Health Technology Assessment of CD19 CAR T-Cell Therapies in the Irish Healthcare Setting

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Niamh Carey
BSc (Pharm), MPharm, MSc

National Centre for Pharmacoeconomics
&
Department of Pharmacology and Therapeutics
University of Dublin, Trinity College
2022

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Niamh Carey
31 March 2022

Summary

The aim of this thesis was to undertake health technology assessments of CD19 CAR T-cell therapies, tisagenlecleucel and axicabtagene ciloleucel. Tisagenlecleucel is licensed for the treatment of paediatric and young adult patients with relapsed or refractory (R/R) B-cell acute lymphoblastic leukaemia (ALL). Tisagenlecleucel is also licensed for the treatment of adult patients with R/R diffuse large B-cell lymphoma (DLBCL), after two or more lines of systemic therapy. Axicabtagene ciloleucel is licensed for the treatment of R/R DLBCL, and R/R primary mediastinal large B-cell lymphoma, after two or more lines of systemic therapy.

A bespoke cost-utility model was developed to examine the cost effectiveness of tisagenlecleucel versus blinatumomab for R/R ALL. A systematic literature review (SLR) was conducted to identify clinical efficacy inputs. Two trials that investigated tisagenlecleucel and one that investigated blinatumomab were included in the final evidence base. All were single-arm. Naïve, unadjusted comparison of overall survival (OS) indicated that tisagenlecleucel had favourable outcomes versus blinatumomab. However, this is uncertain. The quality of evidence for OS was very low. Event-free survival data were not reported for blinatumomab.

An expert elicitation was conducted to elicit judgements (n=5), using a bespoke elicitation tool, regarding key areas of uncertainty in the evidence base of tisagenlecleucel for R/R ALL. Areas of uncertainty examined included: the concept of cure and cure fraction, the proportion of patients expected to receive allogeneic stem cell transplant (alloSCT) following tisagenlecleucel, and the five-year OS of patients treated (i) with and (ii) without subsequent alloSCT. Much uncertainty was observed within and between judgements. Judgements were used in the validation of the cost-utility model inputs and outputs.

Utility data were sourced by SLR. Cost and resource use data were derived from sources in Ireland, where possible. At the defined willingness-to-pay threshold of €45,000 per QALY, tisagenlecleucel was not cost effective versus blinatumomab. The probability of cost effectiveness was 16%. Population expected value of perfect information (EVPI) and

partial EVPI (EVPPI) estimates were low. Further research to decrease decision (parameter) uncertainty, at this threshold, may not be of value. However, uncertainty in the model may not be adequately captured by EVPI.

Bespoke cost-utility models were also developed to examine the cost effectiveness of (i) tisagenlecleucel, and (ii) axicabtagene ciloleucel, for R/R DLBCL. An SLR was conducted to identify clinical efficacy inputs. The utility of Abstrackr, a text-mining tool, in assisting in title and abstract screening of this SLR was evaluated. One trial each for tisagenlecleucel, axicabtagene ciloleucel, and salvage chemotherapy were included in the final evidence base. All were single-arm. Clinical and methodological heterogeneity between the tisagenlecleucel and axicabtagene ciloleucel trials precluded a robust comparison. Naïve, unadjusted comparison of OS indicated that both tisagenlecleucel and axicabtagene ciloleucel had favourable outcomes versus salvage chemotherapy. However, this is uncertain. The quality of evidence for OS was very low. Progression-free survival data were not reported for salvage chemotherapy.

Utility data were sourced by SLR. Cost and resource use data were derived from sources in Ireland, where possible. At the willingness-to-pay threshold of €45,000 per QALY, neither tisagenlecleucel nor axicabtagene ciloleucel were cost effective versus salvage chemotherapy. The probability of cost effectiveness (versus salvage chemotherapy) was 0% for tisagenlecleucel, and 0% for axicabtagene ciloleucel. Population EVPI and EVPPI estimates were €0.00.

Budget impacts (inclusive of VAT) of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL were estimated. Eligible population estimates were derived from the literature and clinical opinion. Assuming that 23 patients would be treated in the first five years, the cumulative five-year gross drug budget impact, of tisagenlecleucel for R/R ALL, is €8.6 million. The cumulative five-year net drug budget impact, assuming displacement of blinatumomab and FLA(G)-IDA (fludarabine, idarubicin, cytarabine, granulocyte colony stimulating factor) is €6.7 million. Assuming a total of 120 patients, and a 50:50 market share, the cumulative five-year gross drug budget impact of tisagenlecleucel (n=60) and axicabtagene ciloleucel (n=60) is €45.6 million. The cumulative five-year net drug budget impact, assuming displacement of

salvage chemotherapy, is €44.9 million. These budget impacts may be underestimated. Affordability of these therapies is a key challenge.

The impact of performance-linked reimbursement agreement scenarios on the cost effectiveness and budget impact of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL was explored. Agreements, which captured outcomes over a longer time horizon, were impactful. In some instances, a decrease in price in combination with a performance-linked reimbursement agreement may be required. Affordability may remain a concern.

The findings of this thesis have implications for decision-makers in Ireland. Based on these analyses, the reimbursement of these CD19 CAR T-cell therapies is unlikely to represent cost-effective use of resources within the Irish healthcare setting. Affordability is a concern. Performance-linked reimbursement agreements are necessary to reduce the associated financial risk.

Acknowledgements

Thank you to Prof Michael Barry for providing me with the opportunity to pursue this research, and for his encouragement and support throughout this journey. I am forever grateful to Dr Laura Mc Cullagh for her dedication to this work. I took great comfort in the knowledge that, no matter what challenge arose, Laura would be on hand with practical and sage advice. Thank you for your encouragement, patience, pragmatism and continued support. You have been a wonderful mentor!

I started my journey at the National Centre for Pharmacoeconomics (NCPE) as an intern – little did I know what lay ahead. I would like to thank Dr Lesley Tilson for giving me that opportunity and for her ongoing support and advice ever since. I have gained a lot from the insightful perspective of Dr Roisín Adams. I am particularly grateful to her for introducing me to the concept of expert elicitation, and for her support and advice in developing the exercise. Thanks to Roisín, I have been involved in several CAR-T related workshops and discussions, which have enhanced this research immensely.

It has been a privilege to conduct this research at the NCPE. I am so grateful to my colleagues for their insightful discussion, knowledge, support and encouragement. Particular thanks to Marie Harte, who has been extremely generous with her time and expertise. Without Marie's knowledge and input, the text-mining project of this research would not have developed as it did. The statistical support I have received from the entire NCPE statistics team has been invaluable. Notable thanks to Dr Lea Trela-Larsen for her helpful guidance throughout, Dr Joy Leahy for her advice and input on the survival analyses, and Conor Hickey for conducting the data analysis of the expert elicitation. With their help, what seemed like very daunting tasks became much simpler.

Sincere gratitude to the clinicians who took part in the elicitation exercise, for providing their time, insight and contribution to this research.

Thank you to my friends. I am very grateful to have friends who have been so understanding and supportive of my commitments (even if they are not entirely sure what it is I do).

I owe a huge thank you to my partner, Harri. I suspect he felt, at times, that the challenges of writing a PhD thesis pale in comparison to the challenge of living with someone who is writing a thesis. I am immensely grateful for his support through moments of self-doubt and worry, and for his unwavering kindness and good-humour.

Finally, to my family. Special thanks to my granny, Ellen, who has lit hundreds of candles for me over the years. Thank you to my parents, Mairéad and Fionnán, for their hard work and sacrifice, affording me the opportunities that I have been so fortunate to have. I have been extremely lucky to have their support, and the support of my wider family, in all my endeavours.

I dedicate this thesis to my maimeó, Máire. A dedicated teacher (and a wonderful granny), she had great enthusiasm for teaching and learning. She would have loved this.

Table of Contents

List of Figures

List of Tables

Abbreviations

AIC	Akaike Information Criteria
ALL	Acute Lymphoblastic Leukaemia
AlloSCT	Allogeneic Stem Cell Transplant
ATMP	Advanced Therapy Medicinal Product
AutoSCT	Autologous Stem Cell Transplant
BEAM	Carmustine, Etoposide, Cytarabine, Melphalan
BFM	Berlin-Frankfurt-Munich
BIC	Bayesian Information Criteria
CADTH	Canadian Agency for Drugs and Technologies in Health
CAR	Chimeric Antigen Receptor
CDS	Community Drugs Schemes
CEAC	Cost-Effectiveness Acceptability Curve
CI	Confidence Interval
CNS	Central Nervous System
CPU	Corporate Pharmaceutical Unit
CR	Complete Response
CrI	Credible Interval
CRS	Cytokine Release Syndrome
CSO	Central Statistics Office
DA-R-EPOCH	Dose-Adjusted; Rituximab - Etoposide, Doxorubicin, Vincristine, Cyclophosphamide, Prednisolone
DLBCL	Diffuse Large B-Cell Lymphoma
DOH	Department of Health
DPS	Drugs Payment Scheme
DRG	Diagnosis Related Group
EBMT	European Society for Blood and Marrow Transplantation
ECOG	Eastern Cooperative Oncology Group
EFS	Event-Free Survival
EMA	European Medicines Agency
EPAR	European Public Assessment Report
EQ-5D	EuroQOL 5 Dimension
EQ-5D-3L	EuroQOL 5 Dimension 3 Level
EQ-5D-5L	EuroQOL 5 Dimension 5 Level
ERG	Evidence Review Group
ESMO	European Society of Medical Oncology
EVPI	Expected Value of Perfect Information
FACT-Lym	Function Assessment of Cancer Therapy- Lymphoma
FLA-IDA	Fludarabine, Cytarabine, Idarubicin
FLAG-IDA	Fludarabine, Cytarabine, Idarubicin, Granulocyte Colony-Stimulating Factor
G-CSF	Granulocyte Colony-Stimulating Factor
GDP	Gemcitabine, Dexamethasone, Cisplatin
GMS	General Medical Services
GVHD	Graft-Versus-Host-Disease
HIQA	Health Information and Quality Authority
HPO	Healthcare Pricing Office
HR	Hazard Ratio
HRQOL	Health-Related Quality of Life
HSCT	Haematopoietic Stem Cell Transplant
HSE	Health Service Executive
HTA	Health Technology Assessment
HTDA	High-Tech Drug Arrangements
ICER	Incremental Cost-Effectiveness Ratio
ICU	Intensive Care Unit
iNMB	Incremental Net Monetary Benefit
IPD	Individual Patient-Level Data
IPHA	Irish Pharmaceutical Healthcare Association

IPI	International Prognostic Index
ITC	Indirect Treatment Comparison
ITT	Intention-to-Treat
IV	Intravenous
LDH	Lactate Dehydrogenase
LTI	Long Term Illness
MAIC	Matching-Adjusted Indirect Comparison
mITT	Modified Intention-to-Treat
MRD	Minimal Residual Disease
NCCN	National Comprehensive Cancer Network
NCCP	National Cancer Control Programme
NCPE	National Centre for Pharmacoeconomics
NCRI	National Cancer Registry Ireland
NE	Not Estimable
NHB	Net Health Benefit
NHL	Non-Hodgkin's Lymphoma
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NMB	Net Monetary Benefit
NoMA	Norwegian Medicines Agency
NOS	Not Otherwise Specified
ODMS	Oncology Drugs Management System
OHE	Office of Health Economics
ORR	Overall Response Rate
OS	Overall Survival
OWSA	One-Way Sensitivity Analysis
PBAC	Pharmaceutical Benefits Advisory Committee
PCRS	Primary Care Reimbursement Service
PedsQL	The Paediatric Quality of Life Inventory
PET	Positron Emission Tomography
PFS	Progression-Free Survival
PMBCL	Primary Mediastinal Large B-Cell Lymphoma
PSA	Probabilistic Sensitivity Analysis
QALY	Quality-Adjusted Life Year
R/R	Relapsed/Refractory
R-CHOP	Rituximab – Cyclophosphamide, Doxorubicin, Vincristine, Prednisolone
R-DHAP	Rituximab – Dexamethasone, Cytarabine, Cisplatin
R-ESHAP	Rituximab – Etoposide, Methylprednisolone, High-Dose Cytarabine, Cisplatin
R-GDP	Rituximab – Gemcitabine, Dexamethasone, Cisplatin
R-GEMOX	Rituximab- Gemcitabine, Oxaliplatin
R-GEM-P	Rituximab- Gemcitabine, Methylprednisolone
R-ICE	Rituximab -Ifosfamide, Carboplatin, Etoposide
R-GIFOX	Rituximab - Gemcitabine, Ifosfamide, Oxaliplatin
RCT	Randomised Controlled Trial
RFS	Relapse-Free Survival
SD	Standard Deviation
SE	Standard Error
SF-36	36-Item Short Form Survey
SF-6D	Short Form- 6 Dimension
SLR	Systematic Literature Review
SMC	Scottish Medicines Consortium
SMR	Standardised Mortality Ratio
SPC	Summary of Product Characteristics
STC	Simulated Treatment Comparison
UK	United Kingdom
US	Unites States of America
VAS	Visual Analogue Scale
VOI	Value of Information

Chapter 1 Introduction

1.1 Introduction

Advanced therapy medicinal products (ATMPs) are an innovative therapeutic class, which encompass gene, somatic-cell, and tissue-engineered therapies (1). CD19 CAR T-cell therapies are cell-based gene therapies, whereby a patient's T-cells are genetically modified to express a protein called chimeric antigen receptor (CAR) (2). Currently authorised CD19 CAR T-cell therapies target advanced, rare diseases. Tisagenlecleucel is indicated for the treatment of paediatric and young adult patients (up to 25 years of age), with B-cell acute lymphoblastic leukaemia (ALL) that is refractory, in relapse post-transplant, or in second or later relapse (herein 'R/R ALL'). Tisagenlecleucel and axicabtagene ciloleucel are licensed for relapsed/refractory (R/R) diffuse large B-cell lymphoma, after two or more lines of systemic therapy, in adult patients (herein 'R/R DLBCL') (3, 4). Axicabtagene ciloleucel is also licensed for the treatment of R/R primary mediastinal large B-cell lymphoma (PMBCL; a subtype of DLBCL), after two or more lines of systemic therapy (4). CD19 CAR T-cell therapies offer treatment options for diseases where limited options exist, and have been proposed to be associated with long-term survival benefit (5). The European Medicines Agency (EMA) accelerated approval of tisagenlecleucel and axicabtagene ciloleucel, was based on phase II, open-label, single-arm data. This was due to their "recognised potential to address, to a significant extent, patients' unmet needs" (6). However, these limited data have raised concerns regarding long-term and relative efficacy and safety (7). Concerns have also been raised regarding their considerable upfront cost and affordability (8-10).

The uncertain evidence base and high cost of CD19 CAR T-cell therapies creates challenges for reimbursement decision-makers. Other treatments and services within the healthcare system may be displaced in order to reimburse them (8). Health technology assessment (HTA) informs evidence-based decision-making in the allocation of scarce healthcare resources within the context of a finite budget. It compares the costs and consequences of new or existing interventions with one or more relevant comparators (11). Consensus is lacking on whether currently available HTA methods are appropriate for the evaluation of innovative CD19 CAR T-cell therapies (12, 13). The need for alternative financing and reimbursement mechanisms, to address affordability issues, has also been proposed (8, 14).

To date, no independent HTAs of these therapies, from the perspective of the state payer in Ireland, have been undertaken. No identified HTAs have conducted value of information (VOI) analysis, to quantify the value of further research to decrease decision uncertainty. Additionally, no identified HTAs have evaluated performance-linked reimbursement agreements, which aim to share financial risk between decision-makers and manufacturers (herein 'Applicants'). The research presented in this thesis can inform Irish policy and future HTA evaluations.

Throughout this thesis, 'R/R ALL' pertains to a paediatric and young adult population (up to 25 years), and 'R/R DLBCL' pertains to an adult population, unless otherwise indicated. Due to the low numbers of patients with R/R PMBCL in Ireland (12 patients with newly diagnosed PMBCL between 1994 and 2015[1]), this subgroup is not explicitly considered in this thesis.

1.2 The Irish Healthcare Setting

In Ireland, healthcare policy provision and expenditure fall under the remit of the Minister for Health, via the Department of Health (DOH). The Health Service Executive (HSE) is responsible for delivery and management of healthcare services. In 2022, over €20 billion was allocated to the HSE for the provision of health services (15).

In 2019, the government funded 74% of health expenditure in Ireland. The remaining expenditure was financed through private funding from health insurance (14%) and household out-of-pocket expenditure (12%) (16). As of December 2019, 46% of the population had private health insurance (17). Total health expenditure (public and private), in 2019, accounted for 6.7% of Gross Domestic Product. This was below the Organisation for Economic Co-operation and Development average (8.8%) (18).

Public health expenditure (capital and revenue) has increased steadily, from a low of €13.4 billion in 2013 to over €20 billion in 2020. The low in 2013 is partly reflective of fiscal measures implemented in response to Ireland's economic bailout of 2010. Preliminary data indicate that approximately €26.4 billion was spent on healthcare in

[1] Written correspondence with the National Cancer Registry Ireland.

2020, representing an increase of 11% on 2019 expenditure (16). The COVID-19 pandemic is known to have had a considerable impact on expenditure.

1.3 Reimbursement of Drugs in Ireland

The HSE Primary Care Reimbursement Service (PCRS) is responsible for the reimbursement of drugs in the primary care setting, through the Community Drugs Schemes (CDS). The Oncology Drugs Management System (ODMS) manages the funding of specific hospital-administered systemic drugs for cancer in public hospitals. All new drugs, for which reimbursement is sought under these schemes, are subject to HTA appraisal (conducted by the National Centre for Pharmacoeconomics (NCPE)). Recommendations by the NCPE are used to inform the reimbursement decision-making process. This is discussed in 1.5.2.

1.3.1 Community Drugs Schemes

The CDS encompass several reimbursement schemes. The General Medical Services (GMS) Scheme provides free or subsidised, at the point of care, healthcare services to patients who are unable, without undue hardship, to arrange such services for themselves and their dependents. It is means-tested. Patients, 70 years and under, pay a co-payment of €1.50 per prescription item (maximum €15.00 per family per month). Patients over 70 years pay a co-payment of €1.00 (maximum €10.00 per family per month). As of December 2020, approximately 35% of the population receive prescription drugs under this scheme (19).

The Drugs Payment Scheme (DPS) applies to those not eligible for the GMS Scheme. Here, the maximum an individual or family unit pays, per calendar month, towards prescription drugs is €100 (20). In 2020, approximately 6% of the population availed of this scheme (19).

Under the Long-Term Illness (LTI) Scheme, patients with 1 or more of 16 specified medical conditions are entitled to free (at the point of care) drugs, medical and surgical appliances relevant to their condition. It is not means-tested and covers conditions such

as multiple sclerosis, haemophilia, and diabetes mellitus. No co-payments apply. In 2020, approximately 4% of the population availed of this scheme (19).

The High-Tech Drug Arrangements (HTDA) cover a category of drugs which are high cost, are generally only prescribed or initiated in hospital, and are dispensed by community pharmacies. Patients pay for HTDA drugs according to their registered scheme (GMS, DPS, LTI).

Several smaller schemes exist under the CDS. These include the European Economic Area (EEA) entitlements and the Opioid Substitution Treatment Scheme (19).

Patients are not obliged to avail of any of the CDS that they are eligible for. Instead, they may choose to pay for drugs privately; no drug costs are incurred by the HSE.

1.3.2 Oncology Drugs Management System

Drugs covered by the ODMS are high-cost drugs for cancer, which have been recommended by the National Cancer Control Programme (NCCP) Technology Review Committee and approved for reimbursement by the HSE. The NCCP base their recommendations on clinical effectiveness, safety, and cost effectiveness (21). The hospital pays for the drug and is reimbursed by the PCRS (22).

1.3.3 Drugs Reimbursed in Hospitals

Drugs dispensed in hospitals, which are not covered under the ODMS, are funded through the hospital block grant. This is not linked to other health service budgets.

1.4 Expenditure on Drugs in Ireland

CDS expenditure was approximately €2.3 billion in 2020 (19). A decline in expenditure on the GMS Scheme, from 2012 onwards, was due to several factors including the introduction of reference pricing, generic substitution, and HSE Medicines Management Programme initiatives (23, 24). In contrast, HTDA expenditure has increased steadily. This increase, Figure 1, is driven by the introduction of new, higher cost drugs, and growth in patient volume (25).

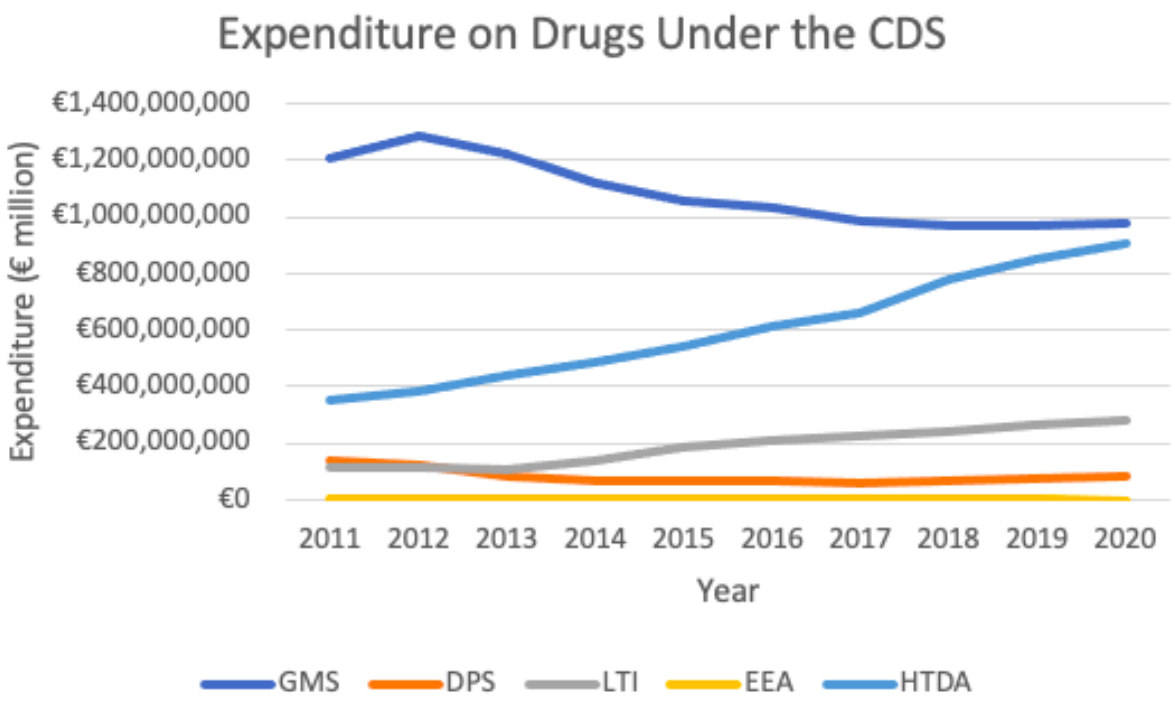

Figure 1 Expenditure on drugs under the CDS 2011 to 2020
CDS: Community Drugs Schemes; **DPS:** Drugs Payment Scheme; **EEA:** European Economic Area; **GMS:** General Medical Services; **HTDA:** High-Tech Drug Arrangements; **LTI:** Long Term Illness.

1.5 Health Technology Assessment

Within a finite healthcare budget, decisions must be made about where to invest and how to allocate resources. HTA is a tool that supports evidence-based decision-making. It aims to inform the formulation of safe, effective, patient-focused health policies and to achieve best value, limit opportunity cost, and maximise population wellness (26).

In the following sections, HTA and economic evaluation are referred to specifically in the context of drug reimbursement decision-making. 'Intervention' is defined as the drug under consideration for reimbursement. 'Comparator' is current routine care (pharmacological and/or non-pharmacological) in the jurisdiction under consideration. The economic evaluation of drugs is referred to a pharmacoeconomic evaluation (27).

1.5.1 Key Concepts in Economic Evaluation

Economic evaluation is one component of HTA, which compares the costs and consequences of new or existing interventions with one or more relevant comparators (11). The key evaluations in healthcare are cost-effectiveness analysis and cost-utility analysis.

1.5.1.1 Cost-Effectiveness Analysis

Cost-effectiveness analysis compares costs and outcomes of the intervention(s) and comparator(s) in natural units (e.g. migraine attacks avoided) (28). The output of this analysis is the incremental cost-effectiveness ratio (ICER). This is a measure of the additional cost for the intervention relative to the comparator, per unit of health outcome (described in 1.5.1.5) (28). The ICER, generated in a cost-effectiveness analysis, cannot be directly compared across interventions and comparators with different outcomes.

1.5.1.2 Cost-Utility Analysis

Cost-utility analysis presents outcomes in terms of quality-adjusted life years (QALYs). The QALY combines length and quality of life; one QALY represents one year of life in perfect health (29). QALYs are weighted using utility values. Utility values represent the preference of individuals for a given health state. To derive a utility value, a description of the health state and a value for it are required. National Guidelines for the Economic Evaluation of Health Technologies in Ireland (herein 'National Economic Evaluation Guidelines') recommend that the health-state description is derived using a generic measure (EQ-5D-3L is the preferred measure of the NCPE (30)) (11). Generic preference-based measures are recommended, as they are widely available, easy to interpret, and use preferences from the general population. A utility value is derived by applying a value to that health state based on societal preferences (11).

The use of the QALY facilitates comparison of outcomes of interventions and comparators across different disease areas. The output of this analysis is the ICER, described in 1.5.1.3 (28). Cost-utility analysis is the preferred type of economic evaluation conducted in the Irish healthcare setting (11). Notably, cost-utility analyses describe the 'cost effectiveness' of an intervention.

1.5.1.3 The Incremental Cost-Effectiveness Ratio

The ICER for Intervention A (drug under consideration) versus Comparator B (current routine care) is calculated as:

$$ICER = \frac{Cost\ A - Cost\ B}{Effect\ A - Effect\ B}$$

Where: **ICER=** Incremental cost-effectiveness ratio; **Cost**= Expected costs; **Effect**= Expected outcomes.

1.5.1.4 The Cost-Effectiveness Threshold

A decision rule may be defined such that the intervention under consideration is deemed cost effective if the ICER falls below a pre-defined willingness-to-pay threshold (λ). In Ireland, most drugs that have been reimbursed to date have been considered under a willingness-to-pay threshold of €45,000 per QALY (31). The empirical basis of this threshold is lacking (32). It is likely based on the £30,000 per QALY threshold used in decision-making by The National Institute for Health and Care Excellence (NICE), in the UK (33).

1.5.1.5 The Incremental Cost-Effectiveness Plane

ICERs are plotted on a cost-effectiveness plane (Figure 2). The comparator is represented by the intersection of the x- and y-axis (the origin); the incremental costs and outcomes of the intervention under investigation are plotted relative to those of the comparator (28). Interventions, which are more costly and more effective, lie in the north-east (NE) quadrant of the plane. The majority of interventions under consideration fall here (28). Interventions that fall below the cost-effectiveness threshold line are considered cost effective, while those that fall above the line are not. Interventions, which are less costly and less effective, lie in the south-west (SW) quadrant. In this instance, a decision-maker must decide whether they are willing to accept lower efficacy for lower cost. ICERs in the NE and SW quadrants are positive.

ICERs may be negative. Interventions, which are less costly and more effective than the comparator, lie in the south-east (SE) quadrant. In this instance, the intervention dominates the comparator and is cost effective. Interventions, which are more costly and less effective, lie in the north-west (NW) quadrant. Such interventions are not cost effective. ICERs in the SE and NW quadrants are not easily interpreted, unless plotted on the cost-effectiveness plane or presented in disaggregated form (i.e. incremental costs, incremental QALYs) (28).

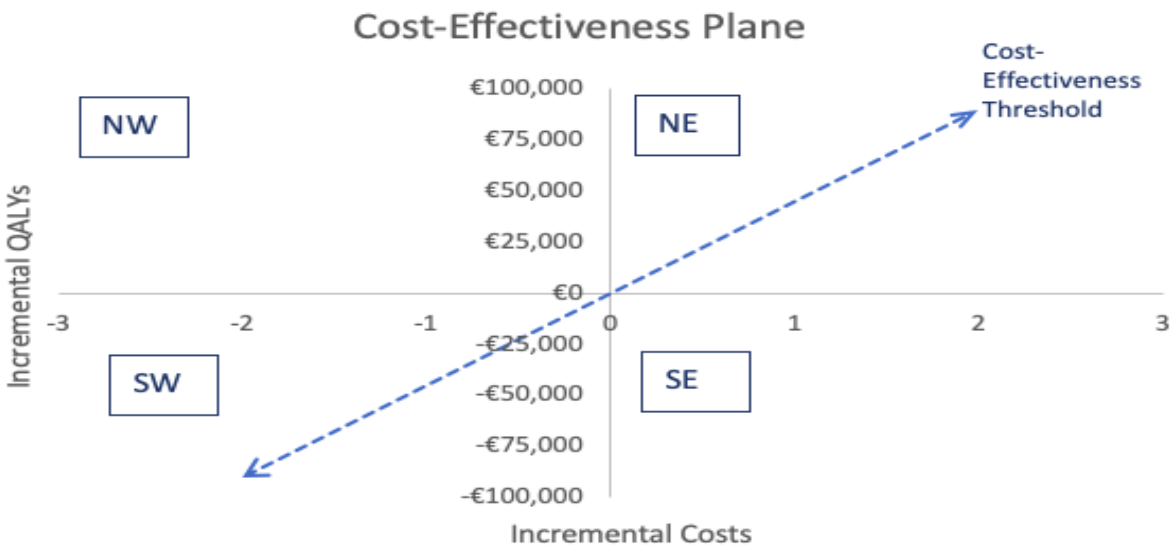

Figure 2 The incremental cost-effectiveness plane*
NE: North-east quadrant; **NW:** North-west quadrant; **QALYs:** Quality-adjusted life years; **SE:** South-east quadrant; **SW:** South-west quadrant.
*In Ireland, the cost-effectiveness threshold for drugs is €45,000 per QALY.

1.5.1.6 Incremental Net Monetary Benefit

Cost-utility analyses outputs can also be presented using the incremental net monetary benefit (iNMB). Interventions with a positive iNMB are cost effective at the threshold under consideration. Challenges in the interpretation of negative ICERs can be avoided using iNMB.

The iNMB is calculated as (34):

$$iNMB = (incremental\ QALYs * \lambda) - incremental\ costs$$

Where: **iNMB**= Incremental net monetary benefit; **Incremental QALYs**= Difference in QALYs between intervention and comparator; λ=Cost-effectiveness threshold; **Incremental Costs**= Difference in costs between intervention and comparator.

1.5.1.7 Handling Uncertainty in Economic Evaluation

Uncertainty in economic evaluation is broadly characterised into two key sources: parameter and structural (28). Parameter uncertainty relates to uncertainty in the precision with which an input parameter is estimated (34). Structural uncertainty relates to the scientific judgements made when constructing a model (28). These judgements include the choice of comparator(s) and evidence synthesis techniques.
Methods to characterise uncertainty are discussed (28).

1.5.1.7.1 One-Way Sensitivity Analysis

One-way sensitivity analysis (OWSA) varies individual parameters across a range of values. All other parameters are held constant at their base case value (35). The impact on the ICER is assessed. Parameters associated with the greatest impact are identified. OWSA outputs are usually presented as a tornado diagram (34).

1.5.1.7.2 Scenario Analysis

Scenario analysis examines the impact of structural and parameter uncertainty on the ICER. It involves employing alternative methods (e.g. evidence synthesis techniques) or varying model parameters (11).

1.5.1.7.3 Probabilistic Sensitivity Analysis

Probabilistic sensitivity analysis (PSA) examines the impact on the ICER when all parameters are varied simultaneously. This involves applying pre-defined distributions to the parameters. Monte Carlo Simulation is then generally used to sample from these distributions, across plausible ranges, at random. The results of the cost-effectiveness model are recorded for each sample (herein 'iteration'). This is repeated thousands of times (35). Results of PSA can be presented in several ways, as described in the following sections.

1.5.1.7.3.1 Probabilistic Scatterplot

A probabilistic scatterplot, plotted on the cost-effectiveness plane, provides a visual representation of the spread of incremental costs and QALYs derived from each PSA iteration (36).

1.5.1.7.3.2 The Cost-Effectiveness Acceptability Curve

The cost-effectiveness acceptability curve is a graphical representation of the probability of a drug being cost effective as a function of the cost-effectiveness threshold. The net monetary benefit (NMB) for each drug under investigation is estimated, using the expected costs and expected QALYs derived from each PSA iteration (34):

$$NMB = (QALYs * \lambda) - costs$$

Where: **NMB**= Net monetary benefit; **QALYs**= Expected QALYs; λ= Cost-effectiveness threshold; **Costs**= Expected costs.

For each PSA iteration, the drug with the maximum NMB is identified. The probability of cost effectiveness for each alternative drug is calculated from the proportion of iterations where this is the case.

1.5.1.7.4 Expected Value of Perfect Information

VOI analysis estimates the impact of reducing uncertainty in parameters (37). Expected value of perfect information (EVPI) estimates the value of simultaneously eliminating all uncertainty of all uncertain parameters (11). It places an upper bound on the value of the additional research that would be required to decrease this uncertainty (34). Each PSA iteration represents a possible future resolution of the existing uncertainty for which the optimal decision can be identified (38). If EVPI exceeds the expected cost of further research to reduce uncertainty, then it is potentially cost effective to conduct this research. If EVPI suggests that additional research would be of value, the parameters that contribute most to the overall decision uncertainty can be identified by partial EVPI (EVPPI) analyses (11). EVPI and EVPPI are discussed in Chapter 5 and Chapter 9.

1.5.2 Economic Evaluation in Ireland

The Health Information and Quality Authority (HIQA) is the statutory authority for HTA in Ireland. The majority of HTA outputs by HIQA to date relate to medical devices, diagnostics, and public health programmes (39). National Economic Evaluation Guidelines are published by HIQA and informed by stakeholders (including the NCPE) (11). The NCPE appraises HTA submissions from Applicants to inform state-payer, drug reimbursement decisions in Ireland (32). A number of academic centres in Ireland also conduct Independent HTAs.

1.5.2.1 The National Centre for Pharmacoeconomics

The NCPE is an independent group comprising pharmacists, clinicians, health economists, statisticians, information specialists, and pharmaco-epidemiologists. The NCPE aim to promote expertise in Ireland for the advancement of pharmacoeconomics through

practice, research, and education (40). The NCPE is commissioned by the HSE Corporate Pharmaceutical Unit (CPU) to appraise Applicant HTA submissions of new and existing technologies for which reimbursement, by the state payer, is sought. The HSE CPU acts as the interface between the HSE and the pharmaceutical industry in relation to drug pricing and reimbursement applications, and the operation of the national pricing framework agreements (41).

1.5.2.2 The Health (Pricing and Supply of Medical Goods) Act 2013

The role of pharmacoeconomic evaluation, in the HTA process, is outlined in Irish legislation under the terms of the Health (Pricing and Supply of Medical Goods) Act 2013 (herein 'Health Act'). This Health Act outlines nine criteria, which the HSE must consider when making reimbursement decisions. Cost effectiveness is one such criterion. Other criteria include the potential or actual budget impact; the efficacy, effectiveness, and expected added therapeutic benefit; the level of certainty in the evidence; the health need of the public; the clinical need for the therapy; the availability of the therapy for supply; the resources available to the HSE; and the level of clinical supervision required to ensure patient safety (42).

1.5.2.3 The Framework Agreement

The framework surrounding processes for reimbursement of new drugs in Ireland, along with supply arrangements, and national pricing agreements is outlined in the Framework Agreement (herein 'Agreement'). This Agreement was drawn up between the HSE, DOH, the Department of Public Expenditure and Reform, and the Irish Pharmaceutical Healthcare Association (IPHA, which represents the pharmaceutical industry). It is renegotiated every four years (41).

1.5.2.4 Reimbursement Process in Ireland

As per the current Agreement (2021 to 2025 inclusive), Applicants seeking reimbursement of drugs must apply to the HSE CPU. All drugs initially undergo a Rapid Review assessment. The Applicant submits a completed NCPE Rapid Review Template, outlining information about the drug including the indication, target population, clinical efficacy and safety, comparators, price and potential cumulative five-year budget impact (43). The NCPE assessment encompasses a targeted literature review and appraisal of the

Rapid Review submission. This is generally completed within four weeks (32). The Rapid Review process is a pragmatic approach to facilitate prioritisation of resources, whilst maintaining the integrity of the evaluation process. The process identifies drugs, which do not require full HTA. If a full HTA is not warranted, a reimbursement recommendation may be made at this stage. An exploratory analysis, conducted by Varley et al., found that, without the Rapid Review process, an additional 15,631 NCPE appraisal days would have been required to evaluate all drugs submitted over a 10-year period (44).

A total of 71 Rapid Review assessments were conducted by the NCPE in 2021. Of these, 27 were for cancer drugs, 10 were for orphan drugs, and 34 were for drugs for other indications. Of these, 37 drugs were recommended for full HTA.

When a full HTA is deemed necessary, the HSE CPU commissions the NCPE to undertake it. The Applicant is required to submit a full HTA to the NCPE, as per the NCPE Applicant Submission Template (30). An NCPE Review Group, generally comprising a lead assessor, a statistician, and an information specialist, critically appraise all submitted documents. The NCPE Senior Management Team oversees the appraisal. This appraisal is comprehensive and evaluates criteria including clinical evidence supporting product registration, comparative effectiveness evidence, evidence synthesis techniques, model input parameters, uncertainty, and budget impact. The NCPE Review Group also validate the submitted economic model (32). The NCPE Review Group may seek clarifications and modifications from the Applicant. The appraisal-time clock is paused until clarifications and modifications are received by the NCPE. The mean appraisal time is 133 days (45). A draft HTA Appraisal Report is sent to the Applicant to check for factual accuracy. The Applicant is required to respond within seven working days. The NCPE Review Group finalise the HTA Appraisal Report (which includes the reimbursement recommendation) and produce a Technical Summary and Plain Language Summary. The reports are sent to the HSE CPU and HSE National Drugs Committee for consideration. In the case of drugs for cancer, the reports are also sent to the NCCP Technology Review Committee. The Technical Summary and Plain Language Summary are published on the NCPE website (32).

Reimbursement decisions are made by the HSE National Drugs Committee and the HSE Leadership Team (31). The HSE National Drugs Committee comprises various heads of HSE Directorates, medical consultants, public interest members, and ethicists. The HSE National Drugs Committee advises the HSE Leadership Team. In instances where funding is not available from existing resources, the HSE can inform the DOH who may bring a memorandum to Government in relation to funding implications (31).

The HSE National Drugs Committee and HSE Leadership Team may consider any patient access schemes (PAS) proposed by the Applicant. Where ICERs exceed the €45,000 per QALY threshold, confidential price negotiations, between the HSE and Applicant may ensue. The NCPE informs these negotiations. It is believed that the payer threshold may not always be met. In specific instances, other Heath Act criteria, such as unmet need, may also be considered key. O'Mahony and Coughlan argue that the €45,000 per QALY threshold in Ireland represents more of a price floor than a price ceiling. Thus, acting as a weak barrier to drugs which pose a net harm (33). When a positive reimbursement decision is made, the Agreement notes that reimbursement is to be implemented within 45 days (41).

1.6 Acute Lymphoblastic Leukaemia (Paediatric and Young Adult)

1.6.1 Disease Overview

ALL is a cancer of the bone marrow, characterised by abnormal proliferation of lymphoblasts. This proliferation inhibits normal production and function of red blood cells, white blood cells, and platelets. These malignant lymphoblasts may spread to and infiltrate other organs such as lymph nodes, spleen, central nervous system (CNS), and testicles. ALL develops rapidly and displays an aggressive course (46). ALL in paediatric and young adult patients relates to patients up to 25 years of age[2].

1.6.2 Disease Categorisation

ALL is broadly categorised according to immunophenotype, B-cell (80% to 85% of cases) and T-cell (10% to 15%) (47, 48). The use of more intensive therapies in T-cell ALL has

[2] Written correspondence with one consultant haematologist in Ireland.

improved outcomes in these patients (49, 50). However, outcomes remain inferior to those of B-cell ALL (47).

1.6.3 Pathophysiology and Aetiology

Genetic risk factors include congenital syndromes such as Down syndrome and Fanconi anaemia. Chromosomal aberrations are notable features in ALL (e.g. rearrangement of *MLL*), yet none are sufficient to generate leukaemia (51, 52). ALL appears as a de novo malignancy in most patients, with no recognised inherited factors (52, 53). Environmental risk factors, such as exposure to radiation and certain chemicals, have been linked to an increased risk in a minority of cases (53).

ALL is likely to arise from an interaction between genetic and environmental risk factors. Evidence suggests that a pre-natal mutation leads to development of a pre-leukaemic clone in utero, which expands postnatally (54, 55). A second mutation is required for overt disease to develop (56). Several hypotheses propose to explain what causes a second mutation. Some propose that ALL arises because of an abnormal immune response to common infections (57-59).

1.6.4 Diagnosis

Patients with ALL may present with fatigue, nausea, fever, shortness of breath, dizziness, palpitations, recurrent infection, and bruising (56). Involvement of extramedullary sites (approximately 20% of patients) can cause lymphadenopathy, splenomegaly, or hepatomegaly (60). CNS involvement, at the time of diagnosis, occurs in approximately 5% to 8% of patients (61).

Diagnosis is generally established by demonstration of 20% or greater lymphoblasts in the bone marrow or peripheral blood. However, there is no definitive limit for the proportion of lymphoblasts required to make a diagnosis (60). Assessment of morphology, flow cytometry, immunophenotyping, and cytogenetic testing can support diagnosis and risk stratification (52).

1.6.5 Disease Staging

There is no universal system to classify ALL according to risk status. Different co-operative groups use various combinations of clinical, biological, and response variables to classify risk (60, 62, 63). Some of these factors are described in 1.6.6. The UKALL 2011 Regimen, used to treat newly diagnosed ALL in Ireland, stratifies patients into standard-risk and high-risk disease. Classification of risk groups, in this instance, is based on age and white blood cell count (64).

1.6.6 Prognostic Factors

Prognostic factors can be broadly categorised into clinical features, biologic and genetic features, and response to treatment (53).

1.6.6.1 Clinical Features

Patient age, white blood cell count, and presence of CNS disease are key factors in determining risk and assessing prognosis. Patients aged between one and nine years are considered to have standard-risk disease. Patients aged 10 years and older, and aged below 1 year, have high-risk disease (53). White blood cell count of 50,000 per cubic milliliter or greater, and presence of CNS disease are associated with worse prognosis (53, 60). Females have been shown, in some studies, to have a better prognosis than males (65, 66).

1.6.6.2 Biologic and Genetic Features

Low hypodiploidy (patients with 30 to 39 chromosomes), near haploidy (patients with less than 30 chromosomes), and *KMT2A (MLL)* translocations are amongst some recognised biomarkers of high-risk disease (60). Philadelphia chromosome-positive ALL is relatively uncommon in paediatric and young adult ALL, but is associated with poor prognosis (60).

1.6.6.3 Response to Treatment

Minimal residual disease (MRD) is the presence of leukaemia cells at a level, which fall below the detection threshold of conventional morphologic methods. A strong correlation has been identified between the presence of MRD and the risk of relapse

(67). MRD has been proposed to be the most powerful prognostic factor in all age groups (67, 68).

1.6.7 Epidemiology

1.6.7.1 Incidence

ALL accounts for approximately 80% of all leukaemia in children. The incidence of ALL rises sharply in children aged between 1 and 4 years and gradually rises again among adults, beginning at approximately 50 years (46, 69). In Ireland, ALL accounted for 12% of leukaemia cases diagnosed between 1994 and 2008; 30% of cases were diagnosed before the age of 5 years, and 55% before the age of 15 years (70).

The crude incidence rate (2010 to 2015 inclusive) of ALL in Ireland for females and males aged 0 to 24 years ranged from 2.7 to 3.9 per 100,000 and 2.4 to 3.3 per 100,000, respectively. The National Cancer Registry Ireland (NCRI) projected that 46 cases of ALL are likely to be diagnosed in people aged 24 years and younger each year from 2020 to 2024 inclusive (71). NCRI present data in terms of five-year age groups. Hence, patients aged 25 years are not included in statistics obtained from the NCRI.

Incidence data on patients with R/R disease are not available from the NCRI. Despite reported long-term survival rates of between 80% and 90% in paediatric and young adult patients with newly diagnosed ALL (72), a subset of patients will be refractory to treatment (approximately 2%) or relapse after initial response to therapy (approximately 10% to 15%) (52, 73). Between 30% and 55% of these patients are expected to experience a second relapse (74-77).

1.6.7.2 Mortality

The NCRI reported that the 1-year relative survival for people aged 24 years and younger, diagnosed with ALL in Ireland between 2011 and 2015 inclusive, was 94.7%. The 5-year relative survival for people aged 24 years and younger, diagnosed with ALL in Ireland between 2011 and 2015 inclusive, was 91.2% (71).

Mortality data on patients with R/R disease are not available from the NCRI. Data from the literature reported that survival following first relapse, in paediatric and young adult patients, ranges from 40% to 70% (74, 78, 79). Limited data regarding survival outcomes in patients who experience a second relapse were identified. However, the complete response (CR) rate in paediatric and young adult patients who experience a second, third, and fourth or later relapse has been reported to be 44%, 27%, and 12%, respectively (80).

1.6.8 Treatment of Acute Lymphoblastic Leukaemia

1.6.8.1 Newly Diagnosed Acute Lymphoblastic Leukaemia

In Ireland, paediatric and young adult patients with newly diagnosed ALL are generally treated according to the UKALL 2011 Regimen, which comprises induction, consolidation, and maintenance. The induction regimen depends on UK National Cancer Institute risk staging (standard-risk versus high-risk). The consolidation regimen is dependent upon MRD level at day 29 of induction. The maintenance regimen is contingent on the consolidation regimen and MRD level at week 14 post-consolidation. These regimens (for induction, consolidation, and maintenance) consist mainly of cytotoxic drugs, and include mercaptopurine, vincristine, pegaspargase, intrathecal methotrexate, oral methotrexate, allopurinol, and dexamethasone (64).

The ALLTogether study (recruiting in Ireland at the time of writing) is a pilot observational study in patients (aged 1 to 45 years) with newly diagnosed ALL. The treatment protocol is based on a personalised risk approach, which uses a novel algorithm to incorporate clinical characteristics, genetic factors, and response to therapy. Patients with high-risk B-cell ALL may be stratified to CAR T-cell therapy (not otherwise specified) (81).

1.6.8.2 Relapsed/Refractory Acute Lymphoblastic Leukaemia

Relapsed disease can refer to first relapse, or subsequent relapses of disease, whereby disease initially responds to treatment but then recurs. Refractory disease can refer to primary refractory or chemo-refractory disease. Primary refractory disease does not respond to initial treatment (for newly diagnosed disease). Chemo-refractory disease does not respond to treatment for relapsed disease (i.e. disease initially responded to treatment but relapse occurs and disease does not respond to new treatment).

1.6.8.2.1 First Relapse or Primary Refractory Disease

For paediatric and young adult patients in first relapse or with primary refractory disease, the ALLR3 Regimen (2007) is widely followed in Irish clinical practice. This regimen stratifies patients according to risk status (standard-risk, intermediate-risk, high-risk), as per adapted Berlin-Frankfurt-Munich classification. The regimen comprises three main phases: induction, consolidation, and intensification. Following intensification, patients with standard-risk disease (regardless of MRD level), and patients with intermediate-risk disease with an MRD level <0.01%, receive maintenance. Patients with intermediate-risk disease with an MRD level ≥0.01%, and all patients with high-risk disease, may proceed to allogeneic stem cell transplant (alloSCT). These regimens (for induction, consolidation, and intensification) consist mainly of cytotoxic drugs and include vincristine, pegaspargase, intrathecal methotrexate, oral methotrexate, cyclophosphamide, cytarabine, etoposide, and dexamethasone (82).

Recently in Irish clinical practice, patients with high-risk relapsed ALL and chemo-refractory ALL are treated according to the R3 guideline (2019), drafted by the UK National Cancer Research Institute Leukaemia Sub-Group. Treatment options include re-induction chemotherapy followed by blinatumomab (as a bridge to alloSCT), or enrolment on the CARPALL (CD19 CAR T-therapy, not otherwise specified), or AMELIA (CD19/22 CAR T-cell therapy, not otherwise specified) trials. For patients with high-risk refractory ALL, recommended treatment options include tisagenlecleucel (CD19 CAR T-cell therapy), FLA(G)-IDA (fludarabine, cytarabine, idarubicin, granulocyte colony stimulating factor; cytotoxic chemotherapy), inotuzumab ozogamicin (antibody-drug conjugate), or enrolment on the AMELIA trial (83). At the time of writing, both CARPALL (84) and AMELIA (85) were closed to recruitment.

1.6.8.2.2 Second Relapse or Chemo-Refractory Disease

Patients who experience a second relapse following chemotherapy or alloSCT, and those who have refractory (either primary refractory or chemo-refractory) disease, have limited treatment options. This is the population relevant to this thesis.

Clinical opinion indicated that the therapeutic landscape of R/R ALL (second relapse, relapse following alloSCT, or refractory) is rapidly evolving. Previously, most patients would have been treated with FLA(G)-IDA, followed by alloSCT (in those deemed eligible). However, blinatumomab, followed by alloSCT (eligible patients), is increasingly used and has become routine care in Ireland[3]. Blinatumomab is a bispecific T-cell engager molecule that binds specifically to CD19 expressed on the surface of cells of B-lineage origin and CD3 expressed on the surface of T-cells (86). Blinatumomab is licensed for the treatment of paediatric and young adult patients (aged one year and older) with Philadelphia chromosome-negative CD19-positive B-cell ALL, that is refractory, in relapse after two prior therapies, or relapse after alloSCT (86). EMA authorisation was granted based on a single-arm, phase I/II, multicentre trial in patients, aged less than 18 years (n=70) (the NCT01471782 trial) (87). Further detail on NCT01471782 is provided in Chapter 2. Blinatumomab is also licensed for the treatment of paediatric and young adult patients (aged one year and older) with high-risk first relapsed Philadelphia chromosome-negative CD19-positive B-cell ALL (as part of consolidation chemotherapy) (86). This indication is not relevant to the research presented here.

In 2018, tisagenlecleucel received EMA marketing authorisation for the treatment of R/R ALL (as defined in Chapter 1). It was approved for reimbursement for R/R ALL in Ireland in July 2021 (88). It is anticipated that tisagenlecleucel will displace both blinatumomab and FLA(G)-IDA at this line of treatment. Further discussion on tisagenlecleucel is presented in 1.8.

1.7 Diffuse Large B-Cell Lymphoma

1.7.1 Disease Overview

DLBCL is the most common form of non-Hodgkin's lymphoma (NHL), accounting for 25% to 30% of all NHL diagnoses (89). DLBCL is a cancer of the lymphatic system, characterised by transformation and abnormal growth and proliferation of large B-cells. These abnormal large B-cells diffuse and accumulate in the lymphatic system. Abnormal large B-cells may also arise in extranodal sites such as the gastrointestinal tract, skin, CNS, bone marrow, salivary gland, lung, kidney, and liver (90). DLBCL changes the normal

[3] Written correspondence with one consultant haematologist in Ireland.

architecture of the lymph node or tissue of origin diffusely (91). It develops rapidly and displays an aggressive course (92).

1.7.2 Disease Categorisation

DLBCL is a heterogeneous disease with many variants and subtypes. This is reflected in the highly variable clinical course (93). DLBCL, not otherwise specified (NOS), is the most common subtype (80% to 85% of cases) (94). DLBCL NOS encompasses DLBCL cases that do not fit into any specific disease subgroups, being described as a diagnosis of exclusion (95). The remaining DLBCL subtypes are differentiated on the basis of clinical findings, morphology, immunophenotypic and molecular/genetic studies (96).

1.7.3 Pathophysiology and Aetiology

The pathogenesis of DLBCL is a complex, multistep process. DLBCL arises from mature B-cells at different stages of differentiation. Gene mutations are responsible for changes in B-cells, altering the gene expression and promoting a neoplastic transformation (90). Mutations can be caused by chromosomal translocations, aberrant somatic hypermutations, sporadic somatic mutations and copy number alterations (97).

Aetiology is multifactorial. DLBCL can arise de novo or can develop from transformation of indolent diseases such as chronic lymphocytic lymphoma, follicular lymphoma, and marginal zone lymphoma. Exposure to pesticides, herbicides, alkylating agents, or ionizing radiation has been identified as a risk. Hepatitis C virus seropositivity, human immunodeficiency virus infection, autoimmune disease, and a family history of lymphoma have all been identified as risk factors (98).

1.7.4 Diagnosis

DLBCL generally presents as a rapidly growing, non-painful mass involving one or more lymph nodes. Approximately 40% of patients present with extranodal disease. Approximately 30% may also present with ‘B symptoms’, which include fever, weight loss, and night sweats (94).

The optimal method of diagnosis has been proposed to be a surgical excision biopsy (98). Immunophenotypic assessment, by means of immunohistochemistry or flow cytometry, is conducted to identify subtype and aid risk stratification (99).

Suspected relapse, based on imagining studies (positron emission tomography (PET)/computerised tomography), is confirmed by needle core biopsy (98).

1.7.5 Disease Staging

A physical examination, performance status evaluation, and assessment of B symptoms assist in disease staging. Laboratory assessments including complete blood count, serum lactate dehydrogenase (LDH) and uric acid levels are also conducted (98, 99).

The Ann Arbor staging system was the first disease-staging tool. This system classifies DLBCL into Stages I (localised disease) to IV (extensive disease). 'A' or 'B' indicate the absence or presence of B symptoms, respectively. An 'X' included in the nomenclature indicates the presence of bulky disease (100).

The Lugano Classification criteria was introduced in 2014. Ann Arbor Stage I and Stage II disease are grouped; patients with these stages and are considered to have Limited Stage disease. Ann Arbor Stage III and Stage IV disease are grouped; patients are considered to have Advanced Stage disease. These criteria do not differentiate between the absence and presence of B symptoms. The 'X' is replaced by the recording of the largest nodal diameter by computerised tomography (101-103).

1.7.6 Prognostic Factors

Due to heterogeneity of disease, the prognostic value of disease staging is limited. As such, additional prognostic factors are also assessed (98, 103).

1.7.6.1 The International Prognostic Index

The International Prognostic Index (IPI) is used to predict outcomes in patients with DLBCL. Factors considered include age over 60 years, elevated LDH, Eastern Cooperative Oncology Group (ECOG) performance status of two or greater, clinical disease stage III or IV, and more than one extranodal disease site. IPI assigns 1 point for the presence of

each factor out of a total of 5 points. An IPI score of 0 or 1 indicates low-risk disease. A score of 2 indicates low-intermediate risk. A score of 3 indicates high-intermediate risk. A score of 4 or greater indicates high-risk disease (104).

1.7.6.2 Cell of Origin

Identification of cell of origin, through gene expression profiling, is used to divide DLBCL into three unique subtypes. Germinal centre B-cell, activated B-cell, and unclassifiable. Patients with activated B-cell disease have poorer prognosis, when treated with first-line rituximab-based regimens, compared to patients with germinal centre B-cell disease (105, 106). The prognostic impact of cell of origin in patients with relapsed disease is less clear (107).

1.7.6.3 Molecular Features

DLBCL with a *MYC* rearrangement concurrent with a rearrangement in *BCL2* or *BCL6* is referred to as double-hit lymphoma. DLBCL with a *MYC* rearrangement concurrent with a rearrangement in both *BCL2* and *BCL6* is referred to as triple-hit. Patients with double-hit and triple-hit lymphoma generally have poor prognosis (94, 107).

1.7.7 Epidemiology

1.7.7.1 Incidence

DLBCL is the most common type of lymphoma globally (108, 109). The probability of developing DLBCL increases with age; from 0.13% for men and 0.09% in women under 39 years, to 1.77% in men and 1.4% in women over 70 years (110). Median age at diagnosis is 64 years (91).

The crude incidence rate in Ireland (2010 to 2015 inclusive) ranged from 4.8 to 5.4 per 100,000 in women and from 5.7 to 7.6 per 100,000 in men. In the same period, the number of cases of DLBCL diagnosed each year (in people aged over 20 years) ranged from 238 to 290 cases. These data are inclusive of patients with PMBCL (12 cases diagnosed between 1994 and 2015) (111).

Incidence data on R/R disease are not available from the NCRI. The proportion of patients who experience first relapse or refractory disease ranges from 30% to 40% (108, 112-114). The proportion of patients receiving second-line therapy ranges from 11% (115) to 33% (112). The proportion who proceed to third-line therapy ranges from 45% (116, 117) to 71% (118).

1.7.7.2 Mortality

Data from the NCRI indicate that the 1-year relative survival for people aged 18 to 99 years, diagnosed with DLBCL in Ireland between 2011 and 2015, was 74.1%. The 5-year relative survival was 61.9% (111).

Mortality data pertaining to patients with R/R disease are not available from the NCRI. For patients with first relapse or primary refractory disease, who receive salvage chemotherapy followed by autologous stem cell transplant (autoSCT), the 5-year overall survival (OS) ranges from 31% to 53%. The 5-year OS, of patients with relapsed or primary refractory disease, who receive salvage chemotherapy but do not proceed to autoSCT, ranges from 17% to 32% (119, 120). Age, comorbidities, or poor fitness may prevent patients from receiving autoSCT (121).

Mortality data on patients who have chemo-refractory disease or who experience relapse after second-line therapy are limited. These patients have poor prognosis (122). Median OS of patients who receive third-line therapy ranges from 4.4 to 10 months (123, 124).

1.7.8 Treatment of Diffuse Large B-Cell Lymphoma

1.7.8.1 Newly Diagnosed Diffuse Large B-Cell Lymphoma

There are no up-to-date published national treatment guidelines for the treatment of patients with DLBCL in Ireland. The HSE NCCP develops and publishes National Chemotherapy Regimens to support safe, evidence-based and cost-effective cancer treatment (125). These provide guidance on dosage and eligibility; they do not indicate preferred regimens (126). R-CHOP (rituximab – cyclophosphamide, doxorubicin,

vincristine, and prednisolone; cytotoxic chemotherapy) is generally considered routine care in Ireland, for patients with newly diagnosed DLBCL (91).

1.7.8.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

1.7.8.2.1 First Relapse or Refractory Disease

For patients aged less than 70 years, who have good ECOG performance status (0 to 1), and no major comorbidities, salvage chemotherapy with rituximab-based regimens, followed by high-dose chemotherapy (e.g. BEAM – carmustine, etoposide, cytarabine, melphalan) and haematopoietic stem cell transplant (HSCT; autoSCT (relapsed disease) or alloSCT (refractory disease)), in eligible patients, is generally considered routine care in Ireland (91). Commonly used regimens include R-DHAP (rituximab – dexamethasone, cytarabine, cisplatin), R-ESHAP (rituximab – etoposide, methylprednisolone, high-dose cytarabine, cisplatin), and R-ICE (rituximab -ifosfamide, carboplatin, etoposide). No regimen has demonstrated superior efficacy over others. Up to 50% of patients who are eligible for HSCT will proceed to HSCT (usually autoSCT), with 25% to 35% of patients having favourable long-term survival outcomes (127, 128).

Patients aged 70 years and older, those with poor ECOG (2 or above) performance status and those with considerable comorbidities may be treated, in Irish clinical practice, with the rituximab-based salvage chemotherapy regimens described above. Less toxic regimens such as R-GEMOX (rituximab- gemcitabine, oxaliplatin) may also be used. However, these patients do not proceed to high-dose chemotherapy and HSCT (91). The majority of patients with relapsed DLBCL fall into this category (129).

1.7.8.2.2 Second Relapse or Refractory Disease

Patients experiencing a second relapse, or those who have refractory disease, may be considered for salvage chemotherapy followed by HSCT (if eligible), usually alloSCT. This is the population relevant to this thesis. There is no universal routine care for these patients. In Irish practice, R-GEMOX or R-GDP (rituximab - gemcitabine, dexamethasone, cisplatin) are commonly administered. Approximately 15% of patients receive HSCT in the third-line setting. Prognosis of patients who do not receive HSCT is poor and salvage chemotherapy is usually administered with palliative intent (98, 130).

In 2018, tisagenlecleucel and axicabtagene ciloleucel received EMA marketing authorisation for the treatment of R/R DLBCL (as defined in Chapter 1). Tisagenlecleucel was approved for reimbursement in Ireland in July 2021 (10), while axicabtagene ciloleucel was approved in January 2022 (131). It is anticipated that tisagenlecleucel and axicabtagene ciloleucel will displace salvage chemotherapy followed by HSCT (in eligible patients) at this line of therapy.

Treatment of patients with R/R PMBCL is generally aligned with that of R/R DLBCL (132).

1.8 CAR T-Cell Therapy

CAR T-cell therapy is adoptive cell therapy, whereby T-cells from a patient, are genetically engineered to express CARs on their surface. CARs are engineered receptors that redirect the specificity, function, and metabolism of T-cells (133). This allows the genetically modified T-cells (i.e. CAR T-cells) to recognise and eliminate cells expressing a specific target antigen (134). When infused back into the patient, CAR T-cells engraft and undergo extensive proliferation (133). CAR T-cells may promote immune surveillance to tumour recurrence through antigen release, by assisting tumour-infiltrating lymphocytes to attack tumours, or by their own persistence (135, 136).

The structure of CARs comprises four domains (134). The extracellular target antigen binding domain binds to the tumour target antigens (133). The hinge region domain optimises accessibility of the epitope (the part to which an antibody attaches) (137). The transmembrane domain anchors the CAR to the T-cell membrane and may be relevant to CAR T-cell function (134, 138, 139). The intracellular signalling domain activates the effector functions of the CAR T-cell (134). CAR T-cells can be categorised into four generations depending on the structure of the intracellular domain. Further, as compared to first-generation CAR T-cells, second-generation CAR T-cells have an additional costimulatory domain. This domain is proposed to improve proliferation and persistence (133, 134). Tisagenlecleucel and axicabtagene ciloleucel are second-generation CAR T-cell therapies.

1.8.1 CD19 CAR T-Cell Therapy

The frequent expression of CD19 in B-cell leukaemia and lymphoma, and its higher expression relative to other potential target antigens, make it an attractive target (133, 140, 141). Additionally, CD19 is not expressed on haematopoietic stem cells and on-target off-tumour activity is limited to B-cell aplasia, an adverse event considered manageable (133, 142).

As highlighted, one CD19 CAR T-cell therapy has been approved by the EMA for the treatment of R/R ALL, and two have been approved for the treatment of R/R DLBCL. These are outlined in Table 1. The clinical efficacy of tisagenlecleucel and axicabtagene ciloleucel is detailed in Chapter 2 (tisagenlecleucel) and Chapter 6 (tisagenlecleucel and axicabtagene ciloleucel).

Table 1 CD19 CAR T-cell therapies licensed by the European Medicines Agency for the treatment of R/R ALL and R/R DLBCL

CD19 CAR T-Cell Therapy	Indication (3, 4)	Clinical Evidence at the Time of Marketing Authorisation (143, 144)
Tisagenlecleucel	Paediatric and young adult patients, up to and including 25 years of age, with B-cell ALL that is refractory, in relapse post-transplant or in second or later relapse	**ELIANA** (n=75): Phase II, single-arm, open-label, multicentre. Median follow up: 13.1 months **ENSIGN** (n=29): Phase II, single-arm, open-label, multicentre (US only) **B2101J** (n=56): Phase I/II, single-arm, single-site, open-label
Tisagenlecleucel	Adult patients with R/R DLBCL, after two or more lines of systemic therapy	**JULIET** (n=99): Phase II, single-arm, open-label, multicentre. Median follow up: 11.4 months
Axicabtagene Ciloleucel	Adult patients with R/R DLBCL and PMBCL, after two or more lines of systemic therapy	**ZUMA-1** (n=101): Phase II, single-arm, open-label, multicentre. Median follow up: 11.3 months

ALL: Acute lymphoblastic leukaemia; **DLBCL:** Diffuse large B-cell lymphoma; **PMBCL:** Primary mediastinal large B-cell lymphoma; **R/R:** relapsed/refractory; **US:** United States.

Tisagenlecleucel and axicabtagene ciloleucel have the same mechanism of action. Their main difference lies primarily in the costimulatory domain. Tisagenlecleucel has a 4-1BB costimulatory domain, while axicabtagene ciloleucel has a CD28 costimulatory domain. It

has been proposed that the 4-1BB costimulatory domain is associated with enhanced persistence (145, 146). The CD28 costimulatory domain has been proposed to be associated with more rapid expansion of CD19 CAR T-cells, which may correlate with more rapid onset of cytokine release syndrome (CRS) (139, 147). However, the evidence supporting this is inconsistent (147). Tisagenlecleucel has a CD8-α transmembrane domain, while axicabtagene ciloleucel has a CD28 transmembrane domain (148).

1.8.1.1 Manufacture of CD19 CAR T-Cell Therapies

The manufacture of CD19 CAR T-cell therapies is a multistep process. T-cells are harvested from the patient via leukapheresis, a process that withdraws blood from the body and removes the required components from the blood (herein 'autologous material'). The autologous material is shipped to the manufacturing facility. In Europe, tisagenlecleucel is manufactured in Germany (137), while axicabtagene ciloleucel is manufactured in the Netherlands (149). At the manufacturing facility, the autologous material is thawed and processed to enrich in T-cells. The T-cells undergo activation with antibody-coated paramagnetic beads. They are transduced with a viral vector containing the anti-CD19 CAR transgene (150). A lentiviral vector is used in the manufacture of tisagenlecleucel, while a gamma-retroviral vector is used for axicabtagene ciloleucel (151). The T-cells are cultured, allowing expansion, until the minimum number of CAR-positive T-cells have been produced (150). Once the culture is complete, the paramagnetic beads are removed. The cells are concentrated and washed (137). The cells are cryopreserved and following product release, the cryopreserved CAR T-cells are shipped to the treating hospital (152). As per the Summary of Product Characteristics (SPC) of tisagenlecleucel, manufacture and release of tisagenlecleucel should take three to four weeks (3). This information is not provided in the SPC of axicabtagene ciloleucel (4). Real-world data, from a cohort of patients receiving either tisagenlecleucel (n=30) or axicabtagene ciloleucel (n=30) for R/R DLBCL in France, indicated that median duration of time between patient selection for CAR T-cell therapy and receipt of CD19 CAR T-cells was 47.5 days (range 30 to 190) (153).

1.8.1.2 Administration of CD19 CAR T-Cell Therapies

During the manufacturing process, patients may receive bridging chemotherapy to maintain disease control without compromising organ function or causing considerable toxicity (150, 154). Although bridging chemotherapy was not permitted in the ZUMA-1 trial of axicabtagene ciloleucel (144), real-world evidence suggests that the majority of patients receive this, regardless of CD19 CAR T-cell therapy (153).

Prior to administration of CD19 CAR T-cell therapy, patients may receive lymphodepleting chemotherapy to facilitate in vivo T-cell expansion (137) and reduce disease burden (154). The SPC of tisagenlecleucel specifies that lymphodepleting chemotherapy should be administered unless white blood cell count, within one week prior to tisagenlecleucel infusion, is 1,000 cells per microlitre or less (3). The SPC of axicabtagene ciloleucel does not specify a white blood cell count threshold (4). Fludarabine in combination with cyclophosphamide is the recommended lymphodepleting regimen for patients with R/R ALL or R/R DLBCL (3, 4). For patients receiving tisagenlecleucel and who experienced previous grade 4 haemorrhagic cystitis with cyclophosphamide, or demonstrated a chemo-refractory state to a cyclophosphamide-containing regimen, cytarabine in combination with etoposide (R/R ALL), or bendamustine monotherapy (R/R DLBCL) may be administered (3).

Both tisagenlecleucel and axicabtagene ciloleucel are administered as a once-off intravenous (IV) infusion (3, 4).

1.8.1.3 Adverse Event Profile of CD19 CAR T-Cell Therapies

Tisagenlecleucel and axicabtagene ciloleucel are associated with a distinctive and considerable adverse event profile. Many adverse events are on-target and reverse when the CD19 CAR T-cells have finished expanding, are eradicated, or exhausted (133).

1.8.1.3.1 Cytokine Release Syndrome

CRS is a systematic inflammatory response caused by release of cytokines during T-cell activation and proliferation. Clinical features include fever, nausea, myalgia, fatigue, tachycardia, hypotension, cardiac dysfunction, renal impairment, and hepatic

impairment. CRS typically occurs within the first week following CD19 CAR T-cell infusion (143, 144). Severity is correlated with peak CD19 CAR T-cell therapy and serum interleukin-6 (IL-6) levels (155). Risk factors for severe CRS include high tumour burden, comorbidities, and development of CRS within three days of infusion (156, 157).

Tocilizumab, an IL-6 receptor antagonist, is licensed for the treatment of CAR T-cell therapy-induced severe or life-threatening CRS (158). Tocilizumab can induce rapid reversal of CRS (159). Corticosteroids may also be administered, due to their suppression of inflammatory response (160). However, corticosteroids may impair the function and durability of CD19 CAR T-cells due to their suppression of T-cell function. Thus, use is generally reserved for patients who do not respond to tocilizumab (3, 4). Patients with grade ≥3 CRS, those with rapid deterioration, and those with vital sign instability are generally transferred to the intensive care unit (ICU) (159).

1.8.1.3.2 Neurotoxicity

Neurotoxicity is a well-documented adverse event of CD19 CAR T-cell therapy. However, its pathophysiology is unclear (137). Several hypotheses have been proposed. These include passive diffusion of inflammatory cytokines into the brain (161, 162) and the presence of CAR T-cells in the CNS (156, 157). Clinical manifestation is diverse. Patients may present with decreased attention, dysphasia, and impaired handwriting. Other symptoms may include confusion, disorientation, agitation, aphasia, somnolence, and tremors (143, 144, 159). More severe cases can result in depressed level of consciousness, coma, seizures, motor weakness, and cerebral oedema (159, 163). Of note, neurotoxicity is also referred to as immune effector cell-associated neurotoxicity syndrome (ICANS) and CAR T-related encephalopathy syndrome (CRES) (137, 160).

Patients who experience neurotoxicity are generally managed with supportive care for low-grade toxicity. Corticosteroids may be administered in more severe cases (159). Tocilizumab may be administered where there is concurrent CRS. However, tocilizumab is not expected to cross the blood-brain barrier and has limited efficacy in patients who experience neurotoxicity in isolation (160). In most cases, neurotoxicity is self-limiting and reversible (159).

1.8.1.3.3 B-Cell Aplasia

B-cell aplasia is characterised by extremely low B-cell counts, which may lead to hypogammaglobulinaemia, making patients susceptible to recurrent infections. It occurs when CD19 CAR T-cells attack normal B-cells (164). It can be used as an indicator of CD19 CAR T-cell persistence (156). Patients may experience B-cell aplasia for a prolonged period. Management includes IV immunoglobulin until B-cell recovery and antibiotics to treat infection (164).

1.8.1.4 Risk Management Plan

All healthcare providers expected to prescribe, dispense, or administer tisagenlecleucel or axicabtagene ciloleucel require training. Training should encompass topics such as the identification and management of adverse events, CD19 CAR T-cell thawing and administration, and patient counselling. Additionally, at least one dose of tocilizumab must be available prior to infusion, in the event of CRS occurring following infusion. Patients should be monitored daily for the first 10 days following infusion. After this period, patients can be monitored at the clinician's discretion. Patients should remain within two hours of travel of where the infusion was administered for at least four weeks following infusion (143, 144).

1.9 Challenges in the Evidence Base of CD19 CAR T-Cell Therapies

National Economic Evaluation Guidelines stipulate that, where available, evidence from high-quality randomised controlled trials (RCTs) should be used to quantify efficacy (11, 165). Considering the advanced stage of disease and small population size in patients with R/R ALL and R/R DLBCL, conducting adequately powered RCTs is problematic (12).

Both tisagenlecleucel and axicabtagene ciloleucel received accelerated approval under the EMA PRIME scheme. This aims to accelerate assessment of therapies that target an unmet need (166). It has been highlighted that accelerated approval may come at the expense of quality of evidence (14). EMA marketing authorisation, of tisagenlecleucel and axicabtagene ciloleucel, was granted based on single-arm trials, with a surrogate endpoint (overall response rate; ORR) as the primary endpoint. Median duration of follow up in the respective trials was short and populations were small (143, 144). There

is a requirement for data to be collected on patients for 15 years post-infusion (143, 144); however, this long duration may result in high levels of drop out.

Single-arm trials may be appropriate when the natural history of the disease is well established, the effect size of the therapy is large, and the study population is homogenous (167). This may not be the case for R/R ALL or R/R DLBCL, particularly in the context of heterogeneity of disease. The single-arm nature of these trials requires the use of historical controls to generate comparative effectiveness estimates. Bias in estimates of effectiveness from historical controls can add uncertainty to comparisons (12). Methodological challenges exist in determining the most appropriate approach to indirect treatment comparisons (ITCs), especially in the face of poor reporting of patient characteristics in historical control cohorts (14). There may also be differences in study design or populations, which cannot be adjusted for (the use of bridging chemotherapy in JULIET (tisagenlecleucel), for example). Bias may also arise when not all confounders that affect outcomes are known (14). The short-term evidence requires extrapolation to a lifetime horizon, leading to a high degree of uncertainty in long-term outcomes.

The use of surrogate outcomes as the primary endpoint can be unreliable without sufficient validation (12). Evidence that the technology improves both the surrogate and the final outcome in several clinical trials, is one criterion required to validate the surrogate outcome (168). This level of evidence is not available for CD19 CAR T-cell therapies. A recently published (2017) trial-level meta-analysis in DLBCL found that CR rates were poorly correlated with progression-free survival (PFS) and OS rates at fixed time endpoints in patients with newly diagnosed DLBCL. However, interpretation of results is limited by the fact that median OS was not reached in many of the studies (169). Additionally, correlations in newly diagnosed disease may not be applicable to R/R disease.

There are also uncertainties regarding the requirement for an additional dose at a later time point (14), optimal bridging chemotherapy regimens, and optimal management of adverse events (170).

The availability of long-term efficacy data from second-generation CAR T-cell therapies may be superseded by the availability of next generation therapies.

1.10 Health Technology Assessment of Potentially 'Curative' Therapies

In anticipation of the challenges associated with CD19 CAR T-cell therapies, NICE (UK) commissioned a 'mock' technology appraisal. The aim was to assess whether its existing HTA approaches were appropriate. The mock appraisal comprised SLRs and development of an exemplar case study of CAR T-cell therapy for the treatment of R/R ALL. Hettle et al. concluded that, although the clinical evidence of 'regenerative medicines' (medicines which replace or regenerate human cells, tissues or organs to restore or establish normal function) is expected to be highly uncertain, existing approaches are appropriate (12). It has been proposed elsewhere that the standard methods of cost-effectiveness analysis are appropriate for therapies such as CD19 CAR T-cell therapy, as the challenges encountered are similar to those for other disease areas, particularly rare diseases (171, 172).

The Office of Health Economics (OHE) (Marsden and Towse (13)) appraised the work of Hettle et al. and concluded that this research did not seek to identify the most appropriate approach for assessing regenerative medicines. Rather, the OHE report states, Hettle et al. tested whether regenerative medicines could fit into the existing methods developed for conventional medicines. Marsden and Towse question whether current methods are the most appropriate approach (13).

1.10.1 Adaptions to the Assessment Process

Both the Institute for Clinical and Economic Review (ICER) in the US (2019 (173)) and NICE, UK (2020 (174)) have recently reviewed their assessment methods, highlighting adaptions to their assessment framework when evaluating ATMPs. Other publications have also proposed additional investigations that should be conducted when assessing therapies in this class (14, 175-177). An overview of some of these is provided below.

1.10.1.1 Survival Extrapolation

For survival modelling, cure proportion modelling (most commonly, 'mixture cure modelling') has been recommended to better represent the long-term survival of patients who are potentially cured (173-176). Mixture cure models may be appropriate when it is expected that different subpopulations of the cohort will have different hazard and survival profiles (178). These models consider two subpopulations, those who are 'cured' and those who are not. 'Cured' patients will not die from their disease and are considered to have a mortality equivalent to that of the age- and sex-matched general population. 'Cure' is defined at the population level. A standard parametric survival model defines the survival of patients who are 'not cured'.

Restricted cubic spline modelling has also been recommended (176). Standard parametric models are limited in the types of hazard function they represent. They may not accurately model survival when multiple changes occur in the hazard function. Flexible parametric models were developed to adequately capture such complex hazard functions (178). Restricted cubic spline models (herein 'spline models') are flexible parametric models, defined by piecewise polynomial distributions fitted sequentially to segmented portions of the data (179). These segments are intersected at a pre-defined number of points; known as 'knots' (180). At each knot, the modelled hazards are smoothed, by imposing constraints, where the distributions change. Spline models can capture more complex shapes and enable more realistic hazard and survival functions to be estimated (178). The complexity of the function depends on the number and location of the knots (179). Mixture cure modelling and spline models are examined in Chapter 5 and Chapter 9.

