accepted to be the case. In the analyses presented in this report, we fit survival models to the SACT data to estimate SACT-based OS, but we demonstrate that SACT follow-up is consistently substantially shorter than the trial follow-up, meaning that SACT-based extrapolations may be considered to be unacceptably uncertain. Instead, SACT-based scenario analyses could be undertaken by applying a factor to trial-based OS curves, such that the shape of these curves is retained but absolute survival is reduced towards the level observed in the SACT data (with comparator OS also reduced, based on the relative treatment effect observed in the trial). We have not undertaken such

analyses in this report, but a learning point from our analysis is that this alternative approach could represent a viable option.

- Scenario 3. When SACT data indicates that both ToT and OS are reduced in the RWD. This scenario was observed in five of our case studies. This scenario is the most difficult to deal with, because it implies that a SACT-based analysis must amend each of the key outcomes included in the economic model, as we have done for each case study included in this report. As previously stated, SACT data on ToT is often relatively mature, and therefore incorporating this into the economic model is relatively unproblematic. For OS, we again suggest that the approach described for Scenario 2 represents a viable option – whereby applying a factor to the trial-based OS estimates to reduce survival to levels observed in SACT may represent a preferable option compared to fitting survival models directly to the SACT data, as we have done in the analyses contained in this report.

It is important to note that SACT-based analyses may not always be informative for decision-making. For example, in circumstances where the SACT sample size is very small, resulting analyses may not be useful.

5.2.2. Data limitations

Whilst we believe that the analyses suggested above may usefully inform decision-making in CDF reviews, it is crucial to highlight the deficiencies in the SACT data provided to NICE in the CDF review process, and how these deficiencies could be alleviated. The absence of data on the comparator is key and prevents any assessment of comparative effectiveness using the real-world data. SACT collects data on all systemic anti-cancer therapies. Therefore, in the vast majority of cases, where the comparator is a systemic anti-cancer therapy, data on the comparator will be available within the SACT dataset. We acknowledge that fewer patients may be treated with the comparator once the intervention is available in the NHS through the CDF. However, historic data for the comparator (with much longer-term follow-up) would be available. Extracting this data from SACT should not be problematic, and providing such data in a CDF review would be likely to reduce decision uncertainty. Decision uncertainty has an opportunity cost in terms of foregone health, and therefore

given that comparator data are available within SACT, all efforts should be made to make these data available for use in NICE appraisals.

Analyses of comparative effectiveness using SACT data (i.e., if data on the intervention and the comparator were provided) would be prone to the limitations of observational data analyses such as selection bias and confounding, necessitating access to patient-level data in order to adjust for these. The NICE RWE framework discusses methods for analysing these data using approaches to minimise biases, including using Target Trial methodology[20]. We believe that it would be optimal for NICE to be provided with data on comparators from SACT to inform CDF reviews, and that comparative effectiveness analyses using patient-level data would be valuable. Whilst we acknowledge that comparative effectiveness analyses based on RWD are prone to biases, and we do not suggest that these should replace RCT-based analyses, provision of patient-level data from SACT would allow higher quality SACT-based analyses to be included in CDF reviews.

It is clear from the analyses provided in this report that survival outcomes in the SACT dataset are often worse than those observed in trials. Whilst summary statistics provide some indication of why this might be (e.g., older age, worse performance status), providing access to patient-level data for patients included in SACT would allow much more robust analyses of these factors. In addition, other analyses could also provide information on relative treatment effects that might be expected in the real world compared to those observed in overall trial populations. Subgroup analyses of trial data often show whether the treatment effect differs in patient groups with different characteristics, and it may be possible to identify subgroups within trials who have characteristics similar to those expected in the real world. This could provide information on whether the treatment effect observed in the overall trial population can be expected to be transferable to the real-world population. It would be useful to encourage companies to conduct such analyses to inform CDF reviews. Furthermore, it may be useful if the SACT reports included survival and ToT curves stratified by relevant subgroups.

A final key limitation associated with the SACT data is the lack of information on PFS. Data on disease progression are not currently collected in SACT, so this data cannot simply be made available to NICE (unlike data on the comparator). This limitation is

due to the data not being collected, rather than not being provided. In this report we have demonstrated methods for estimating PFS based on assumed relationships with OS or ToT. These methods are sub-optimal, and if analyses were being undertaken to inform NICE decision-making careful consideration of the most appropriate approach should be made on a case-by-case basis.

5.2.3. Real-world survival

Discrepancies in survival outcomes between patients treated in clinical trials and those treated in the real world are well documented. These differences have been attributed to patient selection such as differences in baseline characteristics and prognostic factors between trial participants and patients seen in clinical practice[21, 22]. Older patients, those with a higher comorbidity burden, and patients from lower socioeconomic backgrounds are often under-represented in trials[23]. Therefore, it may not be surprising that nine of the eleven case studies presented in our current and previous report showed worse survival outcomes for the new interventions based on the SACT data, compared to survival curves based on relevant trial data.

Of note, this reduction in survival was not observed in TA975, in which the CAR T-cell therapy tisagenlecleucel was shown to be associated with substantially improved survival in the SACT dataset compared to the trial dataset. The reason for this is unclear, but it is possible that very rigorous patient selection and learning effects for CAR T-cell treatments could have contributed to this.

5.2.4. Limitations of our analyses

The methods we used to analyse and incorporate the SACT data into the economic models for each case study rely on strong assumptions, thus should be treated with caution, and alternative approaches could also be used. For OS, we assumed that time-dependent relative treatment effect estimates derived from the parametric models preferred by the AC in the original appraisal would be replicated in the SACT data. This may or may not be true, and it is conceivable that violations to this assumption could differ depending on the nature of the treatment. As stated in Section 5.2.2, access to comparator data would allow this to be investigated further.

We also assumed that the underlying parametric distributions preferred by the AC remain appropriate when fitted to the SACT data, rather than undertaking a rigorous model selection process. In practice, this should be investigated further, in case real-world survival distributions appear to differ to those observed in trials. As described in Section 5.2.2, in the context of CDF reviews, extrapolations from survival models fitted to SACT data are likely to be more uncertain than those from models fitted to the trial data, due to the limited SACT follow-up period compared to the trial. To alleviate this, alternative approaches to derive survival more consistent with that observed in SACT could be used – such as by applying factors to trial-based survival curves to bring them closer to those observed in SACT.

If SACT data were provided for comparators over a longer timer period (i.e., including data collected before the intervention was placed in the CDF), survival curves could be fit to SACT-based comparator OS, with the treatment effect observed in the trial then applied to derive estimates of SACT-based intervention survival. These analyses would not be subject to short follow-up periods (assuming the comparator has been provided in the NHS for a long period of time), thereby removing a key problem with the SACT data currently provided for CDF interventions. In fact, this approach could (and possibly should) be used in all NICE appraisals – not just CDF reviews – to provide estimates of the cost-effectiveness of new interventions in the real-world NHS population.

The lack of SACT data on PFS described in Section 5.2.2 represents an important data limitation that results in important limitations in our analyses. We demonstrated that SACT-based estimates of PFS could be derived in different ways – using an assumed relationship with OS, or an assumed relationship with ToT (including constraints). Each approach involves strong assumptions and can result in markedly different PFS estimates – which in some cases can have an appreciable impact on ICERs. We recommend that different approaches are investigated on a case-by-case basis, with the sensitivity of results reported.

It is important to note that all but one of our case studies included economic models in the form of partitioned survival models. The methods for deriving SACT-based cost-effectiveness estimates described in Section 3 are better suited this type of model, where the parametric extrapolations drive the cost-effectiveness results. State

transition Markov models are less common in NICE appraisals, but we included one such case study – TA766 – to demonstrate how this type of model could be dealt with. In state-transition models, overall survival is a function of transition probabilities between other health states, and therefore the provision of SACT data only for OS and ToT, and the lack of provision of data on PFS, is more of a limitation for state-transition models than for partitioned survival models. We used a simple and pragmatic calibration approach to adjust the economic model based on the SACT OS data, but other more sophisticated calibration techniques could have been used[24-27].

Finally, in the SACT-based analyses provided in this report, due to the time constrained nature of this project and its illustrative nature, we have only undertaken deterministic analyses. For analyses to inform decision-making within a NICE CDF review, it would be best practice to conduct probabilistic analyses.

5.2.5. Further research

The Newcastle upon Tyne Hospitals NHS Foundation Trust (NuTH) group has obtained access to the patient-level data for the SACT cohort for a number of the case studies that we have included in our current and previous report and is exploring patient characteristics and outcomes. Research is ongoing, but the findings may usefully supplement those included in our report.

In Section 5.2.1 we highlighted an alternative method for obtaining SACT-based survival estimates in the context of a CDF review (described in 'Scenario 2'). It may be valuable to conduct additional analyses of our case studies, illustrating and testing this approach.

5.2.6. Conclusions

In this report we aimed to demonstrate how data from SACT – as it is currently provided to NICE for the purpose of CDF reviews – could be used to more routinely inform decision-making, and to show what impact this could have on cost-effectiveness estimates. We present pragmatic methods for conducting these analyses that could be conducted on a routine basis and show that SACT-based analyses could affect NICE decision-making. Whilst we believe that it is unlikely that SACT-based analyses would replace trial-based analyses as the primary analysis

upon which NICE ACs base their decisions – due to the limitations associated with SACT data – we believe that cost-effectiveness scenarios based on SACT data could usefully inform AC decisions, helping to demonstrate when trial-based analyses are likely to be overly optimistic (or pessimistic). We believe that improvements in the SACT data provided to NICE for CDF reviews are important and should be achievable. In particular, data on the comparator should be provided, as should access to patient-level data. If data on disease progression could be added to the variables collected within SACT, this would represent an extremely valuable addition to the dataset.

It is important to note that Trigg *et al.* highlighted that uncertainty around long-term survival estimates is the most common reason for technologies to be placed in the CDF[28]. It is unlikely that survival data from SACT for the duration that a treatment is in the CDF will resolve this uncertainty. However, if long-term comparator data were provided from the period prior to an intervention being placed in the CDF, uncertainty around long-term outcomes could be reduced to some extent. This would even be possible at the time of an initial appraisal of any new cancer treatment, rather than only for reviews of treatments placed in the CDF.