1.10.1.2 Scenario Analysis and Uncertainty

Although the exploration of scenario analyses and EVPI is stipulated in National Economic Evaluation Guidelines, the importance of conducting such analyses for ATMPs is highlighted. Scenario analyses should investigate optimistic and conservative assumptions regarding the benefit of the therapy (173), and explore the effects of different assumptions about long-term benefits (174). The routine use of EVPI has been

proposed to capture and communicate uncertainty (14, 175). EVPI is discussed in further detail in Chapter 5 and Chapter 9.

1.10.1.3 Time Horizon

In cost-utility analysis, a lifetime horizon is usually considered in order to capture any meaningful differences in future costs and outcomes (11). Drummond et al. propose that analyses evaluating gene therapies such be conducted using a range of time horizons (14). A shorter time horizon excludes the impact of costs and outcomes that occur in the long term, where uncertainty is likely to be greatest (12).

1.10.1.4 Discount Rate

Discounting is conducted to reflect the fact that individuals – and by extension, society – value future costs and outcomes less than current costs and outcomes (181). In Ireland, at the time of this analysis, a standard rate of 4% is applied to both costs and outcomes (11). This rate is set by the Department of Public Expenditure and Reform (182).

Discounting health outcomes and costs at the same rate is the dominant practice in economic evaluation (183). This is supported by the consistency thesis (184) and the postponement paradox (185). The consistency thesis argues that inconsistencies may occur when discounting occurs at two different discount rates. This relies on the assumption that QALYs are stable over time (183). The postponement paradox indicates that when health outcomes are discounted at a lower rate than costs, the cost-effectiveness ratio will improve by delaying the introduction of the technology (185). It therefore, becomes optimal to delay introduction of the technology infinitely.

In the case of CD19 CAR T-cell therapies, there is a time divergence between the high upfront costs and potential long-term benefit (181). Some propose that differential discounting, whereby health outcomes are discounted at a lower rate than costs (175), should be applied. This is because societal income is expected to grow at a faster rate than societal health (184, 186). Gravelle and Smith propose that the discount rate on health outcomes should be 1% to 3.5% lower than that on costs (186). Drummond et al., state that the evidence is not sufficiently strong to depart from the principle of

discounting outcomes and costs at the same rate. However, they highlight that sensitivity analysis should examine the impact of alternative rates (14).

1.10.1.5 Perspective

National Economic Evaluation Guidelines stipulate that the perspective of the healthcare payer (i.e. HSE) is adopted; only direct medical costs to the HSE are considered (11). For technologies that have a potential long-term benefit, there may be wider implications for family/caregivers and for the patient themselves. The adoption of a societal perspective may have a considerable impact on cost-effectiveness (14).

1.10.1.6 Payment Models

The financial risk associated with CD19 CAR T-cell reimbursement decision-making, has prompted an increased focus on performance-based risk-sharing agreements. Such agreements are formal schemes between payers and Applicants for sharing this financial risk (187, 188). Typically, they are linked to further evidence collection, and payments made to the Applicant are based upon agreed milestones (176).

1.11 Aims and Objectives

The aims of this thesis are:

- To undertake independent HTAs of tisagenlecleucel for R/R ALL, tisagenlecleucel for R/R DLBCL, and axicabtagene ciloleucel for R/R DLBCL, in the Irish healthcare setting.
- To examine uncertainty and estimate the value of simultaneously eliminating all uncertainty of all uncertain parameters in the respective cost-utility analyses.
- To explore the impact of performance-linked reimbursement agreements on the cost effectiveness and budget impact of tisagenlecleucel for R/R ALL, tisagenlecleucel for R/R DLBCL and axicabtagene ciloleucel for R/R DLBCL.

The objectives identified to fulfil these aims are:

- Review and synthesise the clinical evidence for tisagenlecleucel for R/R ALL, tisagenlecleucel for R/R DLBCL and axicabtagene ciloleucel for R/R DLBCL.

- Investigate the performance of a text-mining tool, in assisting the title and abstract screening process, in the systematic literature review (SLR) of the clinical evidence for tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL.
- Derive utility values, by means of SLR, for paediatric and young adult patients with R/R ALL, and adult patients with R/R DLBCL.
- Conduct an expert elicitation exercise to derive parameter estimates, and associated uncertainty, of key uncertain parameters in the evidence base of tisagenlecleucel for R/R ALL.
- Construct and populate bespoke cost-utility models for R/R ALL and R/R DLBCL.
- Conduct EVPI analysis to estimate the value of simultaneously eliminating all uncertainty of all uncertain parameters relating to the decision. Conduct EVPPI analysis to identify the parameters, which contribute most to the overall decision uncertainty.
- Estimate the potential five-year gross and net drug budget impact of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL.
- Implement performance-linked reimbursement agreement scenarios in the respective cost-utility and budget impact models.

1.12 Thesis Outline

Chapter 2 presents an SLR of the evidence for tisagenlecleucel for R/R ALL. The feasibility of conducting an ITC is assessed.

Chapter 3 describes an expert elicitation exercise, conducted to investigate key areas of uncertainty in the evidence base of tisagenlecleucel for R/R ALL.

Chapter 4 presents an SLR of utility data in paediatric and young adult patients with R/R ALL.

Chapter 5 describes the cost effectiveness of tisagenlecleucel for the treatment of R/R ALL in Ireland. A bespoke cost-utility model, integrating data from the SLRs presented in Chapter 2 and Chapter 4, and the expert elicitation exercise presented in Chapter 3, is presented. VOI analyses are also presented.

Chapter 6 presents an SLR of the evidence for tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL. The feasibility of conducting an ITC is assessed.

Chapter 7 evaluates the performance of a text-mining tool, Abstrackr, when used to assist in the title and abstract screening process of the SLR of the evidence for tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL (Chapter 6).

Chapter 8 presents an SLR of utility data in adult patients with R/R DLBCL.

Chapter 9 describes the cost effectiveness of tisagenlecleucel and axicabtagene ciloleucel for the treatment of R/R DLBCL in Ireland. Separate bespoke cost-utility models, integrating data from the SLRs presented in Chapter 6 and Chapter 8 are presented. VOI analyses are presented.

The potential five-year gross and net budget impacts of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL are presented in Chapter 10.

Chapter 11 illustrates the impact of different performance-linked reimbursement agreements on the cost effectiveness and budget impact of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL.

Chapter 12 summarises the main findings of this thesis. The potential implications for current policy and practice are outlined.

Chapter 2 Efficacy of Tisagenlecleucel for Relapsed/Refractory Acute Lymphoblastic Leukaemia: Systematic Literature Review and Evidence Synthesis

2.1 Introduction

National Economic Evaluation Guidelines stipulate that evidence on treatment effectiveness should be derived by SLR (11). National Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland (herein 'National Clinical Effectiveness Guidelines') provide methodological guidance on clinical evidence evaluation (165). They suggest that an SLR should be no more than six months old when the cost-effectiveness analysis it informs is undertaken. They provide guidance on assessment of data quality, heterogeneity, and evidence synthesis (165). Challenges arise, however, when evidence is derived from single-arm studies, such as ELIANA (pivotal trial on tisagenlecleucel).

2.1.1 Quality of Studies

No validated tool exists to critically appraise the quality of single-arm studies. Several tools exist to assess non-randomised studies. The Newcastle-Ottawa Scale is a quality assessment tool for non-randomised case-control and cohort (exposed versus non-exposed cohorts) studies (189). It is not directly applicable to single-arm studies. An SLR of tools to critically appraise non-randomised evidence found that the Newcastle-Ottawa Scale was most frequently used, reported in 142 of 686 (21%) identified studies (190). Of the five most frequently used tools, the Newcastle-Ottawa Scale was the only one reported to be validated. Content validity and inter-rater reliability are reportedly established; the construct validity is under evaluation. No further detail is provided regarding how validation was achieved (189). The Cochrane Scientific Committee (2017) indicated that the Newcastle-Ottawa Scale may be used as an alternative to the ROBNS-I Tool (191).

An SLR evaluating the use of tools in SLRs registered on PROSPERO, conducted 2018, found that, although the Newcastle-Ottawa Scale was most frequently cited for non-randomised studies, the use of the ROBINS-I Tool (192) was increasing (193). The ROBINS-I Tool is recommended by the Cochrane Collaboration for non-randomised studies, specifically targeted towards comparative analysis of two or more interventions (192). It is not applicable to single-arm trials. It has been reported to be confusing and difficult to use (194-196).

Studies, which have used an adapted version of the Newcastle-Ottawa Scale, to evaluate the quality of single-arm studies, were identified in the literature during this research (197-199). No studies were identified, which adapted the ROBINS-I Tool. The adapted Newcastle-Ottawa Scale was therefore, chosen to appraise the quality of single-arm studies here (197-200).

2.1.1.1 The Newcastle-Ottawa Scale

The Newcastle-Ottawa Scale assigns up to a maximum of nine stars for the highest quality in three domains:

1. Selection of study groups ('Selection Domain'; four stars)
2. Comparability of study groups ('Comparability Domain'; two stars)
3. Ascertainment of exposure and outcomes ('Outcomes Domain'; three stars)

The Selection Domain has four levels, each eligible to receive one star. These are 'representativeness', 'selection of non-exposed cohort', 'ascertainment of exposure', and 'demonstration that outcome of interest was not present at start of study'. The Comparability Domain has one level ('comparability of cohorts'), which is eligible for two stars. The Outcomes Domain has three levels, each eligible for one star. These are 'assessment of follow up', 'length of follow up', and 'adequacy of follow up'. A detailed coding manual is presented in Appendix A (189). Validated thresholds to distinguish between good and poor quality studies, based on the star scoring system, are lacking (189).

The Cochrane Handbook for Systematic Reviews of Interventions (2011) recommended the Newcastle-Ottawa Scale to assess the quality of non-randomised studies. However, no recommendation was made regarding a preferred tool (201).

For this research, levels relating to control groups were excluded from the Scale. This is in line with the approach taken elsewhere (197-200). The level assessing 'selection of the non-exposed cohort' (Selection Domain) and the entire Comparability Domain were excluded. A maximum of three stars could subsequently be obtained in both the Selection and Outcomes Domains. Studies, using this approach previously, scored studies out of a total of six stars. Studies, which scored four or less stars (out of six), were

considered to be poor quality (197-200). To maintain consistency with the published literature, this approach was adopted here. However, this scoring system does not appear to be validated.

The domains and scoring system, based on this adaption, are presented in Table 2. The 'length of follow up' level (Outcomes Domain) requires the user to pre-specify an adequate follow-up period. For this research, this was specified as 60 months, in line with the literature, which suggests that most patients with ALL are expected to relapse within 24 to 60 months post-treatment (12, 202, 203). Additionally, for this research, one star was attributable to the 'assessment of outcome' level for independent review committee (IRC) assessment. Outcome assessment based on any other method (e.g. record linkage) did not receive a star.

Table 2 Adapted Newcastle-Ottawa Scale quality assessment domains and scoring system (197-200)

	Selection Domain			**Outcomes Domain**		
	Representativeness	**Ascertainment of Exposure**	**Demonstration that Outcome of Interest was not Present at Start of Study**	**Assessment of Outcome**	**Length of Follow Up**	**Adequacy of Follow Up**
Total Stars Achievable	1	1	1	1	1	1
Criteria	Generalisability of trial population to population with R/R ALL in Ireland	'Secure record' (i.e. not patient-reported)	Factors that could influence response to treatment or subsequent outcome (e.g. bridging chemotherapy)	IRC assessment	Minimum 60 months	Complete follow up
			Scoring			
Good Quality			6 stars			
Fair Quality			5 stars			
Poor Quality			4 stars or less			

ALL: Acute lymphoblastic leukaemia; **IRC:** Independent review committee; **R/R:** Relapsed/refractory

2.1.2 Chapter Aim

The aim of this chapter is to identify clinical evidence to inform the relative effectiveness of tisagenlecleucel for R/R ALL in paediatric and young adult patients. Quality of evidence will be assessed. Between-study heterogeneity will be determined. The most appropriate evidence synthesis techniques will be established. Results will be used to inform the

bespoke cost-utility model evaluating tisagenlecleucel for the treatment of R/R ALL in Ireland, presented in Chapter 5.

2.2 Methods

2.2.1 Systematic Literature Review

An SLR protocol was developed with reference to the Cochrane Handbook for Systematic Reviews of Interventions (204). Guidance for creating and running the search strategy was obtained from Trinity College Dublin Medical Librarian, Mr David Mockler. Reporting is conducted in line with PRISMA 2020 (205).

2.2.1.1 Population

The population was paediatric and young adult patients, up to 25 years of age, with B-cell ALL that is refractory, in relapse post-transplant, or in second or later relapse. This is in line with the EMA licensed indication of tisagenlecleucel (3). Participants of any sex and any ethnicity were included. Patients with T-cell ALL were excluded.

2.2.1.2 Intervention

The intervention was tisagenlecleucel used as monotherapy at the EMA licensed dose and indication (3).

2.2.1.3 Comparators

Blinatumomab (with or without alloSCT) (herein 'blinatumomab') was the primary comparator. Other comparators included FLA(G)-IDA (with or without alloSCT), best supportive care, and HSCT. Only results pertaining to blinatumomab are presented here.

2.2.1.4 Outcomes

The primary outcomes were:

- OS
- PFS
- Event-free survival (EFS)
- Leukaemia-free survival (LFS)

Due to the single-arm nature of the studies informing the evidence base, outcomes were reported as Kaplan-Meier curves. This facilitated digitisation of Kaplan-Meier curves and reconstruction of individual patient-level data (IPD).

Outcomes relating to response rates were extracted if they were the primary outcome of the trial. However, limited evidence is available to support their surrogacy for OS, and these outcomes will not be directly used in the cost-utility model (206). Data on the proportion of patients with grade ≥3 adverse events and adverse events of special interest were extracted. Health-related quality of life (HRQOL) data (defined by validated quality of life measures or instruments used in each trial) were extracted. HRQOL data were also considered in a separate SLR, presented in Chapter 4.

2.2.1.5 Study Design

Prospective RCTs, phase II non-randomised or single-arm trials, and prospective observational studies were included. Single-centre trials, expanded access programmes, retrospective studies, and case studies or reports were excluded. These study types were excluded as they were considered to be of lower quality than included study types. This was maninly due to factors such as lack of pre-defined protocols and potential selection bias.

2.2.1.6 Search Methods

Electronic databases EMBASE, MEDLINE (via EBSCO), and CENTRAL (via the Cochrane Library) were searched from 01 January 2000 to 21 November 2020 inclusive. The search strategy is presented in Appendix A (Table A2). Proceedings from the American Society of Hematology (ASH) and European Hematology Association (EHA) Annual Conferences were hand searched for the years 2014 to 2020 inclusive. Terms used in searching of conference proceedings included: ‘tisagenlecleucel’, ‘ELIANA’, ‘ENSIGN’, ‘tisa-cel’, ‘blinatumomab’, ‘acute lymphoblastic leukaemia’, and ‘paediatric’. EMA European Public Assessment Reports (EPARs) of tisagenlecleucel (143, 207) and blinatumomab (87), and clinical trial reports from ClinicalTrials.gov (www.clinicaltrials.gov) were also searched. Articles were restricted to those published in English.

2.2.1.7 Citation Management

Identified citations were imported to Endnote® and transferred to Abstrackr. Duplicates were systematically searched for using software in Endnote® and identified manually. Title and abstract screening was conducted in Abstrackr, by a single reviewer, to identify citations eligible for full-text review. The full texts of potentially relevant citations were obtained and assessed for suitability for inclusion in the final evidence base. For quality assurance purposes, a second reviewer screened 10% of full-text articles in duplicate.

Data extraction was conducted using an adapted Cochrane data extraction form (208). Data recorded included population, intervention, comparators, outcomes, study design, authors, title, and publication date. Extracted outcomes data were checked in duplicate by a second reviewer.

2.2.2 Quality of Studies

Assessment of risk of bias in RCTs was pre-specified to be conducted using the Cochrane Risk of Bias Tool 2 (209). Quality of non-randomised studies of two or more interventions was pre-specified to be assessed using the Newcastle-Ottawa Scale (189). As described in 2.1.1, the adapted Newcastle-Ottawa Scale was chosen to assess the quality of single-arm studies. Quality was assessed in duplicate by a second reviewer.

2.2.3 Heterogeneity

Clinical and methodological between-trial heterogeneity were assessed qualitatively. Assessment of statistical heterogeneity, using I^2 and Q statistics, was pre-specified. Sources of heterogeneity were examined.

2.2.4 Indirect Treatment Comparison

Included trials were assessed for inclusion in an ITC. Factors considered were the type of data identified (IPD or study-level; direct or indirect), the type of studies identified (RCT or single-arm), the number of studies identified, and heterogeneity of studies (165).

2.2.5 Reconstruction of Individual Patient-Level Data

Due to the single-arm nature of the trials and lack of publicly available raw IPD, reconstructed IPD were generated from published Kaplan-Meier curves. Time and survival coordinates from published Kaplan-Meier curves were extracted using Digitizelt software (210). In conjunction with NCPE Statistician, Dr Joy Leahy, the algorithm by Guyot et al. (211) was applied using R® (packages: 'Mass', 'Splines', 'Survival') (212) to map these coordinates back to Kaplan-Meier data using information on number of events and numbers at risk.

Reconstructed IPD, from trials that were considered sufficiently homogeneous, were pooled without adjustment to expand the evidence base.

2.2.6 Comparative Efficacy

Using reconstructed IPD, the 12-month probability of OS and EFS in each trial was estimated (212). A 12-month point was chosen due to the short duration of follow up in the tisagenlecleucel trials. Hazard ratios (HRs) for survival were estimated by fitting Cox proportional hazard models using the 'coxph' function of the 'Survival' package in R® (212).

2.2.7 Quality of Evidence for Outcomes

GRADE (Grading of Recommendations, Assessment, Development and Evaluations) was used to rate the confidence in treatment effect estimates. An initial quality grade was assigned to the body of evidence, across the outcome of interest, depending on the study design; RCTs and observational studies generate high and low quality evidence, respectively. This initial grade was modified using several key domains: methodological limitations, indirectness of evidence, imprecision, inconsistency of evidence, and likelihood of publication bias. The grades and their interpretation are presented in Table 3 (213).

Table 3 GRADE quality of evidence for outcomes grades (213)

Grade	Definition
High	Confident the true effect lies close to that of the estimate of the effect
Moderate	Moderately confident in effect estimate: true effect is likely to be close to estimate of effect, but there is a possibility that it is substantially different
Low	Limited confidence in effect estimate: true effect may be substantially different from the estimate of effect
Very Low	Very little confidence in effect estimate: true effect is likely to be substantially different from estimate of effect

Each outcome was assessed for quality independently. Summary of findings tables were generated using GRADEproGDT® software (214). Grading was assessed in duplicate by a second reviewer.

2.3 Results

Searches were conducted on 21 November 2020. Following exclusion of duplicates, a total of 2,139 titles and abstracts underwent screening. Full-text screening led to inclusion of three trials in the final evidence base, as outlined in Figure 3.

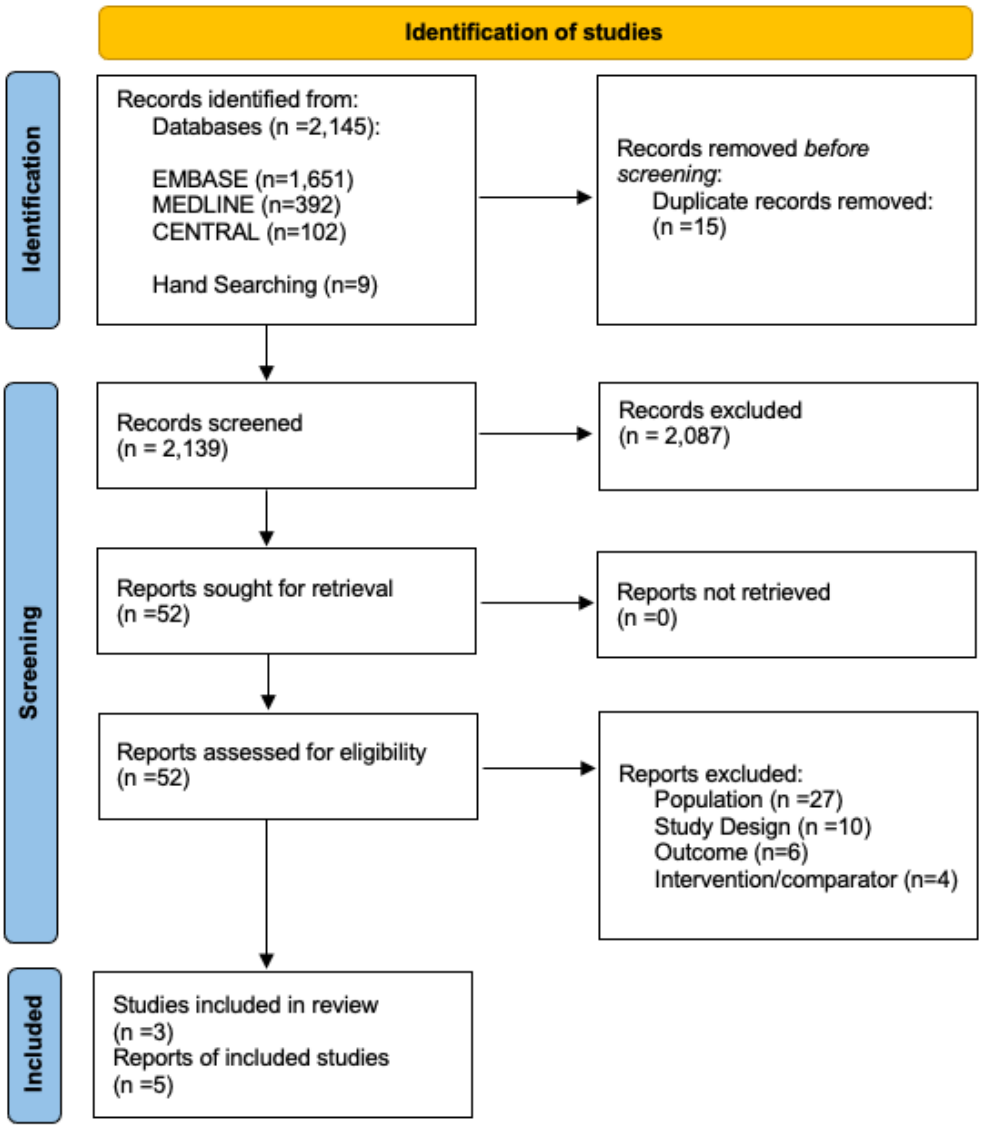

Figure 3 PRISMA Diagram - systematic literature review of efficacy data for relapsed/refractory acute lymphoblastic leukaemia

2.3.1 Excluded Studies

A total of 47 studies were excluded at full-text screening; reasons were population (n=27), study design (n=10), outcome (n=6), and intervention/comparator (n=4). A selected list of studies and reasons for exclusion are presented in Appendix A (Table A3).

2.3.2 Included Studies

The studies included in the final evidence base are summarised in Table 4.

Table 4 Summary of trials that met the inclusion criteria in the systematic literature review of treatments for R/R ALL**

Title (author, year)	Trial Design	Key Eligibility Criteria	Intervention (sample size)	Outcomes
ELIANA (EMA 2018; Maude et al. 2018) (143, 215)	Phase II, single-arm, open-label, multi-centre	Aged 3-21 yrs inclusive CD19+ ALL, Primary refractory*, chemo-refractory†, or relapsed‡ Karnofsky (≥16 yrs) or Lansky (<16 yrs) status ≥50 BM ≥5% lymphoblasts	Tisagenlecleucel; once-off single IV infusion (mITT; n=75‡‡) Patients ≤50kg: 2.0 - 5.0x10^6/kg CAR-positive viable T-cells Patients >50kg: 1.0 - 2.5x10^8 CAR-positive viable T-cells	**Primary:** ORR **Key Secondary:** EFS, OS HRQOL data not reported††
ENSIGN (EMA 2020) (207)	Phase II, single-arm, open-label, multi-centre	Aged 3-21 yrs inclusive Primary refractory*, chemo-refractory†, or relapsed ALL§, Karnofsky (≥16 yrs) or Lansky (<16 yrs) status ≥50 BM ≥5% lymphoblasts	Tisagenlecleucel; once-off single IV infusion (mITT; n=64‡‡) Patients ≤50kg: 2.0 to 5.0x10^6/kg CAR-positive viable T-cells Patients >50kg: 1.0 - 2.5x10^8 CAR-positive viable T-cells	**Primary:** ORR **Key Secondary:** RFS, OS HRQOL data not collected
NCT01471782 (EMA 2018; von Stackelberg et al. 2016) (87, 216)	Phase I dose-finding Phase II: single-arm, open-label, multi-centre	<18 years Relapsed| or refractory¶ ALL Karnofsky (≥16 yrs) or Lansky (age <16 years) performance status ≥50 BM >25% lymphoblasts	Blinatumomab IV infusion (n=70) Phase I: doses between 5mcg/m^2/day and 30mcg/m^2/day Phase II: stepwise 5/15 mcg/m^2/day; 4-week continuous IV infusion, followed by a 2-week treatment-free interval Patients achieving CR within first 2 cycles could receive up to 3 more, or withdraw from treatment to receive chemotherapy or alloSCT	**Phase I:** max. tolerated dose **Phase II: Primary –** CR within the first 2 cycles **Key Secondary:** RFS, OS HRQOL data not collected

ALL: Acute lymphoblastic leukaemia; **AlloSCT:** Allogeneic stem cell transplant**; BM:** Bone marrow**; CR:** Complete response; **EFS:** Event-free survival; **EMA:** European Medicines Agency; **HRQOL:** Health-related quality of life; **IV**: Intravenous; **mITT:** Modified intention-to-treat; **ORR:** Overall remission rate; **OS:** Overall survival; **RFS:** Relapse-free survival; **R/R:** Relapsed/refractory.
*Not achieving CR after two cycles of standard chemotherapy.

†Not achieving CR after one cycle of standard chemotherapy for relapsed ALL.
‡Second or greater BM relapse, or any BM relapse after alloSCT and ≥6 months from alloSCT at the time of infusion.
§Second or greater BM relapse, or any BM relapse after alloSCT and >6 months from alloSCT at the time of infusion.
|Second or later BM relapse, or any BM relapse after alloSCT and >3 months from alloSCT at the time of infusion.
¶Patients in first relapse must have failed to achieve CR following full standard reinduction chemotherapy of at least four weeks duration. Patients who have not achieved a first remission must have failed a full standard induction.
**Comparators not presented as all single-arm trials.
††HRQOL data reported in a separate publication (not identified in this systematic literature review).
‡‡Patients who received infusion with tisagenlecleucel.

2.3.2.1 Tisagenlecleucel

Two single-arm trials, that investigated the efficacy of tisagenlecleucel, were included.

2.3.2.1.1 Survival Outcomes: ELIANA

ELIANA is a phase II trial in which patients received infusion with tisagenlecleucel at a dose of 2.0 to $5.0x10^6$ CAR-positive T-cells per kg (body weight) for patients weighing 50kg or less, or 1.0 to $2.5x10^8$ CAR-positive T-cells for patients weighing more than 50kg (215). ELIANA comprised screening (leukapheresis and cell product acceptance), enrolment, bridging chemotherapy, lymphodepleting chemotherapy, tisagenlecleucel infusion, and primary safety and follow-up phases. Patients were aged 3 to 25 years, with B-cell ALL, and were primary refractory, chemo-refractory, relapsed after alloSCT, or otherwise ineligible for alloSCT. Eligible patients were no older than 21 years at the time of initial diagnosis. At the time of publication, 92 patients were enrolled in ELIANA (the intention-to-treat (ITT) population) and 75 received infusion with tisagenlecleucel (the modified ITT (mITT) population) (143, 215). Patient characteristics are presented in 2.3.4.

Efficacy data were reported primarily for the mITT population (n=75). The primary endpoint ORR (proportion of patients with best overall disease response of CR or CR with incomplete blood count recovery within 3 months) was 81% (95% CI 71 to 89); all patients were MRD negative. Eight patients underwent alloSCT while in remission. EFS (time from infusion to the earliest of the following: no response, relapse before response was maintained for at least 28 days, or relapse after having CR or CR with incomplete haematologic recovery), and OS (time from infusion) were key secondary endpoints. At a median follow up of 13.1 months, EFS was 73% (95% CI 60 to 82) at 6 months and 50%

(95% CI 35 to 64) at 12 months. Median EFS was not reached. OS was 90% (95% CI 81 to 95) at 6 months and 76% (95% CI 63 to 86) at 12 months. Median OS was 19.1 months (95% CI 15.2 to NE) (143, 215). OS and EFS Kaplan-Meier curves were presented only for the mITT population.

Median time from enrolment to infusion was 45 days (range: 30 to 105); 17 patients discontinued prior to infusion. Reasons for discontinuation included death (n=7), tisagenlecleucel product-related issues (n=7), and adverse event prior to infusion (n=3) (215). EFS was not presented for these patients. Median OS for all enrolled patients (ITT; n=92) was 19.4 months (95% CI 14.8 to NE). The 12-month OS was 70% (95% CI not reported) (143).

An updated data cut of ELIANA, presented as conference proceedings and with an additional 11 months of data, was identified in the SLR. Based on this data cut (n=79, mITT population), the 18-month OS was 70% (95% CI 58 to 79) (217). However, OS was not reported as a Kaplan-Meier curve, preventing the reconstruction of IPD. Data on EFS were not reported. As such, it was excluded from this SLR.

2.3.2.1.2 Survival Outcomes: ENSIGN

ENSIGN is a phase II trial; patients received infusion with tisagenlecleucel at the same dose as that in ELIANA. The design (screening phase, enrolment, lymphodepleting chemotherapy, etc.) and eligibility criteria were aligned with that of ELIANA. A total of 75 patients enrolled (ITT population) and 64 received infusion with tisagenlecleucel (mITT population) (207). Patient characteristics are presented in 2.3.4.

Median duration of follow up, in the mITT population, was 32 months (range: 18 to 56). The primary endpoint of ORR (proportion of patients with a best overall disease response of CR or CR with incomplete blood count recovery maintained at two evaluations 28 days or more apart) was 70% (95% CI 53 to 82). Nine patients proceeded to alloSCT. EFS (time from infusion to the earliest of the following: death, relapse, or treatment failure), and OS were key secondary endpoints. In the mITT population, median EFS was 15.6 months (95% CI 6 to NE) and median OS was 29.9 months (95% CI 15 to 42). EFS probability was

48% (95% CI 33 to 61) at 24 months. OS probability was 55% (95% CI 40 to 67) at 24 months (207). Previous data cuts of ENSIGN were presented as abstracts at ASH 2016 (218) and EHA 2018 (219).

Median time from enrolment to infusion was 54.4 days, range: 33 to 182. Eleven patients discontinued prior to infusion; reasons included death (n=6), and tisagenlecleucel product-related issues (n=5). EFS was not presented for these patients. Median OS for all enrolled patients (ITT; n=75) was 26 months (95% CI 10 to 38). OS probability in these patients was 57% (95% CI 44 to 67) at 24 months (207).

2.3.2.1.3 Adverse Events

Safety data were presented for the mITT populations (n=75 ELIANA; n=64 ENSIGN). Grade ≥3 adverse events were reported in 88% of patients in ELIANA and 92% in ENSIGN. CRS (defined as per the Penn Grading Scale (220)) was reported in 77% and 78% of patients in ELIANA and ENSIGN, respectively. Grade ≥3 CRS was documented in 46% and 30% of patients in ELIANA and ENSIGN, respectively. ELIANA reported an ICU admission rate of 47% for management of CRS. This was not reported for ENSIGN. Neurotoxicity was observed in 40% and 30% of patients in ELIANA and ENSIGN, respectively. Grade ≥3 neurotoxicity was documented in 13% in ELIANA and 6.3% (within 8 weeks post-infusion) in ENSIGN. Grade ≥3 infections were reported in 24% of patients in ELIANA and 11% (within 8 weeks post-infusion) in ENSIGN. Grade ≥3 prolonged cytopenias (greater than 28 days) were reported in 32% of patients in ELIANA and 34% (within 8 weeks post-infusion) in ENSIGN. Grade ≥3 febrile neutropenia occurred in 35% and 37.5% of patients in ELIANA and ENSIGN, respectively. IV immunoglobulin, administered at the local investigator's discretion, was administered to 'most' patients in ELIANA; no further detail was published. These data were not published for ENSIGN. Other frequently reported grade ≥3 adverse events included anaemia, white blood cell count decreased, neutrophil count decreased, and hypotension (207, 215).

2.3.2.1.4 Health-Related Quality of Life

HRQOL data were not collected in ENSIGN. HRQOL data were collected in ELIANA; Laetsch et al. describe this (221). Outcomes were collected using the Pediatric Quality of

Life Inventory (PedsQL) and the EQ-5D visual analogue scale (VAS). In total, 58 patients, aged 8 to 23 years inclusive, were included. At baseline, 50 (86%) and 48 (83%) patients completed the PedsQL and EQ-5D VAS, respectively. At day 28 post-infusion, HRQOL data were submitted by 37 of 48 (77%) patients achieving a response for each measure. Of the 10 patients who did not achieve a response, 6 submitted PedsQL data and 7 submitted EQ-5D VAS data. Between 3 and 12 months post-infusion, improvements in HRQOL were reported across all measures and increased over time. Over the 12-month period post-infusion, improvements from baseline were lowest for social functioning and greatest for physical functioning. For the EQ-5D VAS, mean scores increased from baseline (66.8), with increases of 16.8 (95% CI 9.4 to 24.3) at month 3, 17.4 (95% CI 9.0 to 25.7) at month 6, 18.8 (95% CI 7.8 to 29.9) at month 9, and 24.7 (95% CI 13.5 to 35.9) at month 12 post-infusion. Improvements in mean change scores at months 3, 6, 9, and 12 post-infusion were greater than the PedsQL minimal clinically important difference for the total score, psychosocial health summary score, and physical and emotional functioning subscales. Similarly, improvement in mean change scores for the EQ-5D VAS at months 3 to 12 was greater than the minimal clinically important difference (221).

Further critique of ELIANA and ENSIGN is provided in 2.3.3 and 2.3.4.

2.3.2.2 Blinatumomab

2.3.2.2.1 Survival Outcomes: NCT01471782

NCT01471782 was a phase I/II, multi-centre, single-arm trial that evaluated the safety and efficacy of blinatumomab in patients aged less than 18 years of age. Patients had ALL that was refractory (primary and chemo-refractory), in first relapse after full salvage induction regimen, in second or later relapse, or any relapse after alloSCT. Patients had B-cell ALL with greater than 25% bone marrow blasts and had a Karnofsky (age 16 years or older) or Lansky (aged less than 16 years) performance status of 50 or greater (87, 216). Patient characteristics are presented in 2.3.4.

Phase I aimed to determine the maximum tolerated dose. Patients received blinatumomab at doses ranging between 5mcg/m^2/day and 30mcg/m^2/day. Based on the results of phase I, patients in phase II received blinatumomab as a 4-week continuous IV

infusion (stepwise dosage of 5mcg/m^2/day for the first 7 days, and 15mcg/m^2/day thereafter), followed by a 2-week treatment-free interval. Patients achieving CR within the first two cycles could receive up to three additional cycles, or withdraw from treatment to receive consolidation chemotherapy or alloSCT, per investigators' choice. Those who did not achieve CR, within the first two cycles, discontinued treatment. Patients with haematological relapse during their follow-up period could receive up to three additional cycles. The primary outcome of the phase II component was the proportion of patients achieving CR within the first two cycles. Key secondary endpoints included relapse-free survival (RFS; time from first documented CR until first documented relapse or death due to any cause), and OS. The proportion of patients who received alloSCT during blinatumomab-induced remission was also a secondary endpoint. Results were based on central assessment (87, 216).

At the date of study completion (24 May 2016), all patients completed the two-year follow-up, had withdrawn from study, or had died. Seventy patients received treatment at the recommended stepwise dosage; 26 patients in phase I and 44 patients in phase II. The primary endpoint (CR within the first 2 cycles) was reported in 39% (95% CI 27 to 51) of patients across phase I and phase II. Median RFS was 4.4 months (95% CI 2 to 8) among patients who achieved CR (within the first 2 cycles), with a median follow up of 23.1 months. Median OS amongst all 70 patients was 7.5 months (95% CI 4 to 12), at a median follow up of 23.8 months. In total, 36% of patients received alloSCT. Censoring patients who received alloSCT following blinatumomab-induced remission (n=25) resulted in a median OS of 6.5 months (95% CI 4 to 10) (87, 216).

EFS data were not collected. RFS data are not an appropriate proxy for EFS in this research. This is because RFS only accounts for patients who achieved CR and therefore, have more favourable prognosis (compared to the entire cohort). Comparison of RFS from NCT01471782 with EFS collected during ELIANA and ENSIGN would bias outcomes in favour of blinatumomab.

RIALTO Study

RIALTO, an expanded access study that examined outcomes in patients, aged between 29 days and 17 years (inclusive), treated with blinatumomab was also identified in the SLR.

Patients (n=110) had B-cell ALL in second or later relapse, relapse after alloSCT, or refractory to treatment. Study outcomes were presented as correspondence to the editor of Blood Cancer Journal. Median RFS, a secondary endpoint (defined as per NCT01471782), was 8.5 months (95% CI 4.4 to NE), at a median follow up of 11.2 months. Median OS was 13.1 months (95% CI 10.2 to 21.3), at a median follow up of 17.4 months. AlloSCT was received by 65% of patients who achieved CR (222).

Patients in RIALTO may have had better prognosis than those in NCT01471782, based on the proportion of patients who had a bone marrow blast count of <50% (61% versus 26% in RIALTO and NCT01471782, respectively). RIALTO also had a higher proportion of female patients (44% versus 33%), a higher proportion of patients treated in Europe (92% versus 69%), and a lower proportion of patients who had prior alloSCT (41% versus 57%) (87, 216, 222). RIALTO was excluded in this SLR based on study design (expanded access study/correspondence to the editor). Additionally, patients previously treated with blinatumomab were eligible for inclusion in RIALTO (222). There is a possibility that some patients may have received treatment in both NCT01471782 and RIALTO (223). RIALTO was therefore, not considered for pooling with NCT01471782.

2.3.2.2.2 Adverse Events

Grade ≥3 adverse events were reported in 87% of patients in NCT01471782. These were primarily characterised by cytopenias and changes in blood chemistry. CRS was documented in 11% of patients; 6% had grade ≥3 CRS. Neurotoxicity was reported in 24% of patients; 7% were grade ≥3 severity. Grade ≥3 febrile neutropenia was reported in 17%. Grade ≥3 anaemia and thrombocytopenia were reported in 36% and 21%, respectively. Other frequently reported grade ≥3 adverse events included neutropenia, hypokalaemia, alanine aminotransferase increased, platelet count decreased, and pyrexia (216).

2.3.2.2.3 Health-Related Quality of Life

HRQOL data were not collected during NCT01471782.

Further critique of NCT01471782 is provided in 2.3.3 and 2.3.4.

2.3.3 Quality of Included Studies

A detailed description of the levels covered by the Newcastle-Ottawa Scale is presented in Appendix A. As per the adapted Newcastle-Ottawa Scale, all three studies (ELIANA, ENSIGN, NCT01471782) were graded as poor quality; all scored four stars or less.

In the Selection Domain, all studies scored one star in both the 'representativeness' and 'ascertainment of exposure' levels. All studies pertained to populations considered 'somewhat representative of the average R/R ALL population in the community' (one star). The ascertainment of exposure in all studies was considered to be 'secure' (one star), as data were reported by acting clinicians (i.e. not patient-reported). ELIANA and ENSIGN did not score in the 'demonstration that outcome of interest was not present at start of study' level. This was because the possibility of a cross-over effect from bridging chemotherapy could not be ruled out (143).

In the Outcomes Domain, all studies scored one star on the 'assessment of outcome' level. This was due to IRC assessment in these studies. No studies scored on the 'was follow up long enough for outcomes to occur?' level within the Outcomes Domain. This was due to the short duration of follow up in each study, ranging from a median of 13.1 months (ELIANA) (215) to 32 months (ENSIGN) (207). These follow-up durations were not sufficiently long to capture progression of disease. No studies scored on the adequacy of follow up level within the Outcomes Domain. ELIANA and ENSIGN received zero stars, as follow up and reporting of the ITT population was not adequate (143, 207, 215). NCT01471782 scored zero stars, as EFS data were not collected (87, 216).

The scores obtained by each trial are presented in Table 5. The highly subjective nature of this quality assessment is highlighted; results should be interpreted in this context.

Table 5 Quality assessment of studies included in systematic literature review of treatments for relapsed/refractory acute lymphoblastic leukaemia, using the adapted Newcastle-Ottawa Scale*† (197-200)

	Selection Domain			Outcomes Domain			Final Grade‡
	Representativeness	Ascertainment of Exposure	Demonstration Outcome of Interest was not Present at Start of Study	Assessment of Outcome	Was Follow Up Long Enough for Outcomes to Occur?	Adequacy of Follow Up	
ELIANA (143, 215)	1	1	0	1	0	0	Poor Quality
ENSIGN (207)	1	1	0	1	0	0	Poor Quality
NCT01471782 (87, 216)	1	1	1	1	0	0	Poor Quality

*Level assessing 'selection of the non-exposed cohort' in the Selection Domain and the entire Comparability Domain of the Newcastle-Ottawa Scale are excluded (189).

†1= 1 star obtained in this level, 0= 0 stars obtained in this level.

‡**Good quality**= 6 stars; **fair quality**= 5 stars; **poor quality**= 4 stars or less.

2.3.4 Heterogeneity

The single-arm nature of the trials (and thus, naïve nature of the ITC, described in 2.3.5) precluded the use of statistical measures of heterogeneity, such as I^2 and Q statistics. Investigation into sources of heterogeneity, by means of meta-regression, was also ruled out. Statistical measures of heterogeneity are subject to limitations, particularly when sample size is small (224, 225), and so, qualitative assessment of heterogeneity is also recommended (165). Qualitative assessment of clinical and methodological heterogeneity was conducted. A summary of patient characteristics in each trial is presented in Table 6, with a detailed discussion provided below.

Table 6 Baseline characteristics of patients in the trials included in systematic literature review of treatments for relapsed/refractory acute lymphoblastic leukaemia

Characteristic	ELIANA (n=75) (143, 215) (tisagenlecleucel)	ENSIGN (n=64) (207) (tisagenlecleucel)	NCT01471782 (n=70) (87, 216) (blinatumomab)
Median Age, years (range)	11 (3, 23)	12 (3, 25)	8 (<1, 17)
Male, n (%)	43 (57)	30 (47)	47 (67)
Prior Allogeneic Stem Cell Transplant, n (%)	46 (61)	17 (59)	40 (57)
Median Time (months) Since Initial Diagnosis to First Relapse (range)	32.9 (1, 70)	27.6 (1, 108)	NR
Median Time (months) Since Last Relapse to Treatment (range)	3.5 (1.5, 13.8)	2.6 (1.3, 9.8)	2.9 (0.4, 49.8)
Disease Status, n (%):			
Primary Refractory	6 (8)	7 (11)	NR
Chemo-Refractory or Relapsed	69 (92)	57 (89)	NR
Refractory to Last Regimen	NR	NR	39 (56)
Blast Count:			
Mean (SD)	63 (30.9)	61 (30.1)	NR
≥50 n (%)	NR	NR	52 (74)
Geographic Region, n (%):			
European Union	NR	0 (0)	48 (69)
United States	NR	29 (100)	22 (31)

NR: Not reported; **SD:** Standard deviation.

2.3.4.1 Clinical Heterogeneity

NCT01471782 enrolled patients within a lower age range than that of ELIANA and ENSIGN. ELIANA and ENSIGN required patients to be aged at least three years at the time of screening (143, 207, 215). Children under two years accounted for 14% of the population in NCT01471782 (216). Patients aged 18 years and older accounted for 17% of the population in ELIANA (143); this age group was excluded in NCT01471782. Age at diagnosis is a prognostic factor in ALL (53). Patients diagnosed between the ages of 1 and 10 years have improved outcomes compared to those diagnosed beyond this age range (75, 226, 227), with older age associated with worse prognosis (228). The impact of these differences is difficult to determine, due to inconsistent reporting of age ranges in the trials. Additionally, these patients are at an advanced stage of disease and age at trial enrolment may not reflect age at diagnosis.

NCT01471782 had a higher proportion of male patients (216). Females have been shown, in some studies, to have better prognosis than males (65, 66). It is difficult to conclude if this difference is of sufficient magnitude to bias outcomes.

ELIANA and ENSIGN did not include patients who were eligible for alloSCT, while patients in NCT01471782 were treated with the intent to proceed to alloSCT. Prior alloSCT was one reason patients were deemed ineligible for alloSCT in ELIANA and ENSIGN (143, 207). Patients who are ineligible for alloSCT represent a cohort with particularly poor prognosis (87). Of note, eight patients proceeded to alloSCT in ELIANA following tisagenlecleucel infusion (215). This indicates that their alloSCT eligibility changed over the course of the trial. All studies had a similar proportion of patients who received prior alloSCT (207, 215, 216). Patients who relapse following alloSCT also represent a cohort with poor prognosis (229).

For patients with relapsed disease, the time since last relapse to current treatment was similar between all studies (143, 207, 216). A wider range was noted in NCT01471782. A shorter time frame indicates worse prognosis (230).

Reporting on disease status differed between the studies. Patients in ELIANA and ENSIGN were classified as primary refractory, or chemo-refractory/relapsed, while NCT01471782 reported the proportion of patients who were refractory to their last regimen. Those reported to be refractory to their last regimen in NCT01471782 (56%) are likely to encompass patients with primary refractory and chemo-refractory disease (216). Patients with primary refractory disease have poor prognosis (203). The proportion of patients in NCT01471782 who experienced 2 or more relapses was 52%, indicating that a high proportion of these patients had poor prognosis (216). The mean number of previous therapies received by patients was 3.4 (SD 1.55) in ELIANA (143) and 2.9 (range: 1 to 9) in ENSIGN (207). This indicates that patients in these trials also had poor prognosis. In the absence of more granular data, no further conclusions can be drawn regarding heterogeneity in disease status of patients enrolled in each trial.

In terms of blast count in bone marrow, 74% of patients in NCT01471782 had 50% or greater. Mean blast count percentage was 63% (SD 30.9) in ELIANA and 61% (SD 30.1) in

ENSIGN. Trial criteria permitted patients with 5% or greater blasts by morphologic assessment at screening in ELIANA and ENSIGN, and greater than 25% in NCT01471782 (143, 207, 216). Higher blast count is associated with worse prognosis (231).

ENSIGN recruited patients solely in the US (207). Patients in ELIANA were recruited from Europe, the US, Canada, and Japan (143). Patients in NCT01471782 were recruited from Europe and the US (216). Survival outcomes in paediatric and young adult patients with cancer, as assessed using registry data, have been reported to be aligned between Europe and the US (232, 233). However, these studies evaluated outcomes in patients with newly diagnosed disease. Caution should be exercised in extrapolating these findings to patients with R/R disease.

ELIANA had a higher proportion of patients with high-risk genomic abnormalities when compared to patients in NCT01471782; 37% versus 26%, respectively (143, 216). However, it is unclear if trials were accounting for the same genomic abnormalities. These data were not reported for ENSIGN (207). Insufficient reporting of other prognostic factors, including duration of first remission, site of relapse, and MRD at the end of previous therapy (234-236), prevented further analysis of between-trial heterogeneity.

ELIANA and ENSIGN restricted eligibility to patients with a life expectancy of greater than 12 weeks (143, 207). This was not a requirement of NCT01471782 (87). Patients in ELIANA and ENSIGN may be fitter than those expected in Irish clinical practice.

ELIANA and ENSIGN permitted bridging chemotherapy to stabilise disease during the tisagenlecleucel manufacturing period. A total of 87% and 89% of patients received bridging chemotherapy in ELIANA and ENSIGN, respectively (207, 215). In ELIANA, there was a delay of up to three months between staging of tumour burden at study enrolment, for each patient, and administration of tisagenlecleucel (143). Tumour burden, in some patients, may have regressed in response to bridging chemotherapy. As tumour burden was not reassessed prior to infusion of tisagenlecleucel, the potential for a carry-over effect from bridging chemotherapy cannot be ruled out (143). It is difficult to differentiate the influence of patient characteristics and unobserved prognostics factors from the true treatment effect of tisagenlecleucel.

EFS data were not collected during NCT01471782 (87). EFS and OS were reported from the time of tisagenlecleucel infusion for both ELIANA and ENSIGN. This was due to the considerable time lag between enrolment and receipt of tisagenlecleucel infusion. A total of 17 and 11 patients discontinued from ELIANA and ENSIGN, respectively, during this period. Thus, possibly enriching the patient cohort with those who were fit enough to survive the manufacturing period. Notably, death accounted for 7.6% and 8% of patients who discontinued from the study prior to infusion in ELIANA and ENSIGN, respectively (207, 215).

2.3.4.2 Methodological Heterogeneity

ELIANA and ENSIGN differed from NCT01471782 in that they were both phase II trials, while NCT01471782 comprised phase I (dose-finding) and phase II components. Outcomes in ELIANA and ENSIGN were IRC-assessed (207, 215). Outcomes in NCT01471782 were assessed centrally (216). Sample size in all three studies was small. All studies were open-label and single-arm (143, 207, 216). This is a notable limitation.

All three studies were subject to a high degree of censoring towards the end of the follow-up period. For OS, ELIANA had 2 patients left at risk at month 20 (215), ENSIGN had 5 left at month 36 (207), and NCT01471782 had 6 left at month 24 (216). The long-term survival associated with these therapies is highly uncertain.

Patients in ELIANA and ENSIGN were censored at the time of alloSCT (143, 207). This overestimates the relapse rate in these studies. As patients in NCT01471782 were treated with intent to proceed to alloSCT, patients were not censored at the time of alloSCT. Sensitivity analysis, which censored patients at the time of alloSCT in NCT01471782, found that median OS was slightly lower (6.5 versus 7.5 months) (87). Sensitivity analyses, exploring the impact of not censoring for alloSCT, were conducted for ELIANA and ENSIGN. However, results were not presented for OS and EFS (143, 207). As such, it is difficult to determine how this censoring affected outcomes.

2.3.5 Indirect Treatment Comparison Feasibility Assessment

Due to the lack of direct comparative evidence of tisagenlecleucel versus blinatumomab, the feasibility of conducting an ITC, for this research, was assessed.

2.3.5.1 Meta-Analysis and Network Meta-Analysis

Direct meta-analysis statistically combines results of multiple trials to generate an average effect estimate across trials (204). Network meta-analysis is based on the premise of a network of studies involving treatments, which are compared directly, indirectly, or both (in a single analysis) (204, 237). No common comparator arms existed between these single-arm trials (thus, a network could not be formed) and so, neither a direct meta-analysis nor a network meta-analysis were deemed feasible here.

2.3.5.2 Population-Adjusted Comparison Methods

Population-adjusted methods of evidence synthesis, such as matching-adjusted indirect comparison (MAIC) and simulated treatment comparison (STC), rely on the availability of raw IPD (from at least one trial) to adjust for between-trial differences in the distribution of variables that influence outcome (238). As raw IPD were not available from the published literature, no further consideration was given to these methods; reconstructed IPD are not appropriate.

2.3.5.3 Naïve Comparison

Naïve ITCs utilise treatment-level data to compare outcomes between trials. This method combines data from individual arms of separate trials as though they have come from a single trial (11). In the absence of a common comparator arm and considering the lack of publicly available raw IPD, a naïve ITC was considered the only feasible method to compare tisagenlecleucel to blinatumomab.

Naïve ITCs do not adjust for differences between populations and are therefore, highly uncertain. Notably, network meta-analyses, which violate the similarity assumption (that is, trials are not comparable in terms of effect modifiers), are also subject to uncertainty and biased outcomes (237). Additionally, unanchored MAICs and STCs (i.e. based on single-arm studies) require all effect modifiers and prognostic variables to be accounted

for when adjusting. This assumption is deemed largely unreasonable to meet, due to reasons such as inadequate profiling, and measuring and reporting of variables, resulting in biased estimates (238). Differences in trial design, such as the administration of bridging chemotherapy, cannot be adjusted for, using MAIC and STC techniques. These methods may also be limited by the small sample sizes of the included trials. These ITC methods (i.e. network meta-analysis, MAIC, STC) may therefore, also be subject to bias. It is acknowledged, however, that naïve ITCs are subject to the greatest uncertainty and so, a degree of caution should be exercised in interpretation of results presented here.

2.3.6 Reconstruction of Individual Patient-Level Data

IPD of the trials were reconstructed by digitising the published Kaplan-Meier curves, using DigitizeIt software (210), and applying the algorithm by Guyot et al. (211).

ELIANA and ENSIGN were considered homogenous in terms of study design and population. Reported baseline characteristics were closely aligned between the two (Table 6). For this research, these trials were pooled without adjustment to inform the efficacy of tisagenlecleucel; an approach widely adopted in the literature (223, 239-241). However, EFS data in ENSIGN were not reported in the required format (Kaplan-Meier curve). Thus, only the ELIANA data were used to inform EFS of tisagenlecleucel (215).

The reconstructed Kaplan-Meier OS curves, using reconstructed IPD, are presented in Figure 4.

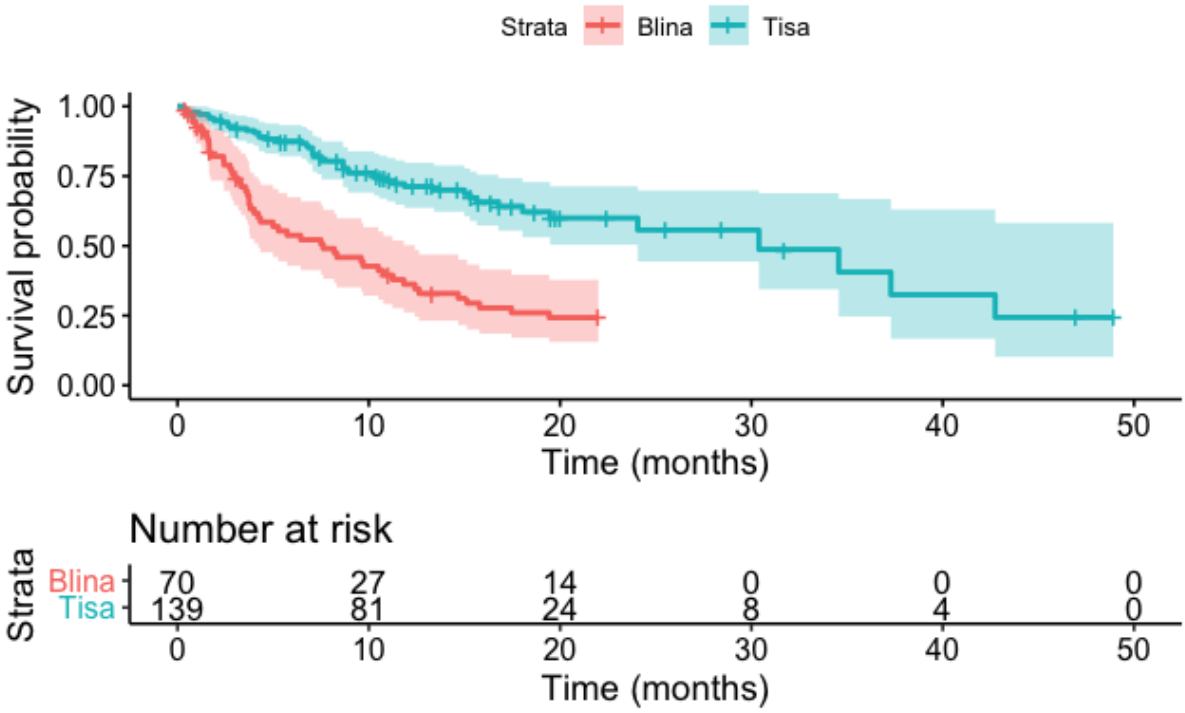

Figure 4 Reconstructed Kaplan-Meier overall survival curves of blinatumomab (NCT01471782) and tisagenlecleucel (pooled ELIANA and ENSIGN)

2.3.7 Comparative Efficacy

The 12-month OS and EFS survival probabilities, and HR for OS of tisagenlecleucel versus blinatumomab, are presented in Table 7. These were estimated using reconstructed IPD. A HR for EFS was not generated, as these data were not available for blinatumomab.

Table 7 Survival estimates of tisagenlecleucel and blinatumomab in relapsed/refractory acute lymphoblastic leukaemia, based on naïve comparison

Treatment	Overall Survival, % (SE)	Hazard Ratio Overall Survival (95% CI)	Event-Free Survival, % (SE)
Tisagenlecleucel (143, 207, 215)	72* (0.05)	0.33 (0.22 to 0.49)|	52† (0.06)
Blinatumomab (87, 216)	36§ (0.06)		NR

NR: Not reported; **SE:** Standard error.
*Pooled ELIANA (data cut 25 April 2017) and ENSIGN (data cut 24 May 2019) data; time from infusion to death due to any cause.
†ELIANA (data cut 25 April 2017); time from infusion to the earliest of the following events: no response, relapse before response was maintained for at least 28 days, or relapse after having complete remission or complete remission with incomplete haematologic recovery.
§Time from infusion to death due to any cause or the date of last follow-up.
|Hazard ratio <1.0 favours tisagenlecleucel.

2.3.8 Quality of Evidence for Outcomes

The quality of evidence for OS, of tisagenlecleucel versus blinatumomab, was graded as very low. As the comparison of tisagenlecleucel versus blinatumomab lacks

randomisation, the quality was initially graded as low. Subsequent downgrading was applied due to very serious concerns regarding risk of bias (downgraded two levels), serious concerns regarding inconsistency (downgraded one level), very serious concerns regarding indirectness (downgraded two levels), and serious concerns regarding imprecision (downgraded one level). Factors considered in reaching these conclusions included: the short duration of follow up; small sample sizes; open-label nature of trials; naïve ITC; and the potential for clinical and methodological heterogeneity based on qualitative assessment (Table 8) (213, 242).

EFS data were not collected during NCT01471782; assessment of the quality of evidence for EFS was therefore, not conducted.

Table 8 Summary of findings table for quality of evidence for overall survival (tisagenlecleucel versus blinatumomab), based on GRADE assessment (213)

Summary of findings:			
Tisagenlecleucel compared to Blinatumomab for R/R ALL in Paediatric and Young Adult Patients			
Patient population: R/R ALL in Paediatric and Young Adult Patients **Setting:** Irish Healthcare Setting **Intervention:** Tisagenlecleucel **Comparison:** Blinatumomab			
Outcomes	Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)
Overall Survival	HR 0.33 (0.22 to 0.49)	209 (3 non-randomised studies)	⊕◯◯◯ Very low[a,b,c,d]
ALL: Acute lymphoblastic leukaemia; **CI:** Confidence interval; **HR:** Hazard Ratio; **R/R:** Relapsed/refractory.			
GRADE Working Group grades of evidence **Very low certainty:** very little confidence in effect estimate: true effect is likely to be substantially different from estimate of effect.			

Explanations

a. Downgraded two levels for risk of bias. High risk of bias; assessment of quality of studies, as per adapted Newcastle-Ottawa Scale, indicated all studies were of poor quality. Reasons included short duration of follow up, small sample sizes, and open-label nature (242).

b. Downgraded one level for inconsistency. Qualitative assessment indicated some degree of clinical and methodological heterogeneity. Statistical assessment of heterogeneity not feasible (243).

c. Downgraded two levels for indirectness. Due to naïve comparison.

d. Downgraded one level for imprecision. Small sample size (<400 'rule-of-thumb') (242, 244).

2.4 Discussion

This SLR identified a limited number of studies examining the efficacy of treatments for R/R ALL in paediatric and young adult patients. The limited number of studies reflects the rarity and advanced nature of this disease. Patients in excluded studies (full-text screening) tended to not be as heavily pre-treated as those specified in the inclusion criteria. All included studies were single-arm. This is unsurprising, as it is difficult to conduct adequately powered RCTs in populations with rare diseases. Additionally, it may be inappropriate to conduct RCTs under such circumstances (12).

2.4.1 Included Studies

Efficacy of tisagenlecleucel was examined in two trials (ELIANA and ENSIGN) (143, 207, 215). The small numbers of patients in these trials, along with the short duration of follow up and single-arm nature, limits the conclusions that can be drawn regarding the efficacy of tisagenlecleucel. The Kaplan-Meier curves of ELIANA and ENSIGN are highly censored towards the end of the follow-up period, adding further uncertainty to long-term efficacy estimates. This uncertainty is compounded by the potential for carry-over effects from bridging chemotherapy and the proportion of patients who did not proceed to tisagenlecleucel infusion. Death prior to infusion was one of the main reasons for not proceeding to infusion (143, 207). Patients enrolled in these trials are likely to be fitter than those expected in clinical practice, limiting the generalisability of observed outcomes. Efficacy data for tisagenlecleucel, chosen for use in the cost-utility model developed as part of this research, will be based on patients who proceeded to infusion (i.e. mITT population). This population may be enriched with patients who were fit enough to survive the manufacturing period (143). Thus, biasing outcomes. Consideration, in the cost-utility model, should also be given to those who did not proceed to infusion.

The efficacy of blinatumomab was examined in one trial (NCT01471782) included in the final evidence base (87, 216). This trial was subject to a number of the same limitations as ELIANA and ENSIGN, in that it included small numbers of patients and was single-arm. The Kaplan-Meier OS curve was also subject to a high degree of censoring. Patients in NCT01471782 have been reported to have high-risk disease, based on tumour load, prior

relapses, and short interval between last treatment and relapse (87, 223). Patients may be treated with blinatumomab with the intent to proceed to alloSCT. Median OS estimates, when censoring for patients who proceeded to alloSCT and not censoring for these patients, were similar (87). Data were not presented separately for patients who did and did not proceed to alloSCT, precluding an in-depth analysis of the survival benefit following alloSCT. Differences in median OS between NCT01471782 and RIALTO (expanded access study; not included in the final evidence base) adds further uncertainty to the evidence base of blinatumomab (87, 216, 222).

The primary endpoints of ELIANA, ENSIGN, and NCT01471782 were surrogate endpoints. The selection of appropriate endpoints for rare diseases and paediatric populations is a challenge faced by regulators, investigators, and decision-makers (245). Although OS is the gold standard in determining efficacy (246), it may not be a practical primary endpoint in clinical trials. Death from disease may occur a long time after start of therapy, and other events (e.g. subsequent therapies) may occur in the interim, influencing the result of the trial. In such instances, ORR may be preferred as it occurs earlier and is usually directly attributable to drug effect (247). This gives better power and hence, smaller sample sizes to detect a difference (248), which is key in trials for rare diseases. However, no published evidence is available to validate ORR as a surrogate for OS in tisagenlecleucel for R/R ALL (245). It has been reported that trials using surrogate endpoints report larger treatment effects than those using final 'hard' outcomes (249).

2.4.2 Quality of Included Studies

Using the adapted Newcastle-Ottawa Scale (197-200), all studies (ELIANA, ENSIGN, NCT01471782) were graded to be of poor quality. It is concerning, yet unsurprising, that EMA authorisation was granted on such limited evidence. This poor quality evidence creates challenges and risk for patients, clinicians, and decision-makers.

The highly subjective nature of the adapted Newcastle-Ottawa Scale is emphasised. Notably, the Newcastle-Ottawa Scale awards the same score (one star) for 'independent blind assessment' and 'record linkage'. For this analysis, it was assumed that only IRC assessment received a star. Other forms of outcome assessment (e.g. investigator

assessment) were pre-specified to be ineligible for a star. This approach was adopted here because independently-assessed endpoints are considered more reliable and less prone to bias than investigator-assessed endpoints (165). It is not reasonable to conclude that independently-assessed and investigator-assessed outcomes are of the same quality.

Despite its previous use in the literature, to assess the quality of single-arm studies (197-200), it is unclear if this adapted Scale is appropriate for this purpose. Important domains that influence the perceived quality of a study, such as sample size and surrogate outcomes as the primary endpoint, are not considered in the adapted Newcastle-Ottawa Scale. As such, caution should be exercised in the interpretation of these findings.

2.4.3 Heterogeneity

NCT01471782 (blinatumomab) failed to report information on key prognostic factors. It is difficult to decipher how heterogeneity observed across the trials impacts on the direction and magnitude of relative treatment effects. This is a challenge faced by independent researchers in conducting comparative effectiveness analyses. This could be eased, in some instances, by minimum reporting standards set out by regulators and journal editors alike. Although the analysis of heterogeneity conducted as part of this research was thorough, interpretation is limited by the qualitative nature of the analysis. The similarity of these trials is of concern, and in light of the naïve ITC, further caution is warranted in interpretation of results.

2.4.4 Comparative Efficacy

A naïve ITC was the only feasible method of comparison. This limits the conclusions that can be drawn from the evidence. This comparison (12-month OS probabilities and HR for OS) indicated that tisagenlecleucel had favourable outcomes compared to blinatumomab. Notably, uncertainty in the ITC was not captured in the 95% CI of the HR. These estimates are highly uncertain.

The difference in reporting of outcomes between studies highlights the lack of a standardised, core outcomes set for clinical trials in R/R ALL. EFS data were not collected during NCT01471782. This limits interpretation of outcomes and precluded a robust

comparative analysis. As highlighted in 2.3.5.3, even with the availability of raw IPD for ELIANA and ENSIGN, population-adjustment methods of comparison, such as MAIC and STC, are unlikely to produce reliable results. Notably, cost-utility analyses, using raw IPD of ELIANA and ENSIGN, have employed naïve ITC for these reasons (88, 250). Considering the notable heterogeneity between the studies, the planning and development of high-quality real-world evidence studies, which aim to mitigate against bias, will be vital in addressing uncertainty in the relative efficacy of tisagenlecleucel.

Tisagenlecleucel was associated with a higher proportion of grade ≥3 adverse events, when compared to blinatumomab. CRS appeared to be more prevalent and severe in patients who received tisagenlecleucel; however, different scales were used to grade CRS and so, results are not directly comparable.

2.4.5 Quality of Evidence for Outcomes

Quality of the evidence for OS, of tisagenlecleucel versus blinatumomab, using the GRADE framework was graded as very low. The true OS benefit of tisagenlecleucel versus blinatumomab is likely to be markedly different from the estimated effect (213). The approach adopted here examines important domains such as sample size, which were not considered in the assessment of quality of studies (using the adapted Newcastle-Ottawa Scale). However, the GRADE framework is subject to a high degree of subjectivity (213). Assessment is further limited by the single-arm nature of the trials. GRADE provides limited guidance for data generated from single-arm studies (213). For example, statistical heterogeneity, in assessing the inconsistency domain, could not be assessed. The utility of GRADE, in assessing the quality of evidence from single-arm studies, is unclear.

Notably, as reported in the NICE Methods Review of Sources and Synthesis of Evidence (July 2020), quality assessment of evidence is difficult to reflect in economic evaluation, unless bias weights are generated to down-weight different forms of evidence. The utility of GRADE for HTA was questioned by NICE, considering the subjective nature. NICE further highlight that individual components of GRADE are addressed through other means in the economic evaluation, such as imprecision in modelling (251).

2.4.6 Comparison with the Published Literature

Aamir et al. conducted an SLR and meta-analysis of CD19 CAR T-cell therapies in R/R ALL in paediatric and young adult patients. Of note, this SLR was not exclusive to tisagenlecleucel and included patients up to 30 years of age. A total of 15 studies were included in the final evidence base; the majority were phase I. Median OS at 12 months ranged from 63% to 84%. Median EFS at 12 months ranged from 46% to 76%. All trials were subject to follow up of less than 15 months. These results are aligned with those obtained as part of this research. Of note, Aamir et al. indicate that in most identified studies, CD19 CAR T-cell therapies were used as a bridge to alloSCT (252). According to expert opinion, this is not expected to occur in Irish clinical practice.

A published MAIC, comparing tisagenlecleucel (ELIANA) to blinatumomab (NCT01471782), indicated that tisagenlecleucel was associated with prolonged OS (HR 0.32; 95% CI 0.16 to 0.64; p=0.0015) (253). However, a number of methodological issues underpin this analysis. The MAIC was unanchored; it is assumed that all effect modifiers and prognostic variables are accounted for. This is a very strong assumption and largely considered impossible to meet, resulting in an unknown amount of bias in the estimates (238). Matching was conducted on a limited number of variables and it is unclear how these were chosen (253). The analysis is also subject to potential bias due to unobserved differences between the trials (254). Notably, the HR obtained via this MAIC was closely aligned with that generated, as part of this research, by naïve ITC (0.33; 95% CI 0.22, to 0.49). Thus, it cannot be concluded that the additional complexity associated with the MAIC leads to less biased estimates.

2.4.7 Limitations

There are a number of limitations associated with this SLR. These are in addition to those discussed thus far. Single-centre studies, retrospective studies, expanded access studies, case studies and case reports were excluded. While these study types may contribute complimentary evidence, they are inherently subject to greater bias than prospective phase II studies. Inclusion of such studies would add further uncertainty to the evidence base.

Outcomes were required to be reported as Kaplan-Meier curves, to allow reconstruction of IPD. This led to the exclusion of an updated data cut of ELIANA. Notably, this data cut was presented as an abstract only and no EFS data were reported.

Data were not reported for key outcomes such as EFS (NCT01471782; blinatumomab). HRQOL data were only collected during ELIANA. These data are limited by the single-arm nature of ELIANA. Assumptions regarding these parameters will therefore, be required for the cost-utility model. Such assumptions will add further uncertainty to the evidence base. This warrants extensive sensitivity analyses regarding the effectiveness and cost-effectiveness of these agents.

The trend of EMA approval of treatments, that are either innovative in their mechanism or treat a rare disease, based on single-arm evidence is likely to continue (255). No validated frameworks, which examine the quality of single-arm studies, were identified. National Clinical Effectiveness Guidelines are vague in their recommendations regarding quality assessment of such evidence, and fail to recognise that this type of evidence is being encountered more frequently (165). It is widely acknowledged that such studies are considered to be of low quality and have a high risk of bias (256). Available frameworks for assessing bodies of evidence, such as GRADE, are subjective (213). However, the availability of a standardised reporting framework facilitates transparent and explicit assessment. It can also improve communication of outcomes (257). Validated measures to assess the risk of bias in individual single-arm studies are required. The absence of such a framework for single-arm studies impedes robust assessment and comparison of outcomes.

Best practice guidelines recommend that screening and assessment of study eligibility are conducted in duplicate (204). This is labour intensive and time consuming. A pragmatic approach (i.e. single reviewer) was adopted here. Bias was limited by conducting a pre-specified quality assurance check on 10% of full-text articles (i.e. screening these in duplicate).

2.5 Conclusion

The evidence base for treatments for R/R ALL in paediatric and young adult patients, identified in this SLR, was limited to single-arm, heterogeneous studies. Naïve ITC indicated that tisagenlecleucel had favourable outcomes when compared to blinatumomab; however, the true magnitude of benefit is unknown. Inconsistency in reporting of patient characteristics and outcomes further limits conclusions that can be drawn. In Chapter 5, the effectiveness estimates derived from this study will be incorporated into a cost-utility model, examining the cost effectiveness of tisagenlecleucel in the Irish healthcare setting. Due to uncertainty in the estimates obtained here, extensive sensitivity analyses will be required.

Chapter 3 Expert Elicitation of Key Uncertain Parameters in the Evidence Base of Tisagenlecleucel for Relapsed/Refractory Acute Lymphoblastic Leukaemia

3.1 Introduction

The SLR, presented in Chapter 2, illustrated the paucity of evidence on medium- to long-term survival in paediatric and young adult patients treated with tisagenlecleucel for R/R ALL. Challenges also arise because clinical experience with a new technology may be limited. In the case of tisagenlecleucel, a first-in-class technology, consensus in the literature, on the order of alloSCT and tisagenlecleucel in the treatment pathway is lacking (258).

Some clinicians have indicated that tisagenlecleucel will be administered with curative intent (223). However, a recognised definition of cancer-related 'cure' is lacking (259). The Erice statement defines cure as the time from diagnosis at which the risk of death from recurrence or metastatic spread of the original cancer becomes very small; this is usually between 2 and 10 years relapse-free since cancer diagnosis (260, 261). Notably, this relates only to the original disease, despite potential for complications and toxicities. The follow-up period of data supporting EMA authorisation of tisagenlecleucel is not sufficiently long to support the concept of cure (143, 207).

Uncertainty in evidence is not a new phenomenon in HTA. However, with the increasing trend of granting EMA marketing authorisation on the basis of single-arm and immature data, this challenge is likely to be encountered more frequently (262, 263). Explicit and transparent judgements must be made regarding the appropriateness of uncertain parameter estimates in cost-utility models. To ensure stakeholders are fully informed, uncertainty in the evidence should be quantified (264). Failure to do this may result in misleading estimates of cost effectiveness and inappropriate reimbursement decisions.

3.1.1 Expert Elicitation

Expert elicitation (herein 'elicitation') may provide parameter estimates when evidence is scare or lacking. It is the process of formulating an individual's judgements about one or more uncertain quantities into a (joint) probability distribution for each quantity (265). It characterises uncertainty both within and between expert judgements. This facilitates ruling out of parameter estimates that are considered implausible. Probabilistic characterisation allows incorporation of judgements into cost-utility models (266). Bojke et al. propose elicitation, conducted appropriately (to minimise bias), to be the 'best'

approach to characterise uncertainty associated with parameters of cost-utility models (264).

Elicitation considers several integral factors including: identifying parameters to be elicited; defining, identifying and recruiting experts; motivating and training experts; the elicitation format; and the elicitation itself (264). Strategies should be implemented to mitigate against bias and heuristics. Post-elicitation procedures include fitting an appropriate distribution to judgements, and determining an appropriate method to aggregate judgements - by mathematical or behavioural methods. A facilitator generally leads the elicitation. Their role is to aid experts in formulating their judgements in probabilistic form (265). The recommended number of experts ranges from 4 to 20 (267, 268); a minimum of 5 has been recommended for elicitation in healthcare decision-making (264).

Distinctions exist between expert opinion and elicitation. Expert opinion involves collecting expert beliefs in a qualitative format. In HTA, this may involve defining care pathways. Expert opinion is generally collected using an unstructured format, with no measure of uncertainty included. Such approaches are susceptible to bias and can result in lack of consistency across assessments (269). In contrast, elicitation involves expressing expert judgements in a quantitative format. As described, this can be used to define point estimates of model parameters and characterise associated uncertainty (266). Elicitation is appropriate when empirical evidence does not exist or is limited. Structured elicitation is less susceptible to bias than expert opinion (264). However, despite efforts to minimise bias, elicitation is situated on the lowest level of the hierarchy of evidence (11). Elicitation is, by nature, highly subjective (270). Although expert opinion is frequently sought for the purpose of HTA, the use of structured elicitation is less common (269, 271).

3.1.2 Expert Elicitation Protocols

Elicitation is widely used in disciplines such as environmental management, food safety, and natural hazards (272-274). Several elicitation protocols have been developed for conducting elicitation research (272, 275, 276). The Sheffield Elicitation Framework (SHELF) includes guidance, templates, and software for conducting elicitation. SHELF is

based around a workshop involving individual-level elicitation, followed by group discussion and then, group elicitation, generating a consensus distribution (275). Alternatively, Cooke's classical method synthesises judgements of multiple experts, whereby experts are assigned a weighting based on their performance in answering a 'seed' question. Experts who perform well on the seed question are assigned a higher weight to their judgements in the overall, pooled judgements (276). The European Food Safety Authority provide guidance on a modified DELPHI technique. In this approach, multiple rounds (usually two or three) of elicitation, at the individual level, are conducted. Following the first round of elicitation, individual experts are provided with the opportunity to revise their judgements based on the feedback of summary results, from all experts, of the first round. This process is repeated. Individual expert judgements from the final round are mathematically aggregated, using equal weighting, to generate one distribution. This differs from the standard DELPHI technique in that the final distribution is not based on consensus (272).

Bojke et al. propose that generic protocols, such as those described above, may not be appropriate for HTA. This is due to the resource and time constraints that are inherent in HTA (264). Additionally, Soares et al. propose that between-expert variation in judgements is a characteristic of clinical experts, which may not be desirable in other disciplines. In other disciplines, between-expert variation is generally linked to varying levels of bias and may be deemed undesirable. Consensus methods of aggregation may discourage such variation. However, in healthcare, between-expert variation may reflect genuine heterogeneity in the populations experts draw upon (271). Of elicitation studies used for the purpose of HTA identified in the literature (n=21), Soares et al. found that heterogeneity existed in the methodology employed. There was also lack of consideration for existing guidance (271).

Bojke et al. recently developed a reference protocol to reflect 'emerging best practice' for elicitation in HTA (herein 'the protocol') (264). Although the protocol offers a degree of flexibility, it clearly and comprehensively defines potential reference methods for all stages of elicitation. Key elements considered in the protocol are presented in Table 9. It has been recommended that further studies, using the protocol, are conducted (264).

Table 9 Key elements of the expert elicitation protocol by Bojke et al. (264) and adherence to this protocol in this study*

Key Element of Elicitation	Suggested Method	Adhered to in this Study
Identification of Parameters and Framing of Questions	1. Simple, observable quantities. 2. Dependent parameters expressed as independent, where necessary. 3. Clear wording. Decompose quantities, where necessary.	1. Yes 2. Yes – independent parameters 3. Yes
Method of Elicitation and Development of Exercise	1. Either fixed interval or variable interval method. 2. Elicit from experts individually, group interaction may follow. 3. Interaction can occur either face-to-face or via a remote Delphi process. 4. Piloting.	1. Yes – fixed interval 2. Yes but no group interaction 3. No. Logistical challenges in gathering experts. Time and resource constraints a factor. Concerns regarding overconfident results due to "groupthink". 4. Yes but not with experts in the field. This was due to the time commitment required for piloting and limited number of experts in the field.
Expert Identification and Recruitment	1. Pursue diversity in judgements. Aim to represent the full range of experts' judgements. Experts should be willing to participate. 2. Recruit experts with substantive expertise. Develop normative skills during training. 3. Document conflicts of interest. Recruit experts not involved in exercise development. 4. At least 5 experts should complete the exercise.	1. Yes 2. Yes 3. Yes 4. Yes
Conduct of Elicitation	1. Conduct training, focusing on expression of uncertainty and minimising bias. 2. Face-to-face, where possible. 3. Provide feedback and opportunity for experts to revise judgements. 4. Collect rationales on how and why expert made judgement.	1. Yes 2. No. Remote elicitation favoured due to time and resource constraints. Facilitator available via phone or email. Background, practice question and explanation video provided. 3. Yes. Graphical feedback provided during

	5.	Document and justify all methodological choices.	elicitation, descriptive statistics and fitted distribution presented after elicitation. 4. Yes 5. Yes
Pooling and Analysis	1. 2. 3. 4.	Fit distributions to individual judgements. Explore between-expert variation. Generate aggregated distribution through linear pooling with equal weighting. Validity can be assessed by internal and external review.	1. Yes 2. Yes 3. Yes 4. Yes

*Adapted from Bojke et al. (264)
(https://www.ncbi.nlm.nih.gov/books/NBK571051/pdf/Bookshelf_NBK571051.pdf)

3.1.3 Chapter Aim

The aim of this chapter is to derive expert judgements, using a bespoke elicitation tool, regarding key areas of uncertainty in the evidence base of tisagenlecleucel for R/R ALL. The elicited judgements will be used to inform input parameters and validate outputs of the bespoke cost-utility model, presented in Chapter 5.

3.2 Methods

This study was conducted in line with the protocol by Bojke et al. (264). Instances where options exist or where there is divergence from the protocol are clearly indicated and choices justified. Reporting is conducted in line with relevant guidelines (266).

3.2.1 Identification of Parameters

Review of published cost-utility analyses indicated that these models are sensitive to assumptions regarding subsequent alloSCT and the proportion of patients alive after five years; much uncertainty exists here (88, 223). These parameters were therefore, chosen for elicitation. Experts were asked whether they expected patients to be 'cured' following treatment with tisagenlecleucel. Here, 'cure' was defined as having mortality equivalent to the age- and sex-matched general population. For experts who deemed cure possible, the cure time-point and proportion of patients cured were elicited. The time-point at which patients' risk of mortality, relative to the general population, is expected to be greatest post-treatment was elicited from those who did not deem cure possible. These

parameters (with the exception of cure time-point and point at which excess risk is expected to be greatest) were also elicited for routine care. This was to establish current clinical practice and survival. The number of parameters elicited was limited, to reduce the burden of the exercise. Parameters were defined to ensure fitness for purpose in the cost-utility model (264). An outline of parameters is presented in Table 10.

In line with previous studies, event probabilities at a single time-point were initially chosen for elicitation (277-279). This was considered a more intuitive way of eliciting judgements, in contrast to a HR, which is not directly observable. Neutral wording was emphasised, to prevent building bias into the questions (270).

3.2.2 Method of Elicitation and Development of Exercise

The protocol indicates that either fixed interval or variable interval methods are appropriate when conducting elicitation. A fixed interval method, known as the histogram method (also the chip-and-bins method (271) and roulette method (278)), was chosen here. This method is well described (280) and has been previously used in elicitations for cost-utility analyses (270, 281, 282). For each parameter, a discrete numerical scale was predefined. The expert was provided with a grid of equally sized bins and asked to place 20 crosses ('chips') between the bins. These chips represented judgements about the distribution of that parameter; each chip represented 5% of the distribution (281). The more chips allocated to a bin, the more certain the expert that the true value lies within that particular range. Clinicians have reported this method to be easier to use than other methods (283). Additionally, fixed interval methods may be more appropriate when training of experts is not conducted face-to-face (264). Independence between parameters was assumed.

The exercise was developed as a bespoke elicitation tool. This approach was chosen as a published review of available elicitation tools found that limited support was available, within the tools, to support experts with the elicitation task (284). Additionally, there were concerns regarding in-hospital access to web-based tools. Bespoke elicitation tools have been widely used in the literature (270, 283, 285). The elicitation tool facilitated remote use, which was favoured due to time and resource constraints, and the requirement to elicit judgements of geographically dispersed experts (given the rare

nature of the disease). Although the protocol recommends face-to-face elicitation, it acknowledges that this may not always be feasible (264). A facilitator was available via phone or email. The exercise was implemented in Microsoft Excel® and programmed so that each worksheet cell changed colour once a chip was placed in it. Thus, the expert (who may have limited experience in expressing judgements in a quantitative format) could visualise the shape of the distribution. This provided a form of instant feedback (271). If the expert was not satisfied with their elicited distribution, they could revise their judgements by clicking 'reset'.

The exercise was piloted. Feedback was requested regarding the guidance provided, definition of the questions, and general outline of the exercise. Members of a multidisciplinary HTA team (n=20) in a national HTA agency were recruited; response rate 70%. Feedback informed modifications. Questions that originally pertained to probability were reworded to describe outcomes in numbers of patients (considered more intuitive). The exercise was condensed to improve flow. The modified exercise was piloted with practising nurses (n=5) not familiar with HTA. Final modifications included addition of an instruction video, the use of less complex language, and the inclusion of a 'not disease-specific' practice question (so the expert would not feel compelled to provide the "right" answer). Of note, the protocol recommends piloting on a small sample of experts from the field under investigation (264). This approach was not adopted here due to the required time commitment, and the limited number of experts in the field. Piloting with experts in the field would reduce the number of experts eligible to participate in the actual elicitation.

3.2.3 Expert Identification and Recruitment

Experts were identified by purposive and snowball sampling, using the NCPE clinician database, NCCP Designated Cancer Centres in Ireland, Principal Treatment Centres in the UK, and published research. Consultant haematologists, from Ireland or the UK, with at least five years' experience were eligible. Experts were required to be experienced in the treatment of paediatric or young adult patients with ALL. Experts were considered to have substantive expertise based on these criteria (264). Prior experience with CD19 CAR T-cell therapy was not required, as this therapy was not reimbursed in Ireland at the time of study. It was desirable to recruit experts from a range of geographically dispersed

practices in order to sufficiently capture heterogeneity in clinical practice and patient populations, and a range of beliefs (269).

Contact details of experts were obtained through the NCPE database, and websites of hospitals and professional organisations. Experts were invited to participate via an email, which outlined the format of the exercise. Experts were requested to reply, either accepting or declining the invitation, or suggesting a suitable colleague, within seven days of receipt. Experts who did not respond were sent a reminder email one week after the seven-day period had elapsed; no further emails were sent thereafter.

3.2.4 Conduct of the Elicitation

The exercise, containing one practice question and eight to ten (depending on responses) elicitation questions (presented in Appendix B), was sent to participating experts via email. In accordance with good practice, background to the elicitation (rationale, parameters, process), background evidence pertaining to tisagenlecleucel (pivotal information from ELIANA and ENSIGN), and general guidance were presented at the start (264, 265, 286). Potential biases and ways to mitigate these were explained (287). Experts were prompted to focus on uncertainty in their judgements. The importance of expressing their own judgements and associated uncertainty was highlighted – there were no "right" or "wrong" answers. To mitigate against anchoring and overconfidence bias, experts were asked to begin answering each question by placing one chip at the upper limit and one chip at the lower limit of their estimate, such that they expected that the true value would not lie above or below these values (288). Of the ten elicitation questions, nine were suitable for pooling. The first question comprised a yes/no (binary) response; statistical pooling methods were therefore, not required. Each question was accompanied by a free-text box for the provision of information deemed relevant. This aimed to understand the rationales underlying the experts' judgements and to facilitate appropriate interpretation of results (264).

Experts were asked to provide details regarding the number of patients treated per year and number of years of experience. This was to gain further insight into their clinical experience; however, this information was not incorporated into the judgements.

Experts were informed that all exercises returned to the facilitator would be pseudonymised, ensuring compliance with General Data Protection Regulation (GDPR). Conflicts of interest were documented; however, these did not preclude participation.

Experts were guided through the exercise using prompts to click on to the next section. They were first presented with a practice question (to elicit the distribution of the expected temperature at 1pm in Dublin) to aid familiarisation with the format (281). Then, they were asked to watch the instruction video. If required, they were invited to contact the facilitator for further clarifications. Once satisfied, they were asked to begin. Experts were requested to return the completed exercise within ten days. Any experts who did not respond within this period were sent up to two reminder emails at two-week intervals.

3.2.5 Pooling and Analysis

Pooling and analysis of judgements was conducted in collaboration with NCPE Statistician, Mr Conor Hickey. Once the completed exercises were returned, predefined distributions were fitted to each parameter from each expert using the SHELF package in R® (289). All parameters, except for that with a yes/no response, were fitted with a beta distribution due to their bounded nature. Fitted distributions and descriptive statistics were then presented to the expert to ensure they were an accurate reflection of their judgements (290). The expert was given a two-week window to revise judgements; this was repeated until the expert was satisfied.

A mathematical approach to aggregation of final judgements was taken. Consensus was not sought; the aim was to capture heterogeneity. For each parameter, individual distributions were aggregated across experts using linear pooling with equal weighting (273). Linear pooling uses an arithmetic mean of the distributions from each expert (264). This was conducted using the 'plinearpool' function of the SHELF package in R®. This function requires a vector of cumulative probabilities for each expert, for each response (289). The multimodal distributions generated were then examined. Differential weighting was dismissed; the most appropriate and informative method of generating weights was unclear (270). Also, knowledge about known parameters does not

necessarily infer knowledge about unknown parameters. The use of seed-derived weights has proved challenging within the scope of elicitation (283). Elsewhere, the use of seed-derived weights resulted in only 2 of the 18 participating experts being represented in the aggregate distribution (291).

3.2.6 Validity of Judgements

Internal consistency was assessed by asking the experts to provide rationales for their judgements. Inconsistencies between the rationales and elicited judgements were assessed. Graphical feedback of the elicited judgements and providing experts with the opportunity to revise their judgements were also implemented to ensure validity (264).

3.3 Results

3.3.1 Participants

A total of 19 experts were invited to take part; 8 agreed to participate and 5 completed the exercise. Only one expert returned the exercise within the specified time period (10 days); reminder emails were sent to the remaining experts. Of those who did not complete the exercise, one reported difficulty due to the need to consider all patient variables; no response was received from the other two. The number of paediatric or young adult patients with R/R ALL treated by each expert ranged between one and five per year. No conflicts of interest were declared. None had previously completed an elicitation exercise.

3.3.2 Elicited Parameters

The individual and pooled judgements of the experts are presented in Table 10. Accompanying rationales are summarised in Appendix B (Table A5).

Table 10 Parameters elicited: individual and pooled judgements

	Parameter	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Pooled† (n=5)
Q1	Assumption of Cure	No	Yes	No	No	Yes	No; n=3‡ Yes; n=2‡
Q2a §	Time-Point of Greatest Risk of Mortality, years (95% CrI)	5 (0.0, 17.7)	N/A	3 (0.0, 14.9)	4 (0.0, 11.5)	N/A	Responded 'no' to Q1: 4 (0.0, 15.2)
Q2b ¶	'Cure' Time-Point, years (95% CrI)	N/A	1.8 (0.9, 2.9)	N/A	N/A	1.6 (0.5, 3.1)	Responded 'yes' to Q1: 2 (0.6, 3.0)
Q2c ¶	Cure Fraction – Tisagenlecleucel, % (95% CrI)	N/A	44 (25.8, 61.8)	N/A	N/A	45 (30.1, 60.3)	44 (27.6, 61.1)
Q2d ¶	Cure Fraction – Routine Care, % (95% CrI)	N/A	24 (5.5, 47.1)	N/A	N/A	19 (12.2, 39.2)	21 (7.6, 43.7)
Q3	Proportion of Patients Proceeding to AlloSCT Following Tisagenlecleucel, % (95% CrI)	40 (8.9, 76.0)	39 (19.8, 59.4)	40 (16.9, 65.0)	36 (15.4, 59.1)	46 (21.4, 71.5)	40 (15.2, 68.4)
Q4|	5-year OS of Patients Proceeding to AlloSCT Following Tisagenlecleucel, % (95% CrI)	22 (8.9, 37.8)	25 (1.0, 64.7)	17 (1.5, 39.2)	21 (5.9, 39.7)	29 (7.6, 54.9)	23 (2.4, 52.0)
Q5|	5-year OS of Patients who do not Proceed to AlloSCT Following Tisagenlecleucel, % (95% CrI)	46 (30.5, 61.3)	20 (3.1, 39.7)	25 (11.0, 42.0)	32 (13.1, 52.6)	44 (29.5, 57.9)	33 (8.7, 56.9)
Q6	Proportion of Patients Proceeding to AlloSCT Following Routine Care, % (95% CrI)	44 (25.2, 62.8)	84 (53.1, 100.0)	68 (35.3, 93.7)	28 (7.9, 52.9)	20 (3.8, 40.5)	49 (8.0, 98.1)
Q7|	5-year OS of Patients Proceeding to AlloSCT Following Routine Care, % (95% CrI)	14 (2.4, 30.6)	15 (4.9, 28.3)	14 (3.4, 27.2)	15 (1.0, 35.5)	18 (3.4, 37.3)	15 (2.7, 32.8)
Q8|	5-year OS of Patients who do not Proceed to	0-10	0-20	0-20	0-20	0-20	Cumulative probability

AlloSCT Following Routine Care, % (95% CrI)		distribution not below 0.4; SHELF cannot fit a distribution to this#

AlloSCT: Allogeneic stem cell transplant; **N/A:** Not applicable; **OS:** Overall survival.
†Linear pooling with equal weighting.
‡Binary outcome; linear pooling not applied.
§Only applicable to participants who responded "no" to Question 1.
¶Only applicable to participants who responded "yes" to Question 1.
|This reports the proportion of patients judged to be alive at 5 years.
#SHELF is unable to fit a distribution when more than 40% of chips are allocated to the first bin. As such, neither individual nor pooled distributions could be generated. This is a limitation of the software.

3.3.3 Validity of Judgements

There were no missing responses to any questions; however, the optional free-text boxes were not completed in all instances. No expert provided revisions to their judgements following the fitting of distributions.

An apparent inconsistency was noted with one expert. For question 5, Expert 2 (who indicated a cure fraction of 44% in question 2c) had the lowest elicited judgement for the five-year OS of patients treated with tisagenlecleucel (20%). Expert 2 provided no rationale in response to question 5. In a sensitivity analysis, Expert 2 was removed from the pooled analysis of question 5 (five-year OS of patients treated with tisagenlecleucel). The pooled judgement for this parameter increased to 36% (95% CrI 13.8 to 57.8). Additionally, for Expert 2, the elicited cure fraction for routine care (question 2d) is higher than the five-year OS of patients proceeding to alloSCT following routine care (question 7). Expert 2 cites 'high-risk disease/relapse' in their rationales to each question. Removal of Expert 2 from the pooled analysis of question 7 had negligible impact on the pooled judgement (15%; 95% CrI 2.6 to 34.9).

3.4 Discussion

The results of this study indicate that much variability exists between the judgements of this cohort of experts and in some cases, much uncertainty exists within these judgements. Variability around the concept of cure was observed. The judgements of those who deemed cure possible (n=2) were closely aligned. The 'cure' time-point of two years suggests that these experts do not consider patients to be at prolonged risk. This is

in line with some estimates in the literature (260). These experts appear to define cure in line with the Erice statement (260, 261); citing cure in terms of negligible risk of relapse from disease. Neither provided evidence of consideration of long-term complications and toxicities. Their judgements regarding the cure fraction of tisagenlecleucel were aligned, with a slightly higher degree of uncertainty associated with the judgement of Expert 2. The pooled judged cure fraction (tisagenlecleucel) (44%, 95% Crl 27.6 to 61.1) is considerably lower than the 18-month OS in ELIANA (70%, 95% CI 58 to 79) (217). Considering a cure point of two years was elicited from these experts, it appears they do not expect tisagenlecleucel to be as effective in clinical practice as in the trial setting. This is important from both a clinical and cost-effectiveness perspective. Patients should be provided with a realistic view of expected treatment outcomes with this therapy. Cost-effectiveness estimates, based on immature trial data, which are not reflective of expected outcomes in clinical practice, will be misleading.

Those experts who did not deem cure plausible (n=3) indicated that the excess risk of mortality is likely to be greatest immediately post-treatment, with risk waning over the patient lifetime. Chips allocated to later time-points may skew point estimates, in this instance. All three experts highlighted that mortality risk never returns to baseline (i.e. general population), due to long-term toxicities, meaning these patients will not be cured. This is in contrast to the Erice statement, which states that cure refers to the original disease, despite the potential for, or presence of, remaining disabilities of side effects of treatment (261). There appears to be fundamental disagreement amongst experts regarding the meaning of ‘cure’. The excess risk of mortality of patients diagnosed with ALL in childhood and who have survived long-term is documented in the literature (292, 293).

The term ‘cure’, in this exercise, was explicitly defined as being subject to age- and sex-matched general population mortality. It is unclear if the experts who deemed cure possible did not expect long-term toxicities to impact on long-term survival, or if they did not consider them at all in their judgements. Both cited risk of relapse from disease, which leaves room for interpretation. Background evidence provided to the experts aimed to ensure that all experts were using the same evidence base. Experts were also drawing upon their own experiences, which will vary. This observation has important

implications for future elicitation exercises. Group interaction, by means of a face-to-face elicitation or a Delphi-style remote process, is expected to prompt discussion around the impact of long-term toxicities. This would have allowed deeper insight into the basis upon which experts were making their judgements and ensured that all experts were considering the same factors. Ultimately, it is important that clinicians are aligned on the interpretation of 'cure', as this has implications for communicating expected outcomes. Until long-term data become available and in light of the contrasting judgements seen here, the impact of applying alternative assumptions in cost-utility analyses (i.e. cure versus excess mortality) should be explored.

In response to eliciting the proportion of patients who proceed to alloSCT following treatment with tisagenlecleucel (question 3), expert rationales (n=4) indicated that patients would proceed to alloSCT if they relapsed following treatment with tisagenlecleucel. The pooled estimate of 40% is therefore, reflective of the proportion of patients who relapse (not all patients who relapse receive alloSCT in the next line). The wide credible intervals illustrate the uncertainty associated with these judgements. Despite the high proportion of patients expected to relapse, this finding indicates that experts do not intend on using tisagenlecleucel as a bridge to alloSCT. However, uncertainty in the long-term OS of patients treated with tisagenlecleucel raises questions over the need to consolidate response with alloSCT. Dissimilar to our results, clinicians consulted by Schulthess et al., stated that they use tisagenlecleucel when alloSCT is not available (294). Current consensus on the order of alloSCT and CAR T-cell therapy in the treatment pathway is lacking (258); the likely place in therapy of tisagenlecleucel may vary between jurisdictions. The positioning of tisagenlecleucel in the treatment pathway will also have an impact on budget impact estimates.

The five-year OS of patients who do not proceed to alloSCT following treatment with tisagenlecleucel (question 5) reflects long-term OS. ELIANA was cited by one expert in their rationale, one cited long-term remission/response; three did not provide a rationale. The variation in judgements and wide credible intervals indicates a large degree of uncertainty. The pooled judgement of 33% (95% Crl 8.7 to 56.9) is much lower than the 18-month OS from ELIANA, 70% (95% CI 58 to 79) (217). This indicates that a plateau in survival, proposed by some (295, 296), is not expected by our experts by year

five. This has important implications for cost-effectiveness modelling and performance-based risk-sharing agreement negotiations. In the NICE HTA appraisal of tisagenlecleucel (TA554), the ERG concluded that a curative approach to modelling was plausible (223). We suggest that the judgements observed here prompt a more conservative approach. The suggested declining OS and associated uncertainty may warrant performance-based risk-sharing agreements to be implemented over a longer time horizon than has been agreed in some European countries (i.e. one to two years) (297).

As highlighted, for question 5, Expert 2 (who indicated a cure fraction of 44% in question 2c) had the lowest elicited judgement for the five-year OS of patients treated with tisagenlecleucel (20%, question 5). Removal of Expert 2's judgement from the pooled judgements of question 5 had a relatively small impact on the pooled estimate. The protocol is limited in guidance regarding inconsistent judgements, recommending that adjustment should only be applied to improve coherence and consistency, and not to reduce variability (264). Interpretation of this judgement is hampered by the lack of an accompanying rationale. Inconsistency is likely a more prominent feature in remote elicitation; group interaction or behavioural aggregation methods have greater potential to identify and reconcile such anomalies.

Also, dissimilar to all other experts, Expert 2 indicated that, after tisagenlecleucel treatment, the five-year OS would be lower in patients who do not proceed to alloSCT (question 5) versus those who do (question 4). Experts were instructed to consider the overall paediatric and young adult population with R/R ALL. However, it is possible that the experts drew upon their own personal experiences, and differences between judgements may be due to heterogeneities occurring between the populations seen by the experts. The influence of the availability heuristic cannot be ruled out in this instance, as experts were not presented with background evidence for routine care and may have relied solely on their recall of events. Judgements may have been influenced by very recent or prominent events (298). Without further justification from the experts, and detail on the patient populations seen by each expert, it is difficult to draw conclusions on this judgement.

Reassuringly, the point estimates and credible intervals of judgements relating to routine care (questions 7 and 8) were aligned between experts and the literature (216). However, much variability existed between experts in the proportion of patients judged to proceed to alloSCT with routine care (question 6). A wide credible interval was observed for the pooled distribution. Expert 2 indicated that they would expect all patients to be eligible to proceed to alloSCT based on achieving a sufficiently good response. Other experts indicated that not all patients would achieve a sufficiently good response or be fit enough to proceed to alloSCT. Notably, the line of therapy was not explicitly stated in this question (unlike questions 7 and 8). As such, there is a possibility that experts were not basing their judgements on the same populations. The risk of question misinterpretation cannot be ruled out.

Of note, is the expected improved OS in patients who receive alloSCT following treatment with tisagenlecleucel (question 4) compared to those who receive alloSCT following routine care (question 7). To our knowledge, there are no studies examining the impact of CAR T-cell therapy on the efficacy of alloSCT at later lines of therapy. The rationales provided by the experts did not provide any insight. However, the credible intervals associated with the judgements are wide, indicating a high degree of uncertainty.

A mathematical approach to aggregation, with equal weighting, was favoured here. Behavioural methods of aggregation (i.e. consensus), require face-to-face elicitation with the input of an experienced facilitator. As elicitation is an underutilised methodology in HTA, there is a limited supply of experienced facilitators. Behavioural methods of aggregation may lead to overconfident results ("groupthink") and are at the risk of being reflective of the judgements of dominant personalities within the group (298). However, as highlighted, group interaction amongst the experts may have facilitated the sharing of knowledge, a more in-depth insight into the experts' reasoning, and highlighted any inconsistencies that arose.

The Microsoft Excel®-based nature of the exercise posed the risk that only experts experienced with this platform responded. Piloting with groups who did not have experience in elicitation was intended to mitigate against the risk of responder bias. In

light of this limitation, further research was conducted, within the NCPE, to develop a user-friendly elicitation application to gather and synthesise results from elicitation exercises. This application was developed by a final year Management Science and Information System Studies (MSISS) student from Trinity College Dublin, in collaboration with members of the NCPE. The application developed, using R® and Shiny, generates descriptive statistics in real time and uses the SHELF package to elicit probability distributions based on expert judgements, allowing both the expert and the NCPE to understand and interpret responses. It incorporates a front-end questionnaire, feedback graphics and a separate dashboard from which the application can be maintained. Functionality is included to facilitate pooling of multiple judgements. The application is intended for use in face-to-face elicitation (299). Work is ongoing to identify a secure location for the Structured Query Language (SQL) database, which stores the elicitation data. Currently, the database is stored locally on one device. This is problematic, as the database is not backed up securely.

3.4.1 Comparison with the Published Literature

In their work (funded by Novartis), which also examined long-term survival of paediatric and young adult patients treated with tisagenlecleucel (for R/R ALL), Cope et al. elicited a 5-year OS of 54.9% (95% CI 24.5 to 80.5) (300). This is considerably higher than the 33% (95% Crl 8.7 to 56.9) elicited here, but also associated with much uncertainty. A possible explanation for this are the different methodologies employed to derive judgements. Survival estimates in Cope et al. were based on Kaplan-Meier curves. Cope et al. propose that experts may have been influenced by the 'flat tail' of the ELIANA OS curve. The authors highlight that it may also be the case that some experts were not clear on the definitions of the upper and lower bounds used in the study. It was further highlighted that variation was observed in the judgements and stated uncertainty was quite large in most cases; the final estimates were a "middle ground" based on consensus (300).

A validation exercise to assess the accuracy of judgements was conducted by Cope et al. This indicated that judgements of survival at two years were generally very close to the observed survival at that time point (300). This may indicate that the experts who took part in that study were well calibrated. It may also be reflective of anchoring based on

availability of up to 18 months of data at the time of study. If this is the case, judgements of survival at five years may not be well calibrated. A limitation of our study is that no such validation exercise was conducted.

There is a paucity of literature investigating the most appropriate elicitation method to accurately capture expert judgements. Little is currently known about how different methods affect results (301). Elicitation aims to devise experts' judgements in an accurate manner; the quality of an elicitation measures the extent to which this aim has been met (265). Based on this, and assuming all precautions are taken to mitigate against heuristics and biases, divergent outputs from different elicitations could be considered to be accurate if they are a true reflection of the experts' judgements, even if these judgements are not well calibrated with reality (286). This study provided experts the opportunity to revise their judgements. This aimed to ensure that the elicited judgements were an accurate reflection of the experts' underlying beliefs. However, it is not possible to objectively measure if the experts' beliefs have been accurately captured (283).

3.4.2 Limitations

One expert, who agreed to take part in the study, opted out due to difficulty in accounting for all patient variables. This issue has been encountered elsewhere (283). Piloting the exercise with experts in the area (264), may have highlighted this issue and identified ways to minimise this risk. Populations, with rare diseases, may be heterogeneous. In cases where the population under consideration is sufficiently heterogeneous, it may be more appropriate to elicit judgements for different subgroups of patients. However, this increases the burden and complexity of the exercise (264).

Limited background information, regarding the experts, was collected during the exercise. Information on the characteristics of patients treated in clinical practice (based on risk status, for example) was not collected. This precludes in-depth investigation into the reasons for heterogeneity between judgements.

It was anticipated that the remote, self-administered method would increase sample size. However, response rate was low. Low response rates are a common feature of

elicitation exercises (283, 302). The low response rate here may be a function of the rarity of the disease. The exercise may also have been perceived as being complicated. Additionally, it has been reported that clinicians are more reluctant than other professionals to express their opinion as probability judgements (303). Further research should be conducted examining how to appropriately incentivise increased participation and retention of experts.

In response to question 8, all experts allocated all of their chips to the first (0% to 10%) and second (10% to 20%) bins. SHELF specifies that chips should be placed in at least three of the bins, and to avoid placing more than 40% of chips in the first bin. This is to allow the fitting of a distribution; SHELF is unable to fit a distribution to judgements when more than 40% of chips are allocated to the first bin (275). When a high proportion of chips (proportion not specified) are placed in the first and last bins, the defined limits may not be small or large enough (272). As such, neither individual nor pooled distributions could be generated. This is a limitation of the software. This may have been prevented in a face-to-face elicitation session. Alternatively, the experts could have been asked to specify the minimum and maximum parameter values. The number of bins could have then been set according to these values.

3.5 Conclusion

Given the increase in treatments being introduced at an early point of evidence generation, the challenge for HTA assessors and decision-makers also increases. Although elicitation is not a substitute for robust RCTs, when conducted using a structured protocol, it can aid in the characterisation of the uncertain data that are available. Challenges that arose, in the conduct of the study and interpretation of judgements, may have been eased in a face-to-face elicitation. The findings of this study indicate that the 'curative' potential of tisagenlecleucel and associated long-term OS is highly uncertain. The clinical benefit observed in clinical trials may not be realised in clinical practice. This warrants conservative approaches to the communication of expected outcomes between clinicians and patients. It is not until long-term data become available that these uncertainties will truly be addressed. Our results highlight that performance-based risk-sharing agreements may need to be implemented over longer time horizons than have been agreed in a number of jurisdictions. This can mitigate against the financial risk that

is associated with potentially inappropriate decisions based on immature data. The outputs of this study will help inform the structural uncertainty associated with the extrapolation of long-term OS outcomes based on limited data. They will be used to validate OS outputs of the cost-utility model, presented in Chapter 5.

Chapter 4 Utility Data in Relapsed/Refractory Acute Lymphoblastic Leukaemia: Systematic Literature Review

4.1 Introduction

National Economic Evaluation Guidelines indicate that utility data should be identified by SLR (11). The NCPE preferred method of HRQOL measurement is the EQ-5D-3L, with utilities derived using the UK valuation set (30, 304). Several child- and adolescent-specific measurement instruments have been developed, including the EQ-5D-Y (EQ-5D-Youth; 'EQ-5D child-friendly version'), CHU9D (Child Health Utility 9 Dimension), HUI2 (Health Utilities Index Mark 2), and HUI3 (Health Utilities Index Mark 3). However, accompanying value sets have not been derived for all instruments, precluding their use in generation of utility values (305). Challenges in collecting HRQOL data in children and adolescents are further compounded by the rarity of R/R ALL. Small, heterogeneous populations may preclude the collection of accurate and generalisable data (306).

4.1.1 Chapter Aim

The aim of this chapter is to derive utility values for use in the bespoke cost-utility model of tisagenlecleucel for R/R ALL (Chapter 5). Relevant data will be identified by a comprehensive SLR.

4.2 Methods

An SLR protocol was developed, in line with the Cochrane Handbook for Systematic Reviews of Interventions (204). Guidance was obtained from NCPE Information Specialist, Ms Marie Harte. Reporting is conducted in line with PRISMA 2020 (205).

4.2.1 Systematic Literature Review

4.2.1.1 Population

The population was in line with the described licensed population of tisagenlecleucel; paediatric and young adult patients, up to 25 years, with R/R ALL (3). No lower age limits were specified.

4.2.1.2 Intervention and Comparators

Studies reporting utility data during or following treatment with any licensed therapy (in Europe) for R/R ALL in paediatric and young adult patients were included.

4.2.1.3 Outcomes

Outcomes were required to be reported as a utility value; a format that allowed use as an input parameter in the cost-utility model. The following varieties of utility value were included:

1. Health-state utility values for event-free survival and progressed disease
2. Utility values associated with long-term survival
3. Disutility values associated with treatment and associated administration or hospitalisation
4. Disutility values associated with short-term (eight weeks or less (143)) and long-term (greater than eight weeks) adverse events of treatment
5. Disutility values associated with HSCT (alloSCT and autoSCT)

4.2.1.4 Study Design

Any study providing the required outcome was included, with the exception of case studies or studies providing data on a single patient.

4.2.1.5 Search Methods

The search strategy is presented in Appendix C (Table A6). Electronic databases EMBASE, MEDLINE (via EBSCO), and CENTRAL (via the Cochrane Library) were searched from 01 January 2000 to 09 January 2021 inclusive. Articles were restricted to those published in English. Proceedings from the ASH and EHA Annual Conferences were hand searched for the years 2014 to 2020 inclusive. Terms used in searching of conference proceedings included: 'tisagenlecleucel', 'ELIANA', 'ENSIGN', 'tisa-cel', 'blinatumomab', 'acute lymphoblastic leukaemia', 'paediatric', 'health-related quality of life', 'quality of life', 'utility', 'QOL', and 'HRQOL'.

4.2.1.6 Choice of Utility Values

The utility values chosen for use in the cost-effectiveness model were selected using several criteria. Preference was given to utility values, which were derived using the EQ-5D-3L with the UK valuation set applied (as per the NCPE preferred approach). If data collected using the EQ-5D-3L were not identified, preference was given to data collected using an alternative generic measure. Subsequent criteria used to select utility data were

prioritised in the following order: relevance of population (including age) from which utility data were derived, sample size, and level of detail provided in the publication.

4.2.1.7 Citation Management

Identified citations were imported to Endnote® and transferred to Abstrackr. Duplicates were systematically searched for using software in Endnote® and identified manually. Title and abstract screening was conducted by a single reviewer to identify citations eligible for full-text review. The full texts of potentially relevant citations were obtained and assessed for suitability for inclusion in the final evidence base. For quality assurance purposes, a second reviewer screened 10% of full-text articles in duplicate. Any disagreements were resolved by a third reviewer.

Data extraction was conducted using an adapted Cochrane data extraction form (208). Data recorded included population, intervention, comparators, outcomes, study design, authors, title, and publication date. Extracted outcomes data were checked in duplicate by a second reviewer.

4.3 Results

A total of 259 citations were identified through database searches; 257 citations were screened following removal of duplicates. No additional citations were identified from the ASH and EHA Annual Conference proceedings. Title and abstract screening excluded 236 citations; 21 citations were brought forward for full-text screening. One study was included in the final evidence base. A PRISMA 2020 diagram is presented in Figure 5.

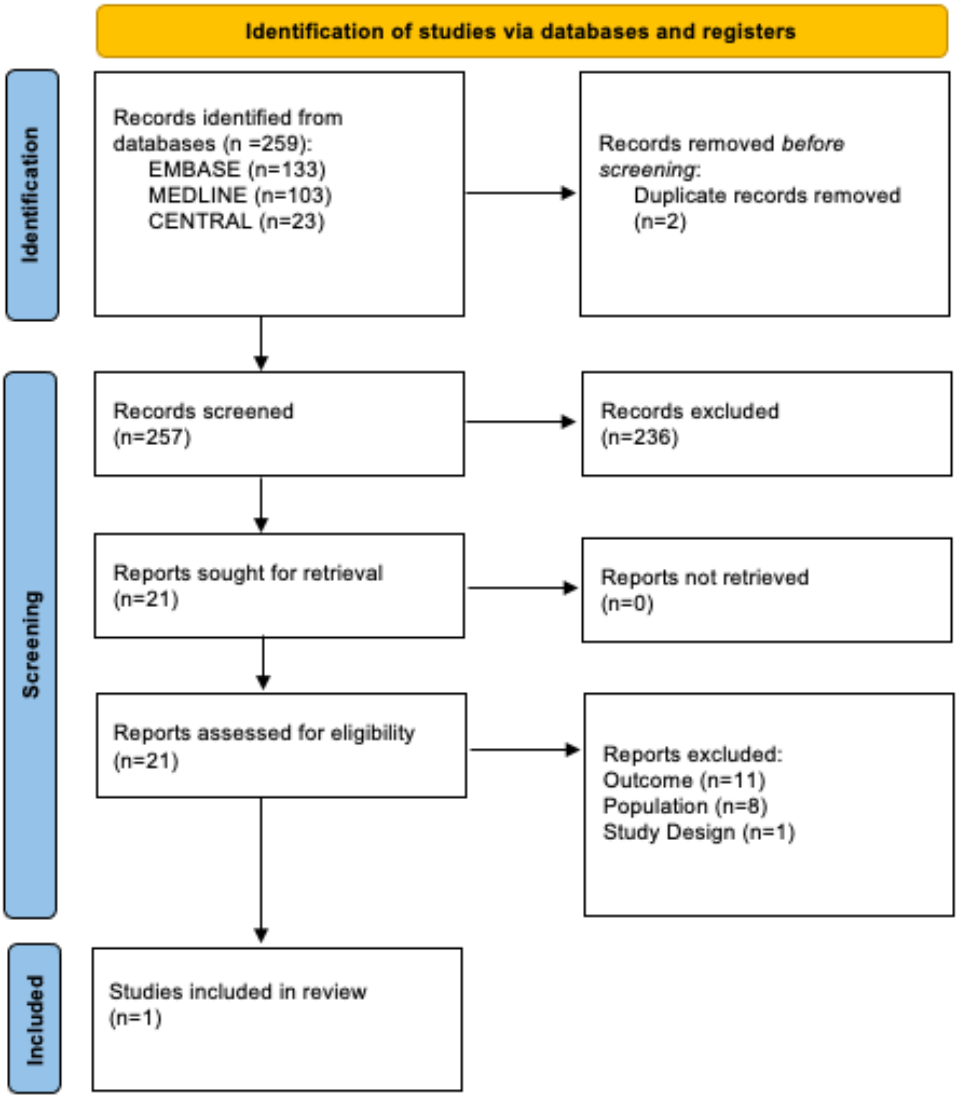

Figure 5 PRISMA diagram - systematic literature review of utility data for relapsed/refractory acute lymphoblastic leukaemia

4.3.1 Excluded Studies

Reasons for exclusion at full-text screening were outcome (n=11), population (n=8), and study design (n=1). Exclusion based on outcome (n=11) was because data were not presented as utility values and were therefore, not suitable for use in the cost-utility model. Despite studies reporting on patients with R/R ALL, many of these employed utility data that were not relevant to this population. These studies were excluded based on population (n=8). One study was excluded as it reported utility data for a single patient. These data were not considered a representative sample. A selected list of studies and reasons for exclusion are presented in Appendix C (Table A7).

4.3.2 Included Studies

The study included in the final evidence base, by Thielen et al. (295), is summarised in Table 11.

Table 11 Study included in systematic literature review of utility data for relapsed/refractory acute lymphoblastic leukaemia in paediatric and young adult patients

Author, year	Intervention	Comparator	Study Design	Source	Utility Values			
					Event-Free (SE)	Progressed (SE)	HSCT	Adverse Event and Age
Thielen et al. 2020 (295)	Tisagenlecleucel	Clofarabine Clofarabine-based combination therapy Blinatumomab	Cost-Utility analysis	EQ-5D-3L data collected in ELIANA with Dutch valuation set Literature	0.83 (0.03)	0.68 (0.05)	**Treatment** -0.21† **6-12 months post-treatment (non-GVHD)** -0.02† **6-12 months post-treatment (GVHD)** -0.173†	**Treatment** -0.202* (0.006) **ICU utility**: 0 **Age-related disutility**: Janssen et al. (307)

GVHD: Graft-versus-host-disease; **HSCT:** Haematopoietic stem cell transplant; **ICU:** Intensive care unit; **SE:** Standard error.
*Assumed to capture disutility associated with active treatment and any associated adverse events.
†Standard error assumed equal to 25% of mean point estimate.

Thielen et al. report health-state utility values of 0.83 and 0.68 for event-free survival and progressed disease, respectively. These were derived from the population of ELIANA, using the EQ-5D-3L, with the Dutch valuation set applied. No further detail regarding their derivation was provided. These values were applied to all treatment arms. Although patients alive after 60 months were considered long-term survivors, the utility accrued by these patients was not reported (295).

Disutility of -0.202 was applied to all treatment arms for the duration of treatment, reflecting disutility associated with active treatment. This was also assumed to capture adverse event disutility. This was applied for 26 and 61 days for tisagenlecleucel and blinatumomab, respectively. ELIANA (tisagenlecleucel) and NCT01471782 (blinatumomab) were cited for these durations (295). The disutility value (-0.202) was sourced from an SLR and meta-analysis of childhood health utilities, conducted by Kwon

et al. This comprised a comprehensive SLR and meta-regression, to determine the effects of factors on health-state utility values (308). However, data appear to be based on a population with newly diagnosed ALL, as opposed to those with R/R disease. Additionally, the disutility value was based on HUI3 (a generic measure of HRQOL) data, which were subject to methodological heterogeneity. This included respondent type (parent versus child), administration mode, and differences in age of respondents (child versus adolescent). It may not be appropriate to combine such data. It was observed, for example, that proxy assessment by parents was associated with an overestimation of children's HRQOL outcomes, as compared with those reported directly by children (308). As such, the validity of this estimate is compromised.

Thielen et al. assumed that patients experiencing grade 3-4 CRS had a utility of 0 for the duration of the event. For tisagenlecleucel, an additional disutility was included for non-CRS ICU admission, by assuming that patients admitted to the ICU, for non-CRS-related events, had a utility of 0. The durations employed were not reported (295).

Patients undergoing HSCT were assumed to experience disutility. Thielen et al. did not differentiate between alloSCT and autoSCT. Disutility associated with HSCT was derived from an SLR of health-state utilities in adult patients with acute myeloid leukaemia, conducted by Forsythe et al. (309). Thielen et al. calculated the relevant disutility value by subtracting the health-state utility value for HSCT (0.613 (309)), from the event-free survival state utility value (0.83) (295). Disutility associated with the 6 to 12 month period post-HSCT, for patients without graft-versus-host-disease (GVHD), was calculated in the same manner.

Health-state utility values for HSCT and the 6 to 12 month period post-HSCT (in patients not experiencing GVHD), were derived by Forsythe et al. by mapping EORTC QLQ-C30 (cancer-specific measure) values, published by Grulke et al. (310), to EQ-5D-3L (309). Grulke et al. conducted an SLR to identify EORTC QLQ-C30 data from patients, aged 14 to 70 years, who received HSCT (n=2,800). Patients received HSCT for acute leukaemia (28%), chronic myeloid leukaemia (5.3%), other haematological diseases (42.1%), and solid tumours (14.8%) (310).

Thielen et al. also accounted for disutility for the 6 to 12 month period post-HSCT in patients experiencing GVHD (-0.173) (295). Forsythe et al. was cited for this; however, the method of derivation was not reported (309).

Data from Janssen et al. were used to account for disutility associated with increasing age. Janssen et al. report EQ-5D-3L data (VAS and index values derived using the time-trade-off method) from surveys of the general population in 24 countries. Utility data and their corresponding age bands are presented (18 to 24, 25 to 34 years, etc.) (307). Using these data, a utility adjustment for age was calculated relative to the starting age in the model. The adjustment at each age in the model was calculated as the ratio between the utility value corresponding to the patient's current age in the model and the utility value corresponding to the starting age of the patient in the model.

Only the health-state utility values (event-free survival and progressed disease), identified in Thielen et al., were in line with the inclusion criteria of this SLR. However, due to paucity of data, all identified utility data will be considered for inclusion in the cost-utility model.

4.3.2.1 Systematic Review and Meta-Analysis by Kwon et al.

The SLR and meta-analysis of childhood health utility values by Kwon et al. (308), identified through Thielen et al. (295), was examined as a source of utility data. In addition to disutility associated with active treatment (-0.202), Kwon et al. derived utility relating to 'survivors of ALL' (0.90). This was derived from meta-analysis of 46 sample utility values. The definition of 'survivor' is not reported (308). This value is subject to the same limitations as those described for disutility associated with active treatment (4.3.2).

4.4 Additional Searches

Due to the paucity of data identified though the SLR, a search of websites of national HTA agencies was conducted. Utility values included in Applicant HTA submissions of tisagenlecleucel, appraised by national HTA agencies, were examined. Websites of relevant national HTA agencies were searched (from inception to 01 Feb 2021) using the terms 'tisagenlecleucel', 'Kymriah', 'acute lymphoblastic leukaemia', 'paediatric', and

'young adult'. Full HTA submission appraisal documents (herein 'HTA appraisal') were reviewed. Summary documents were not included. HTA agencies that published their HTA appraisals in English were selected. The websites of the following HTA agencies were searched:

1. NICE, UK (311)
2. Scottish Medicines Consortium (SMC), Scotland (312)
3. Canadian Agency for Drugs and Technologies in Health (CADTH), Canada (313)
4. Norwegian Medicines Agency (NoMA), Norway (314)

Based on these searches, the following HTA appraisals were reviewed:

1. NICE: Walton et al. Tisagenlecleucel-T for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people aged up to 25 years: A Single Technology Appraisal (TA554; 2018) (223).
2. SMC: Tisagenlecleucel for the treatment of paediatric and young adult patients up to 25 years with B-cell acute lymphoblastic leukaemia (ALL) that is refractory, in relapse post-transplant, or in second or later relapse: Detailed advice (2019) (315).
3. CADTH: Tisagenlecleucel for Acute Lymphoblastic Leukemia: Economic Review Report (2019) (240).
4. NoMA: Single Technology Assessment: Tisagenlecleucel (Kymriah) for the treatment of relapsed/refractory acute lymphoblastic leukaemia (ALL) in paediatric and young adult patients (2018) (250).

The mock technology appraisal conducted by Hettle et al. (12) (herein 'Hettle et al.') (described in 0) and the "Chimeric Antigen Receptor T-Cell Therapy for B-Cell Cancers: Effectiveness and Value Report (2018)", published by ICER (296) (herein 'ICER HTA'), were also examined. The websites of NCPE (Ireland) and Pharmaceutical Benefits Advisory Committee (PBAC), Australia were also searched. However, only summary documents were available (88, 316).

4.4.1 Health-State Utility Values

All six publications (12, 223, 240, 250, 296, 315) published utility values for the event-free survival and progressed disease states, sourced from Kelly et al. (317). Kelly et al. performed a decision analysis of cranial radiation therapy for paediatric patients with T-cell ALL. This included an SLR of utility studies. Kelly et al. mapped SF-36 (Short-Form Survey-36; a generic measure of HRQOL) and CHRI (Child Health Ratings Inventories; a generic measure of HRQOL for children) data to EQ-5D-3L and HUI2 (a generic measure of HRQOL), respectively. Utility data from patients with all subtypes of ALL were included (317).

Only NoMA published utility values derived from ELIANA (250), despite these being explored in the NICE and CADTH HTA appraisals (223, 240). Health-state utility values, published by NoMA, were calculated by the Applicant based on individual dimension scores from ELIANA, using the EQ-5D-3L, and applying the UK valuation set (250). The health-state utility values published by NoMA differ from those published by Thielen et al. due to application of different valuation sets (UK versus Dutch).

Health-state utility values, derived from Kelly et al. and the NoMA HTA appraisal (ELIANA), are presented in Table 12 (250, 317).

An assumption was applied whereby patients who were alive at a specified time point were considered ‘cured’. These were assigned utility associated with the event-free survival state, or an alternative utility value obtained from the literature, regardless of the health state occupied. In the NICE and NoMA HTA appraisals (223, 250), patients considered cured were assigned the event-free survival state utility, derived by Kelly et al. (0.91) (before the cure point, patients in the event-free survival state were assigned utility derived from ELIANA). In the SMC HTA appraisal, Hettle et al., and the ICER HTA, patients considered cured were assigned utility of the event-free survival state (regardless of health state membership) (12, 296, 315). The time point at which patients were considered cured was five years (12, 250, 296, 315), with the exception of the NICE HTA appraisal, where it was assumed to be two years (223).

The NICE and NoMA HTA appraisals (223, 250) also highlighted utility values collected by Aristides et al. This study collected utility values, which were assigned to health states, experienced by adult patients with R/R ALL, by a sample (n=123) of the UK general population using the time-trade-off methodology (318). These values were not considered relevant to the population defined in this SLR; no further consideration was given to this publication.

Table 12 Health-state utility values in relapsed/refractory acute lymphoblastic leukaemia identified from HTA appraisals of tisagenlecleucel (223, 240, 250, 315), Hettle et al. (12), and the ICER HTA (296)

Health State	Mean Utility (Range/SD)	
	Kelly et al. (317) (published by all six sources (12, 223, 240, 250, 296, 315))	ELIANA (published by NoMA (250))
Event-Free Survival	0.91 (0.87, 0.95)	0.80* (0.23)
Progressed Disease	0.75 (0.44, 1.00)	0.63* (0.36)
Long-Term Survival	0.91 (0.87, 0.95)	0.80 (0.23)

HTA: Health technology assessment**; ICER:** Institute for Clinical and Economic Review; **NoMA:** Norwegian Medicines Agency; **SD**: Standard deviation;
*UK valuation set applied to EQ-5D-3L data collected during ELIANA.

Utility values derived from ELIANA were subject to limitations. They were based on small sample size (n=33) and corresponded to patients who responded to treatment. These factors, and the open-label nature of the trial, increase the risk of bias. Data were based only on patients aged 13 years and older, which is not reflective of the entire population in ELIANA. HRQOL data were also collected in patients aged between 8 and 12 years, using the EQ-5D-Y. At the time of ELIANA, and subsequent publication of associated health-state utility data, there was no accompanying valuation set for this instrument. These data were therefore, excluded from the analysis. Despite these limitations, values obtained directly from the population of interest were considered most appropriate for the decision problem. The values from Kelly et al. were deemed less applicable to the population defined here, as described below.

Several limitations were identified in Kelly et al. The event-free survival state utility was derived from SF-36 data collected from patients diagnosed with ALL between 1976 and 2003. These patients were considered cured following relapse and had survived for at least five years. The value of 0.91, used for the event-free survival state, was based on the condition of survival of more than five years (319). As such, patients who initially responded but progressed within five years were likely to be excluded.

The progressed disease state value of 0.75 was derived from patients (aged 5 to 18 years) undergoing myeloablative HSCT (both alloSCT and autoSCT) using the CHRI. The publication cited by Kelly et al. explored HRQOL over a 12-month period (320). Kelly et al. justified the use of these data, stating that myeloablative HSCT is routine care for patients with relapsed T-cell ALL (317). Notably, patients are required to achieve a response before proceeding to HSCT. This value (0.75) may therefore, not be reflective of patients with relapsed disease. Limited data were presented in terms of characteristics of this population and it is unclear if the population is in line with that of ELIANA.

Methodological limitations were also noted in Kelly et al. Different preference instruments were used to generate the event-free survival and progressed disease state utility values. These instruments differ in the dimensions of health they cover, the number of levels defined on each dimension, and the description of each level, resulting in utility values which may not be comparable (28, 250). NoMA raised concerns regarding the mapping algorithm used to map utility data from CHRI to EQ-5D-3L. This mapping exercise could not be validated, raising further concerns regarding the appropriateness of the values (250).

4.4.2 Treatment and Adverse Event Disutility Values

Disutility values, associated with active treatment and short-term adverse events, derived from Sung et al. (321), were outlined in three HTA appraisals (223, 240, 250), Hettle et al. (12), and the ICER HTA (296). The SMC HTA appraisal highlighted that disutility due to adverse events was accounted for; actual values were not reported (315). Sung et al. was not included in this SLR, as it reports on an irrelevant population, as described below.

Sung et al. elicited utility values from clinicians (n=12) who care for patients undergoing HSCT (both alloSCT and autoSCT) at three tertiary care centres in Toronto (Canada). Values were elicited using the VAS, for ‘young adult’ patients (age not specified) with acute myeloid leukaemia in first remission, who survived post-transplantation without recurrent disease. Disutility values of -0.42 and -0.57 were derived for patients receiving

chemotherapy and HSCT, respectively (321). These values were applied in three HTA appraisals (223, 240, 250) and also the ICER HTA (296).

The NICE HTA appraisal applied this disutility value (-0.42) for 25.85 and 9.24 days for tisagenlecleucel and blinatumomab, respectively (223). The CADTH HTA appraisal applied this disutility to patients receiving tisagenlecleucel from the time of lymphodepleting chemotherapy, and for the duration of treatment in the comparator arm (salvage regimen, not otherwise specified). The durations employed were not reported (240). The NoMA HTA appraisal applied this value to those receiving bridging chemotherapy in the tisagenlecleucel arm (21 days); however, it does not appear to have been applied to patients receiving tisagenlecleucel infusion. It was also applied to patients receiving comparator treatment (clofarabine, etoposide, cyclophosphamide) for the duration of treatment (46 days, derivation of duration not reported) (250). The ICER HTA applied this disutility value to patients receiving tisagenlecleucel and those receiving pre-treatment chemotherapy regimens in the tisagenlecleucel arm. The durations were not explicitly stated (296). This value was excluded from the model in Hettle et al., because the disutility estimate for all forms of chemotherapy was the same for all treatments (12).

Disutility associated with HSCT (-0.57; both alloSCT and autoSCT), derived from Sung et al. (321), was applied over various durations in the publications. Durations of one year (12, 240), three months (223), and two months (250) were used. The duration over which this disutility value was applied was not reported in the ICER HTA (296).

The population described in Sung et al. is not in line with the population defined in this SLR. Sung et al. was based on young adult patients who were in first remission and values only accounted for patients who survived without recurrent disease. Values were elicited from clinicians (321). Patient- and clinician-assessed HRQOL, in patients undergoing HSCT, has not been concordant between these groups elsewhere (322). The disutility values, derived by Sung et al., are not in line with National Economic Evaluation Guidelines, which recommend that utility values are derived using a generic preference-based method (11). Sung et al. estimated values using a 10cm VAS and converted to disutility values using the formula: 1-(1- VAS score) (321). VAS measurement is subject to measurement bias (28). Additionally, the VAS does not provide an appropriate measure

of preference for use in HTA as it does not present the respondent with choices (as is the case with standard gamble or time-trade-off) (323).

The NICE HTA appraisal used a different HSCT-associated disutility value (-0.13), from months 4 to 12, to reflect improvement in HRQOL over time (223). This was derived by Felder-Puig et al., who assessed HRQOL, using the HUI2 and HUI3, in paediatric patients (between 4 and 18 years; n=68) who received alloSCT. HRQOL data were collected from patients with a variety of disease types; ALL accounted for 50% of these. The disease stage of patients is not reported. Felder-Puig et al. was subject to a high proportion of patients lost to follow up (27.9%) (324). However, this value was considered more appropriate than those derived by Forsythe et al. (309) and Sung et al. (321). This is because this value was derived using a generic measure and from a population that was more aligned with the population specified here.

All publications accounted for disutility due to grade 3-4 CRS (12, 223, 240, 250, 296, 315). All patients in the tisagenlecleucel arm, with grade 3-4 CRS ICU admission, were assumed to have a utility of 0 for the duration of ICU stay. This duration varied; 7 days ('based on ELIANA', as per NoMA) (12, 250), 8 days ('median duration of ICU stay due to CRS') (296), and 9.8 days ('mean duration ELIANA') (240). This duration was not published in the NICE HTA appraisal. The NICE HTA appraisal also applied a utility of 0 for patients experiencing grade 3-4 CRS in the blinatumomab arm. The duration of utility was assumed to be 11.1 days (223). For comparator treatments, disutility associated with treatment (-0.42, derived from Sung et al. (321)) was assumed to capture all adverse events.

The SMC and CADTH HTA appraisals included disutility to account for non-CRS ICU admission in the tisagenlecleucel arm (240, 315). Detail regarding the derivation and implementation of this disutility value was not provided in the SMC HTA appraisal (315). CADTH assumed a non-CRS ICU admission utility value of 0. This was applied for 1.78 days, based on the mean ICU length of stay for non-CRS adverse events in ELIANA. The proportion of patients requiring non-CRS ICU admission was not reported. CADTH also included a scenario to account for disutility associated with febrile neutropenia;

however, the disutility value employed was not reported (240). A summary of disutility values is presented in Table 13.

No disutility values relating to long-term adverse events were identified.

Table 13 Disutility values in relapsed/refractory acute lymphoblastic leukaemia identified from HTA appraisals of tisagenlecleucel (223, 240, 250, 315), Hettle et al. (12), and the ICER HTA (296)

Parameter	Value (range/SD)	Duration	Identified Through	Source
Active Treatment including Adverse Events	-0.42 (0.16, 0.83)	Duration of treatment	HTA appraisals (223, 240, 250) Hettle et al. (12) ICER HTA (296)	Sung et al. (321)
Grade 3-4 CRS	0	Various assumptions: Tisagenlecleucel: 7 days (12, 250), 8 days (296), 9.8 days (240) Blinatumomab: 11.1 days (223)	HTA appraisals (223, 240, 250) Hettle et al. (12) ICER HTA (296)	Assumption
Non-CRS ICU Admission (Tisagenlecleucel)	0	Mean duration of non-CRS-related ICU admission (ELIANA); 1.78 days	HTA appraisal (240)	Assumption
Haematopoietic Stem Cell Transplant	-0.57 (0.31, 0.87)	Various assumptions: 1 year (12, 240), 3 months (223), 2 months (250)	HTA appraisals (223, 240, 250) Hettle et al. (12) ICER HTA (296)	Sung et al. (321)
	-0.13 (0.16)	9 months	HTA appraisal (223)	Felder-Puig et el. (324)

CRS: Cytokine release syndrome; **HTA:** Health technology assessment; **ICER:** Institute for Clinical and Economic Review; **ICU:** Intensive care unit; **SD:** Standard deviation.

4.4.3 Age-Related Disutility

Three HTA appraisals explicitly indicated that utility values were adjusted to account for decreasing utility associated with increasing age (223, 250, 315). The NICE HTA appraisal (223) adjusted health-state utility values, based on data from Janssen et al. (307) (described in 4.3.2).

The NoMA HTA appraisal (250) adjusted utility based on data from Sun et al. (325) and BurstrØm et al. (326). Sun et al. and BurstrØm et al. report age-specific utility values of the general population in Sweden. An adjustment index, which was set to one at the start of the model, was reduced over time based on age-specific utility values derived from these publications (325-327).

The SMC HTA appraisal and Hettle et al., report that an adjustment to utility was made to account for increasing age; however, no further detail was provided (12, 315). The ICER HTA did not indicate whether an adjustment to utility was made (296).

4.5 Utility Values Selected for Use in the Bespoke Cost-Utility Model

The utility values selected for use in the cost-utility model, developed as part of this research, are summarised in Table 14. The durations employed are presented in Chapter 5 (Table 21).

No treatment-specific utility values were identified for comparator treatments. Therefore, in the cost-utility model, utility will be applied according to health-state occupancy. As described, heath-state utility data from ELIANA, with the UK valuation set applied, were deemed most appropriate for the model (250). It is expected that patients who survive to a certain time post-treatment, usually between 24 and 60 months, will survive long term (202, 203). As such, an assumption will be made whereby all patients alive after 60 months (therefore, considered to be long-term survivors) have utility equivalent to that of the event-free survival state (12, 296, 315). Considering the uncertainty in the HRQOL of patients who are expected to survive long term, and the plausibility of improved HRQOL in these patients, a scenario will be explored whereby patients alive after 60 months are assigned a utility of 0.90, derived by Kwon et al. (308).

A scenario will also be explored, whereby patients alive after 60 months are subject to age- and sex-matched general population utility (328).

The study by Kwon et al. (308) was favoured over Sung et al. (321) to inform disutility associated with active treatment. Disutility derived by Kwon et al. was more closely aligned with the Irish reference case, in that data were collected using a generic measure (HUI3) (308). This disutility value (-0.202) will be applied to patients undergoing apheresis, bridging chemotherapy and lymphodepleting chemotherapy.

Health-state utility values, collected during ELIANA, will be assumed to capture the impact of most adverse events. This is to avoid double-counting. Additionally, no treatment or disease-specific disutility values were identified. Due to the severity of an ICU event, patients who require ICU admission for CRS- and non-CRS-related events will be assigned a utility of 0 for the duration of the event. This is a necessary assumption due to lack of robust, supportive data. No data were identified in this SLR to inform disutility associated with febrile neutropenia. As such, a disutility value of -0.15, identified in the SLR of utility data in R/R DLBCL (presented in Chapter 8), will be used here (329). In ELIANA, 3% of patients experienced grade ≥3 pancytopenia (143). The mean duration of an episode of pancytopenia was not reported. However, based on clinical opinion obtained by the NICE Evidence Review Group (ERG), the impact of pancytopenia, in patients with R/R DLBCL treated with CD19 CAR T-cell therapy, is most notable in the first few months after infusion, gradually improving to resolution within one year (330). For this analysis, it will be assumed that patients experience pancytopenia for six months. This disutility will be assumed equivalent to that of febrile neutropenia (-0.15) (329). Disutility associated with active treatment, and assumed to capture the impact of adverse events on HRQOL, derived by Sung et al. (321), will be explored in scenario analysis.

In line with assumptions employed by Thielen et al. (295) and the NICE HTA appraisal (223), different disutility values will be applied to capture the impact of HSCT on HRQOL, reflecting improvement in the condition of the patient over time (324). Notably, in this analysis, HSCT refers specifically to alloSCT. A disutility value -0.20 will be applied for the first 3 months post-alloSCT. This value was derived from Forsythe et al. (309), using the

same method employed by Thielen et al. (295) (i.e. event-free survival state utility value minus HSCT treatment utility value). This disutility will be reduced to -0.13, derived from Felder-Puig et al., from months 4 to 12 (324). This value was favoured over that sourced from Thielen et al., as it was derived directly from paediatric patients who received HSCT. These chosen values are considerably lower than the HSCT disutility value derived by Sung et al. (-0.57) (321). They are also aligned with the disutility value employed for active treatment (-0.202 (308)). Of note, it may be reasonable to expect that patients undergoing HSCT have worse HRQOL than those undergoing active treatment with chemotherapy. Based on published HTAs, this value is not expected to be a major driver of cost effectiveness in the cost-utility model (12, 223, 240, 250, 296, 315).

The method to derive disutility associated with increasing age, using data by Janssen et al. (307) and the method employed in the NoMA HTA appraisal (250), were the only methods identified to derive such disutility values. However, as described in Chapter 8, using data from Ara and Brazier (328) to derive disutility associated with increasing age is considered a more appropriate approach for the cost-utility models developed as part of this research.

Table 14 Utility values in relapsed/refractory acute lymphoblastic leukaemia used in bespoke cost-utility model: base case and scenario analyses

Parameter	Value (SD)	Source	Justification
Base Case			
Event-Free Survival	0.80 (0.23)	EQ-5D-3L collected in ELIANA with UK valuation set applied, identified through HTA appraisal (250)	Derived from a generic measure of HRQOL, in line with National Guidelines (11)
Progressed Disease	0.63 (0.36)		
All Patients Alive after 60 Months	0.80 (0.23)	HTA appraisals (223, 250, 315), Hettle et al. (12), ICER HTA (296)	Assumption. HRQOL equivalent to that of the event-free survival health state
Disutility Associated with Treatment and Adverse Events	-0.202 (0.006†)	Kwon et al. (308), identified through Thielen et al. (295)	Applied to patients undergoing apheresis, bridging, and lymphodepleting chemotherapy
CRS ICU Admission	-0.80 (0.23)	HTA appraisals (223, 240, 250), Hettle et al. (12), ICER HTA (296)	Assumption (i.e. a utility of 0). Accounts for impact of ICU admission on HRQOL
Non-CRS ICU Admission		HTA appraisal (240)	
Febrile Neutropenia	-0.15 (0.04*)	Lloyd et al. (329) identified through HTA appraisals (331, 332)‡	Febrile neutropenia may require hospitalisation; expected to have an impact on HRQOL
Pancytopenia			Assumption. Based on NICE HTA appraisal (TA677) (330)
Disutility Associated with AlloSCT (first 3 months post-alloSCT)	-0.20 (0.05*)	Forsythe et al. (309), identified through Thielen et al. (295)	Patients may experience a decrease in HRQOL post-alloSCT; this is expected to improve over time
Disutility Associated with AlloSCT (4 to 12 months post-alloSCT)	-0.13 (0.16)	Felder-Puig et el. (324), identified through HTA appraisal (223)	
Age-Related Disutility	Ara and Brazier (328)	Identified through systematic literature review of utility data in R/R DLBCL (Chapter 8); HTA appraisal (333)	Adjustment so that utility is not higher than that of the general population
Scenario Analysis			
All Patients Alive after 60 months	0.90 (0.006†)	Kwon et al. (308), identified through Thielen et al. (295)	Long-term survivors may experience HRQOL improvement
	Age- and sex-matched general population utility (328)	Assumption. Based on systematic literature review of utility data in R/R DLBCL (Chapter 8)	
Disutility Associated with Treatment and Adverse Events	-0.42 (range: 0.16, 0.83)	Sung et al. (321), identified through HTA appraisals (223, 240, 250), Hettle et al. (12), ICER HTA (296)	Health-state utility values may not incorporate disutility due to treatment and adverse events

AlloSCT: Allogeneic stem cell transplant; **CRS**: Cytokine release syndrome; **DLBCL:** Diffuse large B-cell lymphoma; **HRQOL**: Health-related quality of life; **HTA:** Health technology assessment; **ICU:** Intensive care unit; **R/R:** Relapsed/refractory; **SD:** Standard deviation; **UK:** United Kingdom.

*Assumed 25% of the mean point estimate.
†Standard error.
‡Identified in systematic literature review of utility data in R/R DLBCL (Chapter 8).

4.6 Discussion

This SLR highlighted the paucity of utility data in paediatric and young adult patients with R/R ALL. The majority of full-text studies, identified through database searching, were excluded based on outcome (i.e. outcome not presented as a utility value). This may indicate that alternative methods are used to evaluate HRQOL in this population. Of the studies identified through database searching, just one was identified, Thielen et al. (295), which included data that were in line with the inclusion criteria. An additional search of utility data, used in international HTA appraisals of tisagenlecleucel (223, 240, 250, 315), as well as a mock technology appraisal (Hettle et al.) (12) and the ICER HTA (296), further highlighted the scarcity of appropriate data.

4.6.1 Health-State Utility Values

The health-state (event-free survival and progressed disease) utility values in Thielen et al. were obtained from EQ-5D-3L data collected during ELIANA. However, the values were limited, for the purpose of this analysis, in that they did not meet the NCPE preferred approach. The Dutch valuation set was applied to EQ-5D-3L data (collected during ELIANA) to derive these values. As highlighted in 4.1, in the absence of an Irish valuation set for the EQ-5D-3L, the UK valuation set is generally used.

The health-state utility data, derived from ELIANA (EQ-5D-3L) with the UK valuation set applied (published by NoMA (250)), were aligned with the population of interest to this SLR, and in line with National Economic Evaluation Guidelines (11). Applying the Dutch valuation set to the EQ-5D-3L data (from ELIANA) resulted in higher utility values for both the event-free survival and progressed disease states (when compared to the UK valuation set) (250, 295). Health states are valued differently between the Netherlands (334) and the UK (304). For example, the 'worst' health state (33333) has a value of -0.624 in the Dutch valuation set (334), and a value of -0.594 in the UK valuation set (304). Use of different valuation sets could have important implications for cost-effectiveness estimates. Differences between valuation sets may be due to a number of factors including cultural factors, and study design and conduct (334). For this research, the UK

valuation set was considered more representative of the Irish population (due to cultural and societal similarities between Ireland and the UK).

HRQOL data collected during ELIANA are subject to bias. The small sample size, collected in patients who were responding to treatment, is likely to result in health-state utility values that are higher than those observed if data were collected in the entire cohort (i.e. patients who did and did not respond to treatment). The challenges in deriving HRQOL data in children are reflected in these data. Despite the collection of HRQOL data in patients aged 8 years and older, published health-state utility data, from ELIANA, are derived from patients aged 13 years and older (250, 295). This limits the size and generalisability of the published utility data derived from this trial. Exclusion of data collected from children aged between 8 and 12 years, disregards the full spectrum and experience of health within the population. It also assumes that utility values employed are applicable to all cohorts (i.e. as children transition to adolescents and adults). In the absence of comparative data, it is difficult to verify this assumption.

Kelly et al. (317) was used to inform health-state utility values in publications identified through additional searches (223, 240). Utility values for the event-free survival and progressed disease states were mapped to generic measures, which increases uncertainty in these estimates. The values, by Kelly et al., were not aligned with the population of interest to this SLR, and were subject to a number of methodological limitations. The derivation of the event-free survival and progressed disease state utility values using different instruments may produce systematically different results (11, 335). As such, these may not be comparable and may introduce bias into the cost-utility model.

The health-state utility values derived by Kelly et al. (317) were considerably higher than those derived from ELIANA (250, 295). The higher value for the event-free survival state (0.91), derived by Kelly et al., may be explained by the fact that this was based on the condition of survival of more than five years (317). This value may be more reflective of utility associated with long-term survival (as opposed to event-free survival post-treatment). HRQOL data collection in ELIANA continued for up to 12 months post-

treatment (221). The values derived from ELIANA are therefore, likely to encompass some degree of disutility due to adverse events.

The higher value for the progressed disease state, derived by Kelly et al. (when compared to ELIANA), may reflect the fact that this value was derived from patients undergoing HSCT, and not from patients experiencing disease progression. This value may be more reflective of the utility of patients in response while on active treatment. The wide range reported for this value indicates the high degree of associated uncertainty (317).

The lack of data pertaining to HRQOL of patients with R/R disease, considered to be long-term survivors, is notable. Considering the young age of patients defined in this SLR, and the potential long-term survival benefit associated with tisagenlecleucel, this is an important parameter. The utility value of 0.91 (Kelly et al. (317)) was closely aligned with the long-term survival utility derived by Kwon et al. (0.90) (308). The method of derivation of these values differed. Due to paucity of published data, the populations from which these values were derived could not be exhaustively compared. The definition of 'cure' is not reported in Kwon et al. As such, it is difficult to assess the validity of these estimates. Of note, both values are lower than the general population utility in England for those aged between 18 and 24 years (0.933) (307).

HRQOL of patients considered to be long-term survivors is a key area of uncertainty. An SLR examining HRQOL in patients with ALL, who were considered long-term survivors, found inconsistent evidence. Of the 31 studies identified, 13 reported worse HRQOL in long-term survivors, 8 found no difference, and 3 found better HRQOL outcomes when compared to healthy controls or siblings during survivorship (336). Comparison across studies is limited by differences in populations, measures of HRQOL, and definition of 'impaired HRQOL'.

The variation in time point at which patients are considered to be long-term survivors illustrates the uncertainty in this parameter. In the absence of robust evidence, a conservative approach will be adopted in the cost-utility model developed here.

4.6.2 Treatment and Adverse Event Disutility Values

None of the identified values, relating to disutility associated with active treatment and adverse events, were aligned with the inclusion criteria of this SLR. Identified disutility values were based on assumptions or derived from populations, which were not aligned with the specified population. No adverse event-specific disutility values were identified. It is difficult to validate the generalisability of the assumptions and values identified to the population defined here. No disutility values were identified relating to long-term adverse events.

Disutility associated with active treatment (-0.202), derived by Kwon et al. (308), is in line with National Economic Evaluation Guidelines, in that it was derived using a generic measure (HUI3) (11). An alternative disutility value (-0.42), accounting for disutility associated with active treatment, and assumed to capture disutility due to adverse events, was derived by Sung et al. (321). This value has a notably greater decrement on HRQOL than that derived by Kwon et al. (308). This difference may be explained, in part, by the different methodological approaches to their derivation. The wide range reported for the disutility derived by Sung et al. indicates the high degree of uncertainty associated with this value (321). The associated measure of uncertainty, of the disutility derived by Kwon et al., is notably narrower (308). Concerns regarding disutility values derived by Sung et al. have been outlined (4.4.2). Despite these limitations, these values have been widely used in HTA appraisals by national HTA agencies (223, 240, 250). This further highlights the scarcity of relevant utility data in paediatric and young adult patients with R/R ALL.

The use of a single disutility value to capture disutility associated with treatment and related adverse events, precludes a granular analysis of the impact of adverse events on HRQOL in the cost-utility model. This may result in QALY estimates, which are not reflective of the true impact on HRQOL. This lack of data is also reflected in the assumptions regarding the impact of grade 3-4 CRS and non-CRS-related ICU admission on HRQOL. Although uncertainty exists in disutility associated with adverse events, this parameter was not a driver of cost effectiveness in published HTA appraisals (223, 250, 296).

Both Thielen et al. (295) and NICE (223) used different disutility values to reflect the increase in HRQOL over time post-HSCT. The disutility values employed by Thielen et al. (-0.21 from months 1 to 5; -0.02 from months 6 to 12) (295) had a notably smaller decrement than those employed in the NICE HTA appraisal (-0.57 from months 1 to 3; -0.13 from months 4 to 12) (223). These values are not directly comparable, as they were derived from different populations and disease areas. There are also methodological differences in the derivation of these values. Of note, the disutility value of -0.13, derived by Felder-Puig et al. (used in the NICE HTA appraisal (223)), was derived from patients aged between 4 and 18 years (50% of these had ALL) (324). The other identified disutility values, pertaining to HSCT, were derived from adult patients or clinicians. All HSCT disutility values, identified during this SLR, are subject to limitations. Reassuringly, the value of -0.21, derived by Thielen et al., is closely aligned with that identified in the SLR of utility data in R/R DLBCL (-0.30, derived by Guadagnolo et al. (337)). Despite differences in populations and disease areas, between Forsythe et al. and Guadagnolo et al. (337), the close alignment between values provides some degree of reassurance. Disutility due to HSCT was not a driver of cost effectiveness in published HTA appraisals (223, 250, 296).

4.6.3 Age-Related Disutility

The extrapolation of utility data, over the time horizon of an economic model, without an adjustment for age can result in higher utility than that of the general population. Considering the young age of patients under examination in this SLR, an adjustment to account for age-related disutility is important. Of the publications identified, that explicitly mentioned an adjustment to account for age, none detailed the approach taken (i.e. multiplicative or additive) (12, 223, 250, 315). This may lead to inconsistent adjustments (338). A multiplicative approach will be adopted in the cost-utility model. This assumes constant relative decrement of disease health states on utilities. The multiplicative approach is the preferred approach of NICE (338). This approach is warranted in the cost-utility model due to the lack of longitudinal data on HRQOL in patients treated in ELIANA.

Justification for use of an adjustment factor, based on the formula by Ara and Brazier, over that derived from Janssen et al. is provided in Chapter 8 (307, 328). The limitations of Ara and Brazier are also outlined in Chapter 8.

4.6.4 Limitations

There are several limitations to this SLR. These are in addition to those discussed thus far. The inclusion criteria was restricted to studies reporting on patients with R/R ALL. Studies that could act as appropriate proxy data may have been excluded. Limiting the searching of conference proceedings to just ASH and EHA Annual Conferences may have resulted in undetected relevant sources of data. Included outcomes were specifically required to be reported as utility values. This may have excluded studies, which provided valuable information on the impact of treatment and disease on HRQOL.

The paucity of data regarding disutility relating to active treatment and adverse events, in the population specified in this SLR, resulted in the selection of proxy data for use in the cost-utility model. In the absence of comparative data, it is difficult to determine if these proxy data are an accurate reflection of the true disutility experienced by patients.

No measures of associated uncertainty were presented for the values derived by Forsythe et al. (disutility associated with HSCT - first 3 months) or Lloyd et al. (febrile neutropenia and pancytopenia) (309, 329). As such, an assumption is required, for the purpose of the PSA, whereby the standard error is equivalent to 25% of the mean point estimate. This is a necessary assumption, and in line with the approach taken by Thielen et al. (295). However, this assumption may not provide a true reflection of the uncertainty associated with these values.

A pragmatic approach was adopted, in that screening and assessment of study eligibility was conducted by a single reviewer. This was due to the time-intensive nature of these processes. As a quality assurance measure and in an attempt to minimise bias, 10% of full-text articles were screened in duplicate.

4.7 Conclusion

There is a deficient evidence base informing the utility of paediatric and young adult patients with R/R ALL. Health-state utility data, derived from ELIANA, were subject to limitations. Identified disutility values were based on proxy data or assumptions. This illustrates the multiple challenges that exist when deriving utility data in children and young adult patients with a rare disease. Increasing emphasis should be placed on the collection of HRQOL data, and future research should focus on generating evidence to support preferred methodologies for collecting and valuing HRQOL data in such populations. The limitations of the identified data may limit their generalisability and result in biased cost-effectiveness estimates. The associated uncertainty of the utility values identified in this SLR warrants extensive sensitivity analysis in the cost-utility analysis, presented in Chapter 5.

Chapter 5 Cost-Utility and Value of Information Analysis of Tisagenlecleucel for Relapsed/Refractory Acute Lymphoblastic Leukaemia

5.1 Introduction

5.1.1 Chapter Aim

The aim of this chapter is to conduct a cost-utility analysis of tisagenlecleucel for the treatment of paediatric and young adult patients with R/R ALL in the Irish healthcare setting. This chapter presents a bespoke cost-utility model, constructed in line with the Irish reference case (11). The estimates of relative efficacy, derived in Chapter 2, will be used to inform efficacy. Output from the expert elicitation, presented in Chapter 3, will be used to inform key areas of uncertainty. Utility data identified through SLR, presented in Chapter 4, will be utilised. Sensitivity analyses will also be conducted, examining the robustness of results to variation and uncertainty in model inputs and assumptions.

5.2 Model Development

5.2.1 Irish Reference Case

National Economic Evaluation Guidelines informed the framework for the cost-utility model (11) (presented in Table 15).