REFERENCES

1. Programme N-CFHTE-TA. Specification for Cancer Drugs Fund data collection arrangements.
2. Team NECDF. Appraisal and Funding of Cancer Drugs from July 2016 (including the new Cancer Drugs Fund) - A new deal for patients, taxpayers and industry; 2016.
3. Service NCRaA. Systemic Anti-Cancer Therapy Dataset (Chemotherapy). http://www.ncin.org.uk/collecting_and_using_data/data_collection/chemotherapy (Accessed 27 Feb).
4. Kang J CJ. "Don't Think Twice, It's All Right": Using Additional Data to Reduce Uncertainty Regarding Oncologic Drugs Provided Through Managed Access Agreements in England. *Pharmacoecon Open* 2023;7:77-91.
5. NR L. PCN317—the Cancer Drugs Fund: key uncertainties, data collection plans, analytical methods and use of the systematic anti-cancer therapy (SACT) real world data set. *Value Health [Internet]* 2018;21.
6. Guyot P AA, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published kaplan-meier survival curves. *BMC medical research methodology* 2012;12:9.
7. Liu N, Zhou, Y. & Lee, J.J. IPDfromKM: reconstruct individual patient data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2021;21.
8. StataCorp. 2023. Stata Statistical Software: Release 18. College Station, TX: StataCorp LLC.
9. National Institute for Health and Care Excellence. Decision Support Unit (DSU) Technical Support Document (TSD) 14: Survival analysis for economic evaluations alongside clinical trials – extrapolation with patient-level data.
10. Salmon D, Melendez-Torres GJ. Clinical effectiveness reporting of novel cancer drugs in the context of non-proportional hazards: a review of nice single technology appraisals. Int J Technol Assess Health Care. 2023 Mar 8;39(1):e16. doi: 10.1017/S0266462323000119. PMID: 36883316.
11. NICE. Pembrolizumab for adjuvant treatment of resected melanoma with high risk of recurrence (CDF Review of TA553) [ID3776]. https://www.nice.org.uk/guidance/ta766/history (Accessed
12. NICE. Abemaciclib with fulvestrant for treating hormone receptor-positive, HER2-negative advanced breast cancer after endocrine therapy.
13. NICE. Obinutuzumab with bendamustine for treating follicular lymphoma after rituximab.
14. NICE. Osimertinib for treating EGFR T790M mutation-positive advanced non-small-cell lung cancer.
15. NICE. Osimertinib for untreated EGFR mutation-positive non-small-cell lung cancer.
16. NICE. Nivolumab for treating recurrent or metastatic squamous cell carcinoma of the head and neck after platinum-based chemotherapy.
17. NICE. Tisagenlecleucel for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people 25 years and under.

18. NICE. Olaparib for maintenance treatment of BRCA mutation-positive advanced ovarian, fallopian tube or peritoneal cancer after response to first-line platinum-based chemotherapy.
19. NICE. Nivolumab for advanced squamous non-small-cell lung cancer after chemotherapy.
20. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016 Apr 15;183(8):758-64. doi: 10.1093/aje/kwv254. Epub 2016 Mar 18. PMID: 26994063; PMCID: PMC4832051.
21. Sherman REea. Real-World Evidence - What Is It and What Can It Tell Us? *N Engl J Med* 2016;375:2293-7.
22. al. Re. Interpreting clinical trial data in multiple myeloma: translating findings to the real-world setting. *Blood Cancer Journal* 2018;8.
23. Booth CMT, I. F. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. *Br J Cancer* 2014;110:551-5.
24. Vanni T, Karnon J, Madan J, White RG, Edmunds WJ, Foss AM, Legood R. Calibrating models in economic evaluation: a seven-step approach. Pharmacoeconomics. 2011 Jan;29(1):35-49. doi: 10.2165/11584600-000000000-00000. PMID: 21142277.
25. Karnon J, Vanni T. Calibrating models in economic evaluation: a comparison of alternative measures of goodness of fit, parameter search strategies and convergence criteria. Pharmacoeconomics. 2011 Jan;29(1):51-62. doi: 10.2165/11584610-000000000-00000. PMID: 21142278.
26. Moriña, D., de Sanjosé, S. & Diaz, M. Impact of model calibration on cost-effectiveness analysis of cervical cancer prevention. Sci Rep 7, 17208 (2017). https://doi.org/10.1038/s41598-017-17215-2.
27. Whyte S, Walsh C, Chilcott J. Bayesian calibration of a natural history model with application to a population model for colorectal cancer. Med Decis Making. 2011 Jul-Aug;31(4):625-41. doi: 10.1177/0272989X10384738. Epub 2010 Dec 2. PMID: 21127321.
28. Trigg LA, Barnish, M.S., Hayward, S. et al. An Analysis of Uncertainties and Data Collection Agreements in the Cancer Drugs Fund. PharmacoEconomics Open 8, 303–311. 2024; https://doi.org/10.1007/s41669-023-00460-9.