Table 15 Reference case for cost-utility analyses in Ireland, as per National Economic Evaluation Guidelines (11)

Element of Technology Assessment	Reference Case
Evaluation Type	Cost-utility analysis
Perspective on Costs	Health Service Executive (HSE), Ireland
Perspective on Outcomes	Health benefits accruing to individuals
Time Horizon	Sufficient to capture meaningful differences in future costs and outcomes
Comparator	Routine care in Ireland
Synthesis of Evidence	Systematic review
Outcome Measurement	Quality-adjusted life year (QALY)
Discount Rate	4% on costs and outcomes
Sensitivity Analysis	Probabilistic and deterministic sensitivity analysis
Equity Weighting	Equal weighting applied to the outcome measure

5.2.2 Model Structure

5.2.2.1 Short-Term Decision Tree

In line with the ICER HTA (296), the model structure comprised a short-term decision tree (applied to tisagenlecleucel only) and a long-term partitioned survival model (all treatments). The decision tree was applied to partition patients during the pre-treatment phase of tisagenlecleucel. During this phase, events may occur, which prevent patients proceeding to infusion with tisagenlecleucel. All patients in the tisagenlecleucel arm

entered the decision tree, underwent leukapheresis, and subsequently progressed to one of three outcomes, informed by pooled ELIANA and ENSIGN data (207, 215) (Figure 6):

- Node 1 (N1): proceeded to infusion with tisagenlecleucel (83% of patients (207, 215)).
- Node 2 (N2): did not proceed to infusion with tisagenlecleucel, due to manufacturing failure or adverse event (9% of patients (207, 215)). Based on 'Guidance for Treatment of Relapsed ALL 2019 V10' (83), followed by Children's Health Ireland, Tertiary Hospital Crumlin, when manufacturing failure of CD19 CAR T-cell therapy occurs, patients may be considered for treatment with blinatumomab (with or without alloSCT, herein 'blinatumomab'). In this analysis, manufacturing failure or adverse events did not preclude treatment with an alternative therapy (i.e. blinatumomab). This was modelled by assigning the costs and QALYS associated with blinatumomab to the proportion of patients in the tisagenlecleucel arm who did not proceed to infusion due to manufacturing failure or adverse events.
- Node 3 (N3): did not proceed to infusion with tisagenlecleucel, due to death prior to infusion (8% of patients (207, 215)). These patients did not receive any further active treatment. All patients who progressed to N3 incurred the cost of terminal care.

For patients who did not proceed to tisagenlecleucel infusion (17% in total from those who progressed to N2 and N3), it was assumed that 50% received bridging chemotherapy and 50% received lymphodepleting chemotherapy (333).

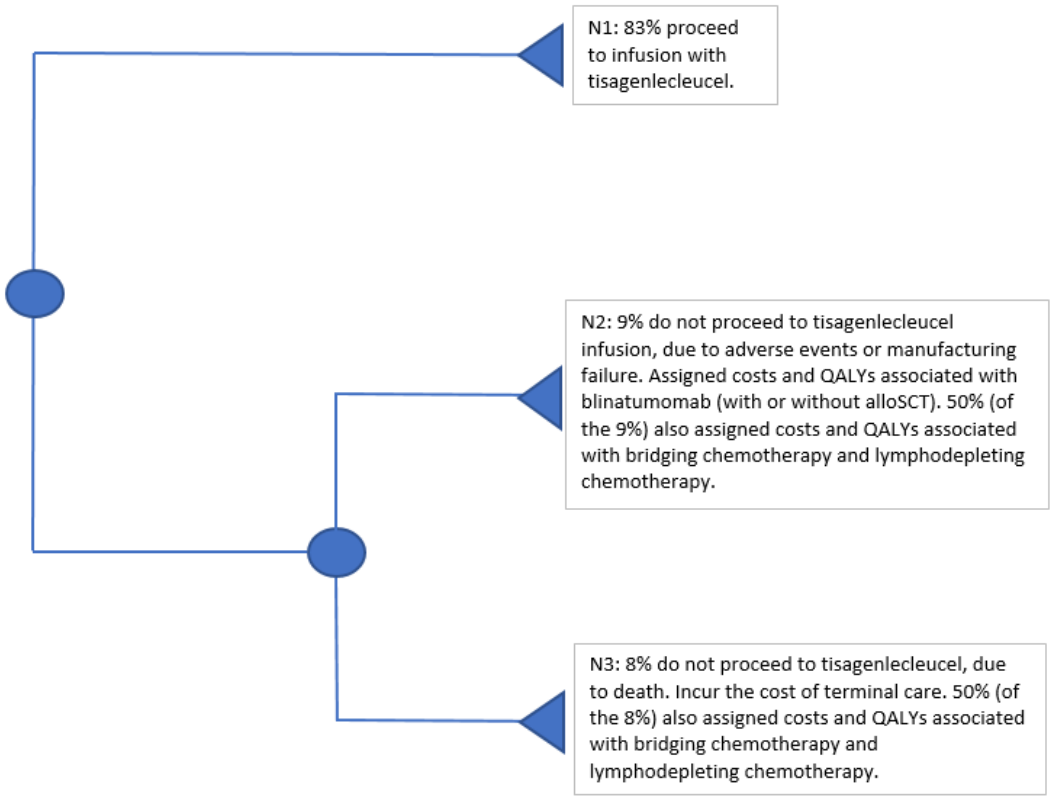

Figure 6 Decision tree depicting outcomes in patients assigned to the tisagenlecleucel arm in cost-utility analysis of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia (207, 215)
AlloSCT: Allogeneic stem cell transplant; **QALY:** Quality-adjusted life year.

5.2.2.2 Partitioned Survival Model

The partitioned survival model, depicted in Figure 7, simulated the progression of patients through three, mutually exclusive health states: event-free survival, progressed disease, and death. The proportion of patients occupying each health state was determined by the area under the curve of the extrapolated EFS and OS curves.

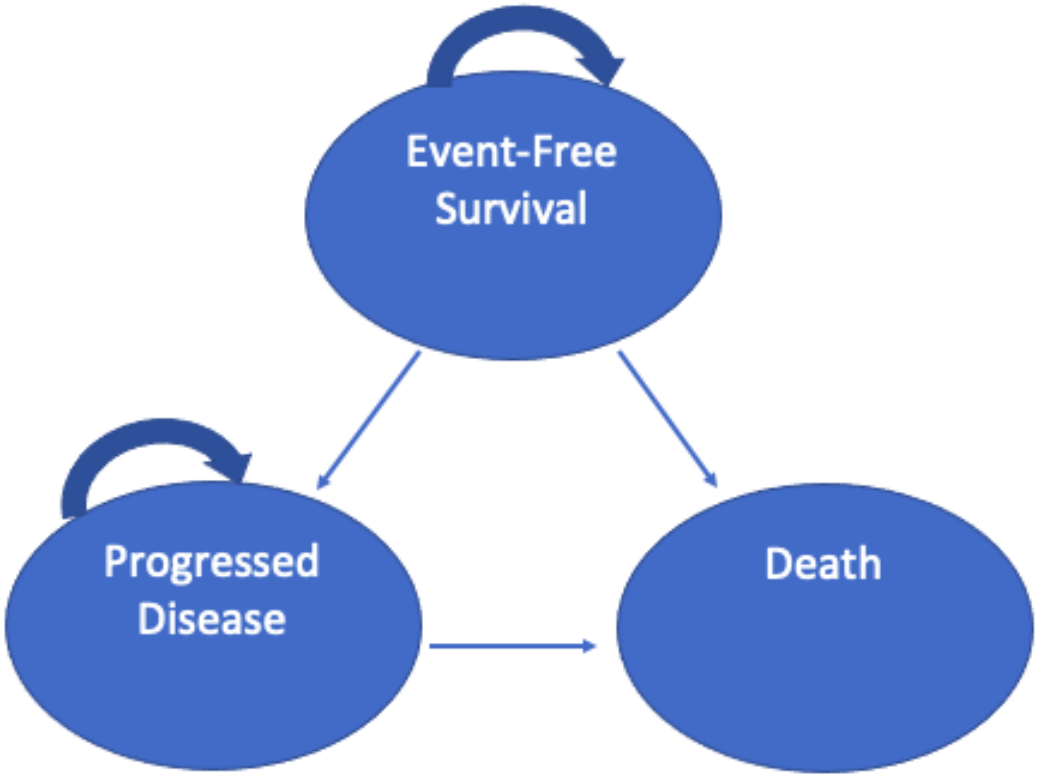

Figure 7 Partitioned survival model depicting outcomes (all treatments) in the cost-utility analysis of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

The area under the extrapolated OS curve provided an estimate of mean survival. Health-state membership of the event-free survival state was provided by the area under the EFS curve. Health-state membership of the death state was estimated by subtracting the area under the OS curve, at each time point, from 1. The proportion of patients in the progressed disease state was derived as the difference between the OS and EFS curve at each time point. Differences between interventions were modelled by using different EFS and OS curves for each treatment (339).

The majority of patients with R/R ALL are expected to relapse within the first 24 to 60 months post-treatment (12, 202, 203); a function of the disease, as opposed to treatment received. Patients who relapse at later time points generally have good outcomes with salvage chemotherapy[4]. It was assumed that patients, who were alive at 61 months (i.e. after 5 years) following treatment with tisagenlecleucel or blinatumomab,

[4] Written correspondence with one consultant haematologist in Ireland.

were long-term survivors. These were subject to age- and sex-matched general population mortality with a standardised mortality ratio (SMR) applied. An SMR was applied to reflect the expectation that mortality risk will never return to age- and sex-matched general population levels, due to late effects and prolonged toxicities. More detail is provided in 5.3.1.3.2.

The partitioned survival approach uses OS and EFS directly from the clinical trials. Time-dependency in underlying risks can be reflected directly. It can also be implemented using reconstructed (Kaplan-Meier) IPD, when the raw IPD from the trial are not publicly available. A partitioned survival model was chosen over a Markov model due to the unavailability of raw IPD for OS and EFS in the public domain. Without the raw IPD from the trials, it was not possible to model transitions, in a Markov model, from the event-free survival state to the death state, as EFS captures both patients who progress or die. This would result in biased estimates of survival, as patients who die prior to documented disease progression are not captured in the analysis. However, an underlying assumption of partitioned survival models is that patients cannot transition to an improved health state. Thus, in the model, patients could not transition from the progressed disease state to the event-free survival state. In clinical practice, such transitions may occur, as patients may respond to subsequent therapy received upon disease progression. The use of a partitioned survival model was a data-driven choice. The implications of this, however, is that the modelled clinical pathway is more simplistic than expected in clinical practice.

A cycle length of one month (30.4 days) (considered sufficient to capture relevant transitions) was applied (223, 239). A half-cycle correction was applied to mitigate against over- or under-predicting state occupancy. The time horizon was 88 years, representing a lifetime horizon (223). In line with National Economic Evaluation Guidelines, a discount rate of 4% was applied to costs and outcomes after the first year of the model (11). Based on recommended adaptations to the HTA of potentially 'curative' therapies, presented in 1.10.1, a number of scenario analyses examining the impact of alternative time horizons and discount rates were explored. In the PSA, the proportion of patients proceeding (from the decision tree) to infusion, comparator therapy, and death were varied according to the Dirichlet distribution. This distribution is

appropriate when varying observed counts of polychotomous events. It is a multinomial equivalent of the beta distribution (constrained between 0 and 1) with one parameter per category (34).

5.2.3 Population

The population was paediatric and young adult patients (up to 25 years of age) with B-cell ALL that is refractory, in relapse post-transplant, or in second or later relapse. This is in line with the EMA licensed indication of tisagenlecleucel (3). Population demographics were based on ELIANA and ENSIGN. Starting age was 12 years, 44% were female, body surface area was 1.32m^2, and weight was 42.2kg (143, 207, 239, 250). No population subgroups were considered due to paucity of data.

5.2.4 Intervention

The intervention was tisagenlecleucel, administered as an IV infusion, at a dose of 0.2 to 5.0x10^6 CAR-positive viable T-cells per kg (body weight) for patients weighing 50kg or less, and 0.1 to 2.5x10^8 CAR-positive viable T-cells for patients greater than 50kg (non-weight based). Tisagenlecleucel was modelled as a single-dose intervention, in line with the ELIANA and ENSIGN protocols (143, 207). However, it is not explicitly stated within the SPC that patients cannot be retreated (3).

5.2.5 Comparator

The comparator was blinatumomab. Blinatumomab is indicated as monotherapy for the treatment of paediatric patients (aged 1 year or older) with Philadelphia chromosome-negative CD19 positive B-cell ALL that is refractory, in relapse after receiving at least two prior therapies, or in relapse following alloSCT (86). Blinatumomab may be administered to patients with the intent to receive alloSCT. It was assumed that 49% of patients received alloSCT following blinatumomab, in line with clinical opinion (n=5, Chapter 3). The dosing regimen of blinatumomab is presented in Table 16, sourced from the NCCP Chemotherapy Regimen (340). As described in 5.3.3.4, it was assumed that 50% of patients received dosing based on body surface area (i.e. dose applied to patients less than 45kg) and 50% received the fixed-dosing regimen (i.e. dose applied to patients weighing 45kg or greater). As per the SPC of blinatumomab, patients who achieve CR after two cycles of treatment may receive up to three additional cycles as consolidation

treatment (86). In Irish clinical practice, patients typically receive one to two cycles of blinatumomab[5]. It was assumed that patients receive up to two cycles of blinatumomab to reflect this. The use of 'blinatumomab' in the following sections pertains to blinatumomab with (49%) or without (51%) alloSCT.

Table 16 Dosing regimen of blinatumomab, as per the National Cancer Control Programme Chemotherapy Regimen (340)

Patient Weight	Cycle 1			Subsequent Cycles	
	Days 1-7	Days 8-28	Days 29-42	Days 1-28	Days 29-42
<45kg (body surface area-based dose)	5 mcg/m^2/day (maximum 9 mcg/day)	15 mcg/m^2/day (maximum 28 mcg/day)	14 day treatment-free interval	15 mcg/m^2/day (maximum 28 mcg/day)	14 day treatment-free interval
≥45kg (fixed-dose)	9 mcg/day	28 mcg/day		28 mcg/day	

5.2.6 Perspective

The perspective was that of the healthcare payer in Ireland, the HSE, in line with National Economic Evaluation Guidelines (11). Direct medical costs borne by the HSE were included.

5.2.7 Outcomes

The primary outcomes were deterministic and probabilistic ICERs, expressed in terms of cost per QALY. Further detail is provided in 5.4.

5.3 Model Inputs

5.3.1 Efficacy Inputs

5.3.1.1 Reconstruction of Individual Patient-Level Data

As described in Chapter 2, IPD from published Kaplan-Meier curves of OS and EFS were reconstructed by digitising the published curves and applying the algorithm by Guyot et al. using R® (211, 212). This facilitated generation of relative efficacy estimates and extrapolation of outcomes, to the time horizon of the model, as described below. The trials included were ELIANA (143, 215) and ENSIGN (207) (tisagenlecleucel), and NCT01471782 (87, 216) (blinatumomab).

[5] Written correspondence with one consultant haematologist in Ireland.

5.3.1.2 Extrapolation of Survival

Treatment effectiveness in the model was based on the effect on OS and EFS. The time horizons of the trials were shorter than that of the model, extrapolation of the data was therefore, required. Extrapolation of outcomes, using reconstructed OS and EFS IPD, was conducted in line with the NICE Decision Support Unit Technical Support Document 14 (herein 'NICE DSU 14') (341). Due to the naïve method of comparison, treatment arms were modelled independently and an assessment of proportional hazards was not conducted. The use of HRs was considered to add additional uncertainty to the analysis. Standard parametric ('parametric'), flexible cubic spline ('spline'), and mixture cure extrapolation models were explored.

The 'Survival', 'Flexsurv' and 'Flexsurvcure' packages in R® (212) were used to derive coefficients for each model fit to the reconstructed IPD for OS and EFS. Several functions were used:

- 'surv': create a survival object, which is used as a response variable in the parametric model formulas.
- 'survfit': compute Kaplan-Meier survival estimates.
- 'flexsurvreg': fit parametric models.
- 'flexsurvspline': fit spline models.
- 'flexsurvcure': fit mixture cure models.

These coefficients were subsequently used to programme different survival models in Microsoft Excel® and populate the health states to the full time horizon of the model. Extrapolations conducted in Microsoft Excel® were cross-checked with extrapolation output from R®, for quality control purposes. In line with the NICE DSU 14, a combination of visual assessment of fit to the Kaplan-Meier data, consideration of the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) statistics, statistical plausibility, clinical plausibility, and comparison with external data were used to guide the choice of most appropriate model (341). OS curves were adjusted to account for general population mortality using Irish data from the Central Statistics Office (CSO) (342). Uncertainty in OS and EFS was captured using Cholesky decomposition to correlate the survival parameters, drawing from the variance-covariance matrix (34).

5.3.1.2.1 Overall Survival Extrapolation

5.3.1.2.1.1 Parametric Survival Extrapolation

A series of parametric models were fit to the OS data of the individual treatment arms (341). These were Gompertz, exponential, Weibull, log-logistic, log-normal, and generalised gamma. Within survival modelling, it is common practice to fit the same type of model to the treatment arms being compared. Fitting different types of models to the treatment arms being compared implies that the underlying hazard differs between arms; this requires substantial justification (341).

The AIC and BIC statistics of each parametric model fitted to the pooled ELIANA and ENSIGN (tisagenlecleucel), and NCT01471782 (blinatumomab) data are presented in Table 17. Lower AIC and BIC statistics indicate better statistical fit relative to other models. When selecting a model based on AIC and BIC statistics, a difference of less than 2 (between lowest and second lowest values) is generally considered negligible and either model could be supported, while a difference of greater than 10 (between lowest and second lowest values) supports selection of the model with the lowest value (343). Based on these statistics, there was no parametric model, which clearly fitted the data better than others.

Tisagenlecleucel: The exponential model had the lowest AIC and BIC statistics for the pooled ELIANA and ENSIGN data; however, the difference between AIC statistics was small. More variation was observed in the BIC statistics. None of the parametric models provided a particularly good visual fit to the pooled ELIANA and ENSIGN data; presented in Appendix D (Figure A1). In terms of clinical plausibility, the log-normal and log-logistic models were most closely aligned with the judgements derived from the expert elicitation (Chapter 3). Based on these judgements, the 60-month OS of patients treated with tisagenlecleucel is expected to be 33% (95% CrI 8.7 to 56.9). The 60-month OS predicted by the log-normal and log-logistics models was 34% and 29%, respectively. Both models appeared to overestimate OS towards the end of the observed Kaplan-Meier data (Figure 8); however, these Kaplan-Meier data were based on 4 patients left at risk from month 39 onwards. The 60-month OS predicted by the other parametric models ranged from 0% (generalised gamma) to 22% (exponential). When compared to

the published update of ELIANA (median follow up 24 months), all parametric models underestimated the 18-month OS; 70% versus 63% (predicted by all models) (219).

The hazard function of the exponential model is constant over time. The innovative mechanism of action of tisagenlecleucel is not expected to be aligned with a hazard function that behaves in this way. The hazard functions of the log-normal and log-logistic models increase initially and then decrease (the log-logistic may also monotonically decrease). They often result in long tails in the survival function (341). It could be hypothesised that patients treated with tisagenlecleucel initially have an increasing hazard function post-infusion, due to the adverse event profile in the immediate period post-infusion. The long tails in the survival function may reflect a cohort of patients who are considered potential long-term survivors.

Blinatumomab: For NCT01471782, the log-normal model had the lowest AIC and BIC statistics, followed by the log-logistic model. However, the difference between any of the parametric models was small. Both the log-normal and log-logistic models presented a reasonable visual fit to the data (Appendix D, Figure A2). The Weibull and exponential models provided the worst visual fit. Due to the small number of patients left at risk from month 14 onwards (n=5), estimates based on model fit to the tail of the NCT01471782 Kaplan-Meier data are unreliable. In terms of clinical plausibility, the 60-month OS predicted by the parametric models ranged from 1% (exponential) to 15% (Gompertz). Based on the elicited judgements, the Weibull (2%) and exponential (1%) models underestimated the 60-month OS.

Table 17 AIC and BIC statistics of parametric models used in the extrapolation of overall survival in cost-utility model of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia *

	Overall Survival			
	ELIANA_ENSIGN Pooled (tisagenlecleucel)		**NCT01471782 (blinatumomab±alloSCT)**	
	AIC	**BIC**	**AIC**	**BIC**
Gompertz	451.8	457.7	342.2	346.7
Exponential	**449.8**	**452.8**	345.3	347.5
Weibull	451.8	457.7	346.2	350.7
Log-Logistic	452.3	458.1	340.9	345.4
Log-Normal	453.7	459.6	**339.3**	**343.8**
Generalised Gamma	453.7	462.5	340.4	347.1

AIC: Akaike information criteria; **AlloSCT:** Allogeneic stem cell transplant; **BIC:** Bayesian information criteria.

*****Lowest AIC and BIC statistics for each data set are highlighted in **Bold.**

5.3.1.2.1.2 Spline Model Survival Extrapolation

Spline models were fit independently to each treatment arm. Knots were distributed as equally-spaced quantiles of the log uncensored survival times, as per the default settings of 'flexsurvspline'. A series of one-, two-, and three-knot spline models were considered. Additional knots were not considered as visual inspection of the plotted models, AIC and BIC statistics, and extrapolation output indicated that there was negligible difference between the three- and four-knot spline models. The additional complexity of additional knots was not warranted. For transformation of the survival function to a linear prediction scale, three options are available: transformation of the Weibull function to the log cumulative hazard scale ('hazard'); transformation of the log-logistic function to the log cumulative odds scale ('odds'), and transformation of the log-normal function to the probit scale ('normal'). All three were considered.

Tisagenlecleucel: The AIC statistics across all scales and number of knots, presented in Table 18, were closely aligned for the pooled ELIANA and ENSIGN data. More variation was observed in the BIC statistics. Based on BIC statistics, the one-knot spline models (across all scales) provided the best statistical fit to the pooled ELIANA and ENSIGN data. Upon visual assessment, all spline models appeared to overestimate OS of tisagenlecleucel towards the end of follow up. The one-knot (hazard) spline model was most closely aligned with the observed Kaplan-Meier OS data towards the end of follow up (Figure 8). As highlighted previously, ELIANA and ENSIGN data were based on considerably small number of patients left at risk towards the end of follow up. The predicted 60-month OS of tisagenlecleucel ranged from 23% (one-knot hazard) to 28% (one-knot odds, two-knot odds, two-knot normal), which is lower than the 33% derived in the expert elicitation.

Blinatumomab: The AIC statistics across all scales and number of knots were closely aligned for the NCT01471782 data. Slightly more variation was observed in the BIC statistics. The one-knot (normal) spline model had the lowest AIC and BIC statistics; however, the difference across the one-knot spline models (all scales) was negligible. All spline models had a good visual fit to the NCT01471782 data. However, good visual fit to the observed data does not imply that a model will result in appropriate extrapolations. The 60-month OS predictions were considered reasonable, ranging from 6% (one-knot

hazard) to 11% (one-knot odds). Based on the expert elicitation, the weighted average 60-month OS of blinatumomab, based on those who receive alloSCT and those who do not, was 9%. Visual fit of the models to the Kaplan-Meier data are presented in Appendix D.

Table 18 AIC and BIC statistics of spline models used in extrapolation of overall survival in cost-utility model of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia*

	Overall Survival			
	ELIANA_ENSIGN pooled (tisagenlecleucel)		NCT01471782 (blinatumomab±alloSCT)	
	AIC	BIC	AIC	BIC
1 Knot Spline (Hazard)	453.7	**462.5**	340.7	347.4
1 Knot Spline (Odds)	454.9	462.7	340.6	347.3
1 Knot Spline (Normal)	**453.2**	463.0	**340.4**	**347.1**
2 Knot Spline (Hazard)	455.6	467.4	342.7	351.7
2 Knot Spline (Odds)	456.2	468.0	342.5	351.5
2 Knot Spline (Normal)	455.9	467.6	342.4	351.4
3 Knot Spline (Hazard)	457.7	472.3	344.4	355.6
3 Knot Spline (Odds)	458.2	472.9	344.0	355.2
3 Knot Spline (Normal)	457.9	472.5	343.9	355.1

AIC: Akaike information criteria; **AlloSCT:** Allogeneic stem cell transplant; **BIC:** Bayesian information criteria.
*Lowest AIC and BIC statistics for each data set are highlighted in **Bold**.

5.3.1.2.2 Event-Free Survival Extrapolation

5.3.1.2.2.1 Parametric Survival Extrapolation

Tisagenlecleucel: EFS data for tisagenlecleucel were based on ELIANA only, as the Kaplan-Meier EFS data were not publicly available for ENSIGN. The AIC and BIC statistics of the Gompertz and exponential models, presented in Table 19, were notably higher than those of the other four parametric models. All parametric models provided a poor visual fit to the data (presented in Appendix D, Figure A5). As expected, the log-normal and log-logistic models exhibited long tails in the extrapolated region. When compared to the OS output provided by the parametric models, EFS was overestimated by four of the parametric models (i.e. the EFS extrapolated curve was above that of OS); the exponential and generalised gamma models were the exception to this. The higher predictions for EFS are likely a function of the less 'stepped' nature of the published EFS Kaplan-Meier curve, when compared to the OS Kaplan-Meier curve. The same issue arose when the ELIANA data alone were used to model OS. As a result, it is difficult to determine the most appropriate parametric model. The 60-month EFS predicted by the parametric models ranged from 5% (exponential) to 43% (log-normal).

Blinatumomab: EFS data were not reported for blinatumomab. As such, EFS was estimated from the OS curve of blinatumomab by assuming that the cumulative hazard function for EFS was proportional to the cumulative hazard function for OS. The ratio between EFS and OS was estimated based on Kuhlen et al. (229). Kuhlen et al. examined OS in patients previously treated according to ALL-SCT-BFM 2003 and ALL-SCT-BFM international 2007 protocols, and experienced subsequent relapse (n=242). Patients were treated with either salvage chemotherapy without alloSCT (48%), salvage chemotherapy with alloSCT (26%), or palliative care (25%) (229).

First, the natural log of OS probability was divided by the natural log of EFS probability at yearly intervals. The overall cumulative HR between OS and EFS was then calculated as the average of the cumulative HR at yearly intervals. This overall cumulative HR was applied to the OS data of NCT01471782 to generate EFS. The predicted EFS for blinatumomab was therefore, contingent upon the model applied to the OS data and no separate model fitting to EFS was required. Given the expectation that patients who are alive after 60 months are considered long-term survivors (described in 5.3.1.3.1), the proportional relationship between EFS and OS is not expected to continue indefinitely. Long-term survivors are expected to be free of progressed disease. As such, after month 60, the cumulative survival probabilities for EFS were assumed to flatten up to the point at which EFS met OS.

The lack of direct evidence for EFS of blinatumomab adds considerable uncertainty to the analysis. The methodology used to derive EFS has been previously justified in the literature, citing the high correlation between EFS and OS (12, 250). It is noted, however, that limited data exist, which validate EFS as a surrogate for OS in paediatric and young adult patients with R/R ALL. The uncertainty is further compounded by differences in populations between NCT01471782 and Kuhlen et al., particularly considering that Kuhlen et al. included patients who received palliative care (229). Despite this, the HR derived from Kuhlen et al. (0.88) was closely aligned with the HR of 0.83, which has been used in the literature and accepted by national HTA agencies (223, 240, 250). The HR of 0.83 was derived from an international RCT examining outcomes in patients (aged 1 to 18 years) with ALL in first relapse (344). Kuhlen et al. was considered a more appropriate source here, as it is more representative of the line of therapy at which blinatumomab is

used. Of note, Kuhlen et al. defined EFS differently to ELIANA. In Kuhlen et al., EFS is defined from the time of first relapse after alloSCT to the time of second relapse after alloSCT, or death due to any cause (229). ELIANA defined EFS from the time of tisagenlecleucel infusion (215).

Table 19 AIC and BIC statistics of parametric models used in extrapolation of event-free survival in cost-utility model of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia†

Event-Free Survival*		
ELIANA (tisagenlecleucel)		
	AIC	**BIC**
Gompertz	217.3	222.0
Exponential	217.0	219.3
Weibull	204.2	208.9
Log-Logistic	206.2	210.8
Log-Normal	208.0	212.6
Generalised Gamma	**200.8**	**208.7**

AIC: Akaike information criteria; **BIC:** Bayesian information criteria.
*Event-free survival of blinatumomab derived by assuming that the cumulative hazard function for event-free survival was proportional to the cumulative hazard function for overall survival. The ratio between event-free survival and overall survival (0.88) was estimated based on Kuhlen et al. (229).
†Lowest AIC and BIC statistics are highlighted in **Bold.**

5.3.1.2.2.2 Spline Model Survival Extrapolation

Tisagenlecleucel: For EFS of tisagenlecleucel, the three-knot spline models (across all scales) had the lowest AIC and BIC statistics. However, upon visual inspection, considerable overfitting to the data was observed. These models were not considered any further. Additionally, the probit ('normal') scale was not considered due to a 'non-finite finite-difference value' error when running the code in R®. This was likely due to the high degree of censoring and low number of events towards the end of the follow-up period in ELIANA. AIC and BIC statistics of the remaining models, presented in Table 20, were very closely aligned across all scales and number of knots. Visual inspection indicated that the one-knot spline models (across all scales) tended towards an improved fit over the two-knot spline models (across all scales). The two-knot spline models appeared to underestimate EFS towards the end of the observed follow-up period. However, as described, caution should be exercised when examining model fit towards the end of the observed Kaplan-Meier data. The issue described in 5.3.1.2.2.1, whereby EFS predicted by the parametric models exceeded predicted OS, was also an issue with the one-knot spline models (when compared to the OS extrapolation of the one-, two-,

and three-knot spline models across all scales). The 60-month EFS predicted by the spline models ranged from 4% (two-knot hazard) to 25% (one-knot odds).

Blinatumomab: As described in 5.3.1.2.2.1, the EFS predicted for blinatumomab was contingent upon the model applied to the OS data, and no separate model fitting to EFS was required.

Table 20 AIC and BIC statistics of spline models used in extrapolation of event-free survival in cost-utility model of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia†

Event-Free Survival*		
ELIANA (tisagenlecleucel)		
	AIC	**BIC**
1Knot Spline (Hazard)	202.6	**209.5**
1 Knot Spline (Odds)	203.6	210.2
2 Knot Spline (Hazard)	201.6	210.9
2 Knot Spline (Odds)	**200.6**	209.8

AIC: Akaike information criteria; **BIC:** Bayesian information criteria.
*Event-free survival of blinatumomab derived by assuming that the cumulative hazard function for event-free survival was proportional to the cumulative hazard function for overall survival. The ratio between event-free survival and overall survival (0.88) was estimated based on Kuhlen et al. (229).
†Lowest AIC and BIC statistics are highlighted in **Bold.**

5.3.1.2.3 Mixture Cure Model Extrapolation

The mixture cure models examined were Gompertz, exponential, Weibull, log-logistic, log-normal, and generalised gamma. The output from the 'flexsurvcure' function consists of a cure fraction parameter (i.e. an estimate of the proportion of patients 'cured') and a parametric distribution representing excess mortality for the 'not cured' cohort. The estimated cure fraction differs according to the type of mixture cure model selected. The estimated cure fraction from each mixture cure model and relevant parametric survival data and coefficients were exported to Microsoft Excel®. Here, the expected survival of the entire cohort ('cured' and 'not cured') was modelled, according to the formula:

$$S(t) = S^*(t)\big(\pi + (1 - \pi)S_u(t)\big)$$

where $S^*(t)$ represents age- and sex-matched general population survival, π represents the cure fraction, and $S_u(t)$ represents the survival function of the 'not cured' cohort (178). General population mortality data for the Irish population were obtained from the CSO (342). These data were used to model survival of the 'cured' cohort.

The fitting of mixture cure models relies on a number of assumptions, mainly that the data are sufficiently mature and robust to reliably estimate a cure fraction. A mixture

cure model is only appropriate in cases where a true cure fraction exists. The NICE Decision Support Unit Technical Support Document 21 (herein 'NICE DSU 21') indicates that in order to reliably estimate the cure fraction, sufficient numbers at risk are required in the tail of the distribution (178). This is a particular concern with ELIANA, whereby 8 patients were at risk at 18 months and 0 were at risk at 22 months (215). These data are highly censored. The same concern arises with NCT01471782, with 14 patients at risk at 22 months, reducing to 6 at 24 months (216).

The estimated cure fractions of the pooled ELIANA and ENSIGN data ranged from 0% to 1%. When just the ELIANA data were assessed, the cure fraction ranged from 0% to 5%. When the cure fraction is 0%, the survival of the overall population is equivalent to that predicted by the associated parametric model. These estimates are in contrast to the cure fraction judged in the expert elicitation (n=2), 44% (95% CrI 27.6 to 61.1) (Chapter 3). The low cure fractions, estimated by the mixture cure models, are likely a reflection of the highly 'stepped' nature of the Kaplan-Meier OS data towards the end of the follow-up period. A clear plateau does not exist in the Kaplan-Meier curves of ELIANA and ENSIGN; the possibility of a non-zero mortality rate cannot be ruled out. The follow-up period of ELIANA and ENSIGN is not sufficiently long to determine whether a true 'cure' fraction exists.

The cure fractions of the NCT01471782 data ranged from 11% to 23%. Concerns exist, however, regarding the duration of follow up and the high degree of censoring in NCT01471782 (216). The use of mixture cure models, to extrapolate OS and EFS of ELIANA and ENSIGN, and NCT01471782, was therefore, deemed inappropriate and no further consideration was given to this method.

5.3.1.2.4 Other Flexible Extrapolation Methods for Survival Analysis

A number of other flexible modelling methods, to capture complex hazard functions, are outlined in the NICE DSU 21 (178). Landmark models assume that response represents an appropriate and strong surrogate for survival. Based on this, a 'landmark' time point is chosen, at which point patients are split into groups according to their response category and separate survival models are fitted to each group. This approach was not considered

here, due to the lack of required Kaplan-Meier data, low patient numbers, and concerns regarding choice of an appropriate 'landmark' time point.

Piecewise models may be appropriate when parametric models do not appear to have a good fit to the data. This approach was not adopted here due to concerns regarding the reduction in sample size as a result of segmentation of the survival function.

5.3.1.3 Summary

5.3.1.3.1 Overall Survival

Tisagenlecleucel: In summary, the log-normal and log-logistic were deemed the most appropriate parametric models, to extrapolate the pooled ELIANA and ENSIGN data. The one-knot spline models (across all scales) were considered the most appropriate spline models, based on marginally more favourable AIC and BIC statistics and visual fit (when compared to the two- and three-knot spline models). The predicted OS (up to month 60) and Kaplan-Meier curve of the pooled ELIANA and ENSIGN data are presented in Figure 8. Of note, the one-knot (normal) spline model is not presented, as the difference in OS predicted by this model and the one-knot (odds) spline model was negligible.

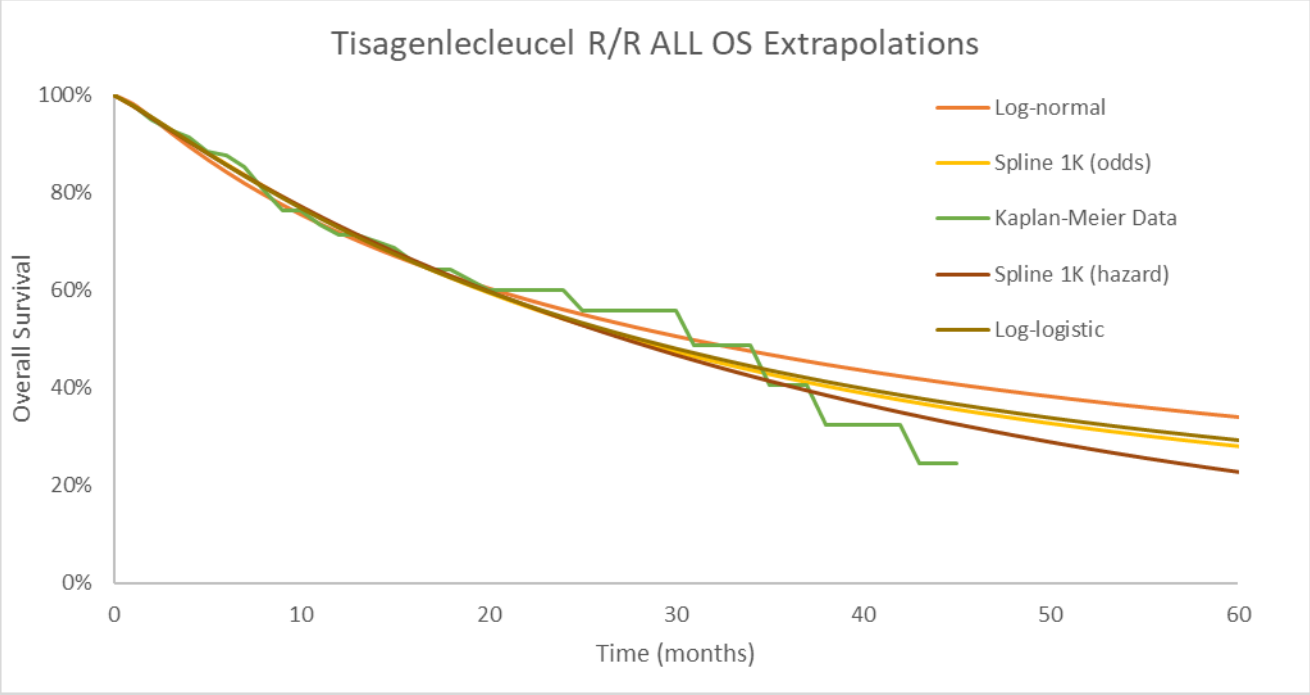

Figure 8 Tisagenlecleucel (R/R ALL) overall survival extrapolation predictions of 'best fitting' parametric and spline models
ALL: Acute lymphoblastic leukaemia; **OS:** Overall survival; **R/R**: Relapsed/refractory; **Spline 1K (odds):** Spline one-knot (odds) model; **Spline 1K (hazard):** Spline one-knot (hazard) model.

Blinatumomab: The log-normal and log-logistic models were considered the most appropriate parametric models to extrapolate the NCT01471782 data. The one-knot spline models (across all scales) were considered the most appropriate of the spline models. All models exhibited long tails in the OS extrapolations (Figure 9).

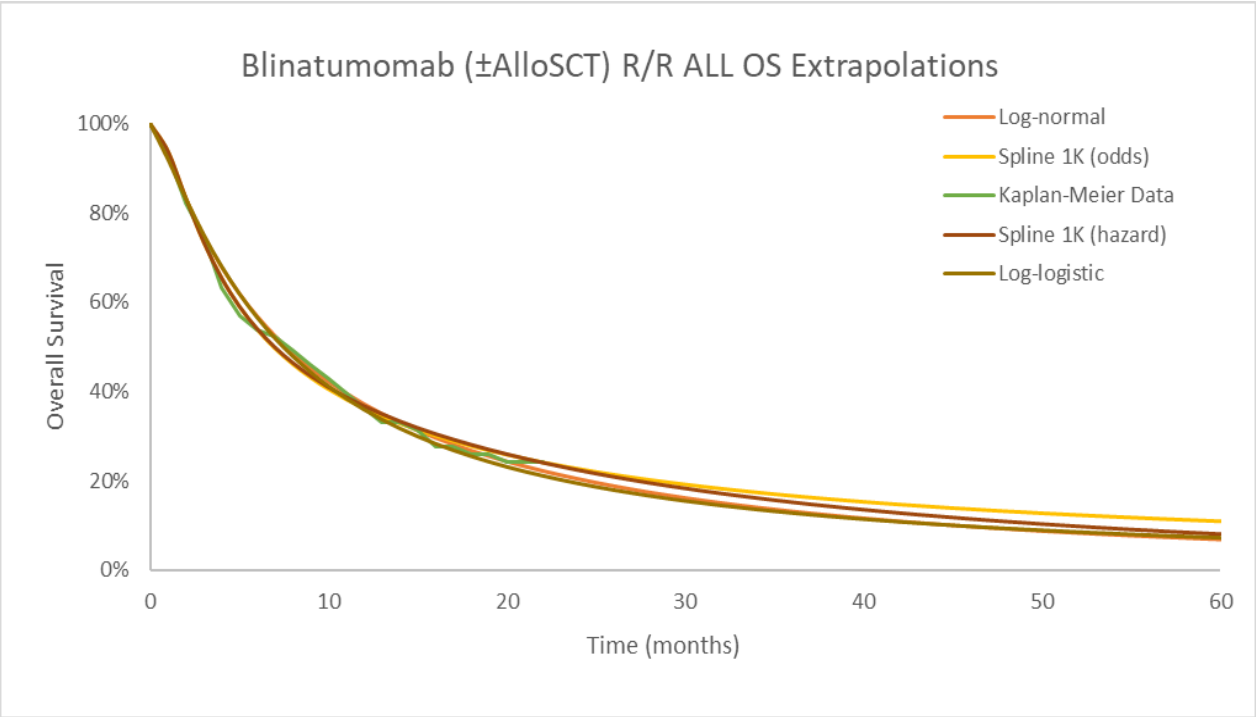

Figure 9 Blinatumomab (R/R ALL) overall survival extrapolation predictions of 'best fitting' parametric and spline models
ALL: Acute lymphoblastic leukaemia; **AlloSCT:** Allogeneic stem cell transplant; **OS:** Overall survival; **R/R**: Relapsed/refractory. **Spline 1K (odds):** Spline one-knot (odds) model; **Spline 1K (hazard):** Spline one-knot (hazard) model.

The tails of the tisagenlecleucel and blinatumomab OS data are heavily influenced by the low numbers of patients left at risk towards the end of follow up. OS extrapolations based on such limited data are unsound. The log cumulative hazard plots of the pooled ELIANA and ENSIGN (tisagenlecleucel), and NCT01471782 (blinatumomab) studies were constructed to examine the hazards observed in these studies. These are presented in Appendix D. In each case, the log cumulative hazard plots were not approximately straight lines, indicating that spline model extrapolation may be more appropriate than parametric extrapolation (341). For the base case, the one-knot (odds) spline model was chosen to extrapolate the pooled ELIANA and ENSIGN, and NCT01471782 OS data. Although this model appeared to overestimate OS of tisagenlecleucel, the model prediction at 60 months was slightly lower than those judged by the experts consulted in

the expert elicitation; 28% and 33%, respectively. The one-knot (odds) spline model provided a reasonable fit to the blinatumomab data, and the 60-month OS prediction was aligned with those judged by the experts.

The uncertainty associated with the extrapolation output must be acknowledged. As such, extensive sensitivity analyses were conducted, examining the impact of alternative models on OS and cost-effectiveness estimates.

5.3.1.3.2 Adjustment for Excess Risk of Mortality

As described in 5.2.2.2, it was assumed that patients who were alive after 60 months were considered to be long-term survivors and subject to age- and sex-matched general population mortality. This assumption has been employed in the literature (12, 250). Although patients with R/R ALL who are alive after 60 months are expected to have long-term survival, it has been reported that the relative risk of mortality in later life among childhood and adolescent cancer survivors is higher than that expected in the general population (293, 345). This was also highlighted by three of the five experts who took part in the expert elicitation; although the excess risk of mortality (compared to age- and sex-matched general population) of patients treated with tisagenlecleucel is expected to reduce over time, mortality risk is not expected to return to age- and sex-matched general population levels, due to late effects and prolonged toxicities. To account for this excess mortality, an adjustment factor (SMR) was incorporated into the model and applied to the age- and sex-matched general population mortality. This SMR (15.5) was derived from Fidler et al. (292) and assumed to remain constant over the time horizon of the model. Although this assumption is not aligned with the experts' judgement that excess mortality will decrease over time, there is a paucity of evidence to indicate the extent of reduction over time.

Fidler et al. examined mortality in paediatric and adolescent patients (less than 15 years) diagnosed (between 1940 and 2006) with ALL and survived 5 years post-diagnosis (n=9,493; obtained from the British Childhood Cancer Survivor Study database (292)). The median follow up was 15.2 years (range: 0.0 to 68.7). An all-cause SMR, as well as cause-specific (e.g. subsequent primary neoplasm, non-neoplastic causes) SMRs were

calculated. Fidler et al. was chosen over other studies employed in cost-utility analyses of treatments for R/R ALL (345-347), due to the large sample size, ALL-specific population, and UK-based population (who were deemed more generalisable to the Irish population than a US cohort). The duration over which the mortality of patients was examined is also more recent than other studies (293). A limitation of Fidler et al. is that it does not relate to patients with R/R disease. The SMR for patients who have experienced R/R disease could be higher due to the effects of multiple lines of chemotherapy.

5.3.1.3.3 Event-Free Survival

As described, the fitting of extrapolation models to the EFS data was only required for tisagenlecleucel.

In summary, the generalised gamma was considered the most appropriate parametric model for tisagenlecleucel. The two-knot spline models (across both scales) were considered most appropriate of the spline models. The EFS output of the generalised gamma model and two-knot spline models (across both scales) were closely aligned for the ELIANA data. The generalised gamma model was chosen to extrapolate the EFS data, based on favourable AIC and BIC statistics. The 60-month EFS predicted by the generalised gamma model was 11%.

Patients who are alive after 60 months are expected to be free of progressed disease (i.e. long-term survivors), as described in 5.2.2.2. To account for this, the cumulative survival probabilities for EFS, after 60 months, were assumed to flatten up to the point at which EFS met OS for all treatment arms.

5.3.2 Health-Related Quality of Life Inputs

HRQOL inputs for use in the cost-utility model were derived through SLR, presented in Chapter 4 (Table 14).

Utility values were applied according to health-state occupancy. It was assumed that all patients alive after 60 months had HRQOL equivalent to that of the event-free survival state. Disutility values were included in the tisagenlecleucel arm to account for disutility

associated with apheresis, bridging chemotherapy and lymphodepleting chemotherapy. The duration of disutility associated with these procedures was assumed to last for the duration of the procedure. Disutility due to ICU admission, febrile neutropenia and pancytopenia was accounted for. These were chosen as they are considered to have a considerable impact on HRQOL. The duration of febrile neutropenia disutility was informed by clinical opinion[6], which indicated that patients experience this event for between five and eight days. The duration of pancytopenia (182.4 days, 6 months) was an assumption.

Patients in the tisagenlecleucel arm, who did not proceed to infusion due to manufacturing failure or adverse event (9%), were assigned utility associated with the blinatumomab arm. It was also assumed that, of patients in the tisagenlecleucel arm who did not proceed to tisagenlecleucel infusion (due to manufacturing failure, adverse event, or death; 17%), 50% were assigned disutility associated with bridging chemotherapy and 50% were assigned disutility associated with lymphodepleting chemotherapy. All patients in the tisagenlecleucel arm were assigned disutility associated with apheresis.

The assumptions employed in the model are presented in Table 21. Assumptions were informed by ELIANA, ENSIGN (143, 207, 215, 219), NCT01471782 (87, 216), and the NICE and CADTH HTA appraisals (223, 240).

[6] Written correspondence with one consultant haematologist in Ireland.

Table 21 Utility values used in the cost-utility model of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

Parameter	Value (SD)	Source	Duration (days)	Proportion
Event-Free Survival	0.80 (0.23)	EQ-5D-3L collected in ELIANA with the UK valuation set applied, identified through HTA appraisal (250)	Duration in health state	100% of patients in state
Progressed Disease	0.63 (0.36)			
All Patients Alive After 60 Months	0.80 (0.23)	Assumption (equal to event-free survival state) based on HTA appraisals (223, 250, 315), Hettle et al. (12), ICER HTA (296)	Duration of survival	100% of patients
Apheresis	-0.202 (0.006)	Kwon et al. (308), identified through Thielen et al. (295)	0.5	100% of patients in tisagenlecleucel arm
Bridging Chemotherapy			21	88% of patients who received tisagenlecleucel (207, 215), and 50% of those who did not proceed to infusion*
Lymphodepleting Chemotherapy			4	95% of patients who received tisagenlecleucel (215), and 50% of those who did not proceed to infusion*
CRS ICU Admission Tisagenlecleucel	-0.80 (0.23)	Assumption (utility of 0) based on HTA appraisals (223, 240, 250), Hettle et al. (12), ICER HTA (296)	8 (143)	47% of patients who received tisagenlecleucel (143, 219)
CRS ICU Admission Blinatumomab			5 (87)	5.7% of patients who received blinatumomab (87)
Non-CRS ICU Admission Tisagenlecleucel		Assumption (utility of 0) based on HTA appraisal (240)	1.78 (240)	90% of patients who received tisagenlecleucel (240)
Febrile Neutropenia	-0.15 (0.04†)	Lloyd et al. (329), identified through HTA appraisals (331, 332)‡	7§	36% of patients who received tisagenlecleucel (207, 215) 17% of patients who received blinatumomab (216)

Pancytopenia			182.4|	3% of patients who received tisagenlecleucel (143)
AlloSCT (first 3 months post-alloSCT)	-0.20 (0.05†)	Forsythe et al. (309), identified through Thielen et al. (295)	91.2 (223)	49% of patients who received blinatumomab¶
AlloSCT (4 to 12 months post-alloSCT)	-0.13 (0.16)	Felder-Puig et el. (324), identified through HTA appraisal (223)	273.6 (223)	
Age-Related Disutility	Ara and Brazier (328)	HTA appraisal (333)	Time horizon of model	100% of patients

AlloSCT: Allogeneic stem cell transplant; **CRS:** Cytokine release syndrome; **HTA:** Health technology assessment; **ICU:** Intensive care unit; **SD:** Standard deviation; **UK:** United Kingdom.
*83% of patients in the tisagenlecleucel arm proceeded to infusion.
†Assumed to be 25% of mean point estimate.
‡Identified in systematic literature review of utility data in R/R diffuse large B-cell lymphoma (Chapter 8).
§Based on clinical opinion.
|Assumption.
¶Based on expert elicitation (Chapter 3).

In the PSA, utility values were varied according to the beta distribution; disutility values were varied according to the normal distribution. The beta distribution was used for utility values as it is constrained between 0 and 1 (34). To estimate parameters of the beta distribution, using the mean and variance, the method of moments approach was used. Alpha and beta were estimated using the following equations:

$$\bar{\mu} = \frac{\alpha}{\alpha + \beta} \qquad s^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

where: $\bar{\mu} = mean;\ s^2 = variance.$

As disutility values are negative, the normal distribution was used (348). Utility values were varied ±25% in the OWSA. To investigate uncertainty associated with the HRQOL of patients considered to be long-term survivors, a scenario was explored, whereby patients alive after 60 months were assigned the long-term survival utility sourced from Kwon et al. (0.90) (308). A scenario was also explored, whereby patients alive after 60 months were subject to age- and sex-matched general population utility (328). An additional scenario, whereby disutility associated with adverse events was removed was also considered. This was to account for the potential for the health-state utility values from ELIANA to capture some degree of disutility due to adverse events.

5.3.3 Cost Inputs

Direct medical costs considered in the model were staff training costs (for tisagenlecleucel), tisagenlecleucel pre-treatment costs (leukapheresis, bridging chemotherapy, lymphodepleting chemotherapy, cryopreservation), drug acquisition costs, hospitalisation costs, initiation and monitoring costs, adverse event costs, costs associated with alloSCT, and terminal care costs. Drug costs were calculated in line with the NCPE Guidelines for Calculation of Drug Costs (349). Resource use estimates were sourced from the clinical trial data (87, 207, 215, 216), and the NCCP Chemotherapy Regimen (blinatumomab) (340). Irish cost data were used, where available. These data were sourced from the Healthcare Pricing Office 'DRG Prices for Inpatients and Daycases 2020' (350) (herein 'HPO DRG List'), the NCPE Internal Cost Database, and tertiary teaching hospitals. Costs were inflated to 2020 using the Consumer Price Index for health (351). Where necessary, costs from non-Irish sources were inflated to 2020 and converted to Euro using purchasing power parities (352). Costs were discounted at a rate of 4% from the beginning of the second year (11). In the PSA, costs were varied according to the gamma distribution. This distribution was chosen as it is constrained on the interval 0 to positive infinity and can be highly skewed to reflect the skew often found in cost data (34). In the OWSA, costs were varied ±25%.

5.3.3.1 Tisagenlecleucel Implementation Costs

Clinical opinion[7] indicated that because Children's Health Ireland, Tertiary Hospital Crumlin is JACIE accredited, no significant infrastructural changes were anticipated for implementation of a CAR T-cell therapy service. As such, no implementation costs were included in the model.

5.3.3.2 Tisagenlecleucel Training Costs

EMA conditional marketing authorisation of tisagenlecleucel stipulates that all healthcare professionals expected to prescribe, dispense, or administer tisagenlecleucel should be provided training on handling of frozen cells, and on associated adverse events (143). In the absence of data, it was assumed that 24 staff in Children's Health Ireland, Tertiary Hospital Crumlin received formal training. This assumption was based on clinical

[7] Written correspondence with one consultant haematologist in Ireland.

opinion[8], which indicated that, for R/R DLBCL, approximately 40 staff members in St James's Tertiary Hospital were expected to receive formal training. Here, it was arbitrarily assumed that due to the fewer patients expected to receive treatment for R/R ALL, staff resource requirements would be approximately 50% of that of R/R DLBCL. In line with the NICE HTA appraisal of axicabtagene ciloleucel (TA559), a training duration of 16 hours per person was assumed (331). Assumptions regarding the distribution of staff members trained is presented in Table 22. Staff costs were estimated as per National Economic Evaluation Guidelines (11).

Table 22 Estimated staff training requirements for implementation of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

Role	Number Trained	Source
Consultant Haematologist	4	All consultant haematologists, specialising in the treatment of leukaemia in Children's Health Ireland, Crumlin receive training
Specialist/Senior Registrar Haematologist	1	Assumption
Consultant Intensivist	1	Based on adverse event profile of tisagenlecleucel*
Consultant Neurologist	1	
Clinical Nurse Specialists	2	Includes CAR T-cell therapy co-ordinator
Nurses	10	Provide most patient monitoring
Pharmacy	3	Pharmacy involved in bridging chemotherapy, lymphodepleting chemotherapy, receipt and adverse event management* (353)
Laboratory	2	Laboratory staff involved in receipt and storage* (353)

*Assumed that these staff members receive 8 hours training (instead of 16) as they are involved only in specific aspects of the process.

To estimate the per patient cost of training, a crude approach was adopted. This considered the cost of training, the number of patients expected to be treated each year, and the number of years before staff require retraining. The cost per centre was estimated using published HSE Salary Scales (354). The midpoint of the lowest and highest points on the scale was selected to calculate an average salary cost (11). The number of eligible patients per centre per year was estimated to be six. The number of years before healthcare professionals require retraining was assumed to be two (331).

[8] Oral correspondence with one consultant haematologist in Ireland.

The formula used to estimate the per patient cost of training, presented below, was derived from the NICE HTA appraisal of axicabtagene ciloleucel (TA559) (331).

$$\frac{Cost\ per\ centre}{Number\ of\ patients\ per\ centre\ per\ year\ \times\ Number\ of\ years\ before\ retraining}$$

Using this formula, the per patient cost of training was estimated to be €1,595.55. This was applied as a once-off cost. Details of the cost breakdown are provided in Appendix D (Table A8).

5.3.3.3 Tisagenlecleucel Pre-Treatment Costs

Tisagenlecleucel-specific pre-treatment costs were applied to the proportion of patients undergoing each procedure as a once-off cost.

5.3.3.3.1 Leukapheresis

Leukapheresis separates T-cells from blood. These are the starting material for the manufacture of tisagenlecleucel (355). The once-off cost of leukapheresis, €1,249 per patient, was sourced from the HPO DRG List (daycase DRG B62Z) (350). All patients in the tisagenlecleucel arm underwent leukapheresis.

5.3.3.3.2 Cryopreservation

T-cells must be cryopreserved during shipping to the manufacturing facility (3). An associated cost of €5,544.68 per patient was sourced from a Tertiary Teaching Hospital. All patients in the tisagenlecleucel arm incurred this cost. Tisagenlecleucel requires cryopreservation following manufacture (3). An additional cost of €5,544.68 per patient was applied to the proportion of patients receiving infusion to account for this.

5.3.3.3.3 Bridging Chemotherapy

The bridging chemotherapy regimen consisted of allopurinol, dexamethasone, vincristine, intrathecal methotrexate and co-trimoxazole (223) (Appendix D, Table A9)). Drug costs were obtained from the PCRS List of Reimbursable Items (356) and the IPHA Price Realignment Files (357-359). Bridging chemotherapy was administered for 1 treatment cycle (21 days), resulting in a total per patient cost of €159.56. Vial sharing was

not assumed. This cost was incurred by 88% of patients (who received infusion) in the tisagenlecleucel arm (207, 215). Of the patients in the tisagenlecleucel arm who did not proceed to infusion, it was assumed that 50% received bridging chemotherapy.

5.3.3.3.4 Lymphodepleting Chemotherapy

As per the SPC, lymphodepleting regimens can consist of fludarabine in combination with cyclophosphamide, or cytarabine in combination with etoposide (3). Here, it was assumed that all patients received fludarabine in combination with cyclophosphamide. This regimen was received by 95% of all infused patients in ELIANA (215). Both regimens comprise off-patent cytotoxics and so, the difference in cost between regimens is expected to be negligible. Drug costs were sourced from the NCPE Internal Cost Database. The total cost of €414.44 per patient, per treatment course (Appendix D, Table A10), was incurred by 95% of patients who received infusion in the tisagenlecleucel arm (207, 215). Vial sharing was not assumed. Of the patients in the tisagenlecleucel arm who did not proceed to infusion, it was assumed that 50% received lymphodepleting chemotherapy.

5.3.3.4 Drug Acquisition Costs

The total drug acquisition costs for tisagenlecleucel and blinatumomab are presented in Table 23. Costs presented are exclusive of VAT; this is not applicable in the cost-utility analysis. Costs were sourced from the NCPE Technical Summary of tisagenlecleucel (88), and the IPHA Price Realignment File 2020 (359).

Table 23 Total drug acquisition costs per patient per treatment course employed in cost-utility model of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

Drug	Total Drug Acquisition Cost Per Patient Per Treatment Course* (€)
Tisagenlecleucel	301,762.13
Blinatumomab (dosing based on body surface area) †§	89,213.55
Blinatumomab (fixed-dosing regimen)‡§	125,381.20

*Including relevant fees and rebates, excluding VAT.
†For patients weighing <45kg, dosing of blinatumomab is based on body surface area. Assumed 50% of patients receive this dosing.
‡For patients weighing ≥45kg, a fixed-dosing regimen of blinatumomab applies. Assumed 50% of patients receive this dosing.
§Does not include cost of allogeneic stem cell transplant.

The cost of tisagenlecleucel was applied as a once-off cost. It was applied only to patients who received infusion (83%). It was assumed that all such patients received just one infusion. It was assumed that 9% of patients in the tisagenlecleucel arm incurred the cost of blinatumomab.

In line with the NCCP Chemotherapy Regimen (340) and clinical opinion, blinatumomab is administered for up to two cycles in Irish clinical practice. It was assumed that 100% of patients received one cycle and 33% received a second cycle, as per NCT01471782 (87). However, in NCT01471782, patients could receive up to five cycles of blinatumomab, provided they achieved CR after the first two cycles. In this trial, 11.4% of patients received a third cycle, 4.3% received a fourth cycle, and 4.3% received a fifth cycle. Applying these treatment durations would result in an overestimation of costs associated with blinatumomab and so, this approach was not taken in the base case. It is noted that the efficacy data were not modified to reflect the fewer number of cycles modelled in the cost-utility model. As such, the impact of administering five cycles of blinatumomab, as per NCT01471782, was explored in scenario analysis.

The dosing of blinatumomab is contingent on patient weight. As presented in 5.2.5 (Table 16), patients weighing less than 45kg receive dosing based on body surface area, while patients weighing 45kg or greater receive a fixed dose. The SPC of blinatumomab details the number of vials required based on patient body surface area or weight, and the duration of infusion (86). Alternate infusion durations of 72- and 96-hours, over the cycle period, were assumed (360). This approach avoids the need to compound the product and change the infusion bag at weekends when outpatient facilities are closed. Patients weighing less than 45kg require 2 vials for the first 7 days of treatment (cycle 1) and 15 vials for the remaining 21 days (cycle 1). A total of 20 vials are required for cycle 2. For patients weighing 45kg or greater, 3 vials are required for the first 7 days of treatment (cycle 1) and 21 vials are required for the remaining 21 days (cycle 1). A total of 28 vials are required for cycle 2 (86). A detailed breakdown is presented in Appendix D. The total cost per patient, based on body surface area dosing, was estimated to be €40,990 in the first cycle and €48,224 in the second cycle. Based on the fixed-dosing regimen, the total cost per patient was estimated to be €57,868 in the first cycle and €67,513 in the second cycle. There is a lack of published data regarding the weight distribution of patients in

ELIANA (143, 215), ENSIGN (207), and NCT01471782 (87, 216). For this analysis, it was arbitrarily assumed that 50% of patients received dosing based on body surface area and 50% received the fixed-dosing regimen. Alternative assumptions regarding dosing were explored in scenario analysis.

5.3.3.5 Outpatient Administration Costs

The vincristine and intrathecal methotrexate components of the bridging chemotherapy regimen require outpatient administration. An administration cost of €346 per patient per day (obtained from a Tertiary Teaching Hospital) was applied. The total bridging chemotherapy administration cost was €692 per patient. The other components of the bridging chemotherapy regimen are administered orally and do not incur an administration cost. Administration costs of lymphodepleting chemotherapy, tisagenlecleucel, and the first seven days of blinatumomab treatment, were assumed to be captured by the cost of hospitalisation, described in 5.3.3.6.

5.3.3.6 Hospitalisation Costs

In the absence of severe adverse events, the duration of hospitalisation for patients receiving tisagenlecleucel (including lymphodepleting chemotherapy) for R/R ALL is expected to be three to four weeks[9]. To account for this, a cost of €37,944 per patient was sourced from the HPO DRG List (DRG R60A) (350). This relates to a case of acute leukaemia with major complexity (mean length of stay 24.5 days). This was applied only to patients receiving infusion with tisagenlecleucel. Patients receiving tisagenlecleucel are required to remain within two hours of travel of the hospital for at least four weeks following infusion (143). It was therefore, assumed that 50% of patients who received infusion were discharged to hospital-associated patient apartments for four nights. The cost per patient per night was €63.90 (295). Patients in the tisagenlecleucel arm, who received lymphodepleting chemotherapy but did not proceed to infusion with tisagenlecleucel, incurred a cost of €5,100 per patient. This cost was sourced from the HPO DRG List (DRG R61B; mean length of stay 4.4 days) (350).

[9] Written correspondence with one consultant haematologist in Ireland.

Patients treated with blinatumomab, in Irish clinical practice, are hospitalised for the first seven days (cycle one) and are usually discharged, with an infusion pump, after this period[10]. As described in 5.3.3.4, infusion durations were assumed to alternate between 72- and 96-hours. This resulted in a total of seven-days inpatient stay (cycle one), seven outpatient visits and nine outpatient visits, in cycle one and cycle two, respectively. The cost of an inpatient stay (€11,826) was sourced from the HPO DRG List (DRG R60B) (350); relating to a case of acute leukaemia of minor complexity (mean length of stay 6.8 days). The cost of an outpatient visit was €136.76 per patient, based on the 2013 HSE Ready Reckoner (R99 'Oncology Repeat Attendance', inflated to 2020) (361). A cost for the infusion pump was included; €118.67 per patient per 28 days. The was sourced from Irish supplier Rockford Healthcare (2018 inflated to 2020).

5.3.3.7 Initiation and Monitoring Costs

All tisagenlecleucel initiation and initial monitoring costs were assumed to be accounted for in the cost of hospitalisation. Costs were included in the blinatumomab arm to account for outpatient monitoring. Resource use was sourced from the NCCP Chemotherapy Regimen of blinatumomab (340); costs were obtained from a number of sources. The total per treatment cycle (42 days) monitoring cost of blinatumomab was estimated to be €198.30 per patient (Appendix D, Table A18).

Health-state-specific follow-up costs were applied for the event-free survival and progressed disease states. Follow-up costs in the event-free survival state varied by time horizon, reflecting the decreased need to monitor patients over time. Follow-up requirements were sourced from the National Comprehensive Cancer Network (NCCN) Clinical Practice Guidelines in Oncology: Paediatric Acute Lymphoblastic Leukaemia (version 2.2021) (60). Additional monitoring requirements, specific to CAR T-cell therapy, were sourced from Yakoub-Agha et al. (353). Yakoub-Agha et al. present best practice recommendations for the management of adults and children receiving CAR T-cell therapy. Further detail is provided in Appendix D (Table A19). For the progressed disease state (both treatment arms), it was assumed that costs incurred were the same as those incurred by patients receiving blinatumomab, in the event-free survival state, in months

[10] Written correspondence with one consultant haematologist in Ireland.

1 to 12. This is in line with assumptions employed elsewhere (223, 239, 250). Patients alive at 61 months incurred the cost of event-free survival month 61 onwards, regardless of health state. This was to align with the assumption that patients who are alive after 60 months are considered to be long-term survivors. All costs were obtained from Irish sources and applied as a per cycle cost to patients in the relevant health state. These are presented in Table 24.

Table 24 Per cycle health-state monitoring costs in cost-utility analysis of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

	Event-Free Survival Months 1-12 (inclusive)	Event-Free Survival Months 13-24 (inclusive)	Event-Free Survival Months 25-60 (inclusive)	Event-Free Survival Month 61 onwards	Progressed Disease
Tisagenlecleucel Per Cycle Cost (€)	115.94	54.85	36.57	18.28	78.74
Blinatumomab Per Cycle Cost (€)	78.74	36.25	24.17	12.08	78.74

5.3.3.8 Adverse Events

For tisagenlecleucel, costs associated with adverse events are likely to be captured, to some degree, by the cost of hospitalisation. To mitigate against double-counting, only select adverse events, considered to be associated with considerable resource use, were included. As patients treated with blinatumomab are discharged from hospital after seven days, additional costs to account for grade ≥3 adverse events, occurring in 5% or greater of the population, were included for blinatumomab. These costs were applied as a once-off at the start of the first cycle. The following tisagenlecleucel-specific adverse events costs were considered: CRS, non-CRS ICU admission, B-cell aplasia, febrile neutropenia and pancytopenia.

5.3.3.8.1 Cytokine Release Syndrome

CRS-associated costs in the model comprised ICU admission and treatment with tocilizumab. It was assumed that 47% of patients treated with tisagenlecleucel were admitted to the ICU for CRS-related events, for 8 days (143, 219). It was further assumed that 28% of patients treated with tisagenlecleucel received treatment with tocilizumab (207, 215). Patients received a mean of 1.24 doses of tocilizumab (239).

It was assumed that 5.7% of patients who received blinatumomab were admitted to the ICU for 5 days. This was based on the proportion of patients in NCT01471782 who experienced 'serious' CRS (5.7%) and median time to resolution of CRS (5 days) (87). As tocilizumab is only licensed for the treatment of CRS associated with CAR T-cell therapy (158), no tocilizumab-associated costs were applied to patients receiving blinatumomab.

ICU admission cost was sourced from O'Brien et al. (€2,797.76; inflated to 2020) (362); the cost of tocilizumab was sourced from MIMS Ireland (January 2020) (363). A total ICU cost of €22,382.08 and €13,988.80 per patient, per stay was estimated for tisagenlecleucel and blinatumomab, respectively. These estimates were applied to the proportion of patients requiring CRS-related ICU admission in each arm. The cost per dose of tocilizumab was €672.84. Vial sharing was not assumed.

5.3.3.8.2 Non-Cytokine Release Syndrome ICU Admission

To account for severity of additional adverse events, it was assumed that 90% of patients treated with tisagenlecleucel were admitted to the ICU for 1.78 days. This represents the mean ICU length of stay for non-CRS adverse events in ELIANA (240).

5.3.3.8.3 B-Cell Aplasia

Patients who experience B-cell aplasia are generally treated with IV immunoglobulin. Median time to B-cell recovery in ELIANA was 11.4 months (250). This estimate has been used in the literature to model the duration of IV immunoglobulin treatment (239, 240). In ELIANA, 47.1% of patients received treatment with IV immunoglobulin (143). However, in ELIANA, approximately 70% of patients with B-cell aplasia had not reached B-cell recovery by 24 months (215). The use of median time to B-cell recovery may therefore, underestimate costs associated with IV immunoglobulin treatment. These data were not reported for ENSIGN (207).

The duration of IV immunoglobulin treatment is a key area of uncertainty. A robust analysis would require IPD on all-cause time to IV immunoglobulin treatment discontinuation. However, these data are not available (143, 207, 215). An alternative approach was therefore, explored. In this approach, it was assumed that all patients, who

received tisagenlecleucel infusion, in the event-free survival state, had B-cell aplasia and 47.1% of these required IV immunoglobulin (250). The cost of IV immunoglobulin was applied over the duration of event-free survival. This assumption was applied in the base case, as B-cell aplasia, being a marker for tisagenlecleucel persistence, is correlated with duration of remission (156).

The total cost per dose was estimated to be €1,365.00 (sourced from a Tertiary Teaching Hospital). The administration cost was assumed to be the same as an outpatient administration of chemotherapy, €346 per patient.

5.3.3.8.4 Febrile Neutropenia and Pancytopenia

Febrile neutropenia is associated with considerable cost, as reported by O'Brien et al. (362), who conducted a microcosting study to estimate the cost of managing febrile neutropenia in the inpatient setting in patients with cancer in Ireland. Based on the pooled ELIANA and ENSIGN data, 36% of patients experienced grade ≥3 febrile neutropenia (207, 215). In NCT01471782, 17% of patients experienced grade ≥3 febrile neutropenia (216). The cost from O'Brien et al. (362), €9,451.31 per patient (inflated to 2020), was applied to these proportions in each arm.

Pancytopenia has been reported to be one of the most impactful adverse events experienced by patients treated with CD19 CAR T-cell therapy for R/R DLBCL (330). It was assumed that patients experiencing grade ≥3 pancytopenia (3% (143)) were treated as a daycase, once per month, for the duration of pancytopenia (six months). A cost of €387 per daycase, sourced from the HPO DRG List (DGR R62B) (350), was applied to these patients.

5.3.3.8.5 Other Adverse Events Associated with Blinatumomab

Grade ≥3 adverse events occurring in 5% or greater of the population in NCT01471782 were included in the blinatumomab arm (216). Costs were sourced from the HPO DRG List and HSE Ready Reckoner (350, 361); see Appendix D (Table A24). These additional adverse events (excluding CRS and febrile neutropenia) cost €943.07 per patient.

5.3.3.9 Allogeneic Stem Cell Transplant

To account for the use of blinatumomab as a bridge to alloSCT in some patients (49% based on expert elicitation), the cost of alloSCT and subsequent follow up was included. The cost of the alloSCT procedure was sourced from the HPO DRG List (DRG A07A); €202,698 per patient (350). Follow-up costs, accounting for 365 days post-discharge following HSCT (both alloSCT and autoSCT), were sourced from a report by Ernst & Young (commissioned by the Anthony Nolan Charity, UK) (364) and converted from 2020 UK Sterling to Euro (352). This report used data from the NHS Digital's Secondary Uses Service to track HSCT hospital activity during the procedure and for 365 days post-discharge (years 2015 to 2016) for adults and children. Data from The Royal Marsden NHS Foundation Trust's Patient Level Information and Costing System were used to approximate costs associated with this activity. Costs were for the first 100 days post-discharge (€64,618.28 per patient), 101 to 200 days post-discharge (€36,524.17 per patient), and 201 to 365 days post-discharge (€40,957.86 per patient). Resource use and associated costs were based on the mean number of surviving patients across the 365-day period. The cost of the alloSCT procedure was applied as a once-off cost, while follow-up costs incurred in each period (i.e. first 100 days, 101 to 200 days, 201 to 365-days post-discharge) were converted to a per cycle cost and applied to the cycle in which they were incurred.

5.3.3.10 Terminal Care

A once-off terminal care cost, sourced from Bourke et al. (365), was applied to patients upon entering the death state. This cost (€7,732.48) was derived through information provided by the Irish Hospice Foundation regarding cost and length of stay, by place of death.

5.4 Model Outputs

5.4.1 Deterministic ICER

The base case analysis considered the ICER, of tisagenlecleucel versus blinatumomab, according to standard decision rules (366). The base case ICER was calculated from deterministic costs and deterministic QALYs.

5.4.2 Probabilistic ICER and Scatterplot

PSA was conducted. Survival, utility, and cost parameters were varied according to their appropriate distributions (as described in the relevant sections). Probabilistic results were generated using Monte Carlo Simulation by running the model for 5,000 iterations. This was run over three successive iterations to ensure the results did not change appreciably. Monte Carlo Simulation involves making random draws of the uncertain parameters from their probability distributions, running the model for each simulated set of parameters and collecting the outputs for each iteration (367). The probabilistic ICER was calculated using the mean incremental costs and mean incremental QALYs of the 5,000 iterations.

A scatterplot of incremental costs and outcomes, generated from each iteration of the PSA, was constructed to illustrate the degree of uncertainty surrounding the estimates.

5.4.3 Cost-Effectiveness Acceptability Curve

For each PSA iteration, the expected costs and QALYs of tisagenlecleucel and blinatumomab were recorded, and combined to estimate the expected NMB for each. NMB was derived by multiplying the total number of QALYs by the cost-effectiveness threshold (€45,000 per QALY), minus the total costs for tisagenlecleucel or blinatumomab (34). From the NMB values, the probability of tisagenlecleucel and blinatumomab being cost effective over a range of willingness-to-pay thresholds (€0.00 per QALY to €350,000 per QALY) was identified. An upper bound threshold of €350,000 per QALY may be unrealistic; this was chosen for illustrative purposes only. These probabilities were then plotted, over the range of thresholds, to produce a cost-effectiveness acceptability curve (368). This curve illustrates the measure of uncertainty in the decision.

5.4.4 One-Way Sensitivity Analysis

OWSA of all model parameters was performed to determine the sensitivity of the model to changes in individual parameters and assumptions. The upper and lower bounds of the 95% CI were used when these were available for point estimates. Otherwise, point estimates were varied ±25%. A tornado plot was then constructed to illustrate the impact of the 10 most influential parameters on the deterministic ICER.

5.4.5 Scenario Analysis

A number of scenario analyses were conducted to assess the impact on the deterministic ICER of employing alternative, plausible assumptions. These scenarios, presented in Table 25, assessed the impact of uncertainty associated with structural and methodological assumptions.

Table 25 Scenario analyses on deterministic ICER of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia

Parameter/Assumption	Base Case	Scenario	Justification
Time Horizon	88 years	2 years	Median follow up was 13.1 months (ELIANA (143)) and 31.7 months (ENSIGN (207))
Proportion of Patients Receiving Tisagenlecleucel Infusion	83%	100%	Efficacy data in ELIANA & ENSIGN are based on patients who received infusion (i.e. mITT) (207, 215)
Clinical Data Informing Tisagenlecleucel OS Efficacy	Pooled ELIANA and ENSIGN	ELIANA	ELIANA informs EFS efficacy
Extrapolation of Pooled ELIANA and ENSIGN (tisagenlecleucel) OS Data	One-knot (odds) spline model	Log-normal model	5-year OS predicted by log-normal most closely aligned with expert elicitation outputs
Extrapolation of NCT01471782 (blinatumomab) OS Data	One-knot (odds) spline model	Log-normal model	Log-normal model was best fit (AIC & BIC) of parametric models
Extrapolation of ELIANA, ENSIGN (tisagenlecleucel), and NCT01471782 (blinatumomab) OS Data	One-knot (odds) spline model	Log-normal model	Combination of the two scenarios above
Extrapolation of ELIANA EFS Data	Generalised gamma model	Two-knot (odds) spline model	Two-knot (odds) spline model also a reasonable option
Time Point at which Patients are Considered Long-Term Survivors	After 60 months	After 24 months	Majority of patients expected to relapse within 24 to 60 months post-treatment (12, 202, 203)
		No long-term survival point	Long-term survival is uncertain, as illustrated in the expert elicitation
HRQOL of Long-Term Survivors	All patients alive after 60 months assumed HRQOL equivalent to event-free survival state (0.80)	All patients alive after 60 months assumed HRQOL derived by Kwon et al. (0.90)	Alternative published utility value (308)
		All patients alive after 60 months assumed HRQOL equivalent to the age- and sex-matched general population (328)	Uncertainty exists regarding HRQOL of long-term survivors (336)

Disutility Associated with Select Adverse Events	Included	Excluded	Potential for disutility due to adverse events to be captured by health-state utility values
Disutility Associated with Treatment and All Adverse Events	Select adverse events included only	-0.42 (derived by Sung et al. (321))	Alternative assumption used in HTA appraisals (223, 240, 250), Hettle et al. (12), ICER HTA (296)
Duration of IV Immunoglobulin Treatment	Duration of event-free survival	11.4 months	Median time to B-cell recovery in ELIANA (250)
Cycles of Blinatumomab Received	100% of patients receive one cycle, 33% receive a second cycle	100% of patients receive one cycle, 33% receive a second cycle, 11.4% receive a third cycle, 4.3% receive a fourth cycle, and 4.3% receive a fifth cycle (87)	Up to five cycles permitted in NCT01471782 (87)
Dosing Regimen of Blinatumomab	50% of patients receive dosing based on body surface area and 50% receive fixed-dosing regimen	100% receive dosing based on body surface area	Lack of published data on weight distribution of patients in ELIANA, ENSIGN (tisagenlecleucel), and NCT01471782 (blinatumomab)
		100% receive fixed-dosing regimen	
Proportion of Patients Receiving AlloSCT in Blinatumomab Arm	49%	35.7%	35.7% of patients received alloSCT in NCT01471782 (87)
Proportion of Patients Receiving AlloSCT in Tisagenlecleucel Arm	0%	12%	12% of patients received alloSCT in pooled ELIANA & ENSIGN data (207, 215)
Discount Rate	4% on costs and outcomes	1.5% on costs and outcomes	NICE may consider a 1.5% discount rate where benefits are likely to be sustained over a long period (369)
		4% on costs and 1.5% on outcomes	Gravelle and Smith propose that the discount rate on health outcomes should be 1% to 3.5% lower than that on costs (186)
		Hyperbolic discounting: 4% (0-30 years), 3.5% (31-60 years), 3% (61-100 years) on both costs and outcomes	Hyperbolic discounting may be applicable when the time horizon exceeds 30 years (182)

AIC: Akaike information criteria; **AlloSCT:** Allogeneic stem cell transplant; **BIC:** Bayesian information criteria; **EFS:** Event-free survival; **HRQOL:** Health-related quality of life; **HTA:** Health technology assessment; **ICER:** Incremental cost-effectiveness ratio; **IV:** Intravenous; **mITT:** Modified intention-to-treat; **NICE:** National Institute of Health and Care Excellence; **OS**: Overall survival; **QALY:** Quality-adjusted life year.

5.4.6 Price Analysis

Tisagenlecleucel was approved for reimbursement in Ireland, following confidential price negotiations, in July 2021. Therefore, the price modelled in this analysis does not reflect the actual price paid by the HSE. An analysis was conducted (using the 'Goal Seek' function in Microsoft Excel®) to determine the decrease in price-to-wholesaler of tisagenlecleucel that would be required for tisagenlecleucel to meet the €45,000 per QALY threshold.

5.4.7 Expected Value of Perfect Information

As described in 0, EVPI estimates the value of simultaneously eliminating all uncertainty of all uncertain parameters relating to the decision (370). To characterise uncertainty, the PSA uses Monte Carlo Simulation to repeatedly sample from the prior distributions assigned to all the uncertain model parameters (θ). For each iteration of the PSA, the expected costs ($E_\theta C(j,\theta)$) and QALYs ($E_\theta Q(j,\theta)$) for each treatment (j) were combined to form a measure of NMB ($E_\theta NMB(j,\theta)$).

The NMB is given by (368, 371):

$$E_\theta NMB(j,\theta) = (E_\theta Q(j,\theta) x\ \lambda) - E_\theta C(j,\theta)$$

where:
NMB: Net monetary benefit
$E_\theta Q(j,\theta)$: Expected QALYs
$E_\theta C(j,\theta)$: Expected costs
λ: Cost-effectiveness threshold
j: All treatment strategies under investigation
θ: All uncertain model parameters

The output of these iterations represents the possible values of the NMB for all possible realisations of the uncertain parameters (θ) (372, 373).

EVPI is the differences between the expected NMB with perfect information and the expected NMB with current information. The optimal decision with current information would choose the intervention that generates the maximum expected NMB ($max_j E_\theta NMB(j,\theta)$). With perfect information, the resolution of uncertainties would be known and the alternative that maximises the NMB for a particular value of

($max_j NMB(j,\theta)$). However, the true values of θ are unknown. The expected value of a decision taken with perfect information is derived by averaging the maximum NMB over the joint distribution of θ.

The EVPI was therefore, calculated as the average of the maximum NMBs across all iterations of the PSA (i.e. the expected NMB with perfect information), minus the maximum of the average expected NMBs across all strategies (i.e. the expected NMB with current information) (371-373).

The EVPI was calculated as (371):

$$EVPI = E_\theta max_j NMB(j,\theta) - max_j E_\theta NMB(j,\theta)$$

where:
$E_\theta max_j$: Expected NMB with perfect information about θ
$max_j E_\theta NMB$: Expected NMB of strategy of choice with current information about θ
NMB: Net monetary benefit

EVPI was calculated on 5,000 iterations of the PSA and over a range of willingness-to-pay thresholds (€0.00 per QALY to €350,000 per QALY). EVPI estimates were scaled up to population according to the incidence of the decision (6 patients per year, total 51 patients over 10 years when discounting is applied) (38). A technology time horizon of 10 years was assumed (12). A discount rate of 4% was applied.

Population EVPI was calculated as (373, 374):

$$Population\ EVPI = EVPI\ x\ \sum_{t=1}^{T} (I_t\ /(1+r)^t)$$

where:
EVPI: Expected value of perfect information
T: Time horizon
I_t: Incidence estimate over time horizon
r: Discount rate

Population EVPI, over a range of thresholds, was plotted and presented graphically as an EVPI curve.

5.4.7.1 Partial Expected Value of Perfect Information

To identify parameters, which contributed most to the overall decision uncertainty, EVPPI was also estimated. The EVPI of a parameter or categories of parameters (φ) (herein 'EVPPI') can be calculated by dividing the uncertain parameters (θ) into two parameter subsets, φ and its compliment ψ. EVPPI is the difference between the expected NMB with perfect information about φ and the expected NMB with current information about φ.

With perfect information, the value of φ is known and the expected NMBs are calculated over the remaining uncertainties ψ ($max_j E_\psi|_\varphi NMB(j, \varphi, \psi)$). However, the true values of φ are unknown and the expected value of a decision taken with perfect information is found by averaging these maximum expected NMBs over the distribution of φ ($E_\varphi max_j E_\psi|_\varphi NMB(j, \varphi, \psi)$).

The expected value with current information is the same as before ($max_j E_\theta NMB(j, \theta)$), since $\varphi \cup \psi = \theta$ (34).

The EVPPI was calculated as (34, 371, 375):

$$Parital\ EVPI\ (\varphi) = E_\varphi max_j E_\psi|_\varphi NMB(j, \varphi, \psi) - max_j E_\theta NMB(j, \theta)$$

where:
$E_\varphi max_j E_\psi|_\varphi NMB(j, \varphi, \psi)$: Expected NMB with perfect information about φ
$max_j E_\theta NMB(j, \theta)$: Expected NMB of strategy of choice with current information about φ

EVPPI was calculated on 5,000 iterations of the PSA and over a range of thresholds (€0.00 per QALY to €350,000 per QALY). A bespoke macro (programmed in Microsoft Excel® Visual Basic for Applications) was written, which generated probabilistic samples of all input parameters (along with probabilistic samples of costs and outcomes) for each PSA iteration. EVPPI was estimated using the Gaussian process regression approach (376, 377). Identified input parameters were categorised; utility values, survival analysis, hospitalisation and monitoring costs, adverse event costs, and alloSCT costs. These categories are generally aligned with those defined elsewhere (378). The total VOI

associated with each category was determined. EVPPI estimates were scaled up to population, using the same method described for EVPI (as described in 5.4.7).

5.5 Results

5.5.1 Deterministic Results

The deterministic model outcomes are presented in Table 26. Tisagenlecleucel was not cost effective, versus blinatumomab, at a willingness-to-pay threshold of €45,000 per QALY (31).

Table 26 Deterministic results of the incremental analysis of cost effectiveness of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia

Drug	Total Costs (€)	Total QALYs	Incremental Costs (€)	Incremental QALYs	ICER (€/QALY)
Blinatumomab*	219,950	2.18			
Tisagenlecleucel	376,878	4.33	156,928	2.15	73,086

ICER: Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year.
*With (49%) or without (51%) allogeneic stem cell transplant.

5.5.2 Probabilistic Results

Expected incremental costs and incremental QALYs in the base case analysis are presented in a scatterplot in Figure 10. Most of the iterations lie in the NE quadrant. Mean expected costs and QALYs are presented in Table 27.

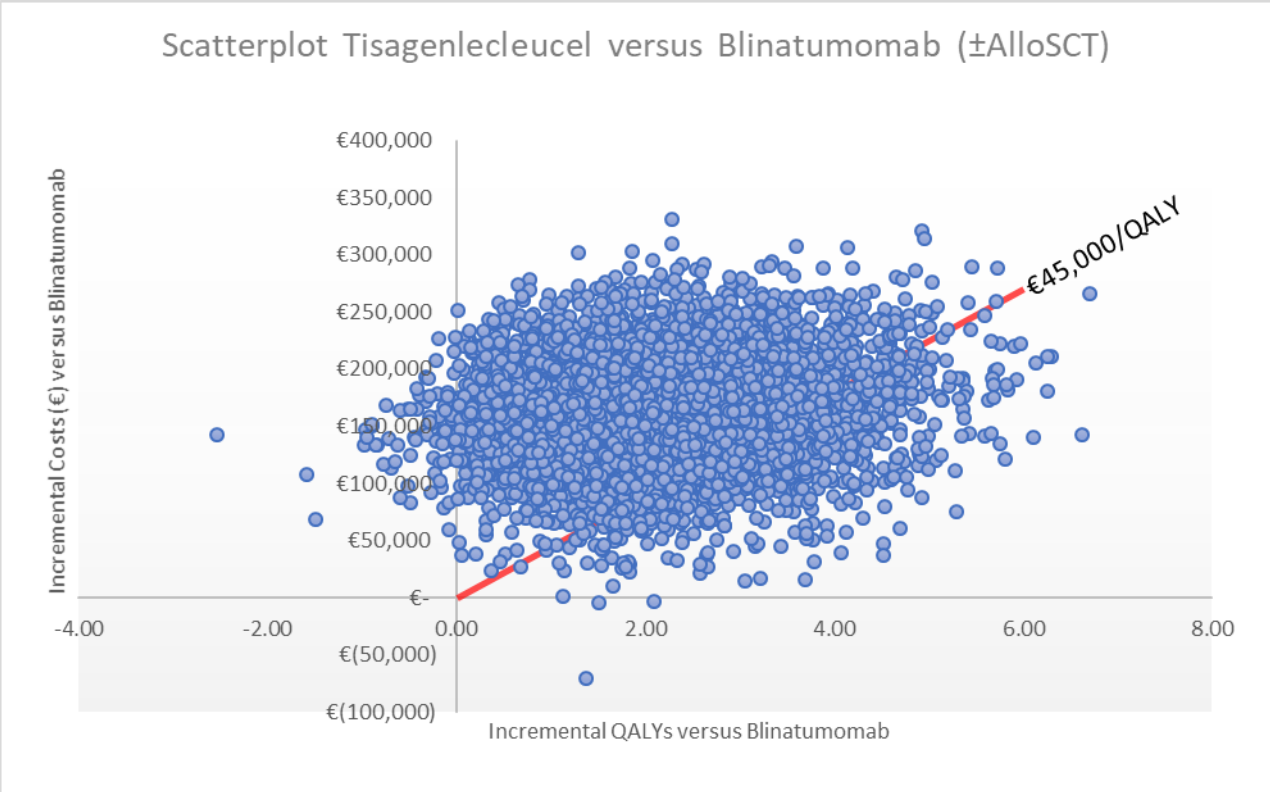

Figure 10 Scatterplot of incremental costs and incremental QALYs from probabilistic sensitivity analysis of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia
AlloSCT: Allogeneic stem cell transplant; **QALY:** Quality-adjusted life year.

Table 27 Mean probabilistic outputs of the incremental analysis of cost effectiveness of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia

Drug	Total Costs (€)	Total QALYs	Incremental Costs (€)	Incremental QALYs	ICER (€/QALY)
Blinatumomab*	219,064	2.31			
Tisagenlecleucel	383,035	4.50	163,971	2.18	75,119

ICER: Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year.
*With (49%) or without (51%) allogeneic stem cell transplant.

The cost-effectiveness acceptability curve is presented in Figure 11. At a threshold of €45,000 per QALY, there was a 16% probability that tisagenlecleucel was cost effective versus blinatumomab. As some iterations lie in the NW quadrant (more costly, less effective), the probability of cost effectiveness of tisagenlecleucel will not reach 100% at any given threshold.

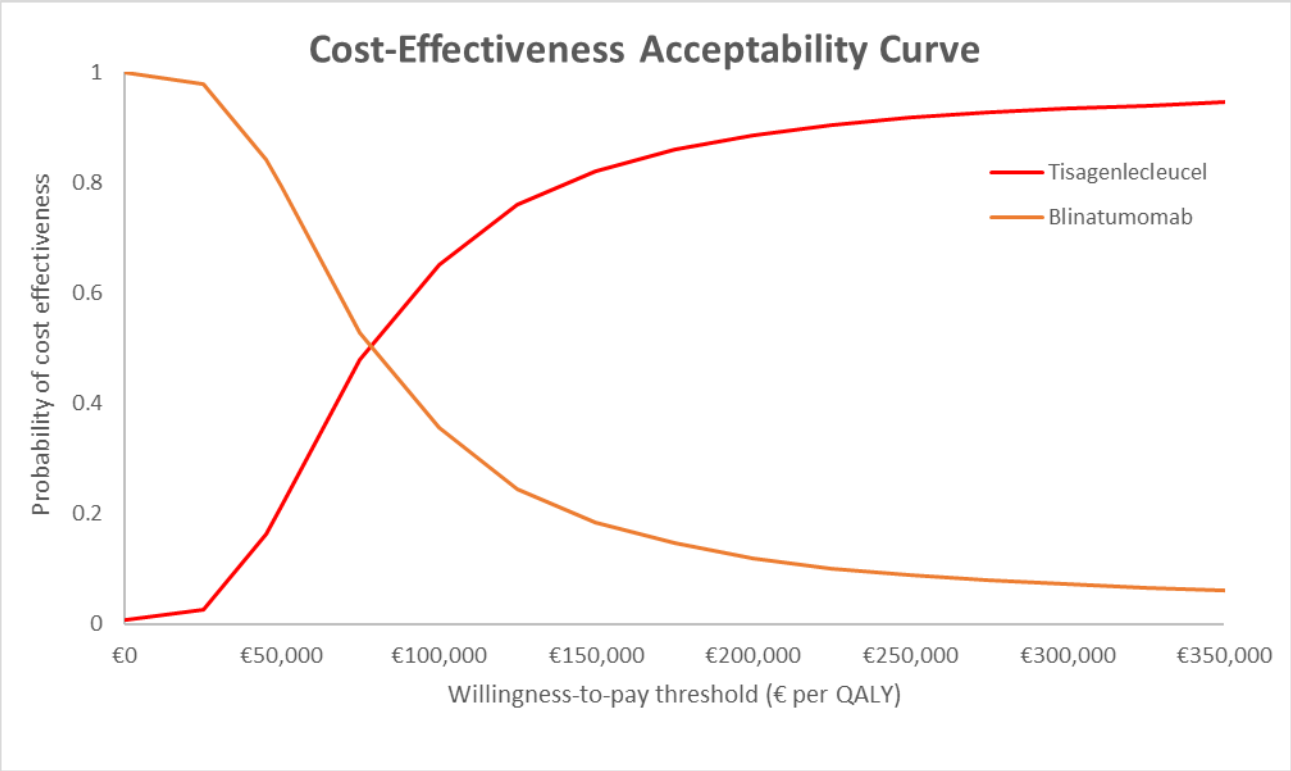

Figure 11 Cost-effectiveness acceptability curve of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia
QALY: Quality-adjusted life year.

5.5.3 One-Way Sensitivity Analysis

Results of OWSA are presented in Figure 12. For pragmatic reasons, OWSA was conducted on deterministic outcomes. Thus, results should be considered indicative only. The main drivers in the model were the rate of alloSCT in the blinatumomab arm, discount rate on outcomes, and tisagenlecleucel infusion cost. The lower bounds on discount rate on outcomes (0%) and tisagenlecleucel infusion cost (€239,494) reduced the ICER to below a threshold of €45,000 per QALY. The upper bound on rate of alloSCT in the blinatumomab arm (98%) reduced the ICER to below a threshold of €20,000 per QALY. However, the efficacy data of blinatumomab were not modified to reflect the higher proportion of patients receiving alloSCT.

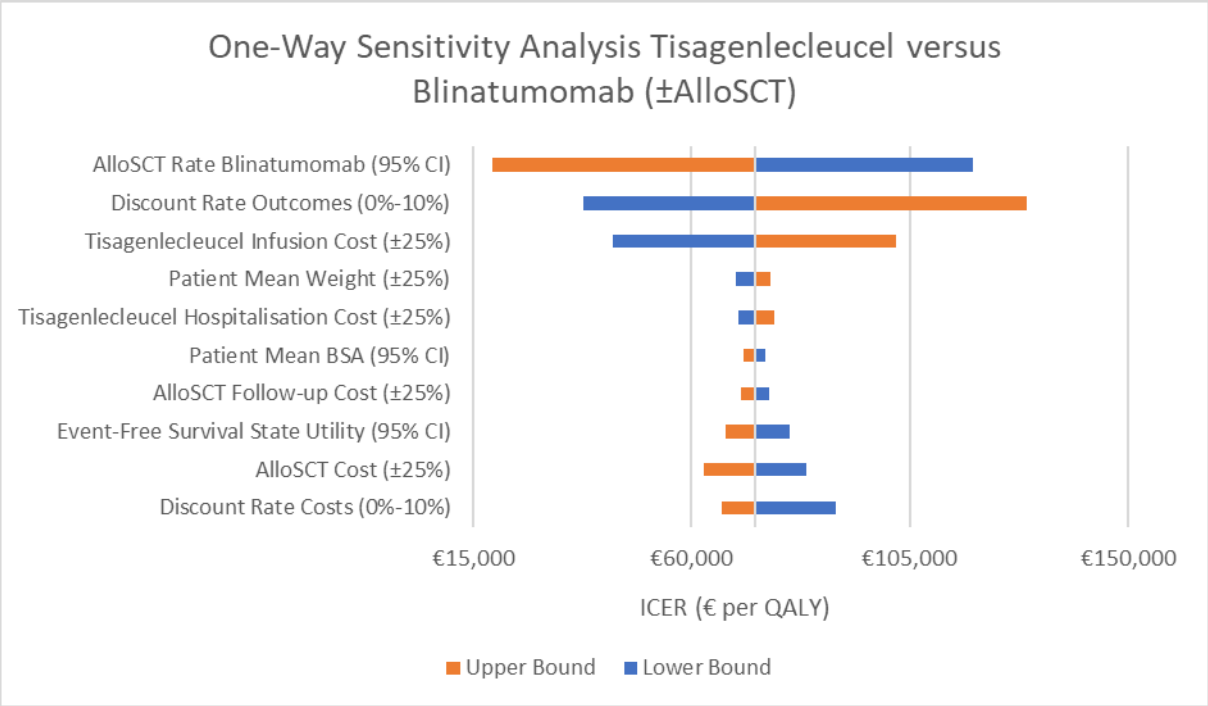

Figure 12 Tornado diagram of one-way sensitivity analysis of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia (base case ICER: €73,086 per QALY)
AlloSCT: Allogeneic stem cell transplant; **BSA:** Body surface area; **ICER:** Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year.

5.5.4 Scenario Analysis

Results of scenario analysis are presented in Table 28. Scenarios, which had the greatest impact on the ICER, are highlighted in bold.

Table 28 Impact of scenario analysis on deterministic ICER of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia*

Parameter/Assumption	Base Case	Scenario	Plausibility of Scenario	Scenario ICER (€/QALY) (base case ICER €73,086/QALY)
Time Horizon	88 years	2 years	Base case most plausible; represents a lifetime horizon	**369,621**
Proportion of Patients Receiving Tisagenlecleucel Infusion	83%	100%	Base case most plausible; represents clinical trial	69,253
Clinical Data Informing Tisagenlecleucel OS Efficacy	Pooled ELIANA and ENSIGN	ELIANA†	Base case most plausible; more mature data	76,115
Extrapolation of Pooled ELIANA and ENSIGN (tisagenlecleucel) OS Data	One-knot (odds) spline model	Log-normal model	Uncertain. More conservative option chosen for base case	56,570
Extrapolation of NCT01471782 (blinatumomab) OS Data	One-knot (odds) spline model	Log-normal model	Uncertain. More conservative option chosen for base case	58,262
Extrapolation of ELIANA, ENSIGN (tisagenlecleucel), and NCT01471782 (blinatumomab) OS Data	One-knot (odds) spline model	Log-normal model	Uncertain. More conservative option chosen for base case	**47,255**
Extrapolation of ELIANA EFS Data	Generalised gamma model	Two-knot (odds) spline model	Negligible impact on ICER	73,290
Time Point at which Patients are Considered Long-Term Survivors	After 60 months	After 24 months	Uncertain. More conservative option chosen for base case	60,090
		No long-term survival point	Base case most plausible; clinical opinion indicated that a cohort of patients survive long-term	**129,379**
HRQOL of Long-Term Survivors	All patients alive after 60 months assumed HRQOL equivalent to event-free survival state (0.80)	All patients alive after 60 months assumed HRQOL derived by Kwon et al. (0.90)	Uncertain; lack of published data. More conservative option chosen for base case	71,817
		All patients alive after 60 months		62,615

		assumed HRQOL equivalent to the age- and sex-matched general population (328)		
Disutility Associated with Select Adverse Events	Included	Excluded	Negligible impact on ICER	72,490
Disutility Associated with Treatment and All Adverse Events	Select adverse events included only	-0.42 (derived by Sung et al. (321))‡	Negligible impact on ICER	72,768
Duration of IV Immunoglobulin Treatment	Duration of event-free survival	11.4 months	Uncertain. Base case likely most plausible; correlation between B-cell aplasia and remission	60,010
Cycles of Blinatumomab Received	100% of patients receive one cycle, 33% receive a second cycle	100% of patients receive one cycle, 33% receive a second cycle, 11.4% receive a third cycle, 4.3% receive a fourth cycle, and 4.3% receive a fifth cycle (87)	Scenario a more accurate reflection, as survival data in base case were not modified to reflect the fewer cycles received	68,037
Dosing Regimen of Blinatumomab	50% of patients receive dosing based on body surface area and 50% receive fixed-dosing regimen	100% receive dosing based on body surface area	Uncertain; lack of published data. True ICER likely to be within this range	78,012
		100% receive fixed-dosing regimen		68,161
Proportion of Patients Receiving AlloSCT in Blinatumomab Arm	49%	35.7%|	Scenario a more accurate reflection, as survival data in base case were not modified to reflect higher rate of alloSCT	88,443
Proportion of Patients Receiving AlloSCT in Tisagenlecleucel Arm	0%	12%§	Base case most plausible; patients censored at time of alloSCT in ELIANA and ENSIGN	85,705

Discount Rate	4% on costs and outcomes	1.5% on costs and outcomes		55,630
		4% on costs and 1.5% on outcomes		50,260
		Hyperbolic discounting: 4% (0-30 years), 3.5% (31-60 years), 3% (61-100 years) on both costs and outcomes	Base case most plausible; reflects current practice	71,887

AlloSCT: Allogeneic stem cell transplant; **EFS:** Event-free survival; **HRQOL:** Health-related quality of life; **ICER:** Incremental cost-effectiveness ratio; **IV:** Intravenous; **OS**: Overall survival; **QALY:** Quality-adjusted life year.
*Scenarios that had the greatest impact on the ICER are highlighted in **Bold**.
†Log-logistic model was used to extrapolate OS of ELIANA. This provided similar extrapolation output to that generated by the one-knot (odds) spline model (used in the base case for ELIANA and ENSIGN). Provided good statistical and visual fit to the data.
‡Disutility (-0.42) applied for 28 days (tisagenlecleucel), 28 days (cycle 1; 100%) and 28 days (cycle 2; 33%) (blinatumomab). Disutility associated with cytokine release syndrome and non-cytokine release syndrome intensive care unit admission accounted for.
|No changes made to efficacy data. Change in ICER reflects decreased costs in the blinatumomab arm.
§No changes made to efficacy data. Change in ICER reflects increased costs in the tisagenlecleucel arm.

5.5.5 Price Analysis

A 28% decrease (including 5.5% rebate) on the price-to-wholesaler of tisagenlecleucel was required to reduce the deterministic ICER to a willingness-to-pay threshold of €45,000 per QALY. The probability of cost effectiveness with this price decrease, at this threshold, was 44%.

5.5.6 Expected Value of Perfect Information

At a threshold of €45,000 per QALY, the 10-year population EVPI was €314,455. The population EVPI of tisagenlecleucel versus blinatumomab, over a range of thresholds, is depicted in Figure 13.

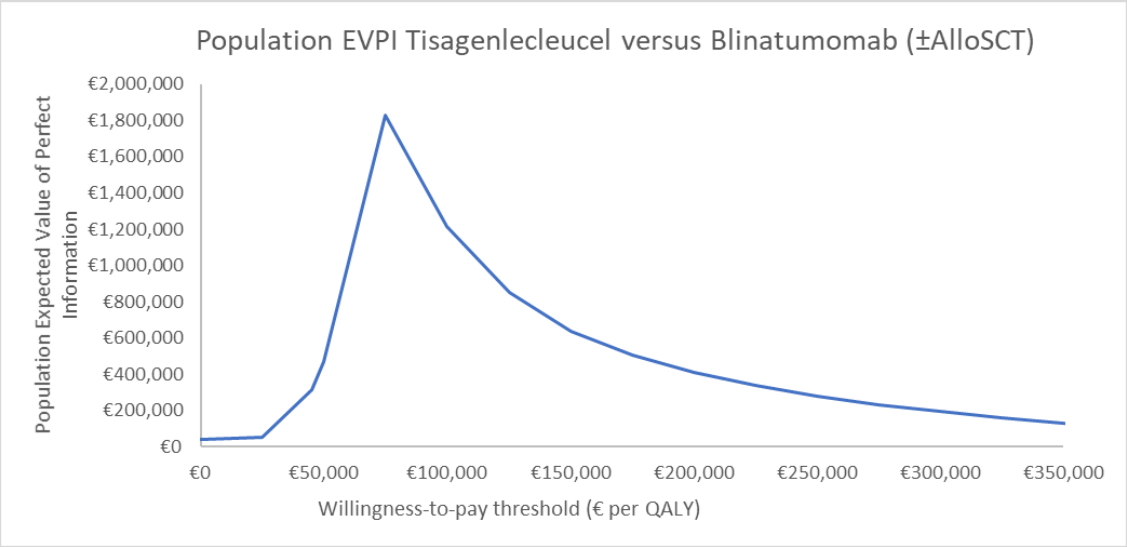

Figure 13 Population EVPI, over various willingness-to-pay thresholds, of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia
AlloSCT: Allogeneic stem cell transplant; **EVPI:** Expected value of perfect information; **QALY:** Quality-adjusted life year.

The population EVPI analysis was re-run at the price that reduced the ICER to €45,000 per QALY (€229,105; representing a 28% price decrease). At this price and threshold, the 10-year population EVPI was €1,149,810 (Figure 14).

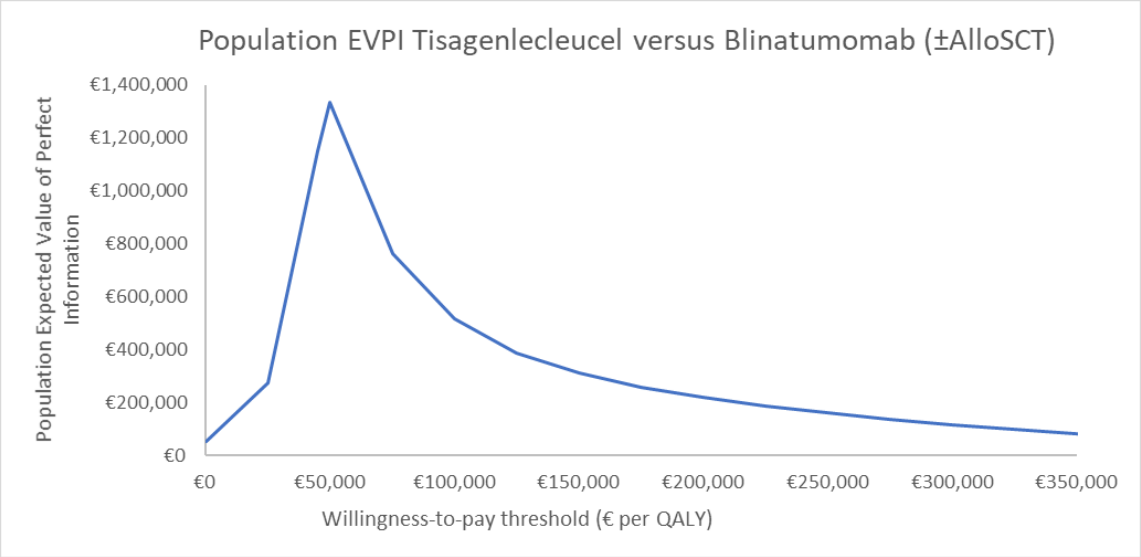

Figure 14 Population EVPI, over various willingness-to-pay thresholds, of tisagenlecleucel (price that reduced the ICER to €45,000 per QALY) versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia
AlloSCT: Allogeneic stem cell transplant; **EVPI:** Expected value of perfect information; **ICER**: Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year.

5.5.6.1 Partial Expected Value of Perfect Information

At the price-to-wholesaler of tisagenlecleucel (€301,762, including 5.5% rebate) and a willingness-to-pay threshold of €45,000 per QALY, the 10-year population EVPPI was

below €100,000 for each category of parameters. Parameters associated with survival analysis had the highest population EVPPI (€67,189), followed by alloSCT costs (€29,338). Population EVPPI of the remaining categories was low; €25,255, €18,649, and €1,215, for parameters associated with utility values, adverse event costs, and hospitalisation and monitoring costs, respectively. Figure 15 depicts the value of uncertainty associated with each parameter category.

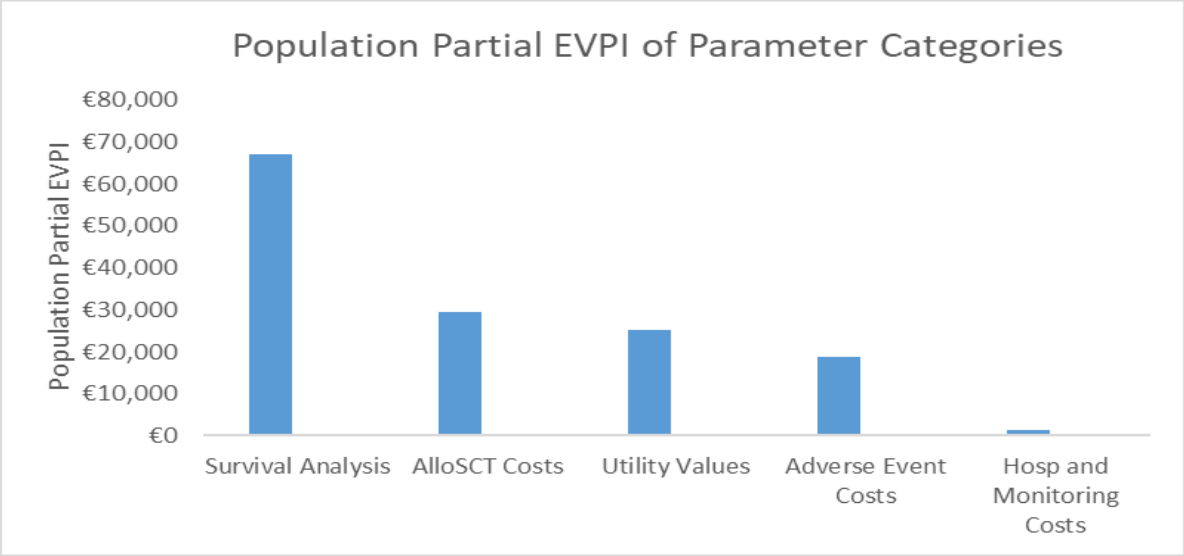

Figure 15 Population partial expected value of perfect information of parameter categories – tisagenlecleucel versus blinatumomab
AlloSCT: Allogeneic stem cell transplant; **EVPI:** Expected value of perfect information.

The population EVPPI analysis was re-run at the price that reduced the ICER to €45,000 per QALY (€229,105). At a threshold of €45,000 per QALY, population EVPPI was below €500,000 for each category of parameters. Parameters associated with survival analysis had the highest population EVPPI (€371,813), followed by alloSCT costs (€272,459). Hospitalisation and monitoring cost parameters had a population EVPPI of €211,894. Parameters associated with utility values and adverse event costs had the lowest population EVPPI; €133,375 and €50,222, respectively. Figure 16 depicts the value of uncertainty associated with each parameter category.

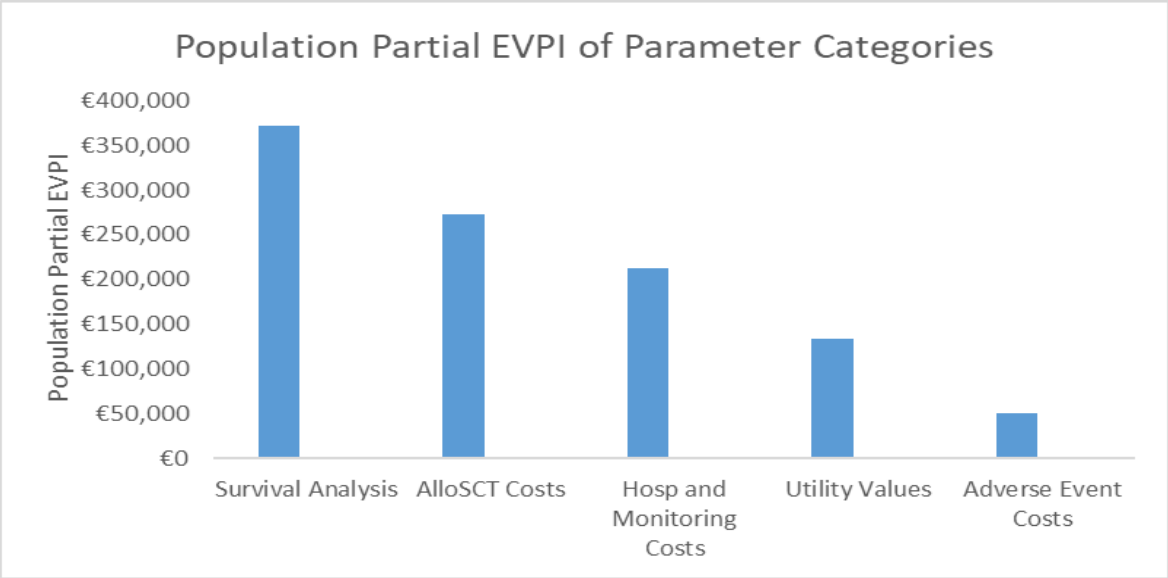

Figure 16 Population partial expected value of perfect information of parameter categories – tisagenlecleucel versus blinatumomab (price that reduced the ICER to €45,000 per QALY)
AlloSCT: Allogeneic stem cell transplant; **EVPI:** Expected value of perfect information.

5.6 Discussion

5.6.1 Deterministic and Probabilistic Results

This cost-utility analysis indicates that, at current prices-to-wholesaler, tisagenlecleucel is not cost effective, versus blinatumomab, at a €45,000 per QALY threshold. It is likely that the true ICER differs to that generated as part of this research, due to commercial patient access schemes that are in place for blinatumomab and tisagenlecleucel.

A high degree of uncertainty exists in the clinical evidence base of tisagenlecleucel for R/R ALL. This translates to uncertainty in cost-effectiveness, which may not be adequately captured by OWSA, PSA and EVPI. Although parameter uncertainty in the model was captured by PSA and EVPI, uncertainty associated with the naïve ITC is difficult to quantify. For immature survival data, such as that used in this analysis, true uncertainty lies in the data extrapolation and appropriate choice of survival model. Such uncertainty is generally not captured in the PSA and thus, EVPI analyses. Instead, variation associated with a particular parametric form is captured (378). As such, caution is warranted in the interpretation of results. The uncertainty captured in the outcomes is illustrated by the wide spread in incremental QALYs in the probabilistic scatterplot. Reassuringly, negligible differences were observed between the deterministic and probabilistic ICERs, due to non-linearity between parameter inputs and model outputs. For tisagenlecleucel (at price-to-wholesaler), the probability of cost effectiveness at the

€45,000 per QALY threshold was 16%. The probability of cost effectiveness of tisagenlecleucel exceeded that of blinatumomab at an approximate threshold of €80,000 per QALY, as illustrated by the cost-effectiveness acceptability curve.

5.6.2 One-Way Sensitivity and Scenario Analyses

Results of OWSA illustrated the sensitivity of the model to certain input parameters. Scenario analyses illustrated the impact of employing alternative, plausible assumptions.

5.6.2.1 Time Horizon

The model was sensitive to variations in the time horizon. This is expected considering the high upfront costs and potential long-term outcomes associated with tisagenlecleucel. Reducing the time horizon from 88 years to 2 years (approximate follow up of ELIANA (215) and ENSIGN (207)), increased the ICER 5-fold. Reducing the time horizon has been proposed to shift the risk associated with particular uncertainties from the health system to the Applicant (12). Although conducting this type of sensitivity analysis has been recommended in the literature (14, 176), relevant costs and outcomes are omitted. It may be difficult for the decision-maker to interpret the implications of a restricted time horizon.

5.6.2.2 Discount Rate

The high sensitivity to the discount rate on outcomes, in the OWSA, reflects the extent to which outcomes are accrued over the long-term in the model. The sensitivity to changes in the discount rate is expected due to the time divergence between costs and outcomes. Reducing the discount rate on outcomes to 0% (whilst maintaining 4% discount on costs), decreased the ICER to less than €45,000 per QALY. Reducing the discount rate to 1.5% on both costs and outcomes reduced the base case ICER by 24%. Applying differential discounting, whereby costs were discounted at 4% and outcomes were discounted by 1.5%, reduced the ICER by 31%. Hyperbolic discounting (reducing the discount rate over time) had negligible impact (2% reduction in ICER). This is in line with findings elsewhere (369). As described in 1.10.1.4, the recent authorisation of ATMPs, which are expected to provide long-term benefit, has renewed debate in the literature regarding the most appropriate discount rate. The overall impact of reducing the discount rate here was a

decrease in the ICER. In an analysis conducted by NICE, to understand the potential quantitative effect of altering the discount rate, appraisals for haematological cancers (n=5) were amongst those whose ICERs were most sensitive to changes in the discount rate (of all appraisals assessed, n=27) (369). Further research is warranted to explore the impact of alternative discounting approaches on a greater sample size, covering a range of technologies and disease areas. Any proposed changes to the discount rate in Ireland need to take into account the potential impact of reduced ICERs (i.e. more cost-effective technologies) on affordability. O'Mahony et al. highlight that, in such instances, the willingness-to-pay threshold must also be reduced to maintain rational allocation of resources (379). O'Mahony argues that the current threshold in Ireland lacks an appropriate evidence base and likely exceeds opportunity costs (380). This renewed interest in determining the most appropriate discount rate may also spark timely discussion on the willingness-to-pay threshold in Ireland.

5.6.2.3 Survival Extrapolation

The choice of the most appropriate model to extrapolate OS was uncertain. The fundamental challenge in this instance is the paucity of long-term data. This is compounded by the small patient numbers in each study. Spline models were chosen to extrapolate the OS data in the base case. However, the log-normal parametric model was also a reasonable option. Employing this model to extrapolate the OS data of tisagenlecleucel decreased the ICER by approximately €16,500 per QALY. This was driven by more favourable OS predictions generated by the log-normal model (for tisagenlecleucel). When the log-normal model was chosen to extrapolate the OS data of blinatumomab, a similar decrease in the ICER was observed. This decrease was driven by less favourable OS predictions generated by this model (for blinatumomab). This resulted in a lower total QALY gain for blinatumomab, with negligible impact on the total costs. Employing the log-normal model to extrapolate the OS data of both tisagenlecleucel and blinatumomab resulted in a notable reduction in the ICER (approximately €25,800 per QALY). Although the log-normal model overestimated OS of tisagenlecleucel, when compared to the observed Kaplan-Meier data (from pooled ELIANA and ENSIGN), the observed Kaplan-Meier data were based on very low numbers of patients left at risk towards the end of follow up.

Using only the ELIANA data to inform the efficacy of tisagenlecleucel increased the ICER by approximately €3,000 per QALY. Notably, the ELIANA data were less mature than ENSIGN, resulting in less favourable survival predictions for tisagenlecleucel.

5.6.2.4 Time Point of Long-Term Survival

Scenario analysis highlighted the impact of changing the time point (post-treatment) at which patients are considered to be long-term survivors. Reducing this time point from 60 months to 24 months had a notable reduction in the ICER (approximately €13,000 per QALY) and could have considerable implications in price negotiations between the HSE and Applicant. This was driven by an increase in incremental QALY gain. The impact on incremental costs was less notable.

The expert elicitation (Chapter 3) illustrated the variability amongst the cohort of experts (n=5) in terms of the long-term survival of patients treated with tisagenlecleucel. A 'worst case' (conservative) scenario, whereby outcomes were derived from full extrapolation of the trial data without additional structural assumptions regarding the point of long-term survival, had a sizeable impact on the ICER (approximately €56,000 per QALY increase). This illustrates the reliance of the cost effectiveness of tisagenlecleucel on the uncertain assumption of a time point of long-term survival. This increase in the ICER was driven by a decrease in the incremental QALY gain. The magnitude of change in the ICER, depending on the assumption employed, creates challenges for decision-makers. The uncertainty underpinning these assumptions creates considerable financial risk.

5.6.2.5 Health-Related Quality of Life Inputs

The event-free survival state utility value was a driver of cost effectiveness in the OWSA. This may partly be due to the fact that all patients alive after 60 months were assumed to have HRQOL of the event-free survival state. The majority of QALY gains in the tisagenlecleucel arm were driven by QALYs accrued in the extrapolation of survival. However, the magnitude of the impact on the ICER was small when compared to the impact of altering other parameters in the model.

Assuming patients, considered to be long-term survivors, had HRQOL equivalent to the age- and sex-matched general population decreased the ICER by approximately €10,000 per QALY. Assuming patients, considered to be long-term survivors, had HRQOL derived by Kwon et al. (0.90) (308), had less impact on the ICER. Notably, the utility derived by Kwon et al. is equivalent to the utility accrued by patients aged 53 years in the general population (derived from Ara and Brazier (328)). Uncertainty in the HRQOL of patients who are considered long-term survivors was highlighted in Chapter 4. Scenario analyses are important in illustrating the impact of this uncertainty. In the absence of robust, supportive data, conservative assumptions regarding HRQOL of long-term survivors are warranted.

5.6.2.6 Cost Inputs

Although a simple price decrease on tisagenlecleucel may reduce the ICER to an acceptable willingness-to-pay threshold (as demonstrated in Price Analysis, 5.5.5), it does not address the risk to payers and patients (due to uncertain clinical evidence).

Decreasing the duration of IV immunoglobulin treatment to 11.4 months, resulted in a decrease in the ICER of approximately €13,000 per QALY. Considering the uncertainty and associated impact on the ICER, the potential for prolonged treatment with IV immunoglobulin is an important consideration for decision-makers. Reimbursement agreements could incorporate this uncertainty by attaching conditions, whereby IV immunoglobulin treatment beyond a pre-agreed period is funded by the Applicant.

Assuming patients received up to five cycles of blinatumomab, as per NCT01471782, decreased the ICER by approximately €5,000 per QALY. This decrease is attributable solely to an increase in costs in the blinatumomab arm, as the efficacy data of blinatumomab were not modified to reflect the reduced number of cycles received (in the base case).

5.6.2.7 Allogeneic Stem Cell Transplant

The upper bound of the 95% CI of the proportion of patients receiving alloSCT in the blinatumomab arm (98%) reduced the ICER to below a threshold of €20,000 per QALY. When the proportion of patients receiving alloSCT in the blinatumomab arm was reduced

from 49% (derived from the expert elicitation), in the base case, to 35.7%, reflecting the proportion of patients who received alloSCT in NCT01471782, a notable increase in the ICER was observed (from €73,086 per QALY to €88,443 per QALY). It should be noted that changes in the ICER were attributable solely to changes in costs, as efficacy data were not modified to reflect the varying levels of patients receiving alloSCT. The subsequent ICERs are therefore, unlikely to be an accurate estimate of the true cost effectiveness. Increasing the proportion of patients who received alloSCT in the tisagenlecleucel arm from 0% (base case) to 12% (based on the proportion who received alloSCT in ELIANA and ENSIGN), increased the ICER by approximately €12,600 per QALY. Based on the expert elicitation, patients are not expected to receive alloSCT following tisagenlecleucel.

5.6.3 Expected Value of Perfect Information

At the price-to-wholesaler of tisagenlecleucel, EVPI indicated that the cost of further research should not exceed €314,455. Population EVPI reached a peak at a threshold of approximately €75,000 per QALY. At this peak, the probability of cost effectiveness of tisagenlecleucel was 48%. As the willingness-to-pay threshold increased (from the peak of €75,000 per QALY), the probability of cost effectiveness of tisagenlecleucel increased. At these higher probabilities of cost effectiveness, the corresponding consequences of decision uncertainty reduce, resulting in a reduction in the population EVPI (12). At low willingness-to-pay thresholds (less than €25,000 per QALY), the probability of cost effectiveness of tisagenlecleucel approached 0%. Consequently, there are no consequences of decision uncertainty at these lower threshold values and the population EVPI is low (approaching zero). Population EVPPI indicated that, at the price-to-wholesaler, parameters associated with survival analysis and alloSCT costs had the highest population EVPPI. If further research is conducted, these areas should be prioritised. However, these estimates were low. Conducting additional research to inform these parameters is unlikely to be of value. However, this does not mean that uncertainty surrounding these estimates is unimportant. Additionally, EVPI and EVPPI analyses examine uncertainty in parameters. Structural uncertainty, associated with the naïve ITC, was not captured.

Population EVPI and EVPPI analyses were re-run at the price of tisagenlecleucel that reduced the ICER to €45,000 per QALY. This analysis was conducted, as the price of tisagenlecleucel (available to the HSE) is likely to be lower than the price used in the base case. The price employed in this scenario represents the 'best case' scenario. Additionally, EVPI is greatest when the ICER is close to the willingness-to-pay threshold, we gain the most from resolving uncertainty when further research could materially impact on the decision. Under this scenario, the 10-year population EVPI, at a €45,000 per QALY threshold, increased. The cost of further research should not exceed €1,149,810. Parameters associated with survival analysis and alloSCT costs had the highest population EVPPI. These parameters were associated with the greatest decision uncertainty. The low population EVPPI estimates for these parameter categories indicate that if further research is conducted on these parameters, the investment made (in terms of cost) should be low. A possible approach to collecting additional information on these parameters could entail the establishment of a long-term registry. This registry could collect data on both the survival of patients and the proportion of patients who require alloSCT. Resources are unlikely to be available to the HSE to fund the required research. However, utilising the established NCRI may be a viable option. This would require communication between the HSE/DOH and the NCRI. Data sharing agreements, for the purpose of reimbursement, may be required.

The ranking of parameter categories in the population EVPPI analysis, conducted at the price of tisagenlecleucel that reduced the ICER to €45,000 per QALY, changed when compared to that conducted at the price-to-wholesaler of tisagenlecleucel. This suggests that uncertainty associated with the model decision becomes driven by different categories depending on the cost of tisagenlecleucel and subsequent estimates of cost effectiveness. The reasons for this change in ranking are not clear. Notably, the top two categories for research prioritisation were consistent between the two analyses.

Of note, modelled alloSCT costs, in the blinatumomab arm, were based on a higher rate of alloSCT than that observed in NCT01471782; efficacy was derived from the trial. This approach favours tisagenlecleucel. In NCT01471782, data were not presented separately for patients who did and did not proceed to alloSCT, precluding an analysis of survival benefit associated with alloSCT. In the absence of a structural link between alloSCT and

survival benefit, it is likely that the EVPPI analysis overstates the impact of uncertainty on alloSCT. This is because stochastic variability on this parameter impacts costs only.

The low population EVPI and EVPPI estimates are likely a reflection, to some degree, of the low patient numbers in the analysis (six patients per year). It should be noted that in this analysis, the sum of the population EVPPI estimates were lower than the total EVPI. This is likely due to correlations between parameters within the model (372).

5.6.4 Comparison with the Published Literature

The findings of the analysis undertaken here (i.e. not cost effective at a threshold of €45,000 per QALY) are in line with those of the NCPE in their HTA appraisal of tisagenlecleucel (88). Due to uncertainty, at the time of NCPE HTA appraisal, in the proportion of patients expected to receive alloSCT following treatment with tisagenlecleucel, a range of ICERs were presented in the NCPE Technical Summary (88). The deterministic ICER, versus blinatumomab, ranged from €75,748 per QALY (incremental costs €321,755; incremental QALYs 4.25) to €116,506 per QALY (incremental costs €457,033; incremental QALYs 3.92), depending on assumptions regarding the rate of alloSCT in the tisagenlecleucel arm (25% and 82%, respectively). Although the deterministic ICER obtained in the analysis undertaken here (€73,086 per QALY, incremental costs €156,928; incremental QALYs 2.15) was close to the lower bound ICER presented in the NCPE Technical Summary (88), the incremental costs and incremental QALYs generated in this analysis were notably lower. Due to the dearth of detail available from the NCPE Technical Summary, the main drivers of differences between the ICERs are unclear. Both the NCPE HTA appraisal and the analysis undertaken here derived estimates by means of naïve ITC. Notably, the discount rate on costs and outcomes reduced from 5% to 4% (employed in this research) since the NCPE HTA appraisal of tisagenlecleucel (88).

Dissimilar to the NCPE HTA appraisal, the impact of uncertainties regarding the rate of alloSCT in the tisagenlecleucel arm have been investigated in this analysis (expert elicitation, Chapter 3). The high rates of alloSCT (assumed to be either 25% or 82%) in the tisagenlecleucel arm of the NCPE HTA appraisal might bias the model against

tisagenlecleucel. This is because these high rates are not expected in clinical practice, inflating the cost associated with treatment. An additional strength of the analysis conducted here is the use of an updated data cut of ENSIGN (207), which was not available at the time of the NCPE HTA appraisal (88). More mature data are expected to generate more robust estimates of OS. However, the analysis presented in this Chapter is subject to limitations, not encountered in the NCPE HTA appraisal, due to the lack of publicly available raw IPD from ELIANA and ENSIGN. Extrapolation of survival outcomes, in this study, was based on reconstructed IPD, generated from digitised Kaplan-Meier curves. Such extrapolations are inevitably less accurate than those generated using the raw IPD from the relevant trial.

Results and conclusions generated from cost-utility analyses conducted in other jurisdictions may not be readily transferable to Ireland (381). Comparison of the results from this analysis with published analyses in the literature is therefore, limited. ICERs identified in the literature ranged from €28,829 per QALY (incremental costs €258,278; incremental QALYs 8.97), versus FLA-IDA (with or without alloSCT), in the Spanish healthcare setting (239), to $213,755 per QALY (US dollars; incremental costs $359,108; incremental QALYs 1.68), versus standard-of-care (not otherwise specified) in the Canadian healthcare setting (241). Assumptions and inputs in these analyses differed to those used in the analysis conducted here, particularly in terms of utility values and costs. Notably, the incremental QALY gain observed in this analysis (2.15), was lower than that estimated in two cost-utility analyses identified in the literature; 9.01 in Thielen et al. (295) and 6.22 in Moradi-Lakeh et al. (382). Both studies were sponsored by the manufacturer of tisagenlecleucel and authors had access to raw IPD (295, 382). Thielen et al. applied a lower discount rate on outcomes (1.5%) and adopted a societal perspective (295). Both studies employed different assumptions regarding HRQOL when compared to this research (295, 382).

A similar pattern of influential parameters was observed in the OWSA and scenario analyses, conducted as part of this research, compared with published cost-utility analyses. The discount rate on outcomes (382, 383), time horizon (295, 382, 383), time point of long-term survival (295), cost of tisagenlecleucel infusion (383), and rate of alloSCT (blinatumomab arm) (295, 382) were all key drivers of published cost-utility

models. This provides some degree of external validity to the results obtained here. None of the identified analyses conducted EVPI or EVPPI analyses.

5.6.5 Limitations

There are several limitations to this analysis. These are in addition to those discussed thus far. The model was highly sensitive to changes in input parameters, which reflected plausible, alternative assumptions. Uncertainty associated with the model inputs exerts considerable influence on the results. This complicates interpretation of results. While longer-term and real-world data will address uncertainties in model inputs, it is generally not reasonable to withhold a reimbursement decision until such data become available. The adoption of performance-based risk-sharing agreements will be valuable in managing the financial risk associated with this uncertainty.

In Chapter 2, the exclusion of study types, such as single-centre trials and expanded access programmes, was justified on the basis that these studies are subject to a greater degree of bias than prospective phase II studies. Excluding these study types may have resulted in the omission of relevant data, and limited the sample size of the data used in the model. The exclusion of these study types contrasts with the decision to include data derived from the expert elicitation exercise. As highlighted in Chapter 3, expert elicitation is inherently subjective and thus, subject to bias. Of note, however, is that data derived from the expert elicitation exercise were used primarily to validate model outputs. Thus, limiting the bias introduced into the model. Additionally, several techniques were implemented in the expert elicitation exercise to mitigate against the impact of bias and heuristics.

There was a high degree of uncertainty in the most appropriate model to extrapolate the OS and EFS data. Model averaging has been proposed as an appropriate approach to address structural uncertainty when several survival extrapolation models, which generate different survival predictions, provide an appropriate fit to the data (384, 385). This approach involves weighting competing, plausible extrapolation models, using measures of accuracy such as AIC or BIC statistics. The final estimation (e.g. OS) is based on a weighted average of each of the model estimates. This has been proposed to better

account for structural uncertainty, as basing the decision on a single “best fitting” model implies with full certainty that this single model is the most appropriate. However, the most appropriate model is rarely known with full certainty. Model averaging has been proposed to lead to better-informed decisions. Of note, however, is that this approach does not address the uncertainty in the long-term survival extrapolations. Weighting, based on AIC statistics for example, will only provide an indication of the appropriateness of each model fit to the observed data (384). It has also been shown that the use of alternative model prediction criteria (e.g. AIC versus BIC) may result in considerable differences in the weights assigned to the models, which may translate to an impact on the reimbursement decision (384, 386). Thus, this approach was not considered in this research.

Patients in the progressed disease state after 60 months were assumed to survive long-term. This was based on clinical opinion. This approach has been accepted by the NICE for reimbursement decision-making (333) and has been employed in the literature (239).

The model structure employed in this analysis improves on some cost-utility analyses in the literature in that it accounts for costs and outcomes of patients in ELIANA and ENSIGN who did not proceed to infusion with tisagenlecleucel (295, 387). This is expected to give a more realistic characterisation of the outcomes expected in clinical practice. This approach was deemed necessary as real-world evidence from the UK indicated that of 60 patients considered eligible for treatment with tisagenlecleucel, by the UK national CAR T-cell panel between November 2018 and July 2020, 49 proceeded to infusion (388). However, in the absence of detailed data on the outcomes of patients who did not proceed to infusion in ELIANA and ENSIGN, a number of assumptions were required. It is unclear if these assumptions are truly reflective of clinical practice.

A pragmatic approach was adopted, in that a beta distribution was applied to utility values in the PSA. This approach was used, as these values were considered to be sufficiently far from zero. However, a limitation of this approach is that it does not capture states considered to be “worse than death” (i.e. values below zero). Additionally, the application of the normal distribution to disutility values in the PSA may result in positive disutility values. As the normal distribution is symmetric about the mean, there

is a non-negligible probability of sampling values above zero. This lacks face validity. In instances where negative utility values are plausible, a more appropriate approach is to transform the data. Transformation of the disutility value, using the formula X=1-disutility, such that X is a disutility value, constrains the value on the interval zero to positive real number. A gamma or log-normal distribution can then be fitted (34).

It was assumed that zero cost was associated with the implementation of a CAR T-cell therapy service. However, it is acknowledged that considerable investment was likely required to establish such a service in Irish clinical practice. The costs estimated here are thus, likely to be underestimated. A consequence of this assumption is that the analysis more closely reflects the marginal costs of treating patients within an existing CAR T-cell therapy service, as opposed to the establishment of a new service.

In the absence of publicly available cost data specific to paediatric and young adult patients, costs relevant to the treatment of adult patients were employed. Paediatric patients may be subject to a greater degree of monitoring than adult patients, as they may not be able to verbalise symptoms (389). This may result in greater resource utilisation. Thus, associated administration, hospitalisation, initiation, monitoring, adverse event, and terminal care costs may be underestimated here. The costs associated with staff training were based on assumptions, and a crude approach to their estimation. It is difficult to conclude whether these are an accurate reflection of the true cost of training staff. It should be noted that, with the exception of the cost of alloSCT and associated follow up, costs associated with resource use were not main drivers of the model.

The inclusion of costs specific only to adverse events, which are expected to have considerable resource requirements, aimed to mitigate against double-counting. As tisagenlecleucel has an innovative mechanism of action and unique resource requirements, the risk of underestimating costs specific to adverse events and hospitalisation cannot be ruled out. However, costs associated with adverse events were not main drivers of the model. Additionally, it was assumed that patients, who experience an adverse event, only experience one incidence of that adverse event. This assumption was required due to the paucity of published relevant data. However, it is

acknowledged that patients may experience more than one incidence. Thus, costs and disutility associated with adverse events may be underestimated.

A limitation of the EVPI analysis is the arbitrary choice of a technology time horizon of 10 years. This assumption was aligned with those employed elsewhere (12, 378). In the absence of evidence, this assumption was a necessary one. Additionally, even if data are available to inform the time horizon, by means of evidence or a formal prior distribution, it will still remain a proxy (374).

5.7 Conclusion

The results of this cost-utility analysis indicate that tisagenlecleucel is not cost effective, versus blinatumomab, for the treatment of paediatric and young adult patients with R/R ALL in Ireland. Although tisagenlecleucel was associated with an incremental QALY gain, the clinical evidence supporting the model was highly uncertain. This uncertainty may not be adequately captured by OWSA, PSA, and thus, EVPI and EVPPI analyses. The model was highly sensitive to assumptions regarding long-term survival, creating challenges for decision-makers in the interpretation of results. EVPI and EVPPI analyses indicated that further research to decrease decision uncertainty (in parameters), at the defined willingness-to-pay threshold, may not be of value. Performance-based risk-sharing agreements may be a valuable approach in managing the financial risk associated with this uncertainty and should be investigated further.

Chapter 6 Efficacy of Tisagenlecleucel and Axicabtagene Ciloleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma: Systematic Literature Review and Evidence Synthesis

6.1 Introduction

The previous four Chapters of this thesis focused on the cost effectiveness of tisagenlecleucel for R/R ALL. The following Chapters will focus on deriving inputs for the bespoke cost-utility models, examining the cost effectiveness of tisagenlecleucel for R/R DLBCL, and axicabtagene ciloleucel for R/R DLBCL.

6.1.1 Chapter Aim

Relative efficacy estimates of tisagenlecleucel, and axicabtagene ciloleucel versus salvage chemotherapy are required to populate the bespoke cost-utility models (Chapter 9). The aim of this chapter is conduct an SLR to identify clinical evidence to inform the relative effectiveness of treatments for R/R DLBCL. Methods will follow those described in Chapter 2. Results will be used to inform input parameters for the cost-utility models evaluating the cost effectiveness of (i) tisagenlecleucel versus salvage chemotherapy, and (ii) axicabtagene ciloleucel versus salvage chemotherapy, for the treatment of R/R DLBCL.

6.2 Methods

6.2.1 Systematic Literature Review

An SLR protocol was developed in collaboration with NCPE Information Specialist, Ms Marie Harte, with reference to the Cochrane Handbook for Systematic Reviews of Interventions (204). Reporting is in line with PRISMA 2020 (205).

6.2.1.1 Population

The population was adult patients (18 to 80 years of age) with R/R DLBCL, who received two or more prior lines of systemic therapy. This is in line with the EMA licensed indications of tisagenlecleucel and axicabtagene ciloleucel (3, 4). Participants of any sex and any ethnicity were included.

6.2.1.2 Interventions

The interventions were tisagenlecleucel and axicabtagene ciloleucel, used as monotherapy at the EMA licensed dose (3, 4).

6.2.1.3 Comparators

Salvage chemotherapy regimens, with or without HSCT (herein 'salvage chemotherapy'), were relevant comparators and included:

- R-ESHAP
- R-DHAP
- R-GDP
- R-GEMOX
- R-GIFOX (rituximab - gemcitabine, ifosfamide, oxaliplatin)
- R-GEM-P (rituximab- gemcitabine- methylprednisolone)
- DA-R-EPOCH (dose-adjusted; rituximab - etoposide, doxorubicin, vincristine, cyclophosphamide, prednisolone)

Studies investigating salvage chemotherapy (not otherwise specified), best supportive care, and any of the included interventions, were also included.

6.2.1.4 Outcomes

The primary outcomes were:

- OS
- PFS
- EFS
- DFS (disease-free survival)

Outcomes were reported as Kaplan-Meier curves, facilitating digitisation of Kaplan-Meier curves and reconstruction of IPD (211).

Outcomes relating to response rates were extracted if they were the primary outcome of the trial. However, there is limited evidence to support their surrogacy for OS and these outcomes will not be directly used in the cost-utility models (206). Data on the proportion of patients with grade ≥3 adverse events and adverse events of specific interest were extracted. HRQOL data (defined by validated quality of life measures or instruments used in each trial) were extracted. These were also considered in a separate SLR, presented in Chapter 8.

6.2.1.5 Study Design

Prospective RCTs, phase II randomised trials, phase II non-randomised or single-arm trials, and prospective observational studies were included. Single-centre trials, expanded access programmes, retrospective studies, and case studies or reports were excluded. As highlighted in 2.2.1.5, these were excluded as they were considered to be of poorer quality than included study types, and subject to greater bias.

6.2.1.6 Search Methods

Electronic databases EMBASE, MEDLINE (via EBSCO), and CENTRAL (via the Cochrane Library) were searched from 01 January 2001 to 25 October 2019, as per the search strategies presented in Appendix E (Table A26). Proceedings from ASH and EHA Annual Conferences were hand searched for the years 2014 to 2019 inclusive. Terms used in searching of conference proceedings included: 'tisagenlecleucel', 'tisa-cel', 'axicabtagene ciloleucel', 'axi-cel', 'JULIET', 'ZUMA-1', 'diffuse large B-cell lymphoma', and 'DLBCL'. EPARs of tisagenlecleucel (143) and axicabtagene ciloleucel (144), and clinical trial reports from ClinicalTrials.gov (www.clinicaltrials.gov) were also searched. Articles were restricted to those published in English.

6.2.1.7 Citation Management

Identified citations were imported to Endnote®. Duplicates were systematically searched for using software in Endnote® and identified manually. Title and abstract screening was conducted by two separate reviewers; Screener 1 used Abstrackr and Screener 2 used Covidence®. A detailed discussion of the title and abstract screening process in presented in Chapter 7.

The full text of citations that were deemed 'relevant' or 'maybe' were obtained and assessed for suitability for inclusion in the final evidence base. For quality assurance purposes, 10% of full-text articles were screened in duplicate by a second reviewer. Data extraction was conducted using an adapted Cochrane data extraction form (208). Data recorded included population, intervention, comparators, outcomes, study design, authors, title, and publication date. Extracted outcomes data were checked in duplicate by a second reviewer.

6.2.2 Quality of Included Studies

Assessment of risk of bias in RCTs was pre-specified to be conducted using the Cochrane Risk of Bias Tool 2 (209). Risk of bias in non-randomised studies of two or more interventions was pre-specified to be assessed using the Newcastle-Ottawa Scale (189). Risk of bias and quality of included studies was assessed in duplicate by a second reviewer.

As described in 2.1.1, the adapted Newcastle-Ottawa Scale was chosen to assess the quality of single-arm studies (Table 29). The 'length of follow up' level in the Outcomes Domain requires the user to pre-specify an adequate follow-up period. For this research, this was specified as 60 months, in line with the literature, which suggests that most patients with DLBCL are expected to relapse within 24 to 60 months post-treatment (113, 390, 391). Additionally, for this research, one star was attributable to IRC assessment on the 'assessment of outcome' level. Outcome assessment based on any other method (e.g. record linkage) did not receive a star.

Table 29 Adapted Newcastle-Ottawa Scale quality assessment domains and scoring system (197-200)

	Selection Domain			**Outcomes Domain**		
Representativeness		**Ascertainment of Exposure**	**Demonstration that Outcome of Interest was not Present at Start of Study**	**Assessment of Outcome**	**Length of Follow Up**	**Adequacy of Follow Up**
Total Stars Achievable	1	1	1	1	1	1
Criteria	Generalisability of trial population to population with R/R DLBCL in Ireland	'Secure record' (i.e. not patient-reported)	Factors that could influence response to treatment or subsequent outcome (e.g. bridging chemotherapy)	IRC assessment	Minimum 60 months	Complete follow up
			Scoring			
Good Quality			6 stars			
Fair Quality			5 stars			
Poor Quality			4 stars or less			

DLBCL: Diffuse large B-cell lymphoma; **IRC:** Independent review committee; **R/R:** Relapsed/refractory.

6.2.3 Heterogeneity

Clinical and methodological between-trial heterogeneity were assessed qualitatively. Assessment of statistical heterogeneity, using I^2 or Q statistics, was pre-specified. Sources of heterogeneity were examined.

6.2.4 Indirect Treatment Comparison

Identified trials were assessed for inclusion in an ITC. Factors considered included the type of data identified (IPD or study-level; direct or indirect), the type of studies identified (RCT or single-arm), the number of studies, and heterogeneity of studies (165).

6.2.5 Reconstruction of Individual Patient-Level Data

To generate relative effectiveness estimates, IPD of the trials were reconstructed using the method described in 2.2.5.

6.2.6 Comparative Efficacy

Using reconstructed IPD, the probability of OS and PFS at 12 months in each trial was estimated (212). A 12-month point was chosen due to the limited duration of follow up in the trials. HRs for survival were estimated by fitting Cox proportional hazard models using the 'coxph' function of the 'Survival' package in R® (212).

6.2.7 Quality of Evidence for Outcomes

The GRADE framework was used to assess the quality of evidence for outcomes (213). Each outcome was assessed for quality independently. Further detail of the domains assessed are provided in 2.2.7. Summary of findings tables were generated using GRADEproGDT® software (214). Grading was assessed in duplicate by a second reviewer.

6.3 Results

Following exclusion of duplicates, 7,723 titles and abstracts underwent screening. Hand searching yielded an additional 32 citations. A total of 544 records were brought forward for full-text screening. Following full-text screening, eight records, reporting on three different studies, were included in the final evidence base, as outlined in Figure 17.

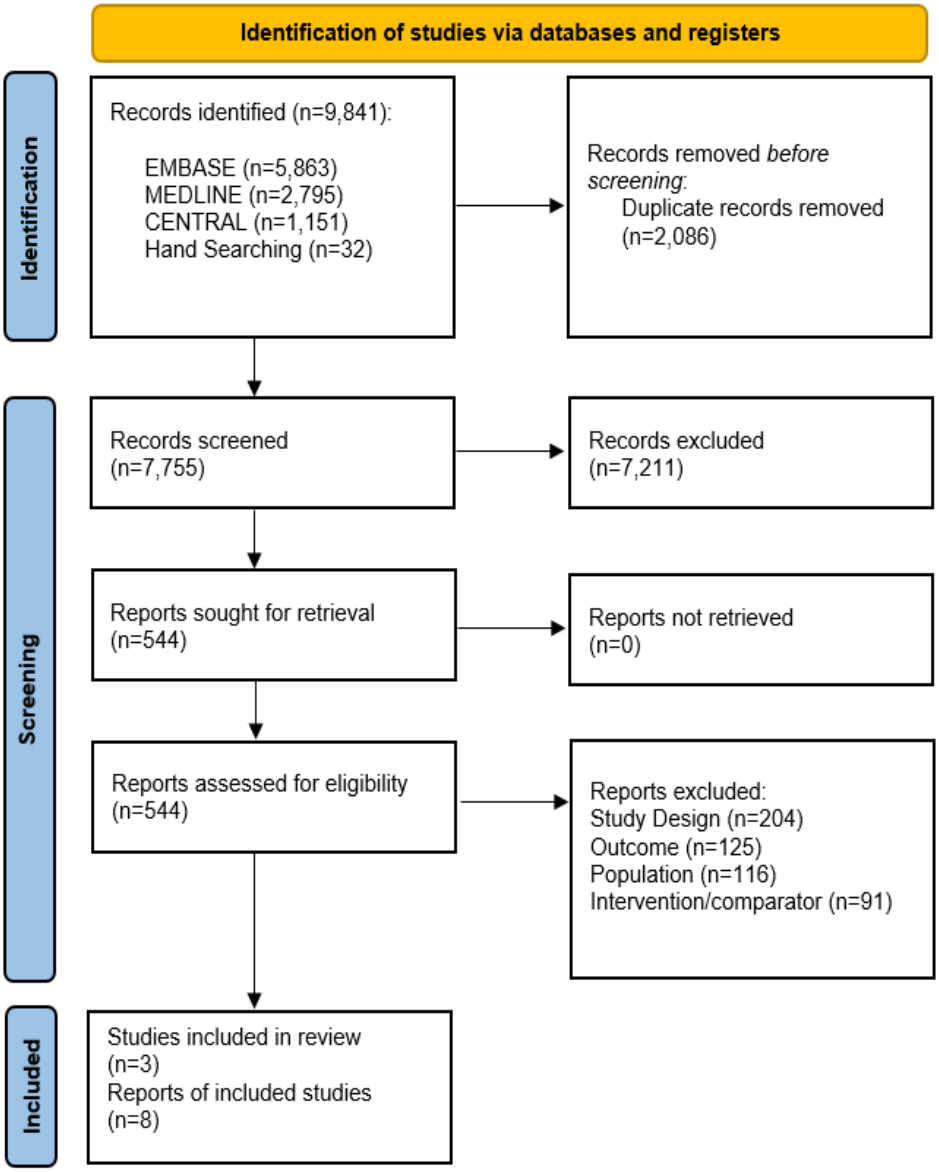

Figure 17 PRISMA diagram – systematic literature review of treatments for relapsed/refractory diffuse large B-cell lymphoma

6.3.1 Excluded Studies

A total of 536 studies were excluded at full-text screening. Reasons were study design (n=204), outcome (n=125), population (n=116), and intervention/comparator (n=91). A selected list of excluded studies and reasons for exclusion are presented in Appendix E (Table A27).

6.3.2 Included Studies

Studies included in the final evidence base for tisagenlecleucel and axicabtagene ciloleucel are summarised in Table 30. A number studies were identified as potential sources to inform the salvage chemotherapy arm. These are presented in Table 31.

Table 30 Summary of trials that met the inclusion criteria in the systematic literature review of treatments for R/R DLBCL|

Title (author, year)	Trial Design	Key Eligibility Criteria	Intervention (sample size)	Outcomes
JULIET (EMA 2018; Schuster et al. 2018; Schuster et al. 2018; Schuster et al. 2019) (143, 392-394)	Phase II, single-arm, open-label, multi-centre	≥18 years Relapsed or refractory* DLBCL or tFL ≥2 prior lines of therapy Failed/ineligible for autoSCT ECOG 0-1	Tisagenlecleucel once-off single IV infusion (mITT; n=115) $5x10^8$ CAR-positive viable T-cells (non-weight based)	**Primary:** ORR **Key Secondary:** PFS, OS HRQOL§
ZUMA-1 (EMA 2018; Locke et al. 2019; Neelapu et al. 2017) (144, 395, 396)	Phase I-II, single-arm, open-label, multi-centre	≥18 years Refractory† DLBCL, PMBCL, or tFL ≥2 prior lines of therapy ECOG 0-1	Axicabtagene ciloleucel IV infusion (mITT; n=101) $2x10^6$/kg Retreatment permitted under several conditions‡	**Primary (phase II):** ORR **Key Secondary (phase II):** PFS, OS HRQOL§

AutoSCT: Autologous stem cell transplant; **DLBCL:** Diffuse large B-cell lymphoma; **ECOG:** Eastern Cooperative Oncology Group; **HRQOL:** Health-related quality of life; **IV:** Intravenous; **mITT:** Modified intention-to-treat; **ORR:** Objective response rate; **OS:** Overall survival; **PFS:** Progression-free survival; **PMBCL:** Primary mediastinal large B-cell lymphoma; **R/R:** Relapsed/refractory; **tFL:** Transformed follicular lymphoma.

*Progressive or stable disease as best response to last line of therapy, or response status unknown.

†Progressed or stable (after at least four cycles of therapy) disease as best response to first-line therapy, progressed or stable (after at least two cycles of therapy) disease as best response to second- or greater lines of therapy, or refractory post-autoSCT.

‡Partial or complete response at month 3 but subsequently relapse; confirmation of CD19 tumour expression after disease progression and prior to retreatment; patient continues to meet eligibility criteria of ZUMA-1; no subsequent treatment for the treatment of lymphoma; no dose-limiting toxicity in phase I or comparable toxicity in phase II; toxicities relating to conditioning chemotherapy resolved (with the exception of alopecia); patient does not have known neutralising antibodies.

§HRQOL data were reported in a separate publication (not identified in this systematic literature review).

|Comparators not presented as all single-arm trials.

6.3.2.1 Tisagenlecleucel: JULIET

6.3.2.1.1 Survival Outcomes

One trial investigated the efficacy of tisagenlecleucel. JULIET is a single-arm, phase II trial in which patients with R/R DLBCL received infusion with tisagenlecleucel, at a dose of $5.0x10^8$ CAR-positive viable T-cells. JULIET comprised screening (leukapheresis and cell product acceptance), enrolment, bridging chemotherapy, lymphodepleting chemotherapy, tisagenlecleucel infusion, and primary safety and follow-up phases. As per Schuster et al. (2018), 165 patients enrolled in JULIET (ITT population), and 111 received infusion with tisagenlecleucel (mITT population). Four patients were awaiting infusion at the time of publication. Patient characteristics are presented in 6.3.4. The primary endpoint was ORR (proportion of patients with CR and partial response (PR), as per the Lugano classification (101)). PFS (time from infusion to date of progression or death from any cause), and OS (time from infusion) were key secondary endpoints. Efficacy data were primarily reported for the mITT population. ORR, in patients with at least 3 months follow up, was 52% (95% CI 41 to 62). No patients proceeded to HSCT (alloSCT or autoSCT) while in response. AlloSCT was received by five patients who did not achieve CR or PR, while one patient received autoSCT followed by alloSCT. At a median follow up of 14 months (range: 0.1 to 26), median OS was 12 months (95% CI 7.0 to NE). The 12-month OS was 49% (95% CI 39 to 59) among patients who received infusion. Median PFS was not reached for patients who achieved CR. The 12-month PFS was 83% among patients who had CR or PR at 3 months (392).

Median time from enrolment to infusion in JULIET was 54 days (range: 30 to 357); 50 patients discontinued from JULIET prior to infusion (manufacturing failure n=12; 'other reasons' n=38). These patients tended to have lower performance status than those who did receive infusion, and a greater proportion of patients with DLBCL that was refractory to the last therapy (392). Median OS from time of enrolment in the ITT population (n=165) was 8.3 months (95% CI 5.8 to 11.7). The 12-month OS was 40% (95% CI 32 to 49). These data were not reported for PFS (392).

An updated data cut of JULIET, presented as a conference abstract (ASH 2018 and European Society for Blood and Marrow Transplantation (EBMT) 2019), with an

additional 5 months of follow up (median follow up 19.3 months), was identified in the SLR (393, 394). Based on this data cut, median OS (presented as Kaplan-Meier) was 11.1 months (95% CI 6.6 to NE) in the mITT population (n=115, including 4 patients who were awaiting infusion at the time of publication by Schuster et al.). Median OS was not reached in patients who achieved CR (95% CI 21 to NE). The 18-month OS was 43% (95% CI 33 to 53), with a maximum follow up of 29 months. ORR was 54% (95% CI 43 to 64). PFS data were not reported (393, 394).

A further update of JULIET, based on a median follow up of 40.3 months, was presented at ASH 2020. Median OS in the mITT population (n=115) was 11.1 months (95% CI 6.6 to 23.9). OS at 12, 24, and 36 months was 48.2%, 40.4%, and 36.2%, respectively. Median OS in patients with CR (n=37) or PR (n=7) was not reached. PFS was presented separately for patients with Myc-positive and Myc-negative disease. Myc-positivity is an independent prognostic factor in DLBCL, associated with worse prognosis (397). PFS in patients with Myc-negative disease (n=38) was 6.2 months (95% CI 2.9 to NE), and 2.5 months (95% CI 1.7 to 3.0) in patients with Myc-positive disease (n=71). Data on ORR were not reported (398). As OS was not reported as Kaplan-Meier curves, these data could not be incorporated into the final evidence base. Additionally, this study was published outside the time frame of this SLR search period.

6.3.2.1.2 Adverse Events

As reported in Schuster et al. (392), grade ≥3 adverse events were reported in 89% of patients (mITT population, n=111). CRS was reported in 58% of patients; grade ≥3 (defined as per the Penn Grading Scale (220)) was noted in 22%. Admission to the ICU was required for 24% of patients experiencing CRS; 14% of patients received treatment with tocilizumab (16% based on updated data cut). Neurotoxicity was observed in 21% of patients; 12% experienced grade ≥3 severity. Concurrent CRS was observed in nine patients with grade ≥3 neurotoxicity. Grade ≥3 infections were reported in 20% of patients. Grade ≥3 prolonged cytopenias (greater than 28 days) were reported in 32% of patients (34% updated data cut). Grade ≥3 febrile neutropenia was reported in 15% of patients. IV immunoglobulin, administered at the local investigator's discretion, was administered to 30% of patients who received tisagenlecleucel (392) (33% based on

updated data cut (399)). Adverse event data reported at ASH 2018 and EBMT 2019 were very closely aligned with data presented by Schuster et al. (392-394).

6.3.2.1.3 Health-Related Quality of Life

Maziarz et al. (not identified in this SLR) describe HRQOL outcomes of patients in JULIET (400). Outcomes were collected using the FACT-Lym (Function Assessment of Cancer Therapy- Lymphoma) and the SF-36 (Short Form 36 Health Survey). Outcomes were collected at baseline (screening phase), and post-infusion at month 3, 6, 12, and 18. Most outcomes were reported by clinical responders (CR or PR). At a median follow up of 19 months, the rates of questionnaire completion (both clinical non-responders and responders) were 94% (108 of 115) at baseline, 76% (47 of 62) at month 3, 81% (35 of 43) at month 6, 86% (31 of 36) at month 12, and 65% (22 of 34) at month 18 post-infusion. Overall, all FACT-Lym domains, among clinical responders, had improved scores above the lower limit, minimal clinically important difference range, compared with baseline scores at all time points. The highest mean change from baseline occurred at the 18-month point for functional, physical, and social/family domains. Among patients who achieved CR or PR (n=57), SF-36 subscale scores surpassed the minimal clinically important difference at month 3, 6, 12, and 18 post-infusion for general health, vitality, physical functioning, role-physical, and social functioning. The SF-36 mental health subscale demonstrated numeric improvement in the mean changes from baseline at month 3, 6, and 12 post-infusion, but did not exceed the minimal clinically important difference (400).

Further critique of JULIET is provided in 6.3.3 and 6.3.4.

6.3.2.2 Axicabtagene Ciloleucel: ZUMA-1

6.3.2.2.1 Survival Outcomes

Efficacy of axicabtagene ciloleucel was evaluated in one trial. ZUMA-1 is a single-arm, phase I-II trial in which 108 patients received infusion with axicabtagene ciloleucel at a dose of $2x10^6$ per kg (body weight). Efficacy data are reported for patients who received infusion in the phase II cohort (mITT; n=101). ZUMA-1 comprised screening, leukapheresis and enrolment, lymphodepleting chemotherapy, axicabtagene ciloleucel

infusion, and primary safety and follow-up phases. At the time of publication (Locke et al. 2019), 111 patients enrolled in phase II of ZUMA-1 (ITT population), and 101 received infusion with axicabtagene ciloleucel (mITT population). Patient characteristics are presented in 6.3.4. Median time from leukapheresis to delivery of axicabtagene ciloleucel to the treatment facility was 17 days. The primary endpoint was ORR (proportion of patients with CR or PR, using the International Working Group Response Criteria for Malignant Lymphoma (401)). PFS (time from infusion to the date of disease progression or death from any cause), and OS (time from infusion) were key secondary outcomes. A total of 10 patients discontinued from ZUMA-1 prior to infusion (death n=3; adverse event n=5; non-measurable disease n=2); these are not included in the analysis. According to investigator-assessment, ORR was 83% (95% CI 72 to 89). AlloSCT was received by 2 of 39 patients with ongoing response. At a median follow up of 27.1 months (IQR 25.7 to 28.8), median OS was not reached (95% CI 12.8 to NE). OS at 24 months was 51% (95% CI 40 to 60). Median investigator-assessed PFS was 5.9 months (95% CI 3.3 to 15.0) (395). Nine patients were retreated with axicabtagene ciloleucel. Of these, five responded (two CR and three PR), and two of these patients had ongoing response (396).

A further update of ZUMA-1 was presented at ASH 2020, representing a median of 39.1 months follow up. Median OS in the mITT population was 25.8 months, with a 36-month OS of 47%. Data on ORR and PFS were not reported (402). OS data were not reported as Kaplan-Meier curves, precluding the reconstruction of IPD. Additionally, this study was published outside the time frame of this SLR search period.

6.3.2.2.2 Adverse Events

Safety data were reported for the mITT population in the phase I (n=7) and phase II (n=101) cohorts of ZUMA-1. Grade ≥3 adverse events were reported in 98% of patients. CRS (defined as per the Lee Grading Scale (403)) was reported in 92% of patients, with grade ≥3 in 11%. The proportion of patients requiring ICU admission or treatment with tocilizumab was not reported. Neurotoxicity was reported in 67% of patients; grade ≥3 in 32%. Grade ≥3 febrile neutropenia was reported in 33% of patients. Grade ≥3 infections were reported in 28% of patients. A total of 17% of patients had grade ≥3 cytopenias at 3

months or later, including neutropenia (11%), thrombocytopenia (7%), and anaemia (3%). Cumulatively, 31% of patients received IV immunoglobulin therapy, per treating investigators' discretion (395).

6.3.2.2.3 Health-Related Quality of Life

An ad hoc HRQOL analysis, based on data collected from the safety management cohort of ZUMA-1, was presented at the 44th Annual Meeting of the EBMT (2018). Outcomes were collected using the EQ-5D-5L, which was administered at screening (n=33), and at week 4 (n=27), month 3 (n=20), and month 6 (n=7) post-infusion. The mean EQ-5D-5L score was 0.80 (SD 0.17) at screening, which decreased to 0.74 (SD 0.15) at week 4 post-infusion. This increased to 0.82 (SD 0.21) at month 6. When grouped by health states, the mean EQ-5D-5L score was 0.80 (SD 0.14) for the progression-free survival state and 0.72 (SD 0.17) for the progressed disease state. A disutility of -0.05 (SE 0.04) at week 4 was associated with the timing of 'axicabtagene ciloleucel-related toxicities' (404).

Further critique of ZUMA-1 is provided in 6.3.3 and 6.3.4.

6.3.2.3 Salvage Chemotherapy

Several studies were identified, which were considered to be potential relevant sources to inform the efficacy of salvage chemotherapy. These are summarised in Table 31, with a detailed discussion provided after. The evidence base of salvage chemotherapy was limited by the number of retrospective and single-centre studies, studies that did not report the outcomes in the required format, and studies, which evaluated outcomes in patients who were not as heavily pre-treated as the population of relevance to this SLR.

Table 31 Summary of studies identified relating to salvage chemotherapy in the systematic literature review of treatments for R/R DLBCL

Title (author, year)	Trial Design	Key Eligibility Criteria	Intervention	Outcome
SCHOLAR-1 (Crump et al., 2017) (108)	Meta-analysis	Refractory* DLBCL, PMBCL, or tFL Prior anti-CD20 monoclonal antibody and an anthracycline	Salvage chemotherapy (±HSCT), not otherwise specified (n=636)	ORR OS§ HRQOL or safety data were not reported
CORAL Extension 1 (van den Neste et al., 2016) (123)	Observational	CD20+ relapsed/refractory DLBCL Enrolled to CORAL RCT but did not proceed to per-protocol autoSCT Candidates for third-line regimen	Salvage chemotherapy (±HSCT) (n=203)†: ICE-based (n=31) DHAP-based (n=30) Gemcitabine-based (n=23) Dexa-BEAM (n=15) CHOP-based (n=14) Miscellaneous (n=53) Treatment doses and duration not reported‡	ORR OS§ HRQOL or safety data were not reported
Mounier et al., 2013 (130)	Phase II, multicentre, open-label	18-75 years CD20+ relapsed (first or second relapse)/refractory DLBCL Not eligible for high-dose therapy ECOG 0-2	R-GEMOX (n=49) 4 treatment cycles (up to 8 cycles if patient achieved PR after 4 cycles)	**Primary:** ORR **Key Secondary:** OS, PFS HRQOL data were not reported
Witzig et al., 2008 (405)	Phase II, multicentre, open-label	Relapsed/refractory CD20+ NHL Suitable for treatment with a platinum-based regimen ECOG 0-2	R-DHAP (n=57) 2 treatment cycles	**Primary:** ORR **Key secondary:** OS, EFS HRQOL data were not reported

AutoSCT: Autologous stem cell transplant; **CHOP:** Cyclophosphamide, doxorubicin, vincristine, prednisolone; **Dexa-BEAM:** Dexamethasone, carmustine, etoposide, cytarabine, melphalan; **DHAP:** Dexamethasone, high-dose cytarabine, cisplatin; **DLBCL:** Diffuse large B-cell lymphoma; **ECOG:** Eastern Cooperative Oncology Group; **EFS:** Event-free survival; **HRQOL:** Health-related quality of life; **HSCT:** Haematopoietic stem cell transplant; **ICE:** Ifosfamide, carboplatin, etoposide; **NHL:** Non-Hodgkin's lymphoma; **ORR:** Objective response rate; **OS:** Overall survival; **PFS:** Progression-free survival; **PMBCL:**

Primary mediastinal large B-cell lymphoma; **PR:** Partial response; **R-DHAP:** Rituximab, dexamethasone, cytarabine, cisplatin; **R-GEMOX:** Rituximab, gemcitabine, oxaliplatin; **R/R:** Relapsed/refractory; **tFL:** Transformed follicular lymphoma.
*Progressive disease (received ≥4 cycles of first-line therapy) or stable disease (received 2 cycles of later-line therapy) as best response to chemotherapy or relapse ≤12 months after autoSCT.
†Data not available for all patients.
‡Three or more cycles were received by 135 patients; 56 received two or less cycles.
§Publication did not differentiate between primary and secondary outcomes (i.e. these were not specified).
|Comparators not presented as all single-arm trials.

6.3.2.3.1 SCHOLAR-1

SCHOLAR-1 is a meta-analysis, which pooled patient-level data from subgroups of two phase III trials (Lymphoma Academic Research Organization-CORAL, n=170; Canadian Cancer Trials Group LY.12, n=219) (127, 128) and two observational studies (MD Anderson Cancer Center, n=165; University of Iowa/Mayo Clinic Lymphoma Specialized Program of Research Excellence, n=82) (113, 406). It included patients with refractory DLBCL, PMBCL, and transformed follicular lymphoma (tFL). Refractory was defined as progressive or stable disease less than 6 months, as best response to last line of chemotherapy, or relapse 12 months or less after autoSCT. Salvage chemotherapy regimens were not reported. OS data were available for 603 patients. ORR (determined by the 1999 International Working Group Response Criteria (407)), among the pooled cohort was 26% (95% CI 21 to 31). Median OS from the start of salvage chemotherapy for refractory disease was 6.3 months (95% CI 5.9 to 7.0). The 1-year and 2-year survival was 28% and 20%, respectively. OS was similar regardless of refractory subgroup, with a slightly lower median OS among patients who were refractory to second- or later-line therapy, or who relapsed 12 months or less after autoSCT (6.1 and 6.2 months, respectively), than among primary refractory patients (7.1 months) (108).

SCHOLAR-1 only included patients who had refractory disease and those who had relapsed within 12 months of autoSCT. Patients who relapse within a short time frame after autoSCT have poor outcomes (408). Additionally, SCHOLAR-1 included patients (14%) with an ECOG performance status between 2 and 4; these patients were excluded from JULIET (392) and ZUMA-1 (395). The SCHOLAR-1 cohort appears to be enriched with patients who have high-risk disease (409). Patients were not as heavily pre-treated as those in JULIET and ZUMA-1. At least 2 prior regimens were received by 96% of patients in JULIET (392) and 99% in ZUMA-1 (40% received at least 4 prior regimens (144)); 49% of patients in SCHOLAR-1 received 2 to 3 prior regimens (less than 1% received 4 or more

prior regimens) (108). The rate of autoSCT in SCHOLAR-1 was higher than is expected for the relevant population in Irish clinical practice (29.9% versus approximately 15%[11]) (108). Additionally, at the line of therapy relevant to this SLR (i.e. after two or more lines of systemic therapy), patients in Irish clinical practice are expected to receive alloSCT (as opposed to autoSCT).

There was a high degree of missing data in SCHOLAR-1. Data on ECOG performance status were missing for 13% of patients, while IPI score was missing for 18% (108). The EMA highlighted a number of concerns regarding heterogeneity and the appropriateness of pooling studies in SCHOLAR-1: retrospective (observational) versus prospective (RCT) collection of data; differences in inclusion criteria (unselected patients in the observational cohorts versus patients eligible for autoSCT in the randomised cohorts); different time point at which patients were included (time of primary refractoriness versus refractory to second or later-line therapy); different response assessment (local versus investigator); potential differences in follow-up schedule (limited information); and potential differences in the management of patients (i.e. who are considered eligible for second HSCT (either alloSCT or autoSCT); limited information) (143).

6.3.2.3.2 CORAL Extension 1

CORAL Extension 1 is an observational study reporting outcomes of a patient cohort (n=203) who were initially enrolled in the CORAL (n=477) phase III, RCT (herein 'CORAL RCT'). CORAL RCT examined R-ICE versus R-DHAP in patients in first relapse, or who were refractory to first-line therapy. Responding patients proceeded to per-protocol autoSCT (127). CORAL Extension 1 comprises patients who did not proceed to per-protocol autoSCT (n=203) in CORAL RCT. Patients in this cohort went on to receive third-line therapy. ORR in this cohort was 39%. OS was measured from the time of failure of induction therapy (either R-ICE or R-DHAP in CORAL RCT) until death due to any cause in CORAL Extension 1. Median duration of follow up was 30.1 months. Median OS in CORAL Extension 1 was 4.4 months; 11.1 months in those who received HSCT (both autoSCT and alloSCT) and 3.3 months in those who did not. The 1- and 2-year survival was 23% and

[11] Oral correspondence with one consultant haematologist in Ireland.

15.7%, respectively. OS was not significantly different according to the type of treatment received (123).

CORAL Extension 1 had comparable patient populations to JULIET and ZUMA-1 in terms of IPI score (data available for 57% of patients in CORAL Extension 1), age, and sex. Patients in JULIET (392) and ZUMA-1 (395) were more heavily pre-treated than those in CORAL Extension 1; 100% of patients in CORAL Extension 1 being treated at third-line. The proportion of patients receiving HSCT (autoSCT and alloSCT) in CORAL Extension 1 was higher than expected in clinical practice; 31.5% versus approximately 15%. A proportion of patients were in response (CR n=26; PR n=30) at the time of withdrawal from CORAL RCT. Patients withdrew from CORAL RCT for reasons other than disease progression (e.g. toxicity; failure to mobilise stem cells); this could potentially confound results. Subgroup analysis found that OS was not significantly different according to the reason for withdrawal from CORAL RCT. However, analysis was based on small patient numbers in each category. CORAL Extension 1 was also subject to a high degree of missing data (123).

A separate study, CORAL Extension 2 (n=75), examining outcomes in patients who proceeded to per-protocol autoSCT in CORAL RCT, but subsequently relapsed and received third-line treatment, was also identified in the SLR. OS, measured from the time of relapse after autoSCT, was the primary outcome. Median OS was 10 months, with a 1-year survival of 39% (124). This study was excluded from this SLR, as it was a retrospective study.

6.3.2.3.3 Mounier et al.

Mounier et al. evaluated the efficacy of R-GEMOX in patients with DLBCL (n=49) in 10 institutions in France. A total of 14% (n=7) of patients were in second relapse; the remaining patients were primary refractory (12%) or in first relapse (74%). Prior rituximab was received by 63% of patients. ORR (defined as the rate of CR, unconfirmed CR and PR), was the primary endpoint. PFS (no definition provided) and OS were key secondary endpoints. ORR was 61% (95% CI 45 to 74). At a median follow up of 65 months, 5-year PFS was 13% (95% CI 5 to 24) and 5-year OS was 14% (95% CI 6 to 26).

Results were not presented separately for patients in first and second relapse. Frequently reported grade ≥3 adverse events included neutropenia (73%) and platelet toxicity (44%) (130).

Rituximab is routine care for first-line treatment of patients with DLBCL (98). As only a proportion (63%) of patients in Mounier et al. received prior rituximab, this cohort are not reflective of current clinical practice in Ireland. Additionally, inclusion of patients in first relapse (74%) results in a high proportion of patients who are not as heavily pre-treated as those in JULIET (392) or ZUMA-1 (395). As such, no further consideration was given to this study.

6.3.2.3.4 Witzig et al.

Witzig et al. evaluated efficacy when four doses of rituximab were added to DHAP (dexamethasone, cytarabine, cisplatin) in patients with NHL that was refractory or in relapse. Patients with DLBCL accounted for 54% of the cohort. Other disease subtypes included: follicular lymphoma (21%), mantle cell lymphoma (7%), and small lymphocytic lymphoma (5%). The mean number of treatments received was 2.2; 33% received prior rituximab. EFS (time from study registration to disease progression, initiation of chemotherapy (other than R-DHAP), or death), and OS were key endpoints. Median EFS was 5.3 months (95% CI 3.9 to 11.0). Median OS was 30.5 months (95% CI 17.8 to 60.6). Commonly reported grade 3-4 adverse events included thrombocytopenia (91%), neutropenia (79%) and febrile neutropenia (23%) (405).

Witzig et al. also included a high proportion (67%) of patients who had not received prior rituximab. Additionally, results were not presented separately for disease subgroups (405). No further consideration was given to this study.

6.3.2.3.5 Choice of Study to Use in Analysis

Both SCHOLAR-1 and CORAL Extension 1 are subject to limitations. Neither report on a patient cohort who are as heavily pre-treated as those in JULIET and ZUMA-1. Both had higher rates of HSCT (autoSCT in SCHOLAR-1; autoSCT and alloSCT in CORAL Extension 1) than is expected in Irish clinical practice, which may result in an overestimation of efficacy in patients at this line of therapy. Publications in the literature urge caution in

the use of SCHOLAR-1 as a benchmark for prospective trials in refractory DLBCL (409, 410). This is due mainly to the poor prognosis of patients included in SCHOLAR-1. Poor reporting of patient characteristics in CORAL Extension 1 limits conclusions that can be drawn regarding comparability with JULIET and ZUMA-1.

CORAL Extension 1 presents results separately for patients who did and did not receive HSCT. In the NICE HTA appraisal of tisagenlecleucel (TA567), the ERG utilised this to explicitly address the uncertainty surrounding the rate of HSCT in clinical practice (333). Conditional survival curves were constructed to estimate survival at different rates of HSCT. Using this method, survival curves can be reconstructed to reflect the rate of HSCT in Irish clinical practice. In their assessment, the EMA concluded that the CORAL Extension Studies (1 and 2) were the most relevant dataset to generate comparisons with tisagenlecleucel. Patients in the CORAL Extension Studies (1 and 2) were considered to have better prognosis, than those in SCHOLAR-1, as they were all initially considered for autoSCT in CORAL RCT (143). An editorial on CORAL Extension 1 concluded that this should constitute the control arm against which new drug combinations should be tested (411). As such, CORAL Extension 1, using the methodology employed by the NICE ERG (TA567) to model the rate of HSCT, was chosen to inform the salvage chemotherapy arm.

Further critique of CORAL Extension 1 is provided in 6.3.3 and 6.3.4.

It is acknowledged that exclusion of CORAL Extension 2 and SCHOLAR-1 results in the omission of a relevant cohort of patients (i.e. those who relapse following autoSCT). Acknowledging the uncertainty in the evidence, a comparison using the SCHOLAR-1 data will be explored as a scenario analysis in the cost-utility models.

6.3.3 Quality of Included Studies

A detailed description of the levels covered by the Newcastle-Ottawa Scale is presented in Appendix A. As per the adapted Newcastle-Ottawa Scale, all three studies (JULIET, ZUMA-1, CORAL Extension 1) were graded as poor quality; all scored four stars or less (Table 29 and Table 32).

All three studies achieved one star in the 'representativeness' and 'ascertainment of exposure' levels (Selection Domain). All studies pertained to populations that were 'somewhat representative of the average R/R DLBCL population in the community' (one star). Ascertainment of exposure in all studies was considered to be 'secure'; data were reported by acting clinicians (i.e. not patient-reported). ZUMA-1 was the only study to obtain a star in the 'demonstration that outcome of interest was not present at start of study' level. This was because a proportion of patients in both JULIET and CORAL Extension 1 were in response when they received interventional treatment (123, 143).

JULIET obtained one star in the 'assessment of outcome' level (Outcomes Domain), due to IRC assessment of outcomes. Outcomes in ZUMA-1 and CORAL Extension 1 were not IRC-assessed and therefore, did not obtain a star in this level. The follow-up period was not considered sufficient in any of the studies. This was due to short follow-up periods, which were all less than 60 months. No study scored a star on the 'adequacy of follow up' level. JULIET and ZUMA-1 reported outcomes for the mITT population (392, 395); follow up and reporting of the ITT population was not adequate. CORAL Extension 1 scored zero stars as PFS data were not collected (123).

The grade obtained by each trial is presented in Table 32. The highly subjective nature of the quality assessment is highlighted; results should be interpreted in this context.

Table 32 Quality assessment of studies included in systematic literature review of treatments for relapsed/refractory diffuse large B-cell lymphoma using the adapted Newcastle-Ottawa Scale*† (197-200)

	Selection Domain			Outcomes Domain			Final Grade‡
	Representativeness	Ascertainment of Exposure	Demonstration that Outcome of Interest was not Present at Start of Study	Assessment of Outcome	Was Follow Up Long Enough for Outcomes to Occur?	Adequacy of Follow Up	
JULIET (143, 392-394)	1	1	0	1	0	0	Poor Quality
ZUMA-1 (144, 395, 396)	1	1	1	0	0	1	Poor Quality
CORAL Extension 1 (123)	1	1	0	0	0	0	Poor Quality

*Level assessing the 'selection of the non-exposed cohort' in the Selection Domain and the entire Comparability Domain of the Newcastle-Ottawa Scale are excluded (189).

†1= 1 star obtained in this level, 0= 0 stars obtained in this level.

‡**Good quality**= 6 stars; **fair quality**= 5 stars; **poor quality**= 4 stars or less.

6.3.4 Heterogeneity

The naïve, unadjusted nature of the comparisons (described in 6.3.5.3) precluded the use of statistical measures of heterogeneity, such as I^2 and Q statistics. Investigation into sources of heterogeneity, by means of meta-regression, was also ruled out. A qualitative assessment of between-trial clinical and methodological heterogeneity was therefore, conducted. A summary of patient characteristics of each trial is presented in Table 33, with a detailed discussion provided below.

Table 33 Baseline characteristics of patients in the trials included in systematic literature review of treatments for relapsed/refractory diffuse large B-cell lymphoma

Characteristic	JULIET (tisagenlecleucel) (143, 392)	ZUMA-1 (axicabtagene ciloleucel) (144, 395)	CORAL Extension 1 (salvage chemotherapy ± HSCT) (123)
Median age, years (range)	56 (22-76)	58 (51-64)	55 (19-65)
Male, %	61	67	61
IPI, %:			
<2	28	27	30
≥2	72	73	70
ECOG Performance Status, %:			
0	55	42	NR
1	45	58	NR
Disease stage, %:			
I–II	24	15	NR
III-IV	76	85	NR
Prior HSCT, %	49	25	NR
Disease status, %:			
Refractory	55	79	NR
Relapse	45	21	NR
Bone marrow involvement, %	7	11	NR

ECOG: Eastern Cooperative Oncology Group; **HSCT:** Haematopoietic stem cell transplant; **IPI:** International Prognostic Index; **NR:** Not reported.

6.3.4.1 Clinical Heterogeneity

Median age was similar in all studies. However, variation was noted in the age range of patients. Patient age range in JULIET was wider than that in ZUMA-1 and CORAL Extension 1. Notably, JULIET included patients up to 76 years, versus 64 years and 65 years in ZUMA-1 and CORAL Extension 1, respectively (123, 392, 395). Older age is associated with worse prognosis (104).

In terms of IPI score, all three studies were closely aligned (123, 392, 395). Limited reporting of patient characteristics prevented comparison of individual components of the IPI score; however, JULIET had a higher proportion of patients with an ECOG performance status of 0, when compared to ZUMA-1. Patients with lower ECOG performance status have better function and prognosis (98). JULIET also had a lower proportion of patients with disease stage III-IV when compared to ZUMA-1 (392, 395).

JULIET and CORAL Extension 1 included patients with refractory disease after two or more lines of therapy (123, 392). ZUMA-1 also included patients with primary refractory disease (i.e. no response to first-line therapy) (144). Patients who have primary refractory disease have been reported to have slightly higher median OS than those who are refractory to second- or later-line therapy (108).

A lower proportion of patients in ZUMA-1 received previous HSCT compared to those in JULIET, likely a reflection of the high proportion of patients with refractory disease in ZUMA-1. HSCT in ZUMA-1 refers to autoSCT (144), no differentiation between autoSCT and alloSCT is provided in JULIET (392). Other reasons for HSCT ineligibility include age and comorbidities (143). Patients who receive HSCT and subsequently relapse have poor prognosis (124). Notably, patients in CORAL Extension 1 did not proceed to per-protocol autoSCT in CORAL RCT (123). It is likely that the proportion of patients in CORAL Extension 1 who received previous HSCT is low. In the absence of more granular data to indicate the reasons why patients did not receive prior HSCT, a more in-depth assessment is not possible.

The proportion of patients with bone marrow involvement was similar between JULIET and ZUMA-1; these data were not reported for CORAL Extension 1 (123, 144, 392). Bone marrow involvement confirms a poor prognosis (412). Of note, this was not assessed in 7% of patients in ZUMA-1 (144).

Notably, CORAL Extension 1 included patients at third-line therapy. Both JULIET and ZUMA-1 included patients who received more than three lines of therapy (123, 392, 395). Patients at later lines of therapy tend to have worse prognosis (98, 130). Subgroup analysis in JULIET indicated that there was no difference in efficacy of tisagenlecleucel, in

terms of ORR, in patients who received two or less lines of therapy versus those who received more than two. However, this was based on small patient numbers (143).

ZUMA-1 included patients with R/R DLBCL (76%), R/R PMBCL (n=8%) and tFL (n=16%) (144). Patients with R/R DLBCL and tFL accounted for 79% and 19% of patients in JULIET, respectively; patients with R/R PMBCL were excluded (392). CORAL Extension 1 appears to have only included patients with DLBCL (123). Patients with PMBCL tend to have better prognosis than those with DLBCL (144).

Differences existed between the subtypes of DLBCL permitted in JULIET and ZUMA-1. JULIET excluded patients with T-cell/histocyte-rich large B-cell lymphoma and Epstein-Barr virus-positive DLBCL of the elderly (143). Both subtypes were included in ZUMA-1 (144). T-cell/histocyte-rich large B-cell lymphoma is a rare histological variant of DLBCL, with limited data regarding clinical outcomes. Retrospective analysis of the National Cancer Database (US) found that patients diagnosed with T-cell/histocyte-rich large B-cell lymphoma (n=622), between 2010 and 2015, had better survival outcomes than those with DLBCL NOS (n=91,588) (HR 0.80; 95% CI 0.67 to 0.94) (413). It is unclear if this is generalisable to patients with R/R disease. Patients with Epstein-Barr virus-positive DLBCL of the elderly have been reported to have inferior survival outcomes to those with Epstein-Barr virus-negative DLBCL (414, 415). The rarity of both T-cell/histocyte-rich large B-cell lymphoma and Epstein-Barr virus-positive DLBCL of the elderly means that the number of patients with these histologic subtypes in ZUMA-1 were likely to have been low and unlikely to have had a meaningful impact on the results.

Patients in JULIET had an expected life expectancy of at least 12 weeks (143), which may indicate a cohort with favourable prognosis. This does not appear to have been a requirement in ZUMA-1 or CORAL Extension 1 (123, 144).

Patients in ZUMA-1, who had an initial response and experienced disease progression at least three months after the first dose of axicabtagene ciloleucel, could be retreated. Nine patients received retreatment. Of these, five patients had a response (396). This biases results as patients improved health status. Patients in JULIET were not retreated.

OS, in CORAL Extension 1, was reported from the time of 'treatment failure' in CORAL RCT. Thus, patients who died in the immediate period following treatment failure were included in the analysis. This is reflected in the initial steep drop in the Kaplan-Meier OS curve of CORAL Extension 1 (123). OS is reported from the time of infusion in JULIET and ZUMA-1, and reflects outcomes in those who were fit enough to survive the manufacturing period (392, 395). This biases relative effectiveness estimates against CORAL Extension 1.

Patients in JULIET were permitted bridging chemotherapy, to maintain disease during the manufacturing period of tisagenlecleucel. Of the 102 patients who received bridging chemotherapy, ORR was 20.6% (95% CI 13.2 to 29.7) (143). Thus, a proportion of patients were in response when they received tisagenlecleucel. A potential carry-over effect from bridging chemotherapy cannot be ruled out. Bridging chemotherapy was not permitted in ZUMA-1 (144). Differences in the duration of manufacturing period between JULIET (median 54 days) and ZUMA-1 (median 17 days) were noted; less patients proceeded to infusion in JULIET; 69% versus 91%, respectively (392, 395). It has been proposed that patients with more rapidly progressing disease were more likely to have been treated in ZUMA-1 (i.e. these patients would not have survived the prolonged manufacturing period in JULIET) (416). A proportion of patients (CR n=26; PR n=30) in CORAL Extension 1 were also in response when they received third-line therapy. Subgroup analysis indicated that OS was not significantly different according to the reason for CORAL RCT withdrawal (i.e. treatment failure, toxicity, protocol violation, 'other'); however, these were based on small patient numbers (123).

There were differences in the lymphodepleting chemotherapy regimens received by patients in JULIET and ZUMA-1. All patients in ZUMA-1 received lymphodepleting chemotherapy consisting of fludarabine and cyclophosphamide (395). JULIET implemented a more flexible protocol, which consisted of fludarabine and cyclophosphamide (administered at a lower dose than in ZUMA-1), bendamustine, or no lymphodepleting chemotherapy (392). A post-hoc analysis of JULIET indicated that patients who received fludarabine and cyclophosphamide (n=85) achieved numerically better outcomes compared to those who received bendamustine (n=22) or no lymphodepleting chemotherapy (n=8) (ORR: 57.6% for fludarabine and

cyclophosphamide versus 40.9% for bendamustine versus 25% for no lymphodepleting chemotherapy) (417). However, this analysis was based on small patient numbers, limiting the conclusions that can be made.

Different grading scales were applied to grade CRS in JULIET and ZUMA-1. The Penn Grading Scale was used in JULIET (143, 220), while the Lee Grading Scale was used in ZUMA-1 (144, 403). These scales are not transferable. The Penn Grading Scale is likely to result in a higher evaluated grade of CRS (418, 419). The reporting period of adverse events in JULIET and ZUMA-1 also differed. ZUMA-1 had a shorter reporting period than JULIET; 3 months versus 12 months, respectively (143, 144). Data on adverse events were not reported in CORAL Extension 1 (123).

6.3.4.2 Methodological Heterogeneity

OS and PFS efficacy data in JULIET and ZUMA-1 were obtained from phase II trials, while CORAL Extension 1 was an observational analysis. There were no per-protocol requirements defined in CORAL Extension 1 (123). CORAL Extension 1 is therefore, subject to a greater degree of bias in outcomes, due to potential inaccurate reporting of outcomes and patient characteristics (420). There may also be selection bias in the type of treatments received by patients. Notably, all studies were single arm with short duration of follow up. Outcomes in JULIET were IRC-assessed (392), while outcomes in ZUMA-1 were investigator-assessed. Concordance between investigator and IRC assessments in ZUMA-1 was 81% for ORR and 90% for CR. Data on concordance were not reported for PFS (395).

A high degree of censoring was noted in all three studies. For OS, JULIET had 2 patients left at risk at month 20 (392), ZUMA-1 had 7 at risk at month 30 (395), and CORAL Extension 1 had 16 at risk at month 30 (123). The long-term survival associated with these therapies is highly uncertain.

6.3.5 Indirect Treatment Comparison Feasibility Assessment

Due to the lack of direct comparative evidence between any of the defined treatments, the feasibility of conducting an ITC, for this research, was assessed.

Considerable heterogeneity was noted between JULIET (tisagenlecleucel) and ZUMA-1 (axicabtagene ciloleucel) study design and populations. Comparing relative efficacy of these CD19 CAR T-cell therapies may produce unreliable estimates, indicating that one therapy has favourable outcomes over the other. Thus, it was not considered appropriate to conduct an ITC between these studies. This is in line with conclusions elsewhere (421).

Comparisons of tisagenlecleucel versus salvage chemotherapy, and axicabtagene ciloleucel versus salvage chemotherapy were considered separately.

6.3.5.1 Meta-Analysis and Network Meta-Analysis

No common comparator arms existed between these single-arm trials. Therefore, neither a direct meta-analysis nor a network meta-analysis was deemed feasible here.

6.3.5.2 Population-Adjusted Comparison Methods

Population-adjusted methods of evidence synthesis (MAIC and STC), rely on the availability of raw IPD from at least one trial, as described in Chapter 2. Raw IPD were not available from the published literature; no further consideration was given to these methods.

6.3.5.3 Naïve Comparison

In the absence of a common comparator arm between studies and considering the lack of publicly available raw IPD, a naïve ITC was considered the only feasible method to compare tisagenlecleucel and axicabtagene ciloleucel to salvage chemotherapy.

As outlined in Chapter 2, naïve ITCs are highly uncertain.

6.3.6 Reconstruction of Individual Patient-Level Data

IPD of identified trials were reconstructed by digitising published Kaplan-Meier curves, using Digitizelt software (210), and applying the algorithm by Guyot et al. (211).

The reconstructed Kaplan-Meier OS curves, using reconstructed IPD, are presented in Figure 18 and Figure 19.

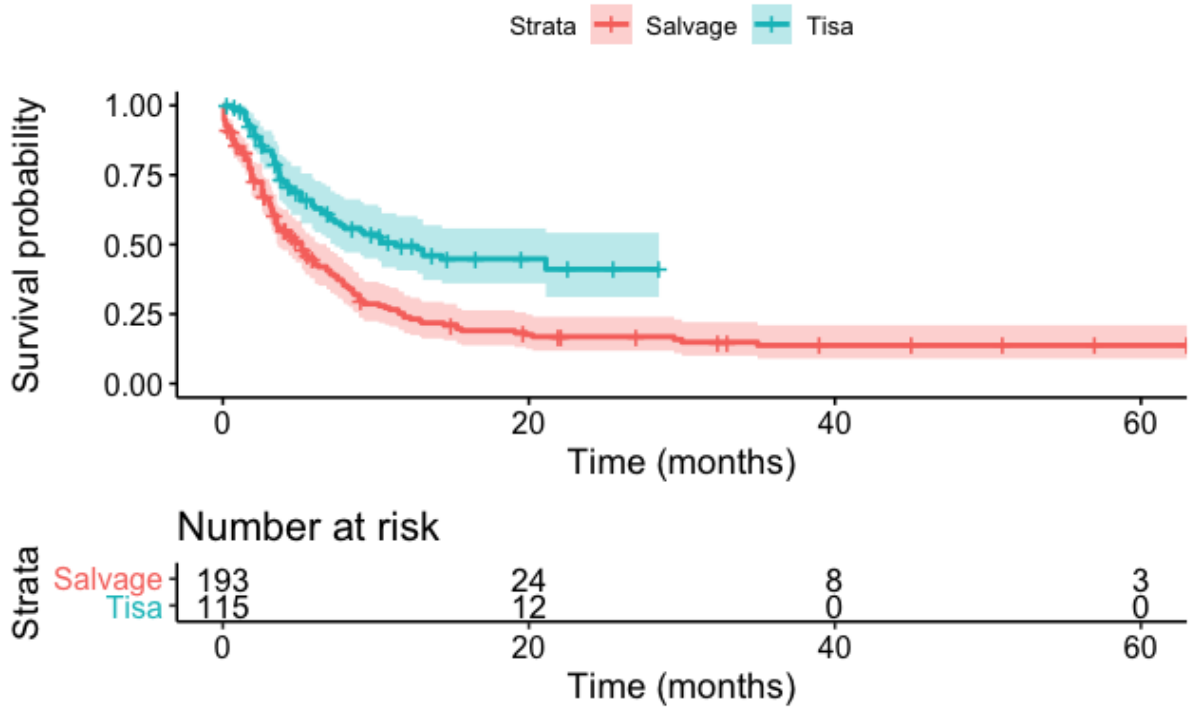

Figure 18 Reconstructed Kaplan-Meier overall survival curves of salvage chemotherapy (CORAL Extension 1) and tisagenlecleucel (JULIET)

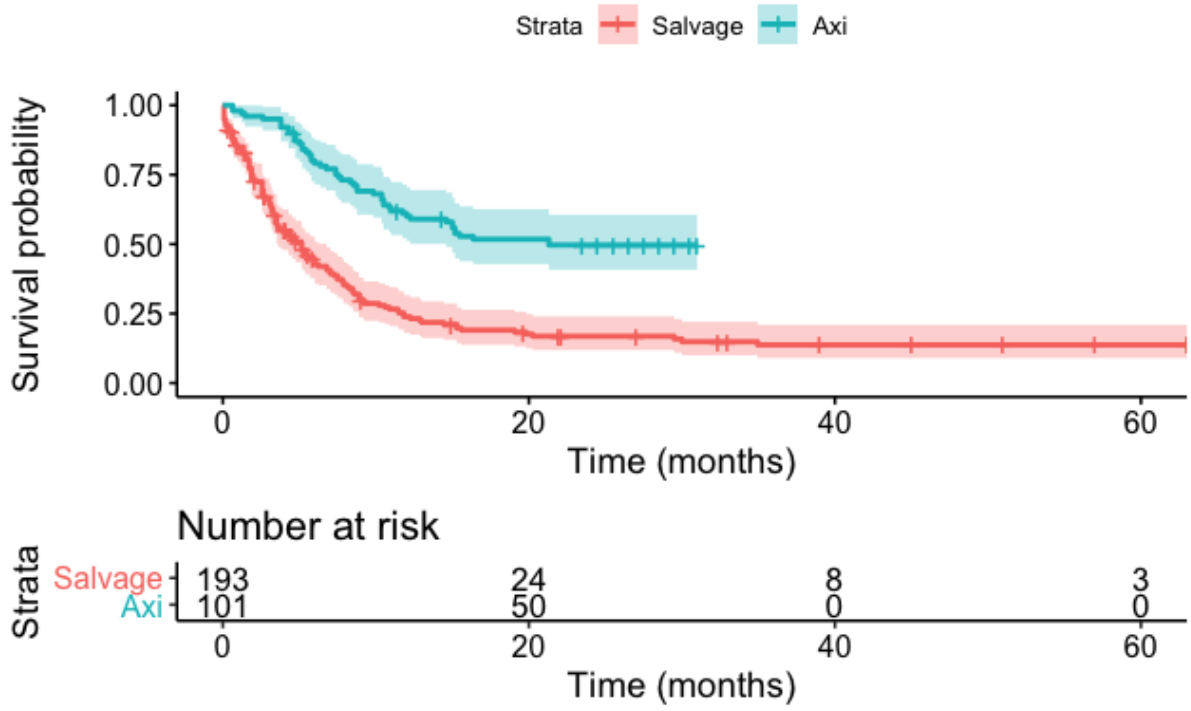

Figure 19 Reconstructed Kaplan-Meier overall survival curves of salvage chemotherapy (CORAL Extension 1) and axicabtagene ciloleucel (ZUMA-1)

6.3.7 Comparative Efficacy

The 12-month survival probabilities are presented in Table 34. OS HRs of tisagenlecleucel versus salvage chemotherapy, and axicabtagene ciloleucel versus salvage chemotherapy are also presented. PFS data were not available for CORAL Extension 1. Thus, HRs are not presented for this outcome.

Table 34 Survival estimates of tisagenlecleucel, axicabtagene ciloleucel, and salvage chemotherapy for relapsed/refractory diffuse large B-cell lymphoma based on naïve comparison

Treatment	Overall Survival, % (SE)	Hazard Ratio Overall Survival (95% CI) ¶	Progression-Free Survival, % (SE)
Tisagenlecleucel (392-394)	48* (0.05)	0.49 (0.36 to 0.67)	34† (0.05)
Axicabtagene Ciloleucel (144, 395)	61‡ (0.05)	0.34 (0.25 to 0.47)	43‡ (0.05)
Salvage Chemotherapy (with and without HSCT) (123)	23 (0.03)		
Salvage Chemotherapy (with HSCT) (123)	39§ (0.07)	-	NR
Salvage Chemotherapy (without HSCT) (123)	16| (0.03)		

HSCT: Haematopoietic stem cell transplant; **NR**: Not reported; **SE:** Standard error.
*Median follow up 19.3 months (May 2018 data cut).
†Median follow up 14 months (December 2017 data cut).
‡Median follow up 27.1 months (August 2018 data cut).
§CORAL Extension 1: Patients who did proceed to HSCT in third-line setting (31.5%); median follow up 30.1 months.
|CORAL Extension 1: Patients who did not proceed to HSCT in third-line setting (68.5%); median follow up 30.1 months.
¶Hazard ratio <1.0 favours tisagenlecleucel or axicabtagene ciloleucel (as appropriate) versus salvage chemotherapy (informed by CORAL Extension 1 entire cohort).

6.3.8 Quality of Evidence for Outcomes

Confidence in the evidence for OS for tisagenlecleucel versus salvage chemotherapy, and axicabtagene ciloleucel versus salvage chemotherapy are presented in Table 35 and Table **36**, respectively. The quality of evidence for OS was graded as very low for both comparisons. Both comparisons were based on observational evidence. Thus, quality was initially graded as low. Subsequent downgrading was applied due to very serious concerns regarding risk of bias (downgraded two levels), serious concerns regarding inconsistency (downgraded one level), very serious concerns regarding indirectness (downgraded two levels), and serious concerns regarding imprecision (downgraded one level). Factors considered in reaching these conclusions included: short duration of follow up, small sample sizes, open-label nature of trials; the naïve ITCs; and the potential for clinical and methodological heterogeneity based on qualitative assessment.

As no PFS data were collected during CORAL Extension 1, an assessment of the quality of evidence for PFS was not conducted.

Table 35 Summary of findings table for quality of evidence for overall survival (tisagenlecleucel versus salvage chemotherapy ± HSCT), based on GRADE assessment (213)

Summary of findings:			
Tisagenlecleucel compared to Salvage Chemotherapy (with or without HSCT) for R/R DLBCL			
Patient or population: R/R DLBCL **Setting:** Irish Healthcare Setting **Intervention:** Tisagenlecleucel **Comparison:** Salvage Chemotherapy (with or without HSCT)			
Outcomes	Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)
Overall Survival	HR 0.49 (0.36 to 0.67)	308 (2 non-randomised studies)	⨁◯◯◯ Very low[a,b,c,d]
CI: Confidence interval; **DLBCL:** Diffuse large B-cell lymphoma; **HR:** Hazard ratio; **HSCT:** Haematopoietic stem cell transplant; **R/R:** Relapsed/refractory.			
GRADE Working Group grade of evidence **Very low certainty** very little confidence in effect estimate: true effect is likely to be substantially different from the estimate of effect.			

Explanations

a. Downgraded two levels for risk of bias. High risk of bias; assessment of quality of studies, as per adapted Newcastle-Ottawa Scale, indicated all studies were of poor quality. Reasons include short duration of follow up, small sample sizes, and open-label nature (242).

b. Downgraded one level for inconsistency. Qualitative assessment indicated some degree of clinical and methodological heterogeneity. Statistical assessment of heterogeneity not feasible (243).

c. Downgraded two levels for indirectness. Due to naïve nature of comparison.

d. Downgraded one level for imprecision. Small sample size (<400 'rule-of-thumb') (242, 244).

Table 36 Summary of findings table for quality of evidence for overall survival (axicabtagene ciloleucel versus salvage chemotherapy ± HSCT), based on GRADE assessment (213)

Summary of findings:			
Axicabtagene Ciloleucel compared to Salvage Chemotherapy (with or without HSCT) for R/R DLBCL			
Patient or population: R/R DLBCL **Setting:** Irish Healthcare Setting **Intervention:** Axicabtagene Ciloleucel **Comparison:** Salvage Chemotherapy (with or without HSCT)			
Outcomes	Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)
Overall Survival	HR 0.34 (0.25 to 0.47)	294 (2 non-randomised studies)	⨁◯◯◯ Very low[a,b,c,d]
CI: Confidence interval; **DLBCL:** Diffuse large B-cell lymphoma; **HR:** Hazard ratio; **HSCT:** Haematopoietic stem cell transplant; **R/R:** Relapsed/refractory.			
GRADE Working Group grade of evidence **Very low certainty:** very little confidence in effect estimate: true effect is likely to be substantially different from the estimate of effect.			

Explanations
a. Downgraded two levels for risk of bias. High risk of bias; assessment of quality of studies, as per adapted Newcastle-Ottawa Scale, indicated all studies were of poor quality. Reasons include short duration of follow up, small sample sizes, and open-label nature (242).
b. Downgraded one level for inconsistency. Qualitative assessment indicated some degree of clinical and methodological heterogeneity. Statistical assessment of heterogeneity not feasible (243).
c. Downgraded two levels for indirectness. Due to naïve nature of comparison.
d. Downgraded one level for imprecision. Small sample size (<400 'rule-of-thumb') (242, 244).

6.4 Discussion

This SLR identified a limited number of studies, aligned with the inclusion criteria, examining the efficacy of treatments for R/R DLBCL. None of the included studies were RCTs, perhaps reflecting the rarity of DLBCL, the advanced stage of disease, and the innovative nature of CD19 CAR T-cell therapies. Outcome and population were common reasons for exclusion at full-text screening. Studies tended to report response outcomes (as opposed to survival) and patients included in identified studies tended to not be as heavily pre-treated as those defined in this SLR.

6.4.1 Included Studies

Efficacy of tisagenlecleucel was examined in one trial (JULIET). The small number of patients, along with the short duration of follow up and single-arm nature of the trial,

limits conclusions that can be drawn regarding the efficacy of tisagenlecleucel. Kaplan-Meier curves were highly censored. Uncertainty is further compounded by the potential for carry-over effects from bridging chemotherapy, and the high proportion of patients who did not proceed to infusion. The population who received infusion (mITT population) may be enriched with patients who were fit enough to survive the manufacturing period. The cohort who did not proceed to infusion had a higher proportion of patients with unfavourable prognostic factors (143). As such, outcomes may be overestimated when compared to those expected in clinical practice. Consideration needs to be given, in the cost-utility models, to the pre-infusion period of these trials (422).

The efficacy of axicabtagene ciloleucel was examined in one trial (ZUMA-1). This trial was subject to a number of the same limitations as JULIET, in that it included small numbers of patients, had short duration of follow up, and was single-arm. Additionally, relevant outcomes were primarily reported for the mITT population. The open-label nature of JULIET and ZUMA-1 increases the susceptibility of outcomes to bias. Data in these studies were immature and it cannot be determined how long-term outcomes will develop.

Patients in JULIET and ZUMA-1 were a highly select cohort. This is illustrated by the high number of patients screened versus those who received tisagenlecleucel (238 screened; 115 infused) (392). These data were not presented for ZUMA-1. Many patients deemed ineligible for these trials would be considered for treatment in clinical practice (423). Real-world data from 17 centres in the US indicated that 49% of patients who received axicabtagene ciloleucel would not have met eligibility criteria for ZUMA-1 at the time of leukapheresis (424). This raises questions over the generalisability of these trial data to patients treated in clinical practice. Although lacking the scientific rigour of a prospective, phase II trial, long-term real-world and registry data will be important in determining the true effectiveness of these therapies.

The exclusion of bridging chemotherapy from ZUMA-1 further limits the generalisability of these data. Real-world evidence suggests that the use of bridging chemotherapy in patients receiving axicabtagene ciloleucel is high, with one centre in Europe reporting that 96% of patients received bridging chemotherapy (425).The use of bridging

chemotherapy in clinical practice has been reported to be associated with poor prognosis and worse survival outcomes (417, 426, 427).

Uncertainty exists in the most appropriate data source to inform the efficacy of salvage chemotherapy. The studies identified were subject to limitations. CORAL Extension 1 was considered the most appropriate source here. However, the limitations of CORAL Extension 1 are acknowledged. Potential for bias in outcomes exist due to the observational nature of the data and the fact that a proportion of patients were in response at the time of study commencement. The SCHOLAR-1 data set will be explored as scenario analysis in the cost-utility models (Chapter 9).

6.4.2 Quality of Studies

Quality of all three studies, as assessed by an adapted Newcastle-Ottawa Scale, was poor. Uncertainty in the evidence base of these therapies creates challenges and risk for patients, clinicians, and decision-makers.

The limitations of the adapted Newcastle-Ottawa Scale are described in Chapter 2 (2.4.2).

6.4.3 Heterogeneity

Due to poor consistency in the reporting of patient characteristics across the studies, it is difficult to conclude how heterogeneity across the trials impacts on the direction and magnitude of relative treatment effects. Between-trial clinical and methodological heterogeneity could only be assessed on a limited number of factors. Of those that were assessed, considerable heterogeneity was noted. Caution is therefore, warranted in interpretation of results. Interpretation of between-trial heterogeneity is further hampered by the qualitative, yet thorough, nature of the analysis.

6.4.4 Comparative Efficacy

Adjusting for differences in patient populations, as per best practice (238), could not be conducted due to the lack of publicly-available raw IPD. The poor availability of raw IPD prevents researchers from conducting robust, independent analyses. Due to the heterogeneity observed between JULIET and ZUMA-1, a comparison of these trials, either

by naïve or adjusted ITC methods, would not provide reliable results (421). Naïve comparison of OS indicated that patients treated with either tisagenlecleucel or axicabtagene ciloleucel had favourable outcomes when compared to patients treated with salvage chemotherapy. Notably, uncertainty in the naïve ITC was not captured in the 95% CI of the HR. These estimates are highly uncertain. Lack of adjustment for differences between populations and study design limits conclusions that can be drawn. Interpretation of results is also hampered by differences in the definition of OS.

The adverse event profile of both tisagenlecleucel and axicabtagene ciloleucel is appreciable and should be given due consideration in cost-utility analyses. No adverse event data were reported in CORAL Extension 1. As a result, the adverse event profile for the salvage chemotherapy arm, in the cost-utility models, will rely on proxy data.

Poor consistency in trial design across CAR T-cell therapies may be an obstacle to patient access. It may result in a situation, whereby patients are denied access to treatments that appear to have unfavourable comparative efficacy based on comparative effectiveness studies that are subject to considerable limitations (421). In contrast, patients and clinicians could also have unrealistic expectations about the effectiveness of a therapy. Increased efforts should be made by both researchers and regulators to harmonise trial design across treatments within a therapeutic class (421). This could cover aspects such as inclusion and exclusion criteria, bridging and lymphodepleting chemotherapy regimens, and subsequent therapies. Although challenging, this may facilitate more robust comparative analyses and will be of value when commercially developed third- and fourth-generation CAR T-cell therapies become the main focus.

In the absence of randomised, comparative data, a number of real-world observational studies examining the use of tisagenlecleucel and axicabtagene ciloleucel have been published (425, 428, 429). However, these studies have short duration of follow up and are based on small sample sizes, precluding robust comparative analyses. Even when mature data become available, the impact of unobserved effect modifiers, such as clinician preference and reimbursement status, cannot be accounted for. The addition of lisocabtagene maraleucel (a CD19 CAR T-cell therapy, approved by the US Food and Drug Administration for the treatment of R/R DLBCL) to the treatment landscape, based on a

single-arm trial (430), adds further complexity to comparative clinical effectiveness assessments of these therapies.

6.4.5 Quality of Evidence for Outcomes

Quality of evidence for OS, as assessed using GRADE, was very low. The true OS benefit of tisagenlecleucel and axicabtagene ciloleucel (versus salvage chemotherapy) is likely to be markedly different from the estimated effect (213). As highlighted in Chapter 2, it is not clear if the GRADE framework is the optimal method to assess the quality of evidence generated using unadjusted, single-arm studies. This approach is also highly subjective. Further research is required to assess what impacts the quality generated from naïve comparison of single-arm studies. Additional discussion regarding this approach is provided in 2.4.5.

6.4.6 Comparison with the Published Literature

Previously published SLRs examining treatments for R/R DLBCL had similar findings to this SLR in that there was a paucity of RCT evidence (431, 432). Inclusion criteria of other published SLRs differ from this study, limiting conclusions that can be drawn regarding consistency between them. Only one SLR identified tisagenlecleucel and axicabtagene ciloleucel as relevant comparators (432).

Oluwole et al. conducted a MAIC to compare outcomes between JULIET and ZUMA-1 (422). Adjusting for patient characteristics between the trials resulted in a HR of 0.51 (95% CI 0.31 to 0.83), for OS of axicabtagene ciloleucel versus tisagenlecleucel. Aside from the large degree of heterogeneity that could not be adjusted for (bridging chemotherapy, for example), the MAIC was subject to additional limitations. The MAIC was unanchored, resulting in an unknown amount of bias in the effect estimates (238). The analysis is also subject to potential bias due to unobserved differences between the trials. It cannot be concluded that the additional complexity associated with this MAIC leads to less biased estimates, when compared with a naïve ITC. In contrast, one real-world study, conducted in a single-centre in Europe, concluded that there was no significant OS difference between axicabtagene ciloleucel and tisagenlecleucel (425). However, slight differences were noted between the baseline characteristics of the

patient cohorts who received axicabtagene ciloleucel and tisagenlecleucel. Additionally, the single-centre nature of the study limits its generalisability.

6.4.7 Limitations

The limitations presented here are in addition to those discussed thus far. Although an updated data cut of ZUMA-1 was identified post-database searching, data were only presented as an abstract and no PFS data were reported. The lack of published Kaplan-Meier curves prevented reconstruction of IPD. Results of this updated data cut will be used to validate survival output in the cost-utility model.

PFS data were not reported for CORAL Extension 1. Assumptions regarding these data will therefore, be required for the cost-utility models, adding further uncertainty to the evidence base.

As with other SLRs in this thesis, a pragmatic approach was adopted; full-text screening and assessment of study eligibility was conducted by a single reviewer. This was due to the time-intensive nature of these processes. As a quality assurance measure, and in an attempt to minimise bias, 10% of full-text articles were screened in duplicate.

6.5 Conclusion

The evidence base of treatments for R/R DLBCL, identified in this SLR, was limited to single-arm, heterogeneous studies. Naïve comparison of OS indicated that treatment with either tisagenlecleucel or axicabtagene ciloleucel is favourable compared with salvage chemotherapy. However, the true magnitude of benefit is unknown. Inconsistency in reporting of patient characteristics and outcomes further limit conclusions that can be drawn. In Chapter 9, the treatment effectiveness estimates derived from this study will be incorporated into cost-utility models, examining the cost effectiveness of tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL in the Irish healthcare setting.

Chapter 7 Investigation of the Performance of Abstrackr, Text-Mining Tool, in Systematic Literature Review of Efficacy of Tisagenlecleucel and Axicabtagene Ciloleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma

7.1 Introduction

As highlighted in previous Chapters, evidence to support treatment effectiveness and its impact on HRQOL requires a comprehensive search of multiple databases to identify relevant data (11). Searches may yield thousands of citations that must be screened for relevance. First, titles and abstracts of identified citations are screened to determine their relevance to the research question. Citations that are deemed relevant, based on this screening, are carried forward for full-text screening, whereby the full texts of citations are retrieved and screened for relevance. Screening should be conducted by two or more independent human screeners, to ensure methodological rigour (11). Emphasis is placed on identifying all relevant studies (i.e. attaining 100% sensitivity), in order to minimise bias (433). Conduct of SLRs requires input from highly skilled researchers and a large time commitment. Associated time commitments, described elsewhere, have ranged between six months and two years (434, 435). This can result in a considerable financial burden. In 2019, the cost of producing an SLR was estimated to be $141,195 (US), when associated labour and costs were quantified (436).

The complexity of SLRs has increased due to growth in the volume of published research, the use of more complex methodologies such as network meta-analysis, and the increasing complexity of new interventions. Inadequate coding of studies indexed in databases generates imprecise search results (437). These factors hinder the production of an SLR in a timely fashion, with updating of the SLR required before or shortly after publication in many instances (438, 439).

SLR management systems have been developed to facilitate title and abstract screening (440). Covidence (https://www.covidence.org/) is one such system, which allows authors to import and screen titles, abstracts, and full-text articles, extract data, populate risk of bias tables, and export results in various formats (440). Despite the use of such systems, researchers are increasingly faced with the challenge of producing a robust SLR, within the confines of time and budget.

7.1.1 Text Mining

These challenges have resulted in growing recognition of the need to develop alternative methods (441). Semi-automating the title and abstract screening process, using text mining, is one proposed solution (442). Text mining is the process of discovering knowledge and structure from unstructured data (i.e. text) (443, 444). Relevant information is identified as patterns, learnt from an initial training sample (i.e. a sample of citations), which are labelled as relevant or irrelevant by the human screener (437). The accuracy of the predictions made improves through interaction with the human screener (442).

Text mining may reduce the workload burden of title and abstract screening in several ways. Using screening truncation, citations that fall beneath a specified prediction score of relevance are excluded (445). This reduces the number of citations that need to be screened. The text-mining tool 'learns' from an initial training sample of citations and subsequently generates a ranked list of the remaining citations, which the human screener may continue to screen. As the human screener continues to screen, the text-mining tool adapts its decision rule to include additional information generated through additional screening. The text-mining tool then generates an updated ranked list of citations. This process continues, with the ranked list of citations updating at pre-specified intervals, until a human screener-specified, pre-defined, stopping point is reached and no further citations are screened by the human screener.

Text-mining tools may replace or supplement the work traditionally undertaken by a second, human screener. Here, citations are screened manually by a human screener and the text-mining tool either acts as an independent check (i.e. predictions generated by the text-mining tool are checked against the judgements of the human screener), or the text-mining tool presents a reduced list of citations for screening by a second, human screener (442). This approach has been proposed to be the most appropriate for HTA (446).

In screening prioritisation, citations most likely to be relevant are screened first by the human screener (447). Identifying the most relevant citations first allows members of the review team to begin full-text screening, reducing the time taken from SLR

commencement to completion. Through this approach, human screeners may become more familiar with the SLR inclusion criteria earlier in the process, ultimately increasing efficiency. This may also address over-inclusiveness, whereby human screeners tend to be cautious and include more citations at early-stage screening (442).

7.1.2 Text-Mining Tools

A practical guide to machine learning tools in research synthesis highlighted six text-mining tools that can semi-automate the screening process (448). These tools are summarised in Table 37. DistillerSR is an additional text-mining tool, summarised in Table 37, identified through the literature (449).

Table 37 Features of text-mining tools for title and abstract screening

Tool	Financial Cost	Screening Truncation	Screening Prioritisation	Additional Key Features	Developed By
Abstrackr (450) http://abstrackr.cebm.brown.edu/	Free	Yes: Prediction score (0-1) and hard predictions (relevant-irrelevant)	Yes	Key term highlighting	Center for Evidence Synthesis in Health, Brown University (US) (451)
Rayyan (452) https://www.rayyan.ai/	Free	Yes: Five-star scale	Yes	Key term highlighting Similarity-based search	Qatar Computing Research Institute (452)
RobotAnalyst (453) http://nactem.ac.uk/robotanalyst/	Free	Yes: Prediction score	Yes	Key term highlighting Similarity-based search Cluster-based screening	UK National Centre for Text-Mining, Machine Learning and Data Analytics (University of Liverpool), and NICE (454)
Colandr (455) https://www.colandrapp.com	Free	No. Citations ranked in order of relevance; human screener decides stopping point	Yes	Key term highlighting Search string generator Full-text screening Data extraction from full text	Science for Nature and People Partnership Conservation International and DataKind (455)

SWIFT-Review (456) https://www.sciome.com/swift-review	Free	No. Citations ranked in order of relevance; human screener decides stopping point	Yes	Key term highlighting Identification of over-represented topics Visualisation of 'themes' identified in evidence through interactive graphs	Sciome LLC, a research and technology consulting company (US) (457)
EPPI-Reviewer* (458) https://eppi.ioe.ac.uk	Subscription fee	Unclear: EPPI-Reviewer website recommends that all citations are screened manually (459)	Yes	Key term highlighting Study classification Cluster-based screening Mapper for visualising 'maps' of research evidence	Evidence for Policy and Practice Information and Co-ordinating (EPPI)-Centre, Social Science Research Unit, University College London Institute of Education, and University of London (UK) (460)
DistillerSR (461, 462) https://www.evidencepartners.com/products/distillersr-systematic-review-software	Subscription fee	Yes. Relevance of citation predicted as a 'hard' outcome (include or exclude)	Yes	Key term highlighting Duplicate detection PubMed integration	Evidence Partners (Canada) (462)

NICE: National Institute for Health and Care Excellence; **UK:** United Kingdom; **US:** United States.
*EPPI-Reviewer is recommended (as an alternative to Covidence) by the Cochrane Collaboration to support authors of Cochrane reviews in the development of systematic literature reviews. The Cochrane Collaboration indicate that EPPI-Reviewer is particularly useful in complex areas such as meta-analysis, framework synthesis, and thematic synthesis (463).

7.1.2.1 Abstrackr

Abstrackr was chosen for this research, due to its widespread use in the literature and reported ease of use compared to other tools. It has also been shown to perform

favourably when compared to other tools (449). Both Abstrackr and Rayyan are freely available and were piloted by an NCPE Information Specialist, experienced in the production of SLRs, for usability. Abstrackr was subsequently determined to be the most user-friendly. RobotAnalyst was not considered during piloting due to its reported poor usability and performance relative to other text-mining tools (449). Colandr and SWIFT-Review do not explicitly include screening truncation functionality and were therefore, not examined. DistillerSR and EPPI-Reviewer were not considered due to their associated subscription fee.

Abstrackr incorporates screening truncation and prioritisation functionality. A key terms list, coded by the human screener to indicate relevance or irrelevance to the research question, can also be uploaded. During screening, Abstrackr highlights these key terms either green ('strongly indicative of relevance'), blue ('indicative of relevance'), red ('strongly indicative of irrelevance'), or purple ('indicative of irrelevance'). This provides a visual aid to the human screener. The first step in the screening process, using Abstrackr (herein 'Abstrackr-assisted screening'), involves uploading the relevant citations. The key terms list may also be uploaded at this stage; additional key terms can be uploaded throughout the screening process. Then, the human screener labels the training sample of citations. The ability of Abstrackr to accurately predict citation relevance depends on the correct labelling of the training sample (464). The prediction algorithm of Abstrackr updates once per day; Abstrackr processes information gained through the labelled training sample. Abstrackr then generates both 'hard' predictions (include or exclude) and a prediction score (between 0 and 1) for each remaining citation. A maximum prediction score of all remaining citations is also presented. The human screener may then choose to continue screening in Abstrackr and thus, improve Abstrackr's learning capacity to generate an updated list of predictions ('hard' predictions, individual prediction scores, and a maximum prediction score).

Some literature suggests that once the maximum prediction score of all remaining citations falls below 0.40, zero citations are generally predicted to be relevant by Abstrackr (451, 465). At this point, cessation of human screening of titles and abstracts may be considered. The maximum prediction score is generally used to guide the stopping point as opposed to 'hard' predictions or individual prediction scores. This is

due to ease of use of the maximum prediction score (just one overall score is presented). Employing a stopping rule based on 'hard' predictions or individual prediction scores requires the human screener to manually assess each prediction, for each citation, every time a new set of predictions is generated. The human screener could alternatively choose to rely solely on the first set of individual predictions generated following the labelling of the training sample. These predictions may be cross-checked against the judgements (relevant or irrelevant) made by a human screener (via manual screening). The individual predictions may also be used to automatically exclude citations deemed irrelevant by Abstrackr, without any cross-check with a human screener. Abstrackr maintains a digital record of the labels (relevant or irrelevant) assigned, by the human screener, to each citation, which can be accessed at any time. Labels can be revised, if necessary.

The performance of Abstrackr-assisted screening, as measured by a variety of metrics, has been examined in the literature. These performance metrics are described in Table 38, with a detailed description of the performance of Abstrackr provided below.

Table 38 Performance metrics used to assess the performance of Abstrackr, text-mining tool, adapted from Gates et al. (434, 449) and Rathbone et al. (437)

Performance Metric	Definition
Sensitivity (True Positive Rate)	Proportion of citations correctly identified as relevant by Abstrackr out of the total deemed relevant by human screener
Specificity (True Negative Rate)	Proportion of citations correctly identified as irrelevant by Abstrackr out of the total deemed irrelevant by human screener
Precision	Proportion of citations predicted as relevant by Abstrackr that were also deemed relevant by human screener
False Negative Rate	Proportion of citations that were deemed relevant by the human screener that were predicted as irrelevant by Abstrackr
Proportion Missed	Total number of studies included in the final evidence base that were predicted as irrelevant by Abstrackr
Workload Savings	Proportion of citations predicted as irrelevant by Abstrackr out of the total number of citations to be screened (i.e. the proportion of citations that would not need to be screened manually)
Time Savings	Time saved based on citations that would not need to be screened (i.e. those predicted as irrelevant by Abstrackr); based on a screening rate of 0.5 minutes per citation and an 8-hour work day

Rathbone et al. examined the performance of Abstrackr on four completed SLRs of varying complexity. Human screening of titles and abstracts continued until the first set of predictions were generated by Abstrackr. Abstrackr's predictions were checked for accuracy against the human-screener judgements; key results are summarised in Table 39 (437). Gates et al. evaluated the performance of Abstrackr on three SLRs and one descriptive analysis. Abstrackr-assisted screening was conducted as per Rathbone et al. (437). The predictions were compared to the human-screener judgements; key results are summarised in Table 39 (434).

The comparative performance of Abstrackr, RobotAnalyst, and DistillerSR was explored by Gates et al. when each tool was used to (i) automatically exclude citations predicted to be irrelevant, and (ii) complement the work of a human screener (449). In the first approach, all citations deemed irrelevant from the first set of predictions generated by Abstrackr were excluded. In the second approach, citations were excluded as per the first approach but a second human screener also screened all citations. Three SLRs were retrospectively evaluated; key results are described in Table 39.

Table 39 Performance of the Abstrackr text-mining tool, as assessed in the literature

Author ,year		Text-Mining Tool	Sensitivity* (%)	Specificity † (%)	Precision ‡ (%)	False Negative § (%)	Proportion Missed (%)	Workload Impact Measured	
Rathbone et al. 2015 (437)		Abstrackr	NR	NR	16.8 - 45.5	2.4 - 14.5	0.00-0.21|	Workload Savings¶ (%): 9 - 57	
Gates et al. 2018 (434)		Abstrackr	79 – 96	19 – 90	14.8 – 64.7	3.5 – 21.2	0.00 – 12.2|	Time Savings# (days): 0.5 - 44	
								Workload Savings¶ (%):	Time Savings # (days):
Gates et al. 2019 (449)	(i) ††	Abstrackr	NR	NR	NR	NR	5§§	90	19
		DistillerSR					97§§	99	23
		RobotAnalyst					70§§	85	20
	(ii) ‡‡	Abstrackr					1§§	40	8
		DistillerSR					2§§	49	11
		RobotAnalyst					2§§	35	8

NR: Not reported.
*Proportion of citations correctly identified as relevant by Abstrackr out of the total deemed relevant by human screener(s).
†Proportion of citations correctly identified as irrelevant by Abstrackr out of the total deemed irrelevant by human screener(s).
‡Proportion of citations predicted as relevant by Abstrackr that were also deemed relevant by human screener(s).
§Proportion of citations that were deemed relevant by human screener(s) that were predicted as irrelevant by Abstrackr.
|Number of citations predicted as irrelevant by Abstrackr that were included in the final evidence base, out of the total number of citations predicted as irrelevant.
¶Proportion of citations predicted as irrelevant by Abstrackr out of the total number of citations to be screened (i.e. the proportion of citations that would not need to be screened manually).
#Time saved based on the citations that would not need to be screened (i.e. those predicted as irrelevant by Abstrackr); based on a screening rate of 0.5 minutes per citation and an 8-hour work day.
**Proportion of citations screened in order to identify all relevant citations.
††Automatically excluded citations predicted to be irrelevant.
‡‡Automatically excluded citations predicted to be irrelevant but all citations also screened by a second, human screener.
§§Proportion of citations included in the final evidence base, after full-text screening, that were predicted to be irrelevant by Abstrackr.

As observed in Table 39, the performance of Abstrackr has varied. It has been widely recommended that further research is required to assess Abstrackr's performance on a diverse range of screening tasks (434, 445). In contrast to the studies presented in Table 39, this study continued screening until a predefined maximum prediction score was reached, and examined a range of maximum prediction scores (as discussed further in 7.4.1).

7.1.3 Barriers to Implementation of Text-Mining Tools

Adoption of text-mining tools in the SLR community has been slow and fragmented (442). A lack of trust by human screeners in these tools has been proposed to be a key barrier

(466). The 'black box' nature of these tools is troublesome due to the requirement for transparency in the conduct of SLRs (467). There may be a perception that these tools cannot reliably perform the required task (466), and apprehension that the high level of recall, required for an SLR, is not achievable by these tools (442).

A key limitation of semi-automated title and abstract screening is that it is not clear at which point it is appropriate for the human screener to stop screening (448). Many of the studies that have assessed the performance of such tools have done so using the initial set of predictions generated once the initial training sample has been labelled (434, 437, 449). Inaccurate labels in the training sample result in unreliable predictions (449). Additionally, a larger number of citations may require a large training sample, which may not be feasible (450). The optimal training sample size is currently unknown (448).

There is a need to develop the evidence base of text-mining tools, to enable evidence synthesis organisations to develop clear guidance on their use (464, 468). For widespread adoption, researchers need to be assured that the use of text mining, to support title and abstract screening, will not compromise the validity of results (469).

7.1.4 Chapter Aim

The aim of this study is to evaluate the performance of Abstrackr-assisted screening, when compared to Single-human screening, in an SLR of treatments for R/R DLBCL. The research question of the SLR was 'what is the efficacy of CD19 CAR T-cell therapies (tisagenlecleucel and axicabtagene ciloleucel) versus salvage chemotherapy in patients with R/R DLBCL, after two or more lines of systemic therapy?'. We aim to investigate the reliability of Abstrackr's predictions once a maximum prediction score of 0.39540 (base case) is reached. The performance metrics described by Gates et al. (434) and Rathbone et al. (437) will be investigated here.

7.2 Methods

7.2.1 Choice of Data Set

The SLR of treatments for R/R DLBCL, described in Chapter 6, was selected here due to the large number of citations identified through database searching. Text-mining tools

have been reported to perform better on large (≥2,500) screening samples (445). The large number of intervention and comparator treatments specified in the inclusion criteria (Appendix E) was also a consideration.

7.2.2 Search Methods

Electronic databases EMBASE, MEDLINE (via EBSCO), and CENTRAL (via the Cochrane Library) were searched from 01 January 2001 to 25 October 2019, as described in Chapter 6.

7.2.3 Citation Management

Identified citations were imported to Endnote®. Duplicates were systematically searched for using software in Endnote® and identified manually throughout. Following exclusion of duplicates, 7,723 citations were included in title and abstract screening. Screening was conducted by two human screeners, both experienced in the production of SLRs. Screener 1 undertook the process using Abstrackr ('Abstrackr-assisted screening'), whilst Screener 2 undertook the process using Covidence (herein 'Single-human screening').

7.2.3.1 Abstrackr-Assisted Screening

Screener 1 uploaded citations to Abstrackr, along with 154 key terms coded by Screener 1 to indicate relevance or irrelevance to the research question. These were mainly informed by the SLR inclusion and exclusion criteria. Screener 1 then screened an initial training sample of 200 randomly selected citations. This is in line with previous training sample sizes in the literature (449, 470). The algorithm was allowed to process the information (from the training sample) overnight.

Once this information was processed and the initial set of predictions were generated by Abstrackr, Screener 1 set the settings to 'single-screen mode'. The order of citations was set to 'most likely to be relevant', so that the most relevant citations, as predicted by Abstrackr, were presented to the human screener in priority order. Screener 1 screened titles and abstracts for relevance. Additional coded key terms were uploaded, as identified, throughout the screening process. Screener 1 continued to screen in Abstrackr until the algorithm indicated that a maximum prediction score of 0.39540 (base case)

was reached (451). Of note, stopping once a maximum prediction score of less than 0.40 was reached was pre-specified; however, due to the time required for Abstrackr's algorithm to update (overnight), the maximum prediction score could not be measured in real time. This resulted in screening until a maximum prediction score of 0.39540 was reached. At this point, Screener 1 assumed that any remaining unscreened citations were irrelevant and did not conduct any further screening in Abstrackr. This inherently assumes that any unscreened citations at this point have been 'screened' and deemed irrelevant by Abstrackr (i.e. Abstrackr is acting as the second, human screener for these citations). Citations that were deemed 'relevant' or 'maybe' by Screener 1, were brought forward for full-text screening. All citations, and their associated labels, were exported from Abstrackr to Microsoft Excel®. It was assumed that Abstrackr deemed all citations with a score of greater than 0.39540 as relevant, despite the label provided by the human screener (Screener 1).

7.2.3.1.1 Sensitivity Analysis

Sensitivity analysis, whereby Abstrackr-assisted screening continued until maximum prediction scores of 0.34458 and 0.29021 were reached, was conducted. Maximum prediction scores of less than 0.35 and 0.30, respectively, were pre-specified; however, as highlighted, the maximum prediction score could not be measured in real time. The aim here was to determine if the trade-off between workload saving and accuracy of Abstrackr could be improved at alternative prediction scores.

7.2.3.2 Single-Human Screening: Covidence

Screener 2 uploaded citations to Covidence screening software and screened all titles and abstracts. Citations deemed 'relevant' or 'maybe' by Screener 2 were brought forward for full-text screening. Citations and their associated labels were exported from Covidence to Microsoft Excel®

7.2.4 Data Analysis to Assess the Performance of Abstrackr-Assisted Screening

Data from 2x2 cross-tabulations, based on the number of citations predicted relevant or irrelevant by Abstrackr-assisted screening versus the number judged relevant or irrelevant by Single-human screening, were used to calculate performance metrics. The

metrics assessed, presented in Table 38, were defined in line with previous studies (434, 437). All metrics relate to title and abstract screening. The formulae used to calculate these metrics are presented in Appendix F (Table A28). Here, it was assumed that Single-human screening identified all relevant citations.

7.3 Results

7.3.1 Included Studies

Of the 7,723 citations, 2,568 (33%; titles and abstracts, including training sample) were screened in Abstrackr before a maximum prediction score of 0.39540 (base case) was reached. In line with previous studies, zero citations were deemed potentially relevant by Abstrackr at this prediction score (451, 465). Of these 2,568 citations, 451 were were brought forward for full-text screening.

Single-human screening (by Screener 2) of all citations on Covidence resulted in 424 citations being brought forward for full-text screening.

7.3.2 Performance of Abstrackr

Data from the 2x2 cross-tabulations used to calculate the performance metrics of Abstrackr-assisted screening (versus Single-human screening) are presented in Table 40. The performance metrics of Abstrackr, based on these data (base case), are presented in Table 41.

The performance metrics, based on sensitivity analysis, are also presented in Table 41. An additional 584 and 1,284 citations required screening (compared to the base case) before reaching maximum prediction scores of 0.34458 and 0.29021, respectively. The 2x2 cross-tabulations used to calculate these metrics are presented in Appendix F (Table A29 and Table A30).

Table 40 2x2 cross-tabulations of Abstrackr predictions versus human-screener (Screener 2) judgements (base case; maximum prediction score of 0.39540)

		Human Screener (Screener 2) Judgements		
		Excl.	**Incl.**	**Total**
Abstrackr Predictions	**Excl.**	5,118* (True Negative)	37‡ (False Negative)	5,155
	Incl.	2,001† (False Positive)	367§ (True Positive)	2,368
	Total	7,119	404	7,523|

*Abstrackr and Screener 2 excluded the same 5,118 citations; the number of true negatives predicted by Abstrackr.
†Abstrackr included 2,001 citations that Screener 2 excluded; the number of false positives predicted by Abstrackr.
‡Abstrackr excluded 37 citations that Screener 2 included; the number of false negatives predicted by Abstrackr.
§Abstrackr and Screener 2 included the same 367 citations; the number of true positives predicted by Abstrackr.
|The total number of citations included in the analysis, excluding the 200 citation training sample.

Table 41 Performance metrics of Abstrackr-assisted screening, when compared to Single-human screening (n=7,523††)

Performance Metric	Sensitivity * (%)	Specificity † (%)	Precision ‡ (%)	False Negative Rate§ (%)	Proportion Missed| (%)	Workload Savings¶ (%)	Time Savings # (days)
Result (stopping point 0.39540)**	91	72	15.5	9	0	67	5.4
Sensitivity Analysis							
Result (stopping point 0.34458)**	97	64	13	3	0	59	4.8
Result (stopping point 0.29021)**	100	54	11	0	0	50	4.0

*Proportion of citations correctly identified as relevant by Abstrackr out of the total deemed relevant by Screener 2.
†Proportion of citations correctly identified as irrelevant by Abstrackr out of the total deemed irrelevant by Screener 2.
‡Proportion of citations predicted as relevant by Abstrackr that were also deemed relevant by Screener 2.
§Proportion of citations that were deemed relevant by Screener 2 that were predicted as irrelevant by Abstrackr.
|Proportion of citations included in the final evidence base after full-text screening that were predicted to be irrelevant by Abstrackr.
¶Proportion of citations predicted as irrelevant by Abstrackr out of the total number of citations to be screened, including the training set (i.e. the proportion of citations that would not need to be screened manually).
#Time saved based on the citations that would not need to be screened (i.e. those predicted as irrelevant by Abstrackr); based on a screening rate of 0.5 minutes per citation and an 8-hour work day.
**Calculations presented in Appendix F.
††Excludes 200 citations included in the training sample.

An additional sensitivity analysis was conducted to explore the impact, on time savings, of assuming a higher screening rate of 1 minute per citation (471). Under this assumption, the time savings were 10.7 days (0.39540 prediction score, base case), 9.5 days (0.34458 prediction score) and 8.1 days (0.29021 prediction score).

7.4 Discussion

7.4.1 Main Findings

In the research question specified here, Abstrackr-assisted screening, conducted until a maximum prediction score of 0.39540 (base case) was reached, identified all relevant citations and reduced title and abstract screening workload by 67% (5.4 days), when compared with Single-human screening conducted on Covidence. Abstrackr demonstrated high sensitivity (91%). Although the false negative rate was 9%, the actual proportion of relevant citations missed was 0%. Abstrackr, in this case, was reliable. No citations that were predicted irrelevant by Abstrackr but relevant by Screener 2 (using Covidence), were included in the final evidence base. Specificity (72%) and precision (15.5%) were low; Abstrackr overestimated citation relevance. However, these were offset, to a large degree, by the workload saving.

Sensitivity analysis, conducted at maximum prediction scores of 0.34458 and 0.29021, resulted in higher sensitivity and lower false negative rates (compared to the base case). However, these came at the expense of decreased specificity and precision, and reduced workload savings. At a maximum prediction score of 0.29021, just one study was predicted to be irrelevant by Abstrackr that was judged relevant by Screener 2. Those producing SLRs may be willing to make the trade-off between this increased sensitivity and reduced workload saving. It should be noted, however, that the proportion missed was zero in both the base case and sensitivity analysis. The results of this sensitivity analysis may stimulate further discussion on what the most appropriate stopping point should be and provides an insight into the trade-offs required to improve sensitivity.

This study contributes to the limited evidence base on the performance of Abstrackr. In contrast to other studies, whereby screening in Abstrackr was conducted until the first set of predictions were available (434, 437), this study continued screening until a

predefined maximum prediction score was reached. The advantage here is that Abstrackr's learning capacity is expected to improve, based on the increased data (and therefore 'learning') provided by the human screener. Screening conducted in line with our study may reduce the impact of poor quality or unrepresentative training samples, as the continued human screening allows Abstrackr to adapt its decision rules. This may give more confidence to human screeners who are sceptical of the reliability of Abstrackr when relying on just the initial set of predictions.

One explanation for the low precision might be that the SLR inclusion criteria contained a number of treatments with lexical similarity. It may have been difficult for the algorithm to differentiate between minor differences in the names of such treatment regimens; the exclusion of rituximab monotherapy but inclusion of other rituximab-based therapies, for example. There was also a high level of imbalance between relevant and irrelevant citations; just 17.5% of screened citations (equivalent to 6% of all citations) in the base case were included in full-text screening. In such instances, the predictions are biased towards the majority irrelevant citations, which produces falsely weighted predictions (i.e. irrelevant citations) (472). These issues have been encountered elsewhere (437).

Previous studies found that SLRs, which have more complex PICOS (population; intervention; comparator; outcome; study design) criteria tend to achieve less magnitude of workload savings (as defined in this study) (437). Also, it has been suggested that text-mining tools perform better for SLRs that only include RCTs (449). In this study, despite the complexity of comparators (due to lexical similarity) and the inclusion of single-arm studies, the workload savings were notable. It is reassuring that these workload savings did not come at the cost of missed citations.

In this study, a single research question was presented. It has been proposed that Abstrackr's predictions are more reliable when fewer research questions are defined (449). When a greater number of research questions are defined, the algorithm may find it more challenging to discern patterns during the training phase. To enhance pattern learning, a larger training sample may be required. However, this may be impractical and may negatively affect workload.

Although, as standard, emphasis is placed on attaining 100% sensitivity in SLRs, it seems unlikely that single-human screening would consistently attain this. An analysis of 280 single-human screeners observed a sensitivity of 86% in this cohort, based on 24,942 screening decisions and 2,000 abstracts. Dual-human screening in the same analysis attained 98% sensitivity. Specificity was 79% and 69% for single-human screening and dual-human screening, respectively (473).

Of importance, only citations deemed relevant by Screener 1 (using Abstrackr, n=451) were brought forward for full-text screening. Thus, mitigating against the negative impact of the high number of predicted false positives (n=2,001), and generating time savings. However, performance metrics are based on Abstrackr's predictions. Based solely on Abstrackr's predictions (thus, ignoring the labels provided by Screener 1), the high number of false positives would add to workload burden at full-text screening. Assuming it takes 4 minutes to retrieve a full text, and 5 minutes for full-text screening, full-text screening of these false positives would require 37.5 days (471). This outweighs workload savings generated at title and abstract screening. It was assumed that all citations screened before reaching the predefined maximum prediction score were deemed relevant by Abstrackr. This may overestimate the number of citations predicted to be relevant by Abstrackr and may partly contribute to the high number of predicted false positives.

7.4.2 Limitations

These findings are based on a single SLR in one disease area. Results are also likely impacted by the sample size, experience, and topic expertise of the human screeners. This limits the generalisability of results. Further research is warranted to investigate if results can be replicated for other research questions and disease areas.

This study assumed that Single-human screening judged all relevant citations with 100% accuracy. However, as described in 7.4.1, this may not be the case. In this study, both screeners were highly experienced. However, to limit any uncertainties associated with this assumption, Screener 2 (Single-human screening) was the more experienced screener. Ideally, however, performance would have been compared to a validated, pre-

screened database of citations. Pre-screening of this database would be conducted by two human screeners, in line with the gold standard approach.

The performance of Abstrackr was determined based on the behaviour of Screener 1 and how well Abstrackr agreed with the judgements of Screener 2. However, the judgements of Screener 1 and Screener 2 were not perfectly aligned. The performance of Abstrackr (given Screener 1's behaviour) and inter-rater reliability, in this study, are conflated. As such, the positive performance of Abstrackr may be underestimated. Consideration may have been given to training and screening in Abstrackr by both Screener 1 and Screener 2. Under this approach, Abstrackr could gain insight from both screeners.

The performance of Abstrackr, based on higher maximum prediction scores, was not evaluated. Higher maximum prediction scores (i.e. earlier stopping points) may result in further workload savings without missing relevant studies. Further research should investigate this, by downloading predictions at pre-specified thresholds. This approach would also give a more realistic indication of the number of false positives predicted by Abstrackr.

The number of times Abstrackr updated, and produced an updated list of predictions, was not recorded during this research. Thus, the number of times Abstrackr had an opportunity to retrain cannot be determined. Downloading Abstrackr's predictions at pre-defined thresholds, and recording the number of times Abstrackr updated, would facilitate an analysis of the pattern of the performance of Abstrackr's text-mining functions throughout screening.

Abstrackr frequently dictated the number of citations that could be screened in a given session. These technical issues may limit the potential for workload savings and may contribute to the proposed lack of trust in some human screeners (466). Such issues should be resolved to facilitate the use of Abstrackr in routine practice.

7.5 Conclusion

Abstrackr-assisted screening generated workload savings that did not come at the expense of omitting relevant citations. However, the importance of conducting further

research to investigate performance at stopping points defined by higher maximum prediction scores is emphasised. Sensitivity analysis at maximum prediction scores of 0.34458 and 0.29021 produced improved sensitivity but came at the expense of workload savings. Although title and abstract screening workload and time savings were notable, the proportion of false positives was high. The associated workload burden of these false positives may have negative workload implications at full-text screening, if relying solely on Abstrackr's predictions. Given that best practice requires two human screeners, a second screener might consider use of Abstrackr to exclude citations below 0.39540, but rely on their own judgements before this threshold. However, further research is warranted before generalising these results to different research questions.

Chapter 8 Utility Data in Relapsed/Refractory Diffuse Large B-Cell Lymphoma: Systematic Literature Review

8.1 Introduction

8.1.1 Chapter Aim

The aim of this chapter is to derive utility values for use in the bespoke cost-utility models, examining the cost effectiveness of tisagenlecleucel and axicabtagene ciloleucel for the treatment of patients with R/R DLBCL. Relevant data will be identified by a comprehensive SLR (11).

8.2 Methods

8.2.1 Systematic Literature Review

An SLR protocol was developed, in line with the Cochrane Handbook for Systematic Reviews of Interventions (204). Guidance regarding the search strategy was obtained from an NCPE Information Specialist. Reporting is conducted in line with PRISMA 2020 (205).

8.2.1.1 Population

The population was in line with those described in the licenses for tisagenlecleucel and axicabtagene ciloleucel; adult patients with R/R DLBCL, who received two or more prior lines of therapy (3, 4). No upper age limits were specified.

8.2.1.2 Intervention and Comparators

Studies reporting utility data in patients treated with any licensed therapy for R/R DLBCL (in Europe) were included.

8.2.1.3 Outcomes

Outcomes were required to be reported as a utility value; a format that allowed use as an input parameter in the cost-utility models. The following varieties of utility value were included:

1. Health-state utility values for progression-free survival and progressed disease
2. Utility values associated with long-term survival
3. Disutility values associated with treatment and associated administration or hospitalisation
4. Disutility values associated with short-term (eight weeks or less (143, 144)) and long-term (greater than eight weeks) adverse events of treatment

5. Disutility values associated with HSCT

8.2.1.4 Study Design

Any study type that provided the required outcome was included, with the exception of case studies or studies providing data on a single patient.

8.2.1.5 Search Methods

The search strategy is presented in Appendix G (Table A31). Electronic databases EMBASE, MEDLINE (via EBSCO), and CENTRAL (via the Cochrane Library) were searched from 01 January 2000 to 05 February 2021 inclusive. Articles were restricted to those published in English. Proceedings from the ASH and EHA Annual Conferences were hand searched for the years 2014 to 2020. Terms used in searching of conference proceedings included: 'tisagenlecleucel', 'tisa-cel', 'JULIET', 'axicabtagene ciloleucel', 'axi-cel', 'ZUMA-1', 'lymphoma', 'diffuse large B-cell lymphoma', 'DLBCL', 'large B-cell lymphoma', 'health-related quality of life', 'quality of life', 'utility', 'QOL', and 'HRQOL'.

8.2.1.6 Choice of Utility Values

The utility values chosen for use in the cost-effectiveness models were selected using several criteria. The hierarchy of criteria was as per that described in 4.2.1.6.

8.2.1.7 Citation Management

Identified citations were imported to Endnote® and transferred to Abstrackr. Duplicates were systematically searched for using software in Endnote® and identified manually throughout. Title and abstract screening was conducted by a single reviewer. The full text of potentially relevant citations were obtained and assessed for suitability for inclusion in the final evidence base. For quality assurance purposes, 10% of full-text articles were screened in duplicate by a second reviewer. Any disagreements were resolved by a third reviewer.

Data extraction was conducted using an adapted Cochrane data extraction form (208). Data recorded included population, intervention, comparators, outcomes, study design,

authors, title, and publication date. Extracted outcomes data were checked in duplicate by a second reviewer.

8.3 Results

A total of 462 citations were identified through database searches. Hand searching of conference proceedings yielded an additional four citations. Following removal of duplicates, 460 citations were screened. Title and abstract screening resulted in the exclusion of 432 citations and 28 were brought forward for full-text screening. Three studies were included in the final evidence base. A PRISMA 2020 diagram is presented in Figure 20.

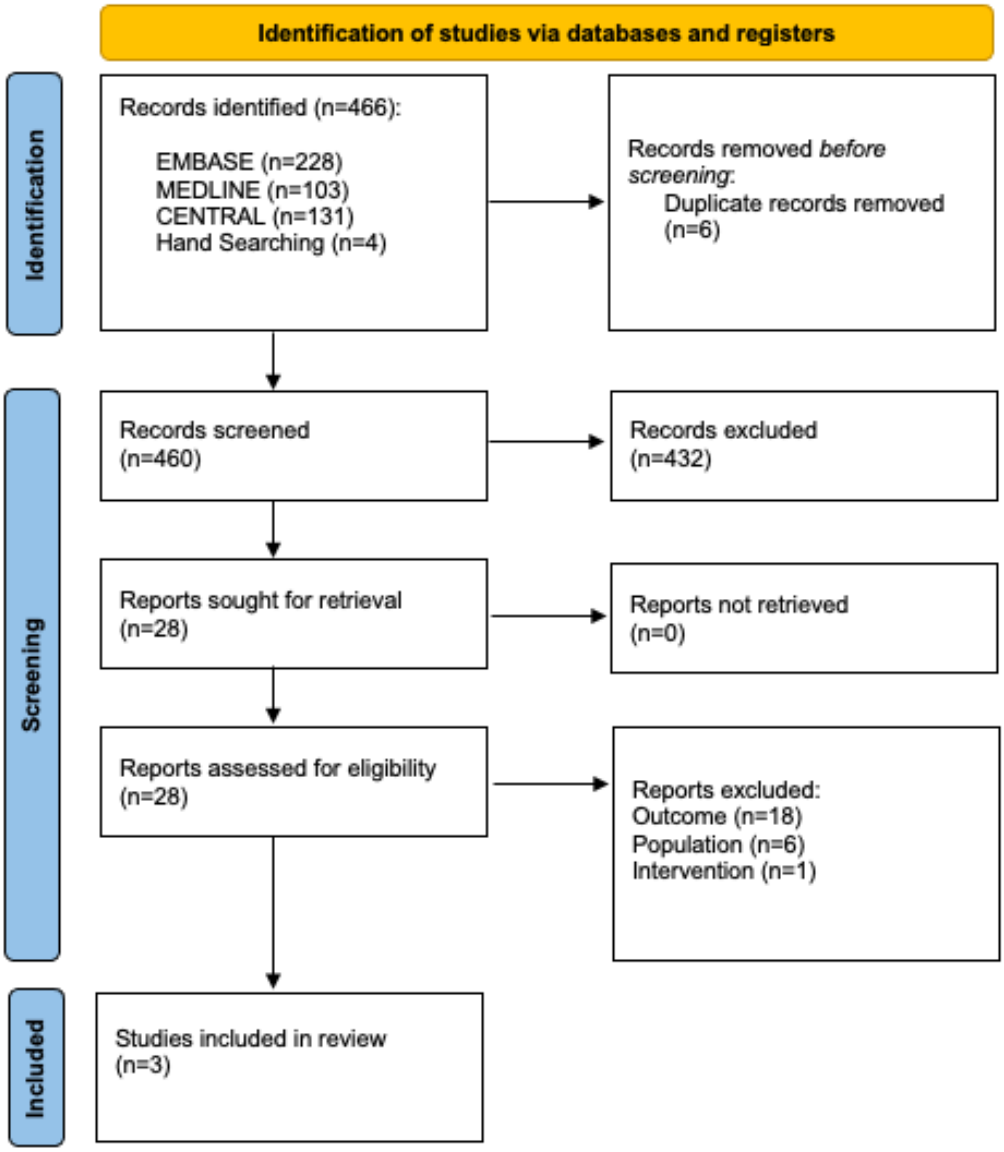

Figure 20 PRISMA diagram - systematic literature review of utility data for relapsed/refractory diffuse large B-cell lymphoma

8.3.1 Excluded Studies

Reasons for exclusion at full-text screening were outcome (n=18), population (n=6), and intervention (n=1). Exclusion based on outcome was generally because data were not presented as utility values and were therefore, not suitable for use in the cost-utility models. Despite studies reporting on patients with R/R DLBCL, many of these employed proxy data, which were not relevant to the population specified here. These studies were excluded based on population (n=6). Utility associated with treatment that was not licensed in Europe led to the exclusion of one study ('intervention'). A list of selected excluded studies is presented in Appendix G (Table A32).

8.3.2 Included Studies

The studies included in the final evidence base are summarised in Table 42.

Table 42 Studies included in systematic literature review of utility data for relapsed/refractory diffuse large B-cell lymphoma, after two or more lines of systemic therapy

Author, year	Intervention	Comparator	Study Design	Source	Utility Values			
					Progression-Free (SD/range)	Progressed (SD/ range)	Long-Term Survival	Adverse Events (SE)
Lin et al. 2018 (404)	Axi-cel	N/A	HRQOL Study	EQ-5D-5L data (ZUMA-1) mapped to EQ-5D-3L US valuation set	0.80 (0.14)	0.72 (0.17)	NR	-0.05 (0.04)
Betts et al. 2020 (474)	Pola+BR	BR	Cost-utility analysis	SF-36 data (JULIET) mapped to SF-6D	0.83*	0.71*	NR	NR
Roth et al. 2018 (475)	Axi-cel	Salvage therapy HSCT	Cost-utility analysis	**Axi-cel:** EQ-5D-5L (ZUMA-1) US valuation set **Long-term survival:** EQ-5D-5L, with US valuation set. US pop.(age 60-69 yrs.) **Salvage therapy:** literature **PD:** literature	**Axi-cel:** 0.74 (0.68,0.80) **Salvage therapy:** 0.67 (0.62,0.77) **In remission <6 months:** 0.78 (0.74,0.83)	0.39 (0.31,0.47)	0.82 (0.74,0.91)	NR

Axi-cel: Axicabtagene ciloleucel; **BR:** Bendamustine, rituximab; **HRQOL:** Health-related quality of life; **HSCT:** Haematopoietic stem cell transplant; **N/A:** Not applicable; **NR:** Not reported; **PD:** Progressed disease; **Pola+BR:** Polatuzumab, bendamustine, rituximab; **Pop:** Population; **SD:** Standard deviation; **SE:** Standard error; **US:** United States.
*SD/range not reported.

All studies reported health-state utility data, which were collected using a generic measure, in line with National Economic Evaluation Guidelines (11). The health-state utility values reported by Lin et al. (404) and Betts et al. (474) were closely aligned for the progression-free survival and progressed disease states. These utility data were derived directly from patients in ZUMA-1 (axicabtagene ciloleucel, Lin et al.) and JULIET

(tisagenlecleucel, Betts et al.). Data reported by Lin et al. were collected using the EQ-5D-5L, which were mapped to the EQ-5D-3L. These data were based on a small sample size (n=33, initially at screening), which decreased at subsequent assessments (404). Data reported by Betts et al. were collected during JULIET, and mapped from the SF-36 to the SF-6D (both generic measures) (474). The valuation set used by Betts et al. is not stated. Additionally, no information was provided regarding sample size, missing data, or frequency of data collection.

Roth et al. derived utility values, for 'axicabtagene ciloleucel' (0.74) and 'in remission with less than 6 months of follow up' (all arms, 0.78), from EQ-5D-5L data collected during ZUMA-1, with the US valuation set applied (475). Lin et al. is cited for these values; however, it is unclear why they differ from the values published by Lin et al. (404, 475).

Roth et al. derived utility for salvage therapy (0.67, active treatment) by applying a disutility value (-0.15, for cytotoxic salvage chemotherapy), sourced from Huntington et al. (476), to the utility value for long-term remission (0.82). Huntington et al. examined the cost effectiveness of routine surveillance imaging in patients with DLBCL in first remission. Utility values were informed by expert opinion; however, no further detail was provided (476). This value cannot be validated.

Utility for progressed disease, in Roth et al. (475), was derived by Doorduijn et al. (477) and adapted by Best et al. (478). Doorduijn et al. analysed data collected from 1996 to 1999, in patients aged 65 to 90 years, with newly diagnosed NHL (n=128). Data were collected using the EQ-5D-3L, the EORTC QLQ-C30, and the MFI-20 (Multidimensional Fatigue Inventory; generic measure) (477). Best et al. subsequently weighted the three-month utility scores, from Doorduijn et al., by the proportion of patients in the GELA study with progressed disease (477, 478). GELA was a phase III trial, examining the efficacy of CHOP (cyclophosphamide, doxorubicin, vincristine) versus R-CHOP in patients aged 60 years and older with DLBCL (n=399) (479). The resultant value of 0.39, for the progressed disease state, was notably lower than those reported by Lin et al. (404) and Betts et al. (474). The utility values, derived by Huntington et al. (476) and Doorduijn et al. (477) (adapted by Best et al. (478)), are not directly applicable to the population specified here.

Roth et al. assumed that patients, who were in remission for greater than six months, accrued utility equivalent to the mean EQ-5D-5L score with tariffs of the US general population, aged between 60 and 69 years (480). These data were collected between 2000 and 2002; both societal preferences and HRQOL research methods are likely to have changed since then and so, these values were considered outdated for this SLR. Of note, Roth et al. is the only study, of the three identified, to explicitly account for HRQOL associated with long-term survivors (404, 474, 475).

Lin et al. was the only study to report disutility related to adverse events. This value (-0.05) was associated with 'axicabtagene ciloleucel-related toxicity'. No further detail was provided (404).

8.4 Additional Searches

Due to the paucity of data identified though the SLR, a search of websites of national HTA agencies was conducted. Similar to the approach taken in Chapter 4, utility values included in Applicant HTA submissions of tisagenlecleucel and axicabtagene ciloleucel, appraised by national HTA agencies, were examined. Websites of national HTA agencies were searched (from inception to 06 March 2021) using the terms 'tisagenlecleucel', 'axicabtagene ciloleucel', and 'diffuse large B-cell lymphoma'. Full HTA appraisal documents were reviewed. Summary documents were not included. HTA agencies that published their HTA appraisals in English were selected. The websites of the following HTA agencies were searched:

1. NICE, UK (311)
2. SMC, Scotland (312)
3. CADTH, Canada (313)
4. NoMA, Norway (314)

Based on these searches, the following HTA appraisals were reviewed:

1. NICE:
    - Corbett et al. Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma: A Single Technology Appraisal (TA567; 2018) (333).

- Corbett et al. Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma, mediastinal B-cell lymphoma and follicular lymphoma: A Single Technology Appraisal (TA559; 2018) (331).

2. SMC:
    - Tisagenlecleucel for the treatment of adult patients with relapsed or refractory diffuse large B-cell lymphoma after two or more lines of systemic therapy: Detailed advice (2019) (481).
    - Axicabtagene ciloleucel for the treatment of adult patients with relapsed or refractory diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after two or more lines of systemic therapy: Detailed advice (2019) (482).
3. CADTH:
    - Tisagenlecleucel for Diffuse Large B-Cell Lymphoma: Economic Review Report (2019) (483).
    - Axicabtagene Ciloleucel for Diffuse Large B-Cell Lymphoma: Economic Review Report (2019) (484).
4. NoMA:
    - Single Technology Assessment: Tisagenlecleucel (Kymriah®) for the treatment of second or later relapsed/refractory diffuse large B-cell lymphoma (2019) (399).
    - Single Technology Assessment: Axicabtagene ciloleucel (Yescarta®) for the treatment of second or later relapsed/refractory diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma (2018) (332).

    Of note, an updated HTA appraisal of axicabtagene ciloleucel by NoMA (2020) was identified (485); however, no additional utility data were included in this appraisal. Only the original HTA appraisal (2018) (332) was considered here.

In addition, the ICER HTA (296) was examined. The websites of NCPE (Ireland) and PBAC (Australia) was also searched. However, only summary documents were available (9, 10, 316).

8.4.1 Health-State Utility Values

All HTA agencies used utility values derived from data collected during JULIET and ZUMA-1 for the appraisals of tisagenlecleucel and axicabtagene ciloleucel, respectively (331-333, 399, 481-484). NoMA was the only agency to publish these values (332, 399).

During JULIET, HRQOL data were collected using the SF-36. Data were collected at screening, months 3, 6, 12, and 18 post-infusion (400). For the NoMA HTA appraisal, these data were mapped to the EQ-5D-3L and converted to utility values (by the Applicant) using the UK valuation set (399). HRQOL data in ZUMA-1 were collected using the EQ-5D-5L. Data were collected in a cohort who had received prophylactic tocilizumab administered on day 2 post-axicabtagene ciloleucel infusion (n=34; data collected from 33 of these). This cohort was distinct from the cohort used to inform the efficacy of axicabtagene ciloleucel. The frequency of data collection was not reported in the NoMA HTA appraisal (332). However, Lin et al. indicated that data were collected at screening, week 4, month 3, and month 6 post-infusion (404). The NoMA HTA appraisal does not indicate whether the ZUMA-1 EQ-5D-5L data were mapped to the EQ-5D-3L (332). Mapping of utility values is an appropriate approach to derive EQ-5D-3L utilities in the absence of such data, but this will increase uncertainty and error around the utility estimates (486). The valuation set applied to the EQ-5D-5L data is not reported. The health-state utility values used in the NoMA HTA appraisals are presented in Table 43.

Patients receiving tisagenlecleucel, who were alive after five years (two years in scenario analysis), in the NICE HTA appraisal (TA567) (333), and three years (five years in scenario analysis) in the SMC HTA appraisal (481) were assumed to have HRQOL equivalent to that of the progression-free survival state, regardless of health-state membership. This assumption does not appear to have been made in the NoMA or CADTH HTA appraisals (399, 483).

The NICE HTA appraisal of axicabtagene ciloleucel (TA559) assumed that patients in the progression-free survival state had HRQOL equivalent to that of the age- and sex-matched general population at month 52 (point of convergence between OS and PFS; 2 years and 5 years examined in scenario analysis) (331). This assumption was applied to patients at two years (five years in scenario analysis) in the SMC, CADTH, and NoMA HTA

appraisals of axicabtagene ciloleucel (332, 482, 484). The NICE and CADTH HTA appraisals obtained general population utility values from Janssen et al. As described in Chapter 4, Janssen et al. published general population utility (according to age and sex), captured using the EQ-5D-3L, for a variety of countries. To estimate general population utility of England, Janssen et al. used data from the Health Survey for England 2010. This survey collected data from 14,763 randomly selected participants (2008), using computer-assisted interviews. The time-trade-off value set derived from an English population was subsequently applied (307). Further detail is provided in 8.4.2.3. The NoMA HTA appraisal sourced general population from Sun et al. (325) and BurstrØm et al. (326), who report age-specific utility of the Swedish general population (captured using the EQ-5D-3L). The source of general population utility was not reported in the SMC HTA appraisal (482).

Assumptions regarding HRQOL of patients considered to be long-term survivors (for both tisagenlecleucel and axicabtagene ciloleucel) were based on a study by Maurer et al. This study evaluated OS and cause-specific survival (conditional on being alive and disease-free at 12 and 24 months post-diagnosis) in patients with DLBCL, who were treated with immunochemotherapy (113). Patients (n=767) were prospectively enrolled onto the Molecular Epidemiology Resource of the University of Iowa/Mayo Clinic Lymphoma Specialized Program of Research Excellence or onto North Central Cancer Treatment Group NCCTG-N0489. Patients with newly diagnosed DLBCL, who were event-free at 24 months, had OS equivalent to that of the age- and sex-matched general population (113).

Utility data derived from JULIET and ZUMA-1 are subject to limitations. The uncontrolled, open-label nature of these trials makes the outcomes susceptible to bias. The small sample sizes further add to the potential for bias. This was a greater concern with ZUMA-1, as utility data were only available for 33 patients. This reduced to 27, 20, and 7 patients at week 4, month 3, and month 6 post-infusion, respectively (404). The progression-free survival state was informed by 49 observations, and the progressed disease state was informed by 5 observations (332). Utility values derived from JULIET were available for 105 patients initially at screening, and decreased in number over time (399). In addition, the utility data were derived largely from patients who were in response after receiving tisagenlecleucel (400). No associated measures of uncertainty were presented for the utility data derived from JULIET and ZUMA-1 (332, 399).

In terms of HRQOL of patients considered to be long-term survivors, there is limited evidence to support the assumption that patients alive at specific time points are subject to age- and sex-matched general population utility. Maurer et al. was not conducted in the R/R setting (113). A follow-up study assessed OS stratified by PFS at 24 months. This study used IPD from patients (n=5,853) with DLBCL who were enrolled in 14 different multicentre, international RCTs (487). The findings of this study echoed those of the first study by Maurer et al. (113). However, a separate study (Howlader et al. (390)), indicated that excess mortality can remain for up to five years following diagnosis. Howlader et al. used data from 18,047 patients included in the US Surveillance, Epidemiology, and End Results database to assess factors associated with DLBCL cancer-specific mortality. This study was conducted in a newly diagnosed population; findings may not be generalisable to R/R disease (390). Additionally, patients who survive long-term may experience long-term effects, impacting their HRQOL.

Table 43 Health-state utility values in relapsed/refractory diffuse large B-cell lymphoma identified from the Norwegian Medicines Agency (NoMA) HTA appraisals of tisagenlecleucel and axicabtagene ciloleucel (332, 399)

Health State	Utility Value‡	
	Derived from JULIET	Derived from ZUMA-1
Progression-Free Survival	0.83*	0.72†
Progressed Disease	0.71*	0.65†
Long-Term Survival	0.83	Equivalent to that of the age- and sex-matched general population

HTA: Health technology assessment.
*UK valuation set applied to EQ-5D-3L (mapped from SF-36) data collected during JULIET.
†EQ-5D-5L data collected during ZUMA-1. It is not stated whether these were mapped to the EQ-5D-3L. The valuation set applied was not reported.
‡No measures of uncertainty presented.

Utility values, which corresponded to those of patients with renal cell carcinoma receiving second-line treatment, used in NICE TA306 (488), were presented in the NICE (TA567 and TA559) and NoMA HTA appraisals of tisagenlecleucel and axicabtagene ciloleucel (331-333, 399). However, these were not deemed relevant to the population specified in this SLR. No further consideration was given to these values.

The ICER HTA employed values of 0.83 and 0.39 for the progression-free survival and progressed disease states, respectively (296). These were derived by Best et al. and Doorduijn et al. (477, 478). As described in 8.3.2, this population was not considered

relevant to the population here. Patients alive after five years, in the ICER HTA, were assumed to be long-term survivors and accrued HRQOL equal to that of the progression-free survival state (296).

8.4.2 Treatment and Adverse Event Disutility Values

In estimating disutility associated with treatment and adverse events, two alternative approaches were taken in the HTA appraisals of tisagenlecleucel and axicabtagene ciloleucel.

8.4.2.1 Tisagenlecleucel

The HTA appraisals of tisagenlecleucel employed a single disutility value (-0.15) to capture disutility of treatment and all associated adverse events (except CRS) (333, 399, 483). This value was derived by Guadagnolo et al., who examined the cost effectiveness of computerised tomography in the routine follow up of patients (aged 25 years old) after primary treatment for Hodgkin's lymphoma (337). This was applied for the duration of treatment for comparators and for the duration of hospitalisation, starting from the time of commencement of lymphodepleting chemotherapy, for tisagenlecleucel. In the NoMA HTA appraisal, this duration was 26 days and 72 days for tisagenlecleucel and salvage chemotherapy, respectively (399).

Disutility associated with grade 3-4 CRS-related ICU admission was assumed to reduce utility to 0 for the duration of stay. A duration of 9.21 days, from JULIET, was reported in the CADTH and NoMA HTA appraisals (399, 483); duration not reported in the NICE HTA appraisal (TA567) (333). It is unclear from the SMC HTA appraisal if this assumption was employed; however, disutility was included to account for grade 3-4 CRS (ICU stay) (481). As published by NoMA, this assumption was applied to 21.6% of patients receiving tisagenlecleucel (399). Utility was also reduced to 0, to account for non-CRS ICU admission. A duration of 0.86 days was assumed, as published by CADTH and NoMA (399, 483); duration was not reported in the other HTA appraisals (333, 481). It was assumed that all patients receiving tisagenlecleucel incurred this disutility (333, 399).

For patients receiving HSCT (both autoSCT and alloSCT), a disutility value of -0.30, derived by Guadagnolo et al., was employed. In the NICE (TA567) and CADTH HTA appraisals, this was applied for 365 days (333, 483). The NoMA HTA appraisal assumed a duration of 72 days (399). The duration was not published in the SMC HTA appraisal (481). The CADTH HTA appraisal also included a scenario, which incorporated disutility due to febrile neutropenia. The value employed was not reported (483).

Data obtained from Guadagnolo et al. are subject to limitations. Guadagnolo et al. state that utility values were derived from the literature or expert opinion (337). However, no further detail was provided. The appropriateness of these values cannot be validated. The patients in Guadagnolo et al. are younger than patients defined in the cost-utility models (developed as part of this research), and the disease is not aligned with the disease defined in this SLR. No associated measures of uncertainty were presented for the utility data derived by Guadagnolo et al. (337).

8.4.2.2 Axicabtagene Ciloleucel

HTA appraisals of axicabtagene ciloleucel employed disutility values derived from the literature, accounting for grade ≥3 adverse events occurring in greater than 10% of the population in ZUMA-1 (331, 332, 484) (detail not provided in SMC HTA appraisal (482)). When a disutility specific to an adverse event was not identified, an assumption was made that this disutility equated to the maximum of the identified non-CRS adverse event disutility. Total disutility was applied as a once-off in the first cycle (331, 332). This disutility was estimated to be -0.03 and -0.01 in the NICE and NoMA HTA appraisals, respectively (331, 332). The reason for the discrepancy between these estimates may be due to the different durations applied for CRS (four days NICE versus eight days NoMA) (331, 332). With the exception of the NoMA HTA appraisal (332), disutility was not applied for adverse events of comparator therapies (331, 482, 484). The disutility value employed in the comparator arm, in the NoMA HTA appraisal, was -0.04. However, it is not clear how this was derived (332). In line with the HTA appraisals of tisagenlecleucel (333, 399, 481, 483), a utility of 0 was assumed for patients experiencing grade 3-4 CRS-related ICU admission (331, 332, 482). It is not stated if this assumption was employed in the CADTH HTA appraisal (484).

The utility values employed for individual adverse events, in the NICE and NoMA HTA appraisals of axicabtagene ciloleucel, are presented in Table 44 (331, 332). There is uncertainty associated with these values as none were derived from the population of interest. There was also variation in the methodology of how these values were derived.

Table 44 Adverse event disutility values employed in the NICE and NoMA HTA appraisals of axicabtagene ciloleucel (331, 332)

Adverse Event	Value*	Duration (days)	Population	Source
Anaemia	-0.12	14	Metastatic renal cell carcinoma	Swinburn et al. 2010 (489)
Cytokine Release Syndrome	0	4/8	Assumption	Assumption based on Hettle et al. (12)
Febrile Neutropenia	-0.15	6	Metastatic breast cancer	Lloyd et al. 2006 (329)
Encephalopathy		9	Assumption	Assumed equal to the maximum of other, non-cytokine release syndrome adverse event disutility values
Hypophosphataemia		16		
Hypotension		5		
Leukopenia		21		
Lymphocyte Count Decreased		64		
Neutrophil Count Decreased		17		
White Blood Cell Count Decreased		40		
Neutropenia	-0.09	47	Non-small cell lung cancer	Nafees et al. 2008 (490)
Pyrexia	-0.11	2	Chronic lymphocytic leukaemia	Beusterien et al. 2010 (491)
Platelet Count Decreased	-0.11	50		Tolley et al. 2013 (492)
Thrombocytopenia		63		

HTA: Health technology assessment; **NICE:** National Institute for Health and Care Excellence; **NoMA:** Norwegian Medicines Agency.
*No measures of uncertainty presented.

In the ICER HTA, disutility associated with treatment was assumed to capture disutility associated with adverse events (296). Treatment-related disutility values, for chemotherapy (-0.42) and HSCT (-0.57), were derived by Sung et al. (321) (described in Chapter 4). Disutility associated with grade 3-4 CRS-related ICU admission was as per the HTA appraisals of tisagenlecleucel and axicabtagene ciloleucel. This was applied for eight days, which was reported to equate to the median duration of CRS-related ICU stay (sourced from Hettle et al. (12)) (296). Disutility values were applied to all treatment arms.

Table 45 Disutility values for relapsed/refractory diffuse large B-cell lymphoma identified in HTA appraisals (331-333, 399, 481-484) and the ICER HTA (296)

Parameter	Value (range)	Duration	Identified Through	Source
Tisagenlecleucel Treatment incl. Adverse Events	-0.15*	Duration of treatment; 26 days (HTA appraisal (399))	HTA appraisals (333, 399, 483)	Guadagnolo et al. (337)
Salvage Chemotherapy Treatment incl. Adverse Events	-0.42 (0.16, 0.83)	Duration of treatment	ICER HTA (296)	Sung et al. (321)
CRS ICU Admission	0*	Duration observed in relevant trial; Tisagenlecleucel: mean duration 9.21 days (399, 483) Axicabtagene ciloleucel: mean duration 8.6 days (332) Comparator treatment (not specified): median duration 8 days (296)	HTA appraisals (331-333, 399, 481-483) ICER HTA (296)	Assumption
Non-CRS ICU admission (Tisagenlecleucel)		Mean duration 0.86 days (399, 483)	HTA appraisals (333, 399, 481, 483)	Assumption
HSCT	-0.30*	365 days	HTA appraisals (333, 483)	Guadagnolo et al. (337)
		72 days	HTA appraisals (332, 399)	
	-0.57 (0.33, 0.87)	Not reported	ICER HTA (296)	Sung et al. (321)
Grade 3-4 Adverse Events Associated with Axicabtagene Ciloleucel	-0.03*	Once-off at start of first cycle	HTA appraisal (331)	Derived from literature (Table 44) (331, 332, 484)
	-0.01*		HTA appraisal (332)	
Grade 3-4 Adverse Events Associated with Salvage Chemotherapy	-0.04*	Once-off at start of first cycle	HTA appraisal (332)	Derived from literature, no further detail provided (332)

CRS: Cytokine release syndrome; **HSCT:** Haematopoietic stem cell transplant**; HTA:** Health technology assessment; **ICER:** Institute for Clinical and Economic Review; **ICU:** Intensive care unit.
*No measures of uncertainty presented.

8.4.2.3 Age-Related Disutility

The NICE HTA appraisal of tisagenlecleucel (TA567) (333) accounted for disutility associated with increasing age, based on Ara and Brazier (328). Ara and Brazier present a formula to estimate general population utility by age. This was derived using ordinary least square regression on EQ-5D-3L responses from a sample of 26,679 participants (aged 16 to 98 years) in the Health Survey for England. A utility adjustment factor was calculated as the ratio between general population utility at each age and the corresponding value at the starting age in the model. The adjustment factor was multiplied by the health-state utility values and applied over the model time horizon (333).

The NoMA HTA appraisal of tisagenlecleucel (399) accounted for disutility associated with increasing age, based on data from Sun et al. (325) and BurstrØm et al. (326). Further detail regarding this approach is provided in Chapter 4. Adjustment to utility to account for increasing age was not mentioned in the SMC and CADTH HTA appraisals of tisagenlecleucel (481, 483). Likewise, the ICER HTA did not comment on whether an adjustment to utility was applied (296).

The HTA appraisals of axicabtagene ciloleucel do not appear to have adjusted utility to account for increasing age. This is because patients alive after a certain time point (either two or five years) were assumed to have utility equivalent to the age- and sex-matched general population (331, 332, 482, 484). General population utility estimates, in the NICE and CADTH HTA appraisals, were sourced from Janssen et al. (331, 484). This study reports general population utility according to 10-year age bands (i.e. 45 to 54, 55 to 64, 65 to 74 years, etc.) (307). This results in a drop in utility every 10 years as opposed to a gradual decline.

8.5 Utility Values Selected for Use in the Bespoke Cost-Utility Models

The utility values that will be used to inform the cost-utility models, developed as part of this research, are summarised in Table 46. The values selected for use in scenario analysis are also presented (11). The durations employed are presented in Chapter 9 (Table 52).

No treatment-specific utility values were identified for comparator treatments. Therefore, in the cost-utility models, utility will be applied according to health-state occupancy and will be treatment independent. The health-state utility values derived from JULIET, with the UK valuation set applied (399), were deemed most appropriate for this research. Although subject to limitations, they were considered a richer data source than those derived from ZUMA-1. They are based on a greater number of observations and are aligned with National Economic Evaluation Guidelines (11). The values derived from ZUMA-1 will be explored in scenario analysis.

In the absence of supportive data, it will be assumed that patients who are considered long-term survivors (after 60 months) have utility equivalent to that of the progression-free survival state. Due to the uncertainty in this assumption, a scenario will be explored, whereby patients who are considered long-term survivors have utility equivalent to that of the age- and sex-matched general population (328).

Disutility derived by Guadagnolo et al. (-0.15) will be applied to patients undergoing apheresis, bridging chemotherapy, and lymphodepleting chemotherapy (337). Although subject to limitations, it is reassuring that this value is closely aligned with the value applied to paediatric and young adult patients with R/R ALL undergoing these procedures. This value (-0.20) was identified through the SLR presented in Chapter 4, and derived by Kwon et al. (308). The value derived by Guadagnolo et al. was favoured over that derived by Sung et al. (also identified during this SLR), due to the considerable methodological limitations associated with Sung et al. (described in Chapter 4) (321).

In terms of adverse events, grade 3-4 CRS-related and non-CRS-related ICU admission will be assumed to have a utility of 0, for the duration observed in the relevant trial. This assumption will be applied to patients receiving tisagenlecleucel and axicabtagene ciloleucel only. In the absence of robust supportive data, this is a necessary assumption and in line with the published literature (331, 333). Due to the potential impact on HRQOL, disutility will also be included, for all therapies, to account for disutility associated with febrile neutropenia (330). This disutility value (-0.15) was derived by Lloyd et al., who reported disutility associated with febrile neutropenia in patients with metastatic breast cancer (329). Disutility due to pancytopenia will also be included for

tisagenlecleucel and axicabtagene ciloleucel. Based on clinical opinion obtained by the NICE ERG, in the assessment of CD19 CAR T-cell therapy for R/R mantle cell lymphoma (TA677), pancytopenia is one of the most impactful adverse events (in terms of HRQOL) experienced by patients treated with CD19 CAR T-cell therapy. The impact of pancytopenia is most notable in the first few months after infusion, gradually improving to resolution within one year (330). For this analysis, it will be assumed that patients experience pancytopenia for six months. In the absence of data, disutility will be assumed equivalent to that of febrile neutropenia (-0.15) (329). No further disutility due to adverse events will be considered in the model due to the paucity of data, and uncertainty in the data that were identified. Based on published HTA appraisals, disutility due to adverse events is not expected to be a major driver of cost effectiveness (331, 333). This is because disutility due to adverse events is likely to be small relative to the overall QALY gain.

With the exception of apheresis, bridging chemotherapy, and lymphodepleting chemotherapy, no treatment-related disutility values will be applied in the base case. It is assumed that the health-state utility values incorporate some degree of disutility due to treatment and adverse events. This is to avoid double-counting. Additionally, none of the identified disutility values, accounting for treatment and adverse events, were treatment- or disease-specific. It is difficult to conclude how generalisable they are to the inclusion criteria defined here. The inclusion of a treatment- and adverse event-related disutility, based on Guadagnolo et al., will be explored in scenario analysis.

A disutility value of -0.30, derived by Guadagnolo et al., will be applied to patients undergoing HSCT (337). The will be applied for 365.25 days (333, 483).

The formula by Ara and Brazier will be used to adjust utility data to account for increasing age, using the multiplicative approach (328).

Table 46 Utility values in relapsed/refractory diffuse large B-cell lymphoma used in bespoke cost-utility model: base case and scenario analyses

Parameter	Value	Source	Justification
Base Case			
Progression-Free Survival	0.83*	EQ-5D-3L (mapped from SF-36) collected in JULIET with UK valuation set applied, identified through HTA appraisal (399)	Derived from a generic measure of HRQOL, in line with National guidelines (11). Data considered more robust than those derived from ZUMA-1
Progressed Disease	0.71*		
All Patients Alive After 60 Months	0.83*	HTA appraisal (333)	Assumption. HRQOL equivalent to that of progression-free survival state
Disutility Associated with Treatment and Adverse Events	-0.15*	Guadagnolo et al. (337), identified through HTA appraisals (333, 399, 483)	Applied to patients undergoing apheresis, bridging chemotherapy, and lymphodepleting chemotherapy
CRS ICU Admission	-0.83*	HTA appraisals (331-333, 399, 482, 483), ICER HTA (296)	Assumption (i.e. a utility of 0). Accounts for impact of ICU admission on HRQOL
Non-CRS ICU admission		HTA appraisals (333, 399, 481, 483)	
Febrile neutropenia	-0.15*	Lloyd et al. (329), identified through HTA appraisals (331, 332)	Febrile neutropenia may require hospitalisation; expected to have an impact on HRQOL
Pancytopenia			Assumption. Based on HTA appraisal (TA677) (330)
HSCT	-0.30*	Guadagnolo et al. (337), identified through HTA appraisals (333, 399, 483)	Patients may experience a decrease in HRQOL post-HSCT. Also accounts for associated adverse events
Age-Related Disutility	Ara and Brazier (328)	Ara and Brazier (328), identified through HTA appraisal (333)	Adjustment so that utility is not higher than that of general population
Scenario Analysis			
Progression-Free Survival	0.72*	EQ-5D-5L data collected during ZUMA-1, identified through HTA appraisal (332)	Data from ZUMA-1 also available
Progressed Disease	0.65*		
All Patients Alive After 60 Months	Age- and sex-matched general population utility (328)	Ara and Brazier (328), identified through HTA appraisals (331, 332, 482, 484)	Long-term survivors may experience HRQOL improvement
Disutility Associated with Treatment and Adverse Events	-0.15*	Guadagnolo et al. (337), identified through HTA appraisals (333, 399, 483)	Applied to all treatments. Health-state utility values may not incorporate disutility due to

	treatment and adverse events

CRS: Cytokine release syndrome; **HRQOL:** Health-related quality of life; **HSCT:** Haematopoietic stem cell transplant; **HTA:** Health technology assessment; **ICU:** Intensive care unit; **UK**: United Kingdom.
*Standard error not reported. Assumed 25% of mean point estimate.

8.6 Discussion

This SLR identified a limited number of studies providing utility data on adult patients with R/R DLBCL. The majority of full-text studies, identified through database searching, were excluded based on outcome (i.e. outcome not presented as a utility value). Of the studies identified through database searching, three met the SLR inclusion criteria (404, 474, 475). A search of utility data, used in international HTA appraisals of tisagenlecleucel (333, 399, 481, 483) and axicabtagene ciloleucel (331, 332, 482, 484), in addition to the ICER HTA (296), identified additional data.

8.6.1 Health-State Utility Values

Health-state utility values for the progression-free survival state ranged from 0.67 (475) to 0.83 (399), while values for the progressed disease state ranged from 0.39 (475) to 0.71 (399). Of note, the value for the progression-free survival state derived from ZUMA-1 (0.72 (332)) is equivalent to that derived for the progressed disease state from JULIET (0.71 (399)). Health-state utility values, derived from data collected during JULIET, were identified in two publications. Values presented in these publications were consistent (399, 474). Health-state utility values, derived from data collected during ZUMA-1, were identified in three publications (332, 404, 475). However, values for the progression-free survival state were not consistent amongst the three publications. Due to the paucity of published detail (Lin et al. is published as an abstract only (404)), the reasons for differences between the published values cannot be determined. This raises concerns over the robustness of these values.

Uncertainty lies in the most appropriate health-state utility values for use in the cost-utility models. Values, derived from JULIET and ZUMA-1, were aligned with National Economic Evaluation Guidelines (11). Both trials collected HRQOL data, using a generic measure, in a population of relevance to this SLR. However, utility data derived from JULIET were mapped from the SF-36 to the EQ-5D-3L, increasing uncertainty in these

estimates. It is not stated if utility data collected during ZUMA-1, using the EQ-5D-5L, were mapped. Notably, the EQ-5D-5L results in a lower utility gain, when compared to the EQ-5D-3L (493). This may partly explain the lower values derived from ZUMA-1. Differences between health-state utility values, derived from JULIET and ZUMA-1, may also be due in part to differences in populations and methodology. Estimates are limited by the small sample sizes and low number of observations. The limited number of patients reporting HRQOL data reflects the challenges in collecting HRQOL data at an advanced stage of DLBCL. The single-arm nature of JULIET and ZUMA-1 prevents an analysis of the relationship between treatment effect and HRQOL.

Data derived from JULIET were chosen for the base case (399); however, the limitations of these data are acknowledged. The disengagement of patients from HRQOL assessments with time, results in data that are not representative of the entire cohort. This limits the generalisability of these data to patients in clinical practice. Additionally, when these data are used in cost-utility analyses, the subsequent cost-effectiveness estimates may be biased.

Variation was observed regarding assumptions of HRQOL of patients who are considered long-term survivors. In the absence of robust data, two different assumptions were identified. The assumption that patients who are alive after a certain time point (usually two or five years) have utility equivalent to that of the age- and sex-matched general population will not be employed in the base case of the cost-utility models. An SLR, examining the evidence to support this assumption, found a limited evidence base. The authors cautioned that care is required in asserting this assumption (494).

The assumption that patients who are considered long-term survivors do not revert to utility equivalent to the age- and sex-matched general population is supported by the finding that long-term survivors of cancer have unmet needs (495, 496). Domains of unmet need, which impact on HRQOL, include psychosocial issues such as fear of cancer recurrence, uncertainty about the future, and worry about family and friends. Long-term effects arising from treatment can also result in loss of productivity and participation in society (497). Uncertainty in the most appropriate methods to model HRQOL of long-term survivors undermines the robustness of QALY estimates. There is a need for

relevant studies (clinical trials, observational studies, registries) to collect long-term HRQOL data. The EORTC QOL cancer survivorship questionnaire is currently in development; this aims to capture the full range of physical, mental, and social HRQOL issues relevant to disease-free long-term survivors of cancer (498). Other HRQOL instruments have been developed to specifically measure HRQOL among patients with cancer who survive long term (499, 500); there is only limited information on the psychometric properties of these (501). It is likely to be some time before a validated measure is available for widespread use. Until robust data become available, a conservative approach is required to model the HRQOL of these patients.

Although patients alive after 60 months in the cost-utility models will be assumed to have utility equivalent to the progression-free survival state, this value (0.83) (399) is slightly higher than general population utility for patients aged 55 to 64 in England (0.82, Janssen et al. (307)). The progression-free survival state utility value, derived from data collected during ZUMA-1 (0.72) (332), is equivalent to general population utility for patients aged 75 years and older in England (0.72) (307). However, when compared to age- and sex-matched general population utility, estimated using the formula by Ara and Brazier, the progression-free survival state utility value (0.83) (399) is slightly lower (0.85 for patients aged 56 years; the median ages of patients in the cost-utility models) (328).

8.6.2 Treatment and Adverse Event Disutility Values

Limited data were identified regarding disutility associated with active treatment and adverse events in patients with R/R DLBCL. None of the identified values were aligned with the inclusion criteria of this SLR. Identified disutility values were based on assumptions, or derived from populations, which were not aligned with the population specified here. It is difficult to validate the generalisability of these assumptions and values to the population of interest.

Different approaches, accounting for disutility due to adverse events, were taken in the HTA appraisals of tisagenlecleucel (333, 399, 483) and axicabtagene ciloleucel (331, 332, 484). It is difficult to determine which approach results in a more accurate reflection of the true disutility experienced by patients. Although uncertainty exists in these values,

this parameter was not a driver of cost effectiveness in published HTA appraisals (331-333).

There was a notable difference in identified HSCT-related disutility values; -0.30 derived by Guadagnolo et al. (337) and -0.57 derived by Sung et al. (321). Both are subject to limitations (described in 8.4.2 and Chapter 4). The value derived by Guadagnolo et al. (337) was most closely aligned with the HSCT-related disutility value (-0.20) identified in the SLR of utility data in R/R ALL, derived by Forsythe et al. (309) (Chapter 4). Despite differences, in populations and disease areas, between Forsythe et al. (309) and Guadagnolo et al. (337), the close alignment between values provides some reassurance. Disutility due to HSCT was not a driver of cost effectiveness in published HTA appraisals (331-333).

8.6.3 Age-Related Disutility

Considering the potential long-term survival benefit associated with tisagenlecleucel and axicabtagene ciloleucel, utility data should be adjusted for age. Not adjusting for increasing age may favour the total QALY gain of tisagenlecleucel and axicabtagene ciloleucel (due to potential long-term survival versus salvage chemotherapy). In the cost-utility models, utility will be adjusted based on Ara and Brazier (using a multiplicative approach) (328). A limitation of this approach is that the adjustment formula was derived from a UK population. Published literature suggests that the population in Ireland value health differently to those in the UK. For example, for the lowest level of health in the anxiety/depression domain of the EQ-5D-5L, the decrement derived in Ireland was -0.65, while that derived in the UK was -0.29 (502). As such, the utility adjustment, derived by Ara and Brazier, may not reflect the true adjustment based on a population in Ireland. In the absence of data pertaining to Ireland, this is a necessary assumption.

General population utility estimates, based on the formula by Ara and Brazier, were chosen over those reported by Janssen et al. for a number of reasons. Ara and Brazier adjust for the proportion of male and female patients in the sample. Male patients in the general population in England have been reported to have higher utility than female (307). Additionally, a gradual decline in utility with increasing age, as suggested by Ara

and Brazier, was considered more appropriate than a decline in utility every 10 years (as per Janssen et al.) (307, 328).

8.6.4 Limitations

There are several limitations to this SLR. These are in addition to those described thus far. No measures of associated uncertainty were presented for any of the utility values chosen for use in the cost-utility models. As such, an assumption is required, for the purpose of the PSA, whereby the standard error is equivalent to 25% of the mean point estimate. This is a necessary assumption, and in line with approaches taken in the literature (333). However, it is unclear if this provides a true reflection of the uncertainty associated with these values.

Outcomes were specifically required to be reported as utility values, to allow incorporation into the cost-utility models. This may have excluded studies, which provided valuable information on the impact of treatment and disease on HRQOL. The inclusion criteria of this SLR, which restricted studies to patients with R/R disease, may have resulted in the exclusion of studies that could have acted as appropriate proxy data.

The paucity of data regarding disutility relating to active treatment and adverse events, in the population specified here, has resulted in the selection of proxy data for use in the cost-utility models. It is difficult to determine if these proxy data are an accurate reflection of the true disutility experienced by patients. Coverage with evidence development risk-sharing agreements implemented in some countries, as discussed in Chapter 11, may provide an opportunity to collect such data.

8.7 Conclusion

The evidence base of utility data for adult patients with R/R DLBCL is limited. This reflects the challenges that exist when collecting utility data in patients with advanced rare diseases. The limitations of the identified data may limit their generalisability and result in biased cost-effectiveness estimates. The uncertainty of the utility values identified in this SLR warrants extensive sensitivity analyses in the cost-utility analyses, presented in Chapter 9.

Chapter 9 Cost-Utility and Value of Information Analysis of Tisagenlecleucel and Axicabtagene Ciloleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma

9.1 Introduction

9.1.1 Chapter Aim

The aim of this chapter is to conduct a cost-utility analysis of (i) tisagenlecleucel, and (ii) axicabtagene ciloleucel for the treatment of adult patients with R/R DLBCL. This chapter presents bespoke cost-utility models of (i) tisagenlecleucel, and (ii) axicabtagene ciloleucel constructed in line with the Irish reference case (11). The estimates of relative efficacy, derived in Chapter 6, will be used to inform efficacy. The utility data identified through SLR, presented in Chapter 8, will be utilised. Sensitivity analyses will be conducted.

9.2 Model Development

9.2.1 Irish Reference Case

The Irish reference case for the HTA of technologies in Ireland, as described in National Economic Evaluation Guidelines (11) (presented in Table 15), informed the framework upon which the models were constructed.

9.2.2 Model Structure

9.2.2.1 Short-Term Decision Tree

Two distinct cost-utility analyses, using cost-utility models of identical structure, were undertaken. As described in Chapter 6, results generated from a comparison of tisagenlecleucel versus axicabtagene ciloleucel would not produce reliable results. As such, these therapies were compared separately to salvage chemotherapy (with or without HSCT). In this analysis, R-GDP (with or without HSCT) was assumed to represent salvage chemotherapy (with or without HSCT), as described in 9.2.5. The cost effectiveness of tisagenlecleucel versus R-GDP (with or without HSCT) was determined. Likewise, cost effectiveness of axicabtagene ciloleucel versus R-GDP (with or without HSCT) was determined. Henceforth, R-GDP (with or without HSCT) will be referenced as R-GDP.

The model structure is aligned with that presented in 5.2.2. All patients in the tisagenlecleucel and axicabtagene ciloleucel arms entered the decision tree, underwent leukapheresis, and subsequently progressed to one of three pathways (similar to Figure 6):

- Node 1 (N1): proceeded to infusion with (i) tisagenlecleucel (69% of patients (393)), or (ii) axicabtagene ciloleucel (91% of patients (144)), as appropriate.
- Node 2 (N2): did not proceed to infusion with (i) tisagenlecleucel (19% of patients (393)), or (ii) axicabtagene ciloleucel (6% of patients (144)) (as appropriate), due to manufacturing failure, adverse event, physician/patient decision, or protocol deviation. Instead, they were assumed to receive R-GDP.
- Node 3 (N3): did not proceed to infusion with (i) tisagenlecleucel (12% of patients (393)), or (ii) axicabtagene ciloleucel (3% of patients (144)) (as appropriate), due to death. These patients did not receive any further active treatment. All patients incurred a terminal care cost.

For patients who did not proceed to tisagenlecleucel infusion (31%), it was assumed that 50% received bridging chemotherapy and also that 50% received lymphodepleting chemotherapy (333). This assumption was also applied to the 9% of patients who did not proceed to axicabtagene ciloleucel.

9.2.2.2 Partitioned Survival Model

Patients treated with R-GDP entered the partitioned survival model directly. Patients in the tisagenlecleucel or axicabtagene ciloleucel arms entered through the decision tree. The partitioned survival model, depicted in Figure 7, simulated the progression of patients through three, mutually exclusive health states: progression-free survival, progressed disease, and death. The proportion of patients occupying each state was determined by the area under the curve of the extrapolated PFS and OS curves. Further detail is provided in 5.2.2.2.

In patients with DLBCL, relapse is expected to occur within 24 to 60 months post-treatment (113, 390, 391). To account for the potential emergence of long-term survivors, it was assumed that patients who were alive after 60 months were subject to age- and sex-matched general population mortality, in all treatment arms. An SMR was applied to reflect the impact of late effects and prolonged toxicities. More detail is provided in 9.3.1.3.2.

A cycle length of one month (30.4 days) (considered sufficient to capture relevant transitions) was applied. A half-cycle correction was applied to mitigate against over- or

under-predicting state occupancy. The time horizon was 44 years, representing a lifetime horizon. A discount rate of 4% was applied to both costs and outcomes after the first year (11). In the PSA, the proportion of patients proceeding to infusion, alternative therapy (R-GDP), and death were varied according to the Dirichlet distribution (described in 5.2.2.2).

9.2.3 Population

The population was adult patients with R/R DLBCL, after two or more lines of systemic therapy (3, 4). Baseline characteristics were sourced from JULIET and ZUMA-1. In JULIET, patients had a body weight of 78.7kg and a body surface area of 1.92m^2 (483, 503). Median age was 56 years and 61% were male (143, 392). In ZUMA-1, median age was 58 years and 67% were male (144, 395). In the absence of publicly available data on weight and body surface area distributions in ZUMA-1, data from JULIET were used.

9.2.4 Intervention

As mentioned, two separate analyses (interventions) were examined in models of identical structure.

- Tisagenlecleucel administered as a single IV infusion at a dose of 0.6 to 6x10^8 CAR-positive viable T-cells (non-weight based).
- Axicabtagene ciloleucel administered as a single IV infusion at a dose of 2x10^6 CAR-positive viable T-cells per kg (body weight).

Both interventions were modelled as single-dose interventions. This is in line with the JULIET protocol (143). Patients in ZUMA-1 could receive retreatment with axicabtagene ciloleucel under a number of pre-specified conditions (144). Clinical opinion indicated that patients in Irish clinical practice are unlikely to receive retreatment[12]. There is no guidance regarding retreatment in the SPC of axicabtagene ciloleucel. Neither the SPC of tisagenlecleucel nor axicabtagene ciloleucel explicitly states that patients cannot be retreated (3, 4).

[12] Oral correspondence with one consultant haematologist in Ireland.

9.2.5 Comparator

There is no universal routine care for patients with R/R DLBCL in Ireland. The most commonly used regimens comprise off-patent cytotoxics; differences in costs between regimens are expected to be negligible. Additionally, the clinical evidence used to inform the efficacy of the comparator arm comprised several salvage chemotherapy regimens (123). For this analysis, R-GDP was defined as the comparator of interest. R-GDP is one of the most widely used salvage chemotherapy regimens for R/R DLBCL, after two or more lines of systemic therapy (9, 10). In Irish clinical practice, 15% of patients are expected to proceed to HSCT (usually alloSCT) following treatment with R-GDP[13]. In this Chapter, 'R-GDP' pertains to R-GDP with (15%) or without (85%) HSCT.

The dosing regimen of R-GDP, presented in Table 47, was obtained from the NCCP Chemotherapy Regimen (504). Patients can receive treatment with R-GDP for up to six cycles, or until disease progression or unacceptable toxicity. However, clinical opinion indicated that clinical practice in Ireland is to administer three cycles[14].

Table 47 Dosing regimen of R-GDP, as per the National Cancer Control Programme Chemotherapy Regimen (504)

Drug	Dose	Frequency	Duration*
Dexamethasone	40mg once daily	Days 1-4	3 cycles
Rituximab	375mg/m^2 once daily	Day 1	
Gemcitabine	1000mg/m^2 once daily	Days 1 and 8	
Cisplatin	75mg/m^2 once daily	Day 1	

*Informed by one consultant haematologist in Ireland.

9.2.6 Perspective

The perspective was that of the healthcare payer in Ireland, the HSE (11). Direct medical costs, borne by the HSE, were included.

9.2.7 Outcomes

The primary outcomes were ICERs (cost per QALY). Deterministic and probabilistic outcomes were examined. Further detail is provided in 9.4.

[13] Oral correspondence with one consultant haematologist in Ireland.
[14] Oral correspondence with one consultant haematologist in Ireland.

9.3 Model Inputs

9.3.1 Efficacy Inputs

9.3.1.1 Reconstruction of Individual Patient-Level Data

As described in Chapter 6, IPD from published Kaplan-Meier curves of OS and PFS were reconstructed (211). This facilitated the generation of relative efficacy estimates and extrapolation of outcomes, to the time horizon of the model, as described below. The trials included were JULIET (tisagenlecleucel) (392, 393) and ZUMA-1 (axicabtagene ciloleucel) (144, 395). These trials were compared to data from CORAL Extension 1 (123), which provided proxy data for R-GDP (as described in Chapter 6).

9.3.1.2 Extrapolation of Survival

Treatment effectiveness was based on the effect on OS and PFS. Time horizons of the trials were shorter than those of the models, and so extrapolation of the data was required. Extrapolation was conducted in line with methods described in 5.3.1.2, and in line with the NICE DSU 14 (341). Parametric, spline, and mixture cure extrapolation models were examined.

9.3.1.2.1 Overall Survival Extrapolation

9.3.1.2.1.1 Parametric Survival Extrapolation: Overall Survival

A series of parametric models were fit to the individual treatment arms (341). These were Gompertz, exponential, Weibull, log-logistic, log-normal, and generalised gamma. The AIC and BIC statistics of each parametric model fitted to the OS data are presented in Table 48. Lower AIC and BIC statistics indicate better statistical fit relative to other models.

Tisagenlecleucel: Based on AIC and BIC statistics, the generalised gamma model had the best statistical fit to the JULIET data. The generalised gamma model also provided a good visual fit. Although, it may underestimate OS towards the end of the observed trial period. The generalised gamma model exhibited a long tail in the extrapolation output, with 27% of patients alive at 60 months and 20% alive at 120 months. The Gompertz model also had a good visual fit to the trial data; all other models had a poor visual fit (Appendix H, Figure A9). The Gompertz model exhibited a long tail; the other parametric

models converged to zero at much earlier time points. When compared to the data cut of JULIET presented at ASH 2020 (median follow up 40.3 months), the generalised gamma and Gompertz models were most closely aligned with the available data. The 36-month OS derived from this data cut was 36% (398); the generalised gamma and Gompertz models predicted 36-month OS of 33% and 38%, respectively. The generalised gamma and Gompertz models predicted 60-month OS of 27% and 36%, respectively.

Axicabtagene ciloleucel: For the ZUMA-1 data, the log-normal model had the lowest AIC and BIC statistics; however, this model provided a poor visual fit to the observed data. With the exception of the Weibull model, the difference in statistical fit between the parametric models was minimal. The generalised gamma and Gompertz models provided the best visual fit to the observed ZUMA-1 data (Appendix H, Figure A10). However, the generalised gamma model appeared to slightly underestimate OS towards the end of the follow-up period. Underestimation of OS towards the end of the observed follow-up period was noted for all parametric models. The generalised gamma and Gompertz models exhibited long tails; all other parametric models converged to zero at much earlier time points. The 60-month OS varied considerably between the parametric models, ranging from 17% (exponential) to 41% (Gompertz). When compared to the data cut of ZUMA-1 presented at ASH 2020 (median follow up 39.1 months), the generalised gamma and Gompertz models were most closely aligned with the available data. The 36-month OS derived from this data cut was 47% (402); the generalised gamma and Gompertz models predicted 36-month OS of 42% and 46%, respectively. The 60-month OS predicted by the generalised gamma and Gompertz models was 33% and 41%, respectively.

R-GDP: As described in Chapter 6, CORAL Extension 1 presented separate Kaplan-Meier OS curves for those patients who did and did not receive HSCT. Separate parametric models were fitted to these Kaplan-Meier curves. To model OS of the overall population (i.e. those with and without HSCT), a weighted OS curve combining the extrapolations from the separate Kaplan-Meier curves was generated. The weight applied corresponded to the expected rate of HSCT in Irish clinical practice (15%).

The Gompertz model provided the best statistical fit to Kaplan-Meier OS data of CORAL Extension 1 (both with and without HSCT). For the CORAL Extension 1 (with HSCT) data, the Gompertz model also provided the best visual fit and was the only model that appeared to accurately capture OS towards the end of the follow-up period. The Gompertz model exhibited a long tail. The generalised gamma model also exhibited a long tail in the CORAL Extension 1 (with HSCT) data. In terms of 36-month OS, the Gompertz model was most closely aligned with that of the available OS data. Both the observed OS data and the Gompertz model had a 36-month OS of 34% (123). The Gompertz model predicted a 60-month OS of 32%. The generalised gamma model predicted a 36-month OS of 33% and a 60-month OS of 26%.

For the CORAL Extension 1 (without HSCT) data, the Gompertz model provided a good visual fit. The log-logistic model also provided a good visual fit to the data. The Gompertz model was the only parametric model to exhibit a long tail in the extrapolation output. In terms of 36-month OS, the Gompertz and log-logistic models were most closely aligned with the available OS data. Both models predicted a 36-month OS of 7%, which is what was observed with the available data (123). The log-normal model was also closely aligned, generating a 36-month OS of 8%. The 60-month OS predicted by the Gompertz, log-logistic and log-normal models was 6%, 4%, and 4%, respectively.

Table 48 AIC and BIC statistics of parametric models used in the extrapolation of overall survival in cost-utility models for relapsed/refractory diffuse large B-cell lymphoma*

	Overall Survival							
	JULIET (tisagenlecleucel)		**ZUMA-1 (axicabtagene ciloleucel)**		**CORAL Extension 1 (with HSCT)**		**CORAL Extension 1 (without HSCT)**	
	AIC	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
Gompertz	455.48	460.97	450.58	455.81	**249.58**	**253.89**	**642.23**	**647.95**
Exponential	465.47	468.21	455.86	458.47	262.68	264.84	694.49	697.35
Weibull	466.04	471.53	457.33	462.56	258.42	262.74	660.90	666.62
Log-Logistic	457.61	463.10	452.09	457.32	253.80	258.12	649.13	654.85
Log-Normal	452.63	458.12	**449.61**	**454.84**	252.70	257.02	655.15	660.87
Generalised Gamma	**443.27**	**451.51**	449.68	457.53	253.25	259.73	654.99	663.57

AIC: Akaike information criteria; **BIC:** Bayesian information criteria; **HSCT:** Haematopoietic stem cell transplant.
*The lowest AIC and BIC statistics for each data set are highlighted in **Bold.**

9.3.1.2.1.2 Spline Model Survival Extrapolation: Overall Survival

One-, two-, and three-knot spline models were considered. Additional knots were not considered as visual inspection of the plotted models, AIC and BIC statistics, and survival output indicated that there was negligible difference between the three- and four-knot spline models. All three transformations of the survival function to a linear prediction scale (hazard, odds, and normal, described in 5.3.1.2.1.2) were considered. The AIC and BIC statistics of each spline model fitted to the OS data are presented in Table 49.

Tisagenlecleucel: The one-knot spline models (across all scales) had the lowest AIC and BIC statistics. All spline models provided a good visual fit to the observed JULIET data; no model provided a more favourable visual fit over others. The extrapolation output from the one-, two-, and three-knot spline models (across all scales) was closely aligned. The one-knot spline models (across all scales) provided the highest long-term OS predictions. When compared to the data cut of JULIET presented at ASH 2020, all spline models were aligned with the observed OS data. The 36-month OS derived from this data cut was 36% (398); the 36-month OS predicted by the spline models ranged from to 37% to 38%. The 60-month OS predicted by all models ranged from 30% to 32%.

Axicabtagene ciloleucel: The two-knot spline models (across all scales) provided the lowest AIC and BIC statistics, when fitted to the ZUMA-1 data. All spline models provided a good visual fit to the data. The extrapolation output from the two- and three-knot spline models (across all scales) was very closely aligned and the difference in long-term OS predictions was negligible. The extrapolation output from the one-knot spline models (across all scales) was closely aligned with that predicted by the two- and three-knot spline models (across all scales) until approximately month 30, after which the predicted OS from the one-knot spline models fell below that predicted by the two- and three-knot spline models. The difference in predicted OS between the one-knot, and two- and three-knot spline models (across all scales) increased as time progressed. The two- and three-knot spline models (across all scales) were aligned with the 36-month OS observed in ZUMA-1; 47% (402). The 36-month OS predicted by the one-knot spline models (across all scales) ranged from 43% (one-knot normal) to 45% (one-knot hazard). The predicted 60-month OS ranged from 32% (one-knot normal) to 42% (two- and three-knot spline models, across all scales).

R-GDP: The AIC statistics of the spline models, fitted to the observed CORAL Extension 1 (with HSCT) data, were similar across all scales and number of knots. Slightly more variation was observed in the BIC statistics; however, differences between the BIC statistics were still minimal. The one-knot spline models (across all scales) provided the lowest BIC statistics. Of note, the AIC and BIC statistics of the three-knot spline models (across all scales) are not presented due to the presence of overfitting upon visual inspection. None of the spline models provided a good visual fit to the data; no model adequately captured the distal portion of the Kaplan-Meier curve towards the end of the follow-up period. The two-knot spline models (across all scales) predicted higher OS when compared to the one-knot spline models (across all scales), with differences in predicted OS between these models increasing as time progressed. In terms of 36-month OS, the two-knot (hazard) spline model was most closely aligned with that of the available data. Both the observed data and the two-knot (hazard) spline model had a 36-month OS of 34% (123). The 60-month OS predicted by the two-knot (hazard) spline model was 27%. The 60-month OS predicted by the other one- and two-knot spline models ranged from 22% (one-knot odds and one-knot normal) to 25% (two-knot odds and two-knot normal).

Greater variation was observed in the AIC and BIC statistics of the spline models fitted to the observed CORAL Extension 1 (without HSCT) data. There was no spline model, which provided an optimal statistical fit. In terms of visual fit, the two- and three-knot spline models (across all scales) provided a marginally better fit than the one-knot spline models (across all scales). The one-knot spline models (across all scales) provided the lowest OS predictions, while the three-knot spline models (across all scales) provided the highest OS predictions. However, differences in predictions between the two- and three-knot spline models (across all scales) were minimal. All spline models were aligned with the 36-month OS observed in CORAL Extension 1 (without HSCT) (7%) (123), ranging from 5% to 8%. The predicted 60-month OS ranged from 2% (one-knot odds and one-knot normal) to 6% (three-knot spline models across all scales).

Table 49 AIC and BIC statistics of spline models used in extrapolation of overall survival in cost-utility models for relapsed/refractory diffuse large B-cell lymphoma*

	Overall Survival							
	JULIET (tisagenlecleucel)		ZUMA-1 (axicabtagene ciloleucel)		CORAL Extension 1 (with HSCT)		CORAL Extension 1 (without HSCT)	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1 Knot Spline (Hazard)	440.55	448.79	446.39	454.24	256.94	263.41	656.36	664.94
1 Knot Spline (Odds)	**440.50**	**448.73**	446.79	454.63	255.31	261.78	647.27	655.85
1 Knot Spline (Normal)	440.92	449.16	449.65	457.50	**254.58**	**261.06**	651.68	660.26
2 Knot Spline (Hazard)	442.49	453.47	**441.64**	**452.10**	256.56	265.20	642.25	**653.69**
2 Knot Spline (Odds)	442.53	453.51	441.75	452.21	256.58	265.22	644.88	656.32
2 Knot Spline (Normal)	442.54	453.53	441.73	452.19	255.93	264.57	645.44	656.88
3 Knot Spline (Hazard)	444.41	458.14	443.19	456.26	Values not presented due to overfitting		642.45	656.75
3 Knot Spline (Odds)	444.46	458.18	443.33	456.40			642.50	656.80
3 Knot Spline (Normal)	444.47	458.20	443.29	456.36			**640.92**	655.22

AIC: Akaike information criteria; **BIC:** Bayesian information criteria; **HSCT:** Haematopoietic stem cell transplant.
*The lowest AIC and BIC statistics for each data set are highlighted in **Bold**.

9.3.1.2.2 Progression-Free Survival Extrapolation

9.3.1.2.2.1 Parametric Survival Extrapolation: Progression-Free Survival

The AIC and BIC statistics of the parametric model fit to the PFS data of JULIET and ZUMA-1 are presented in Table 50. PFS data were not reported for CORAL Extension 1.

Tisagenlecleucel: Based on AIC and BIC statistics, the generalised gamma had the best statistical fit to the PFS data of JULIET. Based on visual inspection, the generalised gamma underestimated PFS towards the end of the observed follow-up period. The Gompertz model had a good visual fit to the PFS data; however, OS extrapolations of the parametric models, with the exception of the Gompertz OS extrapolation, fell below the Gompertz PFS extrapolation at various time points. This crossing of the curves is likely a result of the less 'stepped' nature of the PFS Kaplan-Meier curve of JULIET. The 60-month PFS predicted by the generalised gamma and Gompertz models was 33% and 16%, respectively.

Axicabtagene ciloleucel: The Gompertz model provided the best statistical and visual fit to the observed PFS data of ZUMA-1. However, the OS extrapolations of the parametric models, with the exception of the Gompertz OS extrapolation, fell below the Gompertz PFS extrapolation at various time points. The 60-month PFS predicted by the Gompertz model was 38%. The generalised gamma had the next best statistical fit to the data and provided a reasonable visual fit. Although, PFS appeared to be underestimated towards the end of the follow-up period. The 60-month PFS predicted by the generalised gamma model was 24%.

R-GDP: PFS data were not reported for CORAL Extension 1. As such, PFS was estimated from the OS curve by assuming that the cumulative hazard function for PFS was proportional to the cumulative hazard function for OS. The PFS predicted for CORAL Extension 1 was therefore, contingent upon the model applied to the OS data and no separate model fitting to PFS was required. The HR between PFS and OS (0.65) was based on the mean cumulative HR from the CORAL RCT (described in 6.3.2.3.2) (127, 505), and was identified through the literature (399, 483). This approach introduces additional uncertainty to the model due to differences between CORAL RCT and CORAL Extension 1, and due to the limited evidence to support PFS as a surrogate for OS in R/R DLBCL. Despite these uncertainties, in the absence of data, this HR has been accepted by national HTA agencies in HTA appraisals of tisagenlecleucel (399, 483).

Given the expectation that patients who are alive after 60 months are considered to be long-term survivors of R/R DLBCL, the proportional relationship between PFS and OS is not expected to continue indefinitely. Long-term survivors are expected to be free of progressed disease. As such, after month 60, the cumulative survival probabilities for PFS were assumed to flatten up to the point at which PFS met OS.

Table 50 AIC and BIC statistics of parametric models used in extrapolation of progression-free survival in cost-utility models for relapsed/refractory diffuse large B-cell lymphoma*†

	Progression-Free Survival			
	JULIET (tisagenlecleucel)		ZUMA-1 (axicabtagene ciloleucel)	
	AIC	BIC	AIC	BIC
Gompertz	372.54	377.96	**437.07**	**442.30**
Exponential	416.39	419.10	488.09	490.70
Weibull	406.29	411.71	468.79	474.02
Log-Logistic	387.73	393.15	456.91	462.14
Log-Normal	382.88	388.30	453.84	459.08
Generalised Gamma	**348.28**	**356.41**	449.41	457.25

AIC: Akaike information criteria; **BIC:** Bayesian information criteria.
*The lowest AIC and BIC statistics for each data set are highlighted in **Bold**.
†PFS of CORAL Extension 1 was derived by assuming that the cumulative hazard function for PFS was proportional to the cumulative hazard function for OS. The ratio between PFS and OS (0.65) was based on the mean cumulative hazard ratio from the CORAL RCT (127, 505).

9.3.1.2.2.2 Spline Model Extrapolation: Progression-Free Survival

The AIC and BIC statistics of the spline model fit to the PFS data of JULIET and ZUMA-1 are presented in Table 51.

Tisagenlecleucel: Statistical fit, based on AIC and BIC statistics, to the JULIET PFS data was closely aligned between all spline models. Of note, AIC and BIC statistics are not presented for the three-knot spline models (across all scales) due to overfitting observed upon visual inspection. All spline models fitted to the observed JULIET data provided very similar visual fit. This was reflected in the PFS predictions generated by the one- and two-knot spline models (across all scales). There was minimal difference in the long-term PFS predictions, with the predicted 60-month PFS ranging from 28% to 29%, across all scales and number of knots.

Axicabtagene ciloleucel: Statistical fit, based on AIC and BIC statistics, to the ZUMA-1 data was similar across all spline models. The two-knot spline models (across all scales) exhibited the lowest BIC statistics, while the BIC statistics of the one-knot (normal) spline model were notably higher. All spline models provided a good visual fit, which was closely aligned between all models. The two- and three-knot spline models (across all scales) provided a better fit to earlier portions of the observed Kaplan-Meier data, when compared to the one-knot spline models (across all scales). However, based on visual fit, any of the spline models could be considered a reasonable option. As time progressed, the PFS predictions diverged between the models. In particular, the one-knot (normal)

spline model predicted considerably lower PFS than the two- and three-knot spline models (across all scales). The predicted 60-month PFS ranged from 28% to 34%.

Table 51 AIC and BIC statistics of spline models used in extrapolation of progression-free survival in cost-utility models for relapsed/refractory diffuse large B-cell lymphoma*†

	Progression-Free Survival			
	JULIET (tisagenlecleucel)		ZUMA-1 (axicabtagene ciloleucel)	
	AIC	BIC	AIC	BIC
1 Knot Spline (Hazard)	**341.17**	**349.30**	434.67	442.52
1 Knot Spline (Odds)	343.68	351.81	437.21	445.05
1 Knot Spline (Normal)	344.75	352.88	445.82	453.66
2 Knot Spline (Hazard)	342.11	352.95	428.11	438.57
2 Knot Spline (Odds)	342.56	353.40	427.04	437.50
2 Knot Spline (Normal)	341.44	352.27	**427.03**	**437.49**
3 Knot Spline (Hazard)	Values not presented due to overfitting		427.16	440.23
3 Knot Spline (Odds)			427.99	441.07
3 Knot Spline (Normal)			429.22	442.30

AIC: Akaike information criteria; **BIC:** Bayesian information criteria.
*The lowest AIC and BIC statistics for each data set are highlighted in **Bold**.
†PFS of CORAL Extension 1 was derived by assuming that the cumulative hazard function for PFS was proportional to the cumulative hazard function for OS. The ratio between PFS and OS (0.65) was based on the mean cumulative hazard ratio from the CORAL RCT (127, 505).

9.3.1.2.3 Mixture Cure Model Extrapolation

The following mixture cure models were examined: Gompertz, exponential, Weibull, log-logistic, log-normal, and generalised gamma. Extrapolation was conducted in line with methods presented in 5.3.1.2.3 (178). The estimated cure fractions of JULIET ranged from 39% to 43%, while cure fractions of ZUMA-1 ranged from 44% to 50%. The estimated cure fractions of CORAL Extension 1 (with and without HSCT) ranged from 25% to 33% and 2% to 7%, respectively.

As highlighted in Chapter 5, for mixture cure models, data are required to be sufficiently mature and robust to reliably estimate a cure fraction. Sufficient numbers at risk are required in the tail of the distribution (178). This is a particular concern with JULIET, whereby 12 patients were at risk at 21 months and 6 were at risk at 24 months (393). These data are highly censored. The same concern arises with ZUMA-1, with 7 patients at risk at 30 months, reducing to 0 at 33 months (395). In CORAL Extension 1 (with and without HSCT), 9 patients were at risk at month 24, reducing to 4 by month 42 (123). The data of these trials are not sufficiently robust to reliably estimate a cure fraction. The use of mixture cure models was therefore, deemed inappropriate and no further consideration was given to this modelling method.

9.3.1.3 Summary

9.3.1.3.1 Overall Survival

The Gompertz and generalised gamma models were deemed the most appropriate parametric models to extrapolate the JULIET (tisagenlecleucel) data. The one- and two-knot spline models (across all scales) were considered the most appropriate of the spline models. The OS predicted by these models and the Kaplan-Meier curve of JULIET are presented in Figure 21.

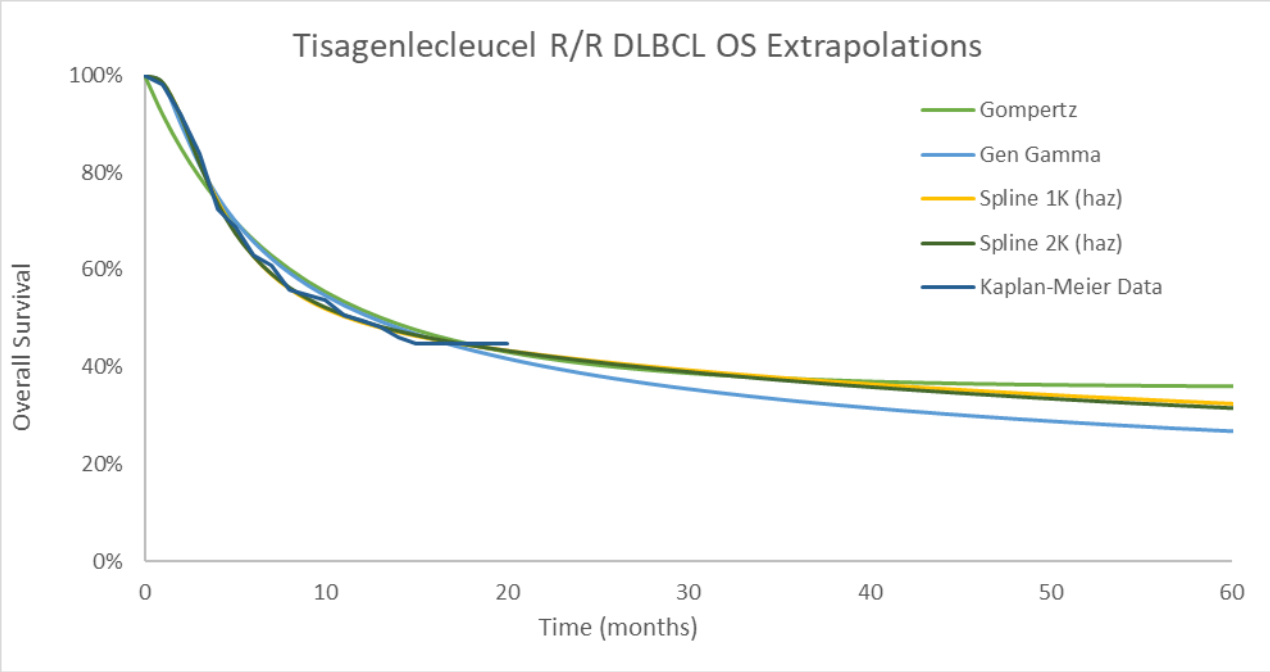

Figure 21 Tisagenlecleucel (R/R DLBCL) overall survival extrapolation predictions of 'best fitting' parametric and spline models
DLBCL: Diffuse large B-cell lymphoma; **Gen. gamma:** Generalised gamma; **OS:** Overall survival; **R/R:** Relapsed/refractory; **Spline 1K (haz):** Spline one-knot (hazard) model; **Spline 2K (haz):** Spline two-knot (hazard) model.

The Gompertz and generalised gamma models were considered the most appropriate parametric models to extrapolate the ZUMA-1 data. The two-knot spline models (across all scales) were considered the most appropriate of the spline models. The predicted OS outcomes and Kaplan-Meier curve of ZUMA-1 are presented in Figure 22.

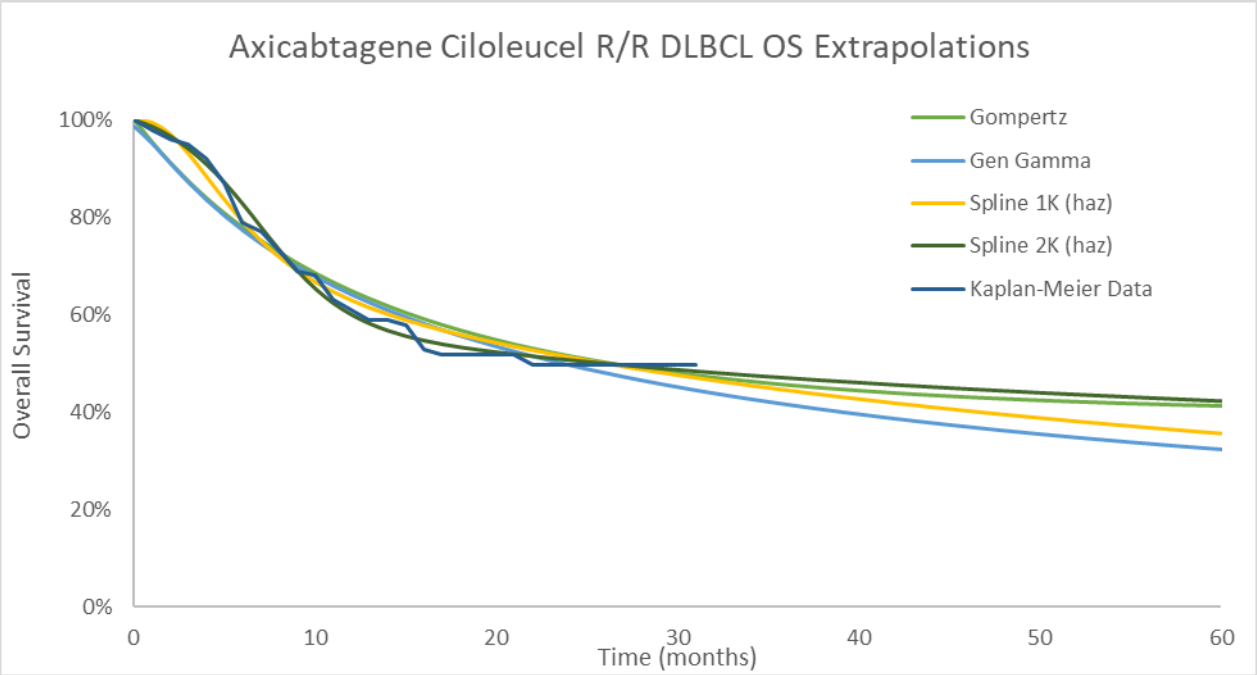

Figure 22 Axicabtagene ciloleucel (R/R DLBCL) overall survival extrapolation predictions of 'best fitting' parametric and spline models
DLBCL: Diffuse large B-cell lymphoma; **Gen. gamma:** Generalised gamma; **OS:** Overall survival; **R/R:** Relapsed/refractory; **Spline 1K (haz):** Spline one-knot (hazard) model; **Spline 2K (haz):** Spline two-knot (hazard) model.

For CORAL Extension 1 (with and without HSCT), the Gompertz model was considered the most appropriate parametric model. For CORAL Extension 1 (with HSCT), the one-knot (normal) spline and the two-knot (hazard) spline models provided the best fit of the spline models. The two-knot (hazard) spline and three-knot (normal) spline models provided the best fit, of the spline models, to the CORAL Extension 1 (without HSCT) data. The model predictions and observed Kaplan-Meier data are presented in Figure 23 (with HSCT) and Figure 24 (without HSCT).

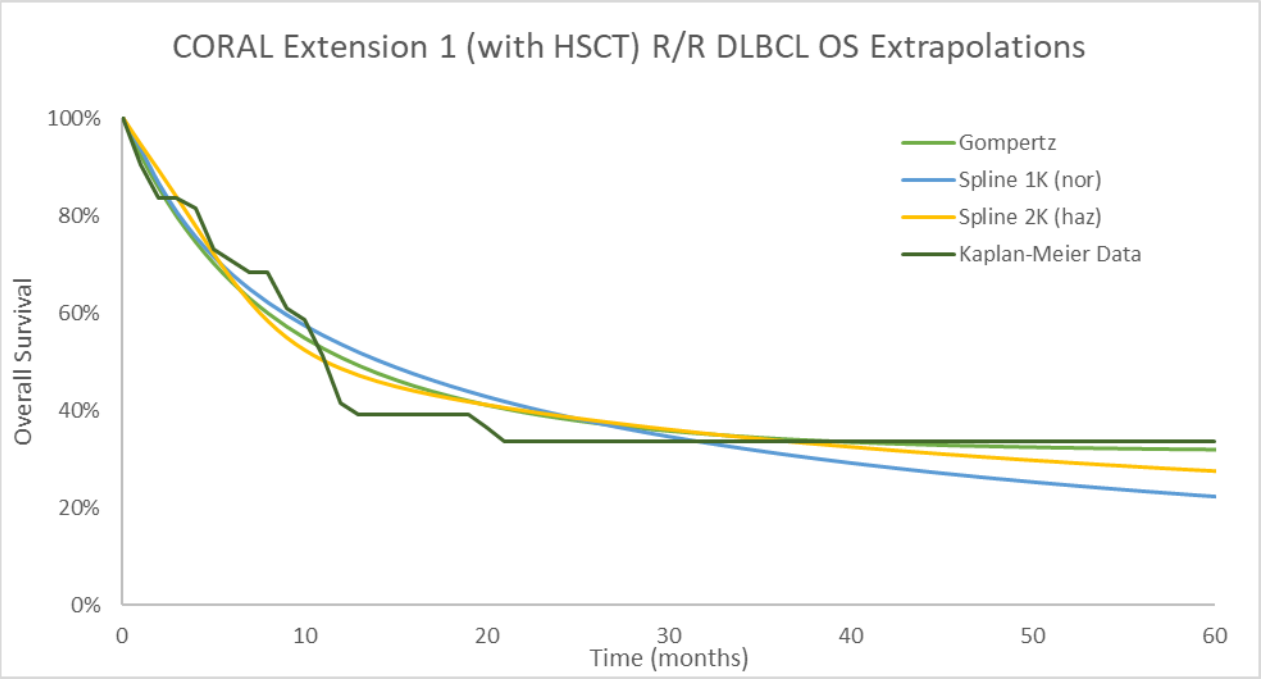

Figure 23 R-GDP with HSCT (R/R DLBCL) overall survival extrapolation predictions of 'best fitting' parametric and spline models
DLBCL: Diffuse large B-cell lymphoma; **HSCT:** Haematopoietic stem cell transplant; **OS:** Overall survival; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin; **R/R:** Relapsed/refractory; **Spline 1K (nor):** Spline one-knot (normal) model; **Spline 2K (haz):** Spline two-knot (hazard) model.

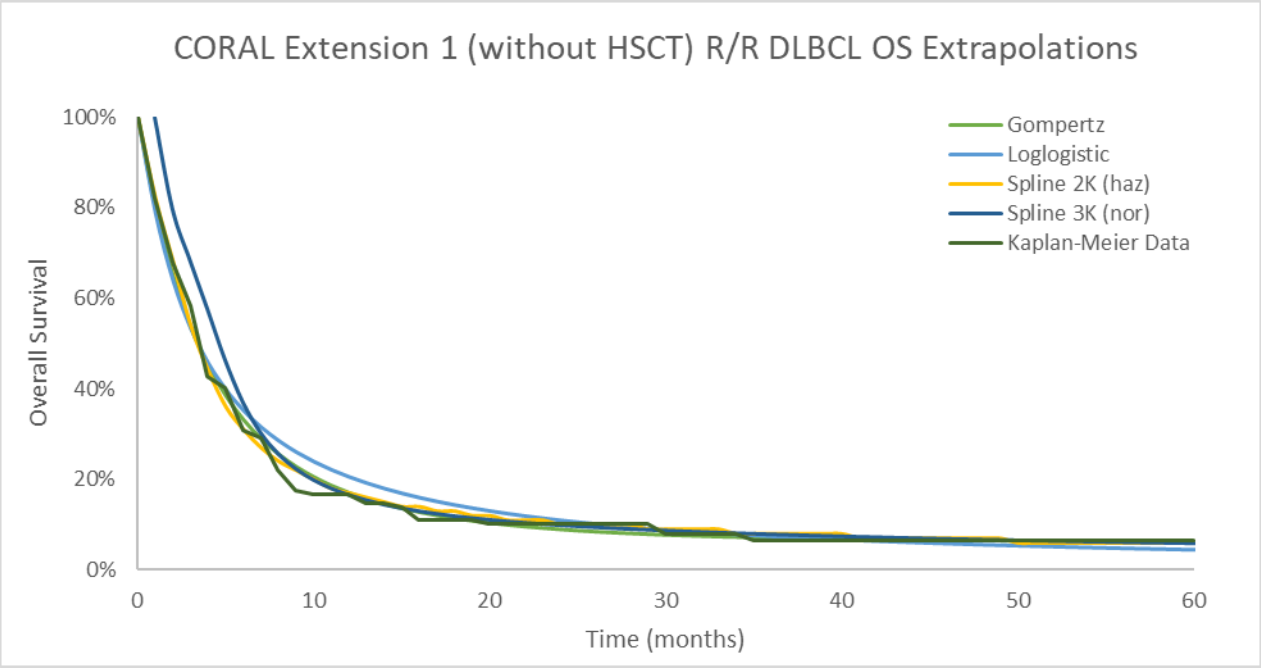

Figure 24 R-GDP without HSCT (R/R DLBCL) overall survival extrapolation predictions of 'best fitting' parametric and spline models
DLBCL: Diffuse large B-cell lymphoma; **HSCT:** Haematopoietic stem cell transplant; **OS:** Overall survival; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin; **R/R:** Relapsed/refractory; **Spline 2K (haz):** Spline two-knot (hazard) model; **Spline 3K (nor):** Spline three-knot (normal) model.

The log cumulative hazard plots of the relevant data were not approximately straight lines (Appendix H), indicating that spline model extrapolation may be more appropriate

than parametric extrapolation (341). The two-knot (hazard) spline model was chosen to extrapolate OS for both tisagenlecleucel and axicabtagene ciloleucel. The spline models had more favourable statistical fit when compared to the Gompertz and generalised gamma models. The Gompertz model was chosen to extrapolate the CORAL Extension 1 (with and without HSCT) data. The Gompertz model was chosen over the two-knot (hazard) spline model due to improved (with HSCT) or similar (without HSCT) visual fit, more favourable AIC and BIC statistics (both with and without HSCT), and clinical plausibility.

As highlighted previously, fitting different types of models to the treatment arms being compared implies that the underlying hazard differs between arms. Such an assumption requires substantial justification (341). Due to the innovative mechanism of action of tisagenlecleucel and axicabtagene ciloleucel when compared to R-GDP, it may be reasonable to assume that the underlying hazard differs between CD19 CAR T-cell therapy and R-GDP. The fitting of different types of models to each treatment arm has been accepted by some national HTA agencies in their appraisals of tisagenlecleucel (333, 399).

It is acknowledged that a number of plausible options exist to extrapolate the OS data of these studies. As such, extensive sensitivity analyses were conducted to explore the impact of alternative model extrapolations and assumptions.

9.3.1.3.2 Adjustment for Excess Risk of Mortality

The assumption, that most patients with DLBCL are expected to relapse within 24 to 60 months post-treatment, is supported by studies based on patients with newly diagnosed disease (113, 390, 391). However, clinical opinion indicated that these data may also be applicable to patients with R/R disease[15]. This is supported by one identified study, which indicated that by at least four years of PFS (post-HSCT), the mortality rate of patients who received HSCT following relapse had stabilised. Additionally, compared with the age- and sex-matched general population, the SMR was appreciably higher until 60 months after HSCT (506). It should be noted; however, that data towards month 60 were limited

[15] Oral correspondence with one consultant haematologist in Ireland.

by the small number of deaths. Other studies indicate that patients may be at risk of relapse for up to 8 (507) or 10 (508) years post-HSCT. To account for the potential emergence of long-term survivors, an assumption was employed, whereby patients who were alive after 60 months were considered to be long-term survivors and subject to age- and sex-matched general population mortality. It has been reported that the relative risk of mortality of these patients is higher than that in the general population (507, 509). Thus, an adjustment factor (SMR) was incorporated into the model. SMRs, identified in the literature, varied from 1.09 (113) to 1.56 (390) for non-cancer related mortality. A higher SMR of 3.4 (95% CI 2.9 to 4.1) was also identified. This encompassed death due to relapsed disease as well as non-cancer related mortality (509); interpretation is limited by the small sample size (n=23).

For this analysis, an SMR of 1.36 (95% CI 1.12 to 1.63), identified in Jakobsen et al. (391), was employed. This was derived from a population-based study of patients with newly diagnosed DLBCL in first remission, whose outcomes were recorded in the Danish National Lymphoma Registry (n=1,621). Patients were diagnosed between 2003 and 2011, aged greater than 16 years at diagnosis, and achieved CR after first-line treatment with R-CHOP. Excess mortality was calculated as the ratio of observed to expected mortality (derived from Danish life tables). Median follow up was 85 months. SMRs, at month 60, were presented for patients aged less than 50 years (1.11; 95% CI 0.22 to 3.25) and those aged 50 years and older (1.36; 95% CI 1.12 to 1.63). The population aged 50 years and older (SMR 1.36) was considered most relevant to this assessment.

9.3.1.3.3 Progression-Free Survival

The fitting of extrapolation models to the PFS data was only required for the tisagenlecleucel and axicabtagene ciloleucel arms of the respective models.

In summary, the generalised gamma model was considered the most appropriate parametric model to extrapolate the JULIET data. Of the spline models, the one-knot (hazard) spline model was chosen based on marginally more favourable AIC and BIC statistics. The one-knot (hazard) spline model had more favourable AIC and BIC statistics when compared to the generalised gamma model. Thus, the one-knot (hazard) spline

model was chosen to extrapolate the JULIET data in the base case. The predicted 60-month PFS of this model was 29%.

The Gompertz model was considered the most appropriate parametric model to extrapolate the ZUMA-1 PFS data. Of the spline models, the two-knot spline models (across all scales) were chosen based on marginally more favourable AIC and BIC statistics. The two-knot (hazard) spline model had more favourable AIC and BIC statistics when compared to the Gompertz model. The two-knot (hazard) spline model was therefore, chosen to extrapolate the PFS data of ZUMA-1. The predicted 60-month PFS of this model was 34%.

After 60 months, the cumulative survival probabilities for PFS were assumed to flatten up to the point at which PFS met OS for all treatment arms.

9.3.2 Health-Related Quality of Life Inputs

Utility estimates were derived through SLR, presented in Chapter 8 (Table 46). Assumptions are generally aligned with those employed in 5.3.2. The assumptions employed in the base case are presented in Table 52.

All utility values were varied in the PSA. Utility values were varied according to the beta distribution. Parameters of the beta distribution were calculated as described in 5.3.2. Disutility values were varied according to the normal distribution. All values were varied ±25% in the OWSA. To investigate uncertainty associated with the HRQOL of patients considered to be long-term survivors, a scenario was explored whereby patients who were alive after month 60 were assigned age- and sex-matched general population utility (328). An additional scenario, whereby disutility associated with adverse events was removed was also conducted. This was to account for the potential for health-state utility values from JULIET to capture some degree of disutility due to adverse events.

Table 52 Utility values used in cost-utility models for relapsed/refractory diffuse large B-cell lymphoma

Parameter	Value (SE)	Source	Duration (days)	Proportion
Progression-Free Survival	0.83 (0.21‡)	EQ-5D-3L (mapped from SF-36) collected in JULIET with UK valuation set applied, identified through HTA appraisal (399)	Duration in health state	100% of patients in state
Progressed Disease	0.71 (0.18‡)			
All Patients Alive After 60 Months	0.83 (0.21‡)	Assumption (equal to progression-free survival) based on HTA appraisal (333)	Duration of survival	100% of patients
Apheresis			0.5	100% of patients in tisagenlecleucel and axicabtagene ciloleucel arms
Bridging Chemotherapy	-0.15 (0.04‡)	Guadagnolo et al. (337), identified through HTA appraisals (333, 399, 483)	5 (504)	90% of patients who received tisagenlecleucel* and axicabtagene ciloleucel† (393) 50% of those who did not proceed to infusion in these arms (333)
Lymphodepleting Chemotherapy			3 (3, 4)	93% of patients who received tisagenlecleucel* (392, 393) 93% of patients who received axicabtagene ciloleucel† (144) 50% who did not proceed to infusion in these arms (333)
CRS ICU Admission	-0.83 (0.21‡)	Assumption (utility of 0) based on HTA appraisals (331-333, 399, 482, 483), ICER HTA (296)	7 (143)	24% of patients who received tisagenlecleucel* (392) 13% of patients who received axicabtagene ciloleucel† (332)
Non-CRS ICU Admission		Assumption (utility of 0) based on HTA appraisals (333, 399, 481, 483)	0.9 (399)	30% of patients who received tisagenlecleucel* (143)

				50% of patients who received axicabtagene ciloleucel† (144)
Febrile Neutropenia	-0.15 (0.04‡)	Lloyd et al. (329), identified through HTA appraisal (331)	6 (332)	15% of patients who received tisagenlecleucel* (393) 33% of patients who received axicabtagene ciloleucel† (395) 9% of patients who received R-GDP (128)
Pancytopenia		Assumption. Disutility from Lloyd et al. (329)	182.4§	14% of patients who received tisagenlecleucel*¶ (143) 17% of patients who received axicabtagene ciloleucel†| (144)
HSCT	-0.30 (0.08‡)	Guadagnolo et al. (337), identified through HTA appraisals (333, 483)	365.25 (333, 483)	15% of patients who received R-GDP**
Age-Related Disutility	Ara and Brazier (328)	Ara and Brazier (328), identified through HTA appraisal (333)	Time horizon of model	100% of patients

CRS: Cytokine release syndrome; **HSCT:** Haematopoietic stem cell transplant; **HTA:** Health technology assessment; **ICU:** Intensive care unit; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin; **UK:** United Kingdom.

*69% proceeded to infusion with tisagenlecleucel.

†91% proceeded to infusion with axicabtagene ciloleucel.

‡Assumed 25% of mean point estimate.

§Assumption.

¶Assumed that the proportion of patients experiencing grade ≥3 unresolved, prolonged cytopenias (greater than 28 days) was representative of the proportion of patients experiencing pancytopenia. The cytopenia with the lowest incidence was assumed to reflect the proportion of patients experiencing pancytopenia (14% for anaemia).

|Assumed that the proportion of patients experiencing grade ≥3 unresolved, prolonged cytopenias (greater than 30 days) was representative of the proportion of patients experiencing pancytopenia. The cytopenia with the lowest incidence was assumed to reflect the proportion of patients experiencing pancytopenia (17% for anaemia).

**Based on clinical opinion.

9.3.3 Cost Inputs

Costs were calculated as described in 5.3.3. Resource use estimates were sourced from clinical trial data (392, 395), the NCCP Chemotherapy Regimen (R-GDP) (504), and the literature. In the PSA, costs were varied according to the gamma distribution. In the OWSA, costs were varied ±25%.

9.3.3.1 Implementation Costs

As per 5.3.3.1, no implementation costs were considered.

9.3.3.2 Training Costs

Clinical opinion[16] indicated that approximately 40 staff members in St James's Tertiary Hospital are expected to receive formal training in relation to CD19 CAR T-cell therapies. The per patient cost of training was estimated to be €568.15 (Appendix H (Table A33)). The formula used to calculate the cost is presented in 5.3.3.2 (331).

9.3.3.3 Pre-Treatment Costs

9.3.3.3.1 Leukapheresis

The once-off cost of leukapheresis, €1,249 per patient (described in 5.3.3.3.1) (350), was applied to all patients in the tisagenlecleucel and axicabtagene ciloleucel arms.

9.3.3.3.2 Cryopreservation

For tisagenlecleucel, the leukapheresis product requires cryopreservation during shipping to the manufacturing facility (3). A cost of €5,544.68 (described in 5.3.3.3.2) per patient was employed for all patients in the tisagenlecleucel arm. Axicabtagene ciloleucel is generated from fresh leukapheresis product, which does not require cryopreservation (503). Both tisagenlecleucel and axicabtagene ciloleucel require cryopreservation following manufacture (3, 4). Thus, a cost of €5,544.68 per patient was applied to the proportion of patients receiving infusion with tisagenlecleucel or axicabtagene ciloleucel.

[16] Oral correspondence with one consultant haematologist in Ireland.

9.3.3.3.3 Bridging Chemotherapy

Bridging chemotherapy was assumed to consist of R-GDP, administered for one treatment cycle (21 days) (333). The total per patient cost was €1,540.96. Based on JULIET, 90% of patients who received tisagenlecleucel infusion received bridging chemotherapy (393).

Bridging chemotherapy was not permitted in ZUMA-1. However, clinical opinion indicated that, in Irish clinical practice, most patients are likely to receive at least one cycle of bridging chemotherapy regardless of CD19 CAR T-cell regimen received[17]. Additionally, real-world evidence suggests that the use of bridging chemotherapy in patients receiving axicabtagene ciloleucel is high, with one centre in Europe reporting that 97% of patients (n=61) received bridging chemotherapy (425). As such, it was assumed that 90% of patients who received infusion in the axicabtagene ciloleucel arm received bridging chemotherapy (in line with the proportion in the tisagenlecleucel arm).

It was also assumed that 50% of patients who did not proceed to infusion (with tisagenlecleucel or axicabtagene ciloleucel, as appropriate) received bridging chemotherapy (333).

9.3.3.3.4 Lymphodepleting Chemotherapy

In JULIET, fludarabine in combination with cyclophosphamide was received by 73% of patients who received infusion; bendamustine was received by 20% (392). Both regimens were included in the model. The cost of fludarabine in combination with cyclophosphamide was €298.40 per patient, per treatment course, while that of bendamustine was €700.50 per patient, per treatment course. As per the SPC of axicabtagene ciloleucel, patients should receive lymphodepleting chemotherapy consisting of fludarabine in combination with cyclophosphamide. Doses differ to those administered to patients receiving tisagenlecleucel (4). The cost of fludarabine in combination with cyclophosphamide, for patients receiving axicabtagene ciloleucel, was €419.04 per patient, per treatment course. Lymphodepleting chemotherapy was received by 93% of patients in ZUMA-1 (144).

[17] Oral correspondence with one consultant haematologist in Ireland.

It was also assumed that 50% of patients who did not proceed to infusion (with tisagenlecleucel or axicabtagene ciloleucel, as appropriate) received lymphodepleting chemotherapy (333).

9.3.3.4 Drug Acquisition Costs

Drug acquisition costs for tisagenlecleucel, axicabtagene ciloleucel and R-GDP are presented in Table 53. These were sourced from the NCPE Technical Summaries of tisagenlecleucel (10) and axicabtagene ciloleucel (9), and the NCPE Internal Cost Database, respectively. Costs presented are exclusive of VAT, as VAT is not applicable in the cost-utility analysis.

Table 53 Total drug acquisition costs per patient per treatment course employed in cost-utility models for relapsed/refractory diffuse large B-cell lymphoma

Drug	Total Drug Acquisition Cost Per Patient Per Treatment Course* (€)
Tisagenlecleucel†	301,762.13
Axicabtagene Ciloleucel‡	309,015.00
R-GDP§	4,622.89

R-GDP: Rituximab, gemcitabine, dexamethasone, cisplatin.
*Including relevant fees and rebates, excluding VAT.
†Applied only to those who received infusion (69%).
‡Applied only to those who received infusion (91%).
§Does not include the cost of haematopoietic stem cell transplant. In line with clinical opinion in Ireland, it was assumed that patients received three cycles of R-GDP[18].

9.3.3.5 Outpatient Administration Costs

The total outpatient administration cost associated with bridging chemotherapy was €692 per patient (€346 per patient per day, as per 5.3.3.5). An administration cost of €692 was applied per patient, per cycle of R-GDP (salvage chemotherapy) received. Administration costs of lymphodepleting chemotherapy, tisagenlecleucel, and axicabtagene ciloleucel were assumed to be captured in the cost of hospitalisation, described in 9.3.3.6.

9.3.3.6 Hospitalisation Costs

In the absence of severe adverse events, the duration of hospitalisation (including lymphodepleting chemotherapy) for patients receiving tisagenlecleucel or axicabtagene

[18]Oral correspondence with one consultant haematologist in Ireland.

ciloleucel is approximately 21 days[19]. For the period of lymphodepleting chemotherapy administration, a cost of €5,100 per patient was applied (described in 5.3.3.6) (350). To account for increased monitoring and resource requirements, a cost of €14,033 per patient was applied to cover the first 13.5 days from the time of CD19 CAR T-cell infusion. This was sourced from the HPO DRG List (DRG R61A) (350). A cost of €5,100 (DRG R61B) per patient was also applied to cover the final stage of hospitalisation, whereby the patient is subject to less monitoring requirements. The total length of hospitalisation covered by these costs equates to 22.3 days (350). In line with 5.3.3.6, it was assumed that 50% of patients who received infusion in the tisagenlecleucel and axicabtagene ciloleucel arms were discharged to hospital-associated patient apartments for a duration of six days. The cost per patient per night was €63.90 (295).

9.3.3.7 Monitoring Costs

Initiation and monitoring costs were included in the R-GDP arm to account for outpatient administration (504). Total initiation costs associated with R-GDP were €186.04 per patient; per cycle monitoring costs were €272.81 per patient (Appendix H (Table A42 and Table A43)). Health-state specific follow-up costs were applied to patients in the progression-free survival and progressed disease states (98, 353). Details are presented in Appendix H (Table A44). These costs are presented in Table 54.

Table 54 Per cycle health-state monitoring costs used in cost-utility models for relapsed/refractory diffuse large B-cell lymphoma

	Progression-Free Survival Months 1-12 (inclusive)	Progression-Free Survival Months 13-36 (inclusive)	Progression-Free Survival Months 37-60 (inclusive)	Progression-Free Survival Month 61 onwards*	Progressed Disease†
Tisagenlecleucel/ Axicabtagene Ciloleucel Per Cycle Cost (€)	96.31	37.68	20.10	18.28	71.75
R-GDP Per Cycle Cost (€)	71.75	25.28	13.90	12.08	71.75

R-GDP: Rituximab, gemcitabine, dexamethasone, cisplatin.
*All patients alive after 60 months were assumed to be long-term survivors and incurred this cost, regardless of health state.
†Assumed that costs incurred were the same as those incurred by patients undergoing R-GDP treatment in the progression-free survival state in Months 1 to 12 (223).

[19]Oral correspondence with one consultant haematologist in Ireland.

9.3.3.8 Adverse Events

CD19 CAR T-cell therapy adverse event costs were applied, in line with 5.3.3.8. R-GDP is administered in the outpatient setting and so, the cost of treating associated adverse events was accounted for separately. Data on grade ≥3 adverse events occurring in 5% or greater of the population were sourced from Crump et al. (128). A limitation of this data source is that it does not account for adverse events associated with rituximab. Adverse event costs were applied as a once-off cost.

9.3.3.8.1 Cytokine Release Syndrome

Based on JULIET, it was assumed that 24% of patients treated with tisagenlecleucel were admitted to ICU, for 7 days (143, 392). It was further assumed that 16% of patients who received tisagenlecleucel received treatment with tocilizumab (393); 36% of these received one dose, while 64% received two doses (392).

Based on ZUMA-1, it was assumed that 13% of patients treated with axicabtagene ciloleucel were admitted to ICU (332). In the absence of data, a 7-day duration was assumed (in line with tisagenlecleucel). As reported in the EPAR, 43% of patients who received axicabtagene ciloleucel were treated with tocilizumab (144). In line with tisagenlecleucel, it was assumed that 36% of patients received one dose and 64% received two doses of tocilizumab (392).

A total ICU cost of €19,584 per patient, per stay was estimated for both tisagenlecleucel and axicabtagene ciloleucel (362). This was applied to the proportion of patients requiring ICU admission for CRS in each arm. The cost per dose of tocilizumab was estimated to be €1,345.68 per patient (363). Vial sharing was not assumed.

9.3.3.8.2 Non-Cytokine Release Syndrome ICU Admission

In JULIET, patients who were hospitalised for non-CRS-related adverse events were subject to a mean ICU stay of 0.9 days (399). Thus, it was assumed that patients who experienced a serious adverse event (with the exception of CRS and febrile neutropenia) in JULIET and ZUMA-1 incurred an ICU stay for a duration of 0.9 days. This was incurred by 30% and 50% of patients who received tisagenlecleucel and axicabtagene ciloleucel, respectively (143, 144).

9.3.3.8.3 B-Cell Aplasia

In JULIET, 33% of patients were treated with IV immunoglobulin (399), while 31% of patients were treated in ZUMA-1 (395). The duration of IV immunoglobulin treatment is not reported in the publications of JULIET or ZUMA-1. As discussed in 5.3.3.8.3, the median time to B-cell recovery in ELIANA (tisagenlecleucel in paediatric and young adults with R/R ALL) was 11.4 months (250). However, as discussed, this duration was considered an underestimate. Of relevance here, 75% of patients in ZUMA-1, who responded to treatment with axicabtagene ciloleucel, had restored B-cell count at 24 months post-infusion; B-cell recovery began at 9 months in some cases. The authors of the ZUMA-1 publication suggest that durable responses in adult patients with R/R DLBCL may not require long-term persistence of functional CAR T-cells (395). Additionally, the approach to treatment of B-cell aplasia differs between paediatric and adult patients. While it has become practice in some centres to administer empiric IV immunoglobulin to paediatric and young adult patients following administration of CAR T-cell therapy, adult patients tend to receive treatment only in the presence of serious or recurrent infection (353, 510).

Assumptions regarding IV immunoglobulin treatment duration in patients with R/R DLBCL treated with tisagenlecleucel or axicabtagene ciloleucel have ranged from 1.43 doses (503) to 3 years (333). Here, it was assumed that patients received treatment with IV immunoglobulin for three years (333). This was applied to patients in the progression-free survival state only. Alternative assumptions were explored in scenario analyses. The total cost per dose, administered once per cycle, was €2,535 per patient. An administration cost was included; this was assumed to be the same as an outpatient administration of chemotherapy, €346 per patient.

9.3.3.8.4 Febrile Neutropenia and Pancytopenia

Grade ≥3 febrile neutropenia occurred in 15%, 33%, and 9% of patients in JULIET (tisagenlecleucel) (393), ZUMA-1 (axicabtagene ciloleucel) (395), and Crump et al. (R-GDP) (128), respectively. The per patient cost, €9,451.31 (inflated to 2020) (362), was applied to the relevant proportion in each arm.

It was assumed that patients experiencing pancytopenia would be treated as a daycase, once per month, for a duration of six months (330). A cost of €387 per daycase (350) was applied to the proportion of patients experiencing pancytopenia (described in 9.3.2).

9.3.3.8.5 Other Adverse Events Associated with R-GDP

Grade ≥3 adverse events occurring in 5% or greater of the population, associated with R-GDP, were included (128). Details are presented in Appendix H (Table A50) (350, 361). Incorporating these adverse events (excluding febrile neutropenia) resulted in a cost of €804.19 per patient.

9.3.3.9 Haematopoietic Stem Cell Transplant

Based on clinical opinion, 15% of patients who receive treatment with R-GDP are expected to receive HSCT[20] (usually alloSCT). The cost of the alloSCT procedure was €116,323 per patient (HPO DRG List (DRG A07B)) (350). Follow-up costs, accounting for 365-days post-discharge, were sourced from a report by Ernst & Young (364). Details are provided in 5.3.3.9.

9.3.3.10 Terminal Care

A once-off terminal care cost was applied to patients upon death (€7,732.48 per patient). This was sourced from Bourke et al. (365), described in 5.3.3.10.

9.4 Model Outputs

The model outputs are aligned with those described in 5.4.

9.4.1 Deterministic ICER

The base case analyses considered the cost effectiveness of (i) tisagenlecleucel versus R-GDP, and (ii) axicabtagene ciloleucel versus R-GDP, using ICERs from deterministic costs and QALYs.

[20] Oral correspondence with one consultant haematologist in Ireland.

9.4.2 Probabilistic ICER and Scatterplot

PSA was conducted, whereby survival, utility, and cost parameters were varied according to their appropriate distributions (as described in the relevant sections). Further description is provided in 5.4.2. A scatterplot of incremental costs and outcomes, generated from each iteration of the PSA, was constructed to illustrate the degree of uncertainty surrounding the estimates.

9.4.3 Cost-Effectiveness Acceptability Curve

A cost-effectiveness acceptability curve was constructed, as per methods described in 5.4.3 (368). The cost-effectiveness acceptability curve illustrates the measure of uncertainty in the decision.

9.4.4 One-Way Sensitivity Analysis

OWSA of all parameters in the model was performed to determine the sensitivity of the model to changes in individual parameters and assumptions, as described in 5.4.4. Tornado plots were constructed to illustrate the impact of the 10 most influential parameters on the deterministic ICER.

9.4.5 Scenario Analysis

A number of scenario analyses were conducted to assess the impact on the deterministic ICER of employing alternative, plausible assumptions. The scenarios explored were aligned with those presented in 5.4.5. These are presented in Table 57 and Table 60.

9.4.6 Price Analysis

Both tisagenlecleucel and axicabtagene ciloleucel were approved for reimbursement following confidential price negotiations (10). An analysis was conducted to determine the price reduction in the price-to-wholesaler of tisagenlecleucel and axicabtagene ciloleucel that would be required to meet the €45,000 per QALY willingness-to-pay threshold (versus R-GDP).

9.4.7 Expected Value of Perfect Information

EVPI analysis was conducted to estimate the value of collecting additional information to eliminate or reduce uncertainty (34), following methods described in 5.4.7.

9.4.7.1 Partial Expected Value of Perfect Information

To identify the parameters, which contributed most to the overall decision uncertainty, EVPPI was also estimated. This was estimated using methods described in 5.4.7.1.

9.5 Results

9.5.1 Results: Tisagenlecleucel versus R-GDP

9.5.1.1 Deterministic Results: Tisagenlecleucel versus R-GDP

The deterministic model outcomes are presented in Table 55. Tisagenlecleucel was not cost effective, versus R-GDP, at a willingness-to-pay threshold of €45,000 per QALY (31).

Table 55 Deterministic results of the incremental analysis of cost effectiveness of tisagenlecleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma

Drug	Total Costs (€)	Total QALYs	Incremental Costs (€)	Incremental QALYs	ICER (€/QALY)
R-GDP*	55,900	1.50			
Tisagenlecleucel	273,992	3.33	218,092	1.82	119,509

ICER: Incremental cost-effectiveness ratio; **QALY**: Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.
*With (15%) or without (85%) haematopoietic stem cell transplant.

9.5.1.2 Probabilistic Results: Tisagenlecleucel versus R-GDP

The expected incremental costs and incremental QALYs are presented in a scatterplot in Figure 25. Most iterations lie in the NE quadrant. Mean expected costs and QALYs are presented in Table 56. Mean probabilistic outputs were similar to those of the deterministic analysis, with a slightly lower ICER generated in the PSA.

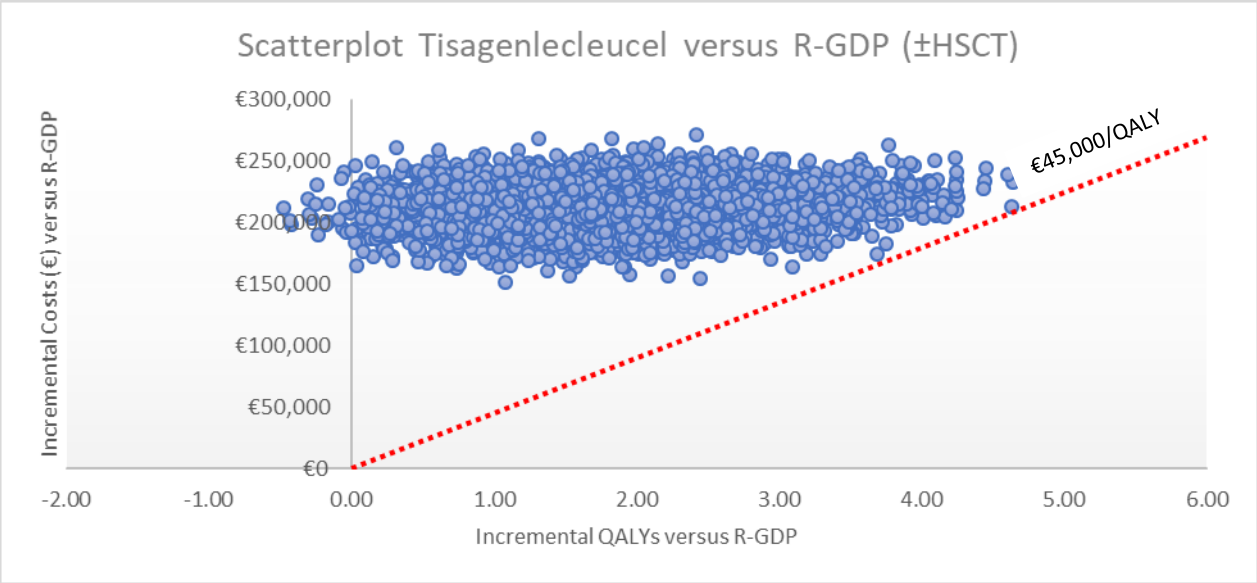

Figure 25 Scatterplot of incremental costs and incremental QALYs from probabilistic sensitivity analysis of tisagenlecleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
HSCT: Haematopoietic stem cell transplant; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

Table 56 Mean probabilistic outputs of the incremental analysis of cost effectiveness of tisagenlecleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma

Drug	Total Costs (€)	Total QALYs	Incremental Costs (€)	Incremental QALYs	ICER (€/QALY)
R-GDP*	55,716	1.65			
Tisagenlecleucel	268,216	3.48	212,499	1.83	116,005

ICER: Incremental cost-effectiveness ratio; **QALY**: Quality-adjusted life year**; R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.
*With (15%) or without (85%) haematopoietic stem cell transplant.

The cost-effectiveness acceptability curve is presented in Figure 26. At a willingness-to-pay threshold of €45,000 per QALY, there was a 0% probability that tisagenlecleucel was cost effective versus R-GDP.

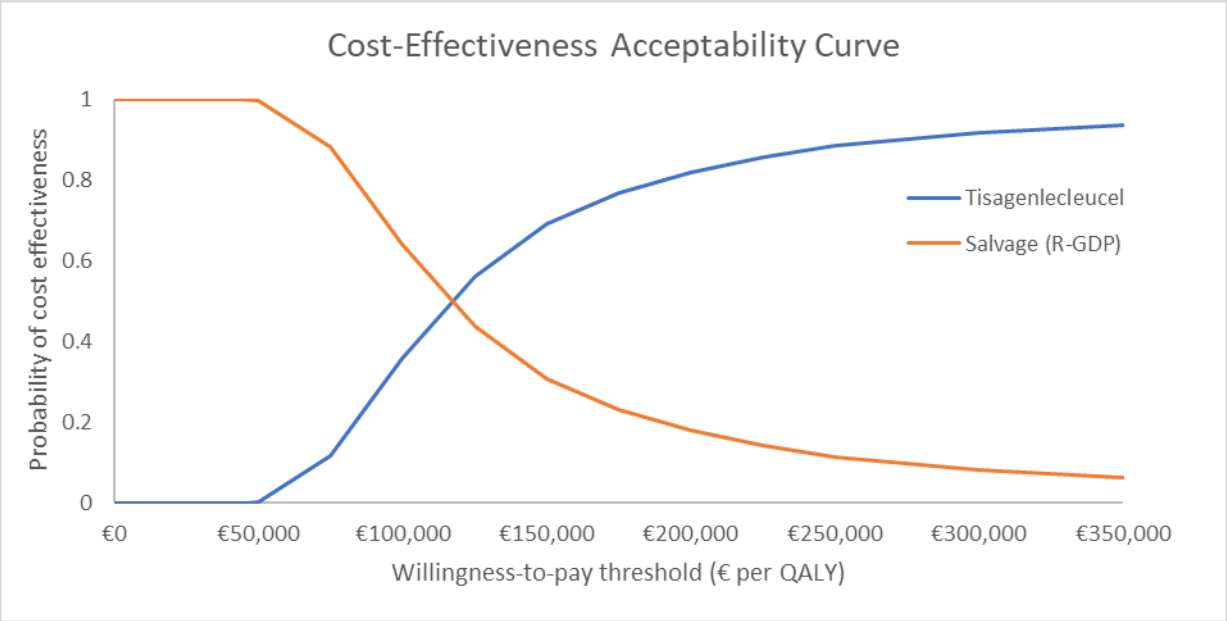

Figure 26 Cost-effectiveness acceptability curve of tisagenlecleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
QALY: Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

9.5.1.3 One-Way Sensitivity Analysis

Outcomes of OWSA, versus R-GDP, are presented in Figure 27. For pragmatic reasons, OWSA was conducted on deterministic outcomes. Thus, results should be considered indicative only. The main drivers in the model were the discount rate on outcomes, tisagenlecleucel infusion cost, and the progression-free survival state utility value. The threshold of €45,000 per QALY was not met in any analyses. The lowest ICER (€72,379 per QALY) occurred when the discount rate on outcomes was reduced to 0%.

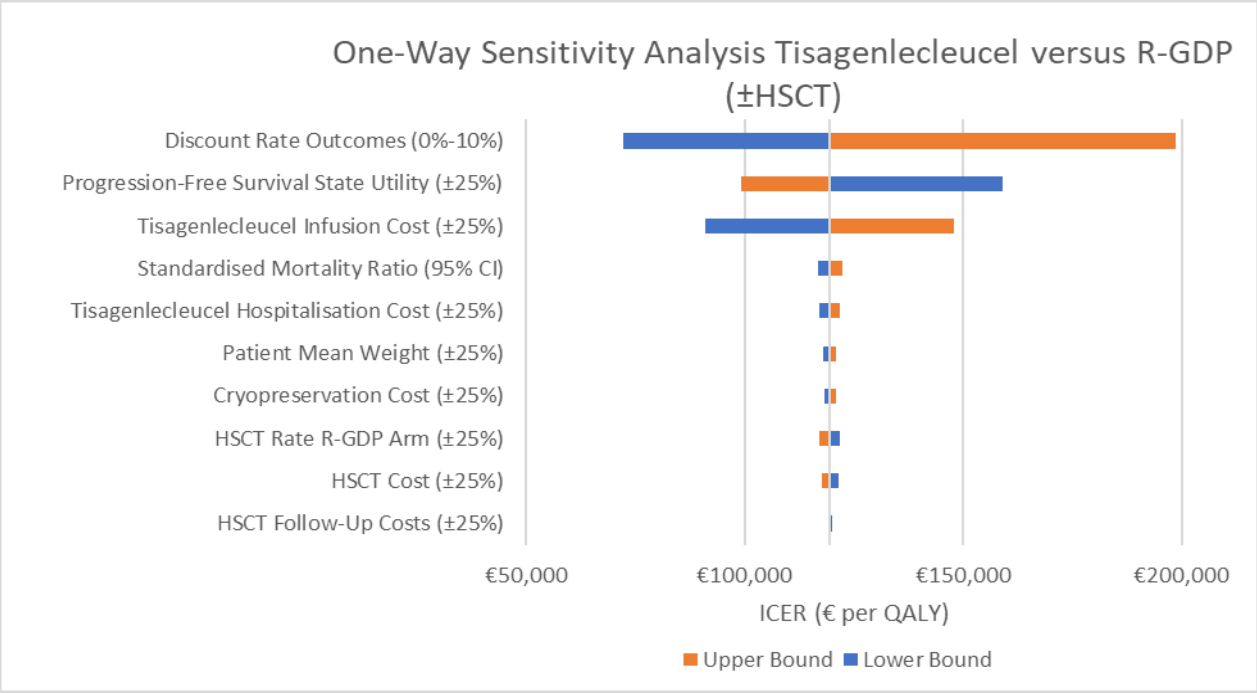

Figure 27 Tornado diagram of one-way sensitivity analysis of tisagenlecleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma (base case ICER: €119,509 per QALY)
HSCT: Haematopoietic stem cell transplant; **ICER:** Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

9.5.1.4 Scenario Analysis

Results of scenario analyses are presented in Table 57. Scenarios, which had the greatest impact on the ICER, are highlighted in bold.

Table 57 Impact of scenario analysis on deterministic ICER of tisagenlecleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma*

Parameter/Assumption	Base Case Assumption	Scenario	Justification	Plausibility of Scenario	Scenario ICER (€/QALY) (base case ICER: €119,509 per QALY)
Time Horizon	44 years	2 years	Median and maximum follow-up periods in JULIET were 19 and 29 months, respectively (393, 394)	Base case most plausible; represents a lifetime horizon	**862,927**
Proportion of Patients who Receive Tisagenlecleucel Infusion	69%	100%	Efficacy data from JULIET derived from those who received infusion (i.e. mITT)	Base case most plausible; represents clinical trial	111,482
Extrapolation of JULIET (tisagenlecleucel) OS Data	Two-knot (hazard) spline model	Gompertz model	Gompertz model was 'best fitting' parametric model	Uncertain. More conservative option chosen for base case	103,744
Extrapolation of CORAL Extension 1 (R-GDP) OS Data	Gompertz model	Two-knot (hazard) spline model	Two-knot (hazard) spline model was a reasonable option and maintains consistency with tisagenlecleucel arm	Base case most plausible; captures tail of the distributions more accurately	123,797
Extrapolation of JULIET and CORAL Extension 1 OS Data	Two-knot (hazard) spline (JULIET) and Gompertz (CORAL Extension 1) models	Gompertz (JULIET) and two-knot (hazard) spline model (CORAL Extension 1)	Combination of the two scenarios above	Uncertain. More conservative option chosen for base case	106,960
Extrapolation of JULIET PFS Data	One-knot (hazard) spline model	Generalised gamma model	Generalised gamma was 'best fitting' parametric model	Negligible impact on ICER	120,535
Time Point at which Patients are Considered Long-Term Survivors	After 60 months	After 24 months	Most patients expected to relapse within 24 to 60	Uncertain. More conservative	103,364

			months post-treatment (113, 390, 391)	option chosen for base case	
		No long-term survival point	Limited evidence that a proportion of patients will be long-term survivors	Base case most plausible; clinical opinion indicated that a cohort of patients survive long-term	143,854
Standardised Mortality Ratio		1.09	Ratio has been employed in the literature (113, 333)	Uncertain. Base case likely more plausible due to longer follow-up data	116,321
	1.36	3.4	Ratio accounts for death due to relapse and non-relapse (509)	Base case most plausible; model assumes that patients will not relapse after 60 months	136,470
Clinical Data Informing OS of R-GDP	CORAL Extension 1	SCHOLAR-1†	SCHOLAR-1 (108) identified in systematic literature review (Chapter 6)	Uncertain. Population of base case more reflective of population of interest	**173,397**
Health-State Utility Values	Progression-free survival: 0.83 Progressed disease: 0.71	Progression-free survival: 0.72 Progressed disease: 0.65	ZUMA-1 utility data available (332); however, not considered as robust as JULIET data	Base case most plausible; more robust data	137,540
HRQOL of Long-Term Survivors	All patients alive after 60 months assumed HRQOL equivalent to progression-free survival state (0.83)	All patients alive after 60 months assumed HRQOL equivalent to the age- and sex-matched general population (328)	Uncertainty exists regarding the HRQOL of long-term survivors (511)	Uncertain. More conservative option chosen for base case	111,289
Disutility Associated with Select Adverse Events	Include	Exclude	Possible that disutility is captured within health-state utility values	Negligible impact on ICER	118,700

Disutility Associated with Treatment and All Adverse Events	Select adverse events included only	Disutility of -0.15 (Guadagnolo et al. (337)) to account for treatment and adverse events‡	Alternative assumption identified through HTA appraisals (333, 399, 483)	Negligible impact on ICER	119,224
Duration of IV Immunoglobulin Treatment	3 years	11.4 months	Median time to B-cell recovery in ELIANA (250)	Base case most plausible; scenario likely an underestimate	115,796
		Duration of progression-free survival	B-cell aplasia may persist while patient is in remission (156)	Uncertain; lack of published data. Scenario is more conservative	**144,220**
Proportion of Patients Receiving HSCT in the R-GDP Arm	15%	30%	Rate of HSCT may be higher	Base case most plausible; as per clinical opinion	124,298
Proportion of Patients Receiving HSCT in the Tisagenlecleucel Arm	0%	6%|	Proportion of patients who received HSCT in JULIET (399)	Base case most plausible; patients censored at time of HSCT in JULIET	126,224
Discount Rate	4% on costs and outcomes	1.5% on costs and outcomes	NICE may consider a 1.5% discount rate where benefits are likely to be sustained over a very long period (369)	Base case most plausible; reflects current practice	89,243
		4% on costs and 1.5% on outcomes	Gravelle and Smith propose the rate on outcomes should be 1% to 3.5% lower than that on costs (186)		89,118
		Hyperbolic discounting: 4% (0-30 years), 3.5% (31-60 years) on both costs and outcomes	Hyperbolic discounting may be applicable when the time horizon exceeds 30 years (182)		118,586

HRQOL: Health-related quality of life; **HSCT:** Haematopoietic stem cell transplant; **HTA:** Health technology assessment; **ICER:** Incremental cost-effectiveness ratio; **IV**: Intravenous; **mITT:** Modified intention-to-treat; **NICE:** National Institute for Health and Care Excellence; **OS:** Overall survival; **PFS:** Progression-free survival; **QALY**: Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.
*Scenarios that had the greatest impact on the ICER are highlighted in **Bold**.
†One-knot (odds) spline model used to extrapolate OS data of SCHOLAR-1. No PFS data reported for SCHOLAR-1. Thus, PFS was derived by assuming the cumulative hazard function for PFS was proportional to the cumulative hazard function for OS. The ratio between PFS and OS (0.65) was based on the mean cumulative hazard ratio from the CORAL RCT (127, 505).
‡Disutility (-0.15) applied for 28 days (tisagenlecleucel) and 15 days (R-GDP). Disutility associated with cytokine release syndrome and non-cytokine release syndrome ICU admission accounted (tisagenlecleucel).
|No changes were made to efficacy data. Change in ICER reflects increased costs in the tisagenlecleucel arm.

9.5.1.5 Price Analysis

A 65% decrease (including 5.5% rebate) on the price-to-wholesaler of tisagenlecleucel was required to reduce the ICER to a willingness-to-pay threshold of €45,000 per QALY. The probability of cost effectiveness with this price decrease, at this threshold, was 54%.

9.5.1.6 Expected Value of Perfect Information

At a willingness-to-pay threshold of €45,000 per QALY, the 10-year population EVPI was €0.00. The population EVPI of tisagenlecleucel versus R-GDP, over a range of thresholds, is depicted in Figure 28.

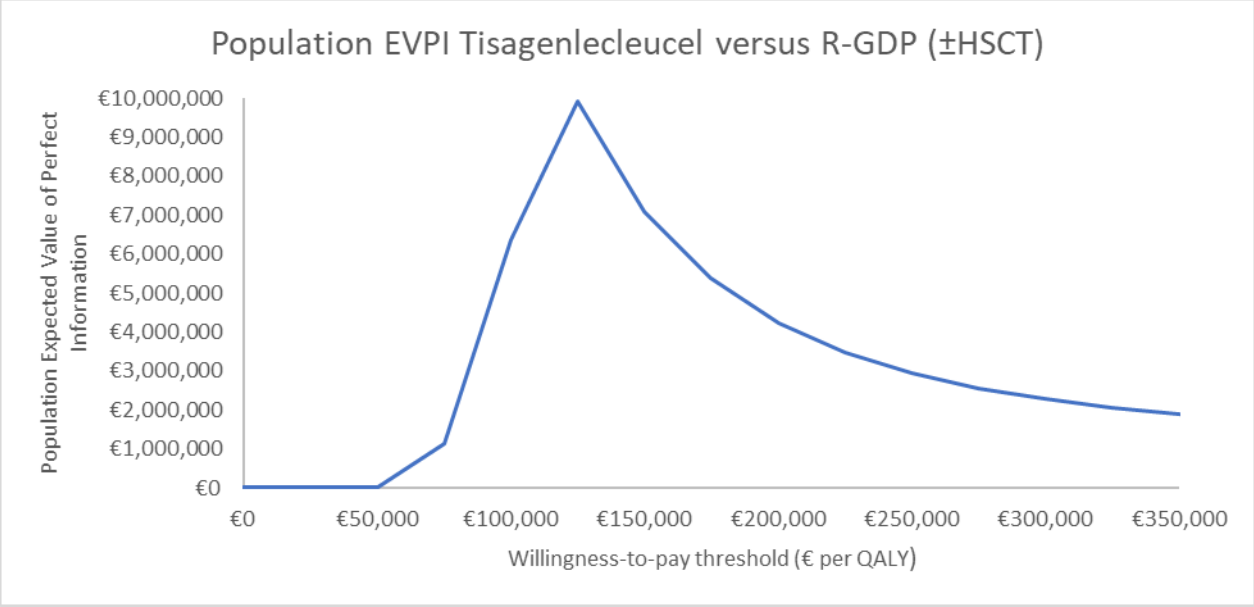

Figure 28 Population EVPI, over various willingness-to-pay thresholds, of tisagenlecleucel (price-to-wholesaler) versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
EVPI: Expected value of perfect information; **HSCT:** Haematopoietic stem cell transplant; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

The population EVPI analysis was re-run at the price that reduced the ICER of tisagenlecleucel to €45,000 per QALY (€104,702; representing a 65% price decrease including 5.5% rebate). At this price and threshold, the population EVPI was €3,989,438. The population EVPI of tisagenlecleucel at this price, over a range of thresholds, is depicted in Figure 29.

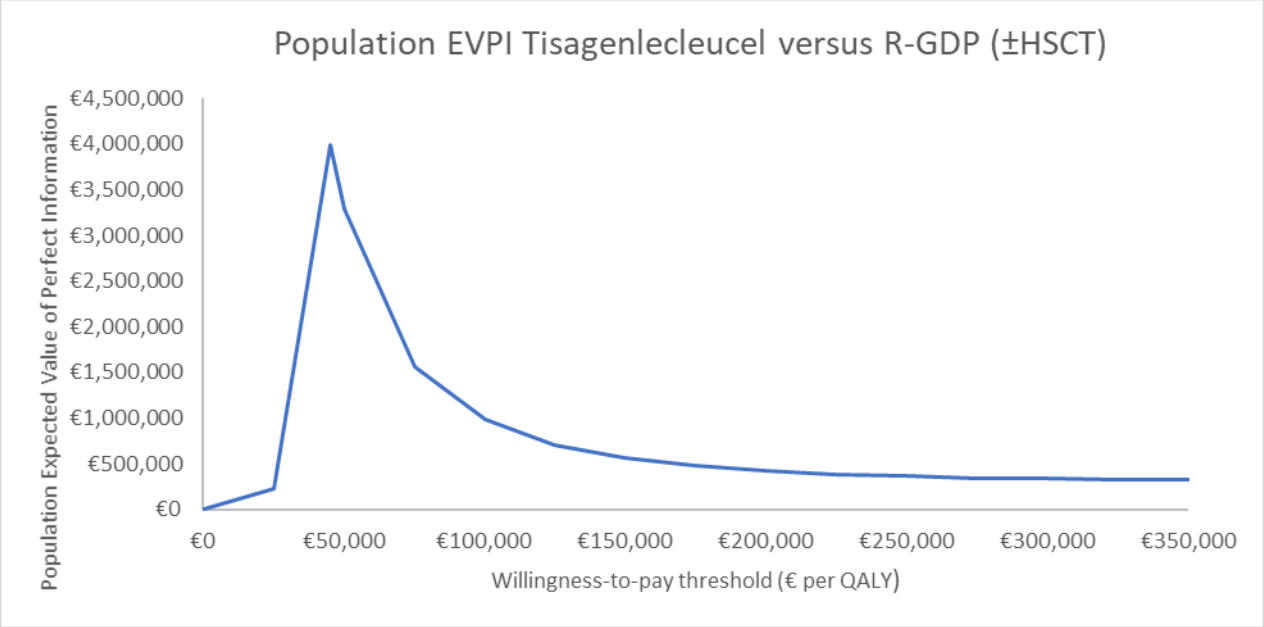

Figure 29 Population EVPI, over various willingness-to-pay thresholds, of tisagenlecleucel (price that reduced the ICER to €45,000 per QALY) versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
EVPI: Expected value of perfect information; **HSCT:** Haematopoietic stem cell transplant; **ICER:** Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

9.5.1.6.1 Partial Expected Value of Perfect Information

At the price-to-wholesaler of tisagenlecleucel (€301,762, including 5.5% rebate), and a willingness-to-pay threshold of €45,000 per QALY, the 10-year population EVPI was €0.00. Thus, population EVPPI was not estimated.

The population EVPPI analysis was run at the price that reduced the ICER of tisagenlecleucel to €45,000 per QALY (€104,702). At a willingness-to-pay threshold of €45,000 per QALY, the survival analysis parameters had the highest population EVPPI; €1,128,053. Utility values had the second highest population EVPPI; €905,809. Costs associated with hospitalisation and monitoring had the third highest population EVPPI; €718,740. HSCT cost, and adverse events cost parameters had population EVPPIs of

€668,497 and €62,932, respectively. Figure 30 depicts the value of uncertainty associated with each parameter category.

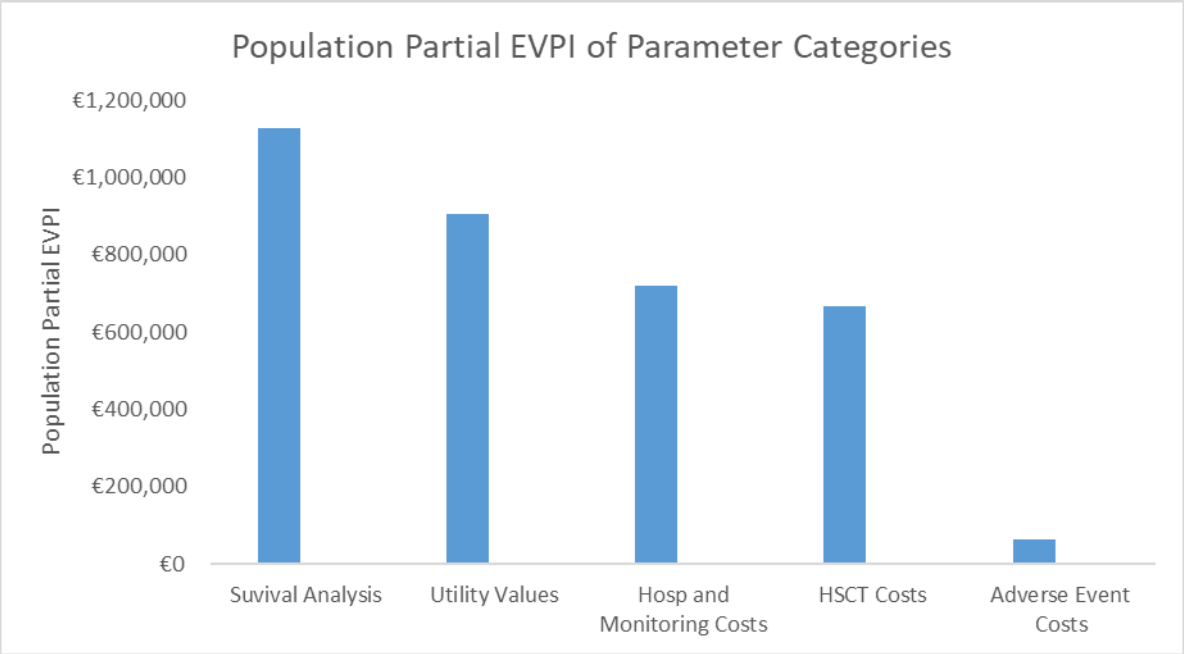

Figure 30 Population partial EVPI of parameter categories - tisagenlecleucel (price that reduced the ICER to €45,000 per QALY)
EVPI: Expected value of perfect information; **HSCT:** Haematopoietic stem cell transplant; **ICER:** Incremental cost-effectiveness ratio.

9.5.2 Results: Axicabtagene Ciloleucel versus R-GDP

9.5.2.1 Deterministic Results: Axicabtagene Ciloleucel versus R-GDP

The deterministic model outcomes are presented in Table 58. Axicabtagene ciloleucel was not cost effective, versus R-GDP, at a willingness-to-pay threshold of €45,000 per QALY (31).

Table 58 Deterministic results of the incremental analysis of cost effectiveness of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma

Drug	Total Costs (€)	Total QALYs	Incremental Costs (€)	Incremental QALYs	ICER (€/QALY)
R-GDP*	55,901	1.50			
Axicabtagene Ciloleucel	344,725	5.17	288,825	3.67	78,634

ICER: Incremental cost-effectiveness ratio; **QALY**: Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.
*With (15%) or without (85%) haematopoietic stem cell transplant.

9.5.2.2 Probabilistic Results: Axicabtagene Ciloleucel versus R-GDP

The expected incremental costs and incremental QALYs are presented in a scatterplot in Figure 31. All iterations lie in the NE quadrant. Mean expected costs and QALYs are

presented in Table 59. Mean probabilistic outputs were similar to those of the deterministic analysis, with a slightly higher ICER generated in the PSA.

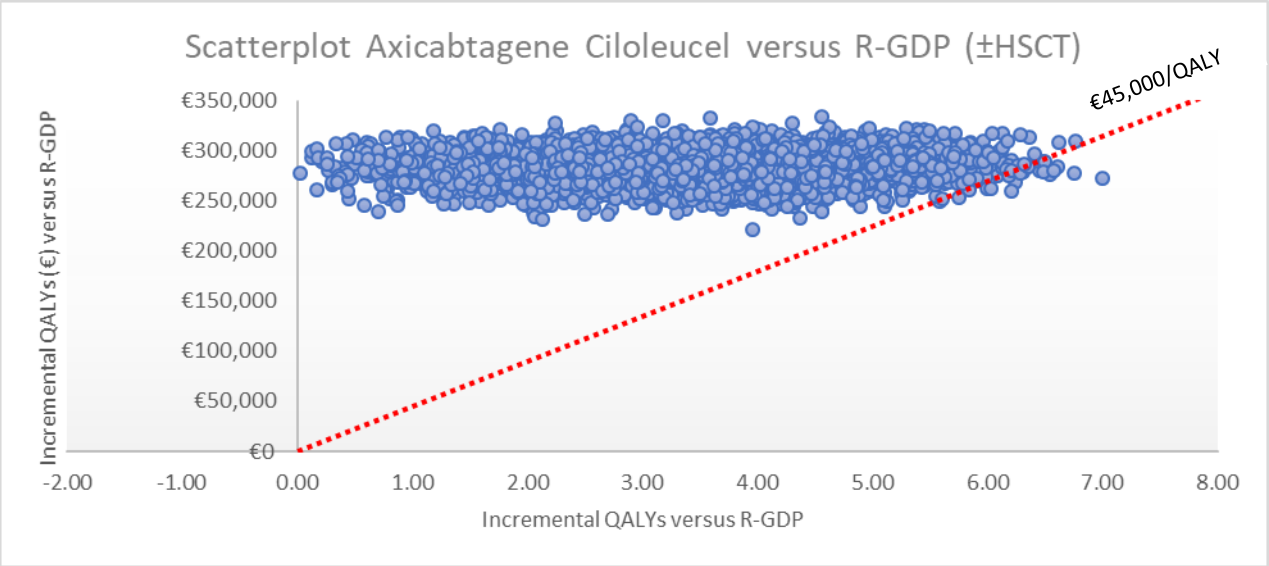

Figure 31 Scatterplot of incremental costs and incremental QALYs from probabilistic sensitivity analysis of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
HSCT: Haematopoietic stem cell transplant; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

Table 59 Mean probabilistic outputs of the incremental analysis of cost effectiveness of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma

Drug	Total Costs (€)	Total QALYs	Incremental Costs (€)	Incremental QALYs	ICER (€/QALY)
R-GDP*	55,716	1.65			
Axicabtagene Ciloleucel	339,556	5.22	283,839	3.57	79,444

ICER: Incremental cost-effectiveness ratio; **QALY**: Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.
*With (15%) or without (85%) haematopoietic stem cell transplant.

The cost-effectiveness acceptability curve is presented in Figure 32. At a willingness-to-pay threshold of €45,000 per QALY, there was a 0% probability that axicabtagene ciloleucel was cost effective versus R-GDP.

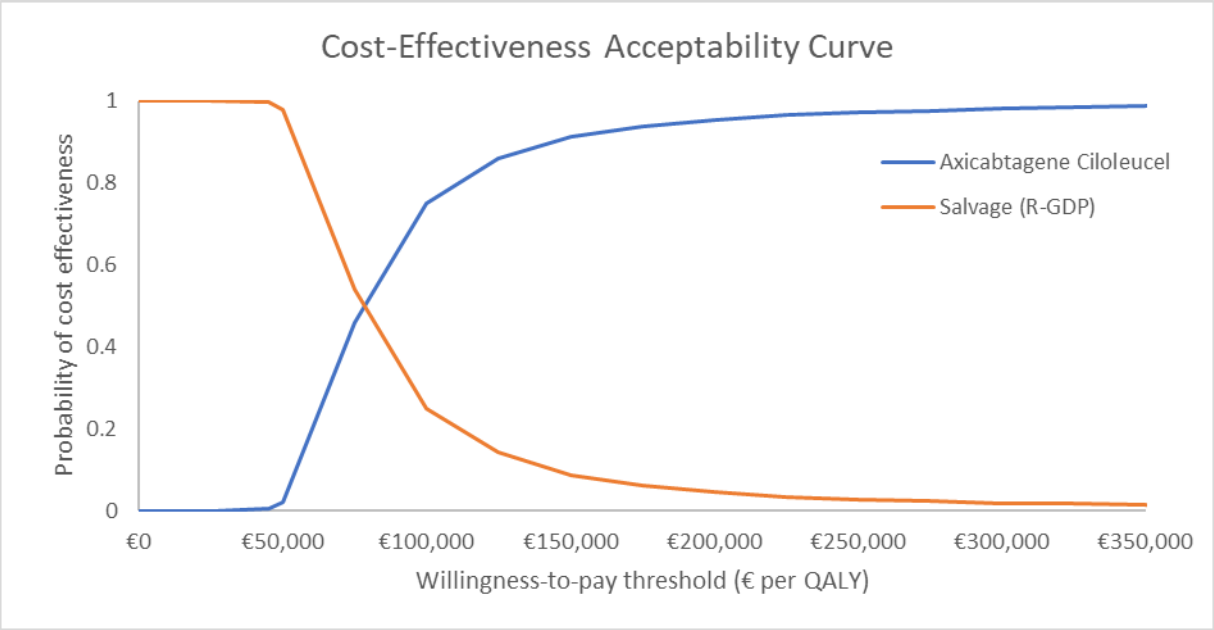

Figure 32 Cost-effectiveness acceptability curve of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
QALY: Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

9.5.2.3 One-Way Sensitivity Analysis

The outcomes of OWSA, versus R-GDP, are presented in Figure 33. For pragmatic reasons, OWSA was conducted on deterministic outcomes. Thus, results should be considered indicative only. The main drivers in the model were discount rate on outcomes, axicabtagene ciloleucel infusion cost, and progression-free survival state utility value. The willingness-to-pay threshold of €45,000 per QALY was not met in any analyses. The lowest ICER (€49,100 per QALY) occurred when the discount rate on outcomes was reduced to 0%.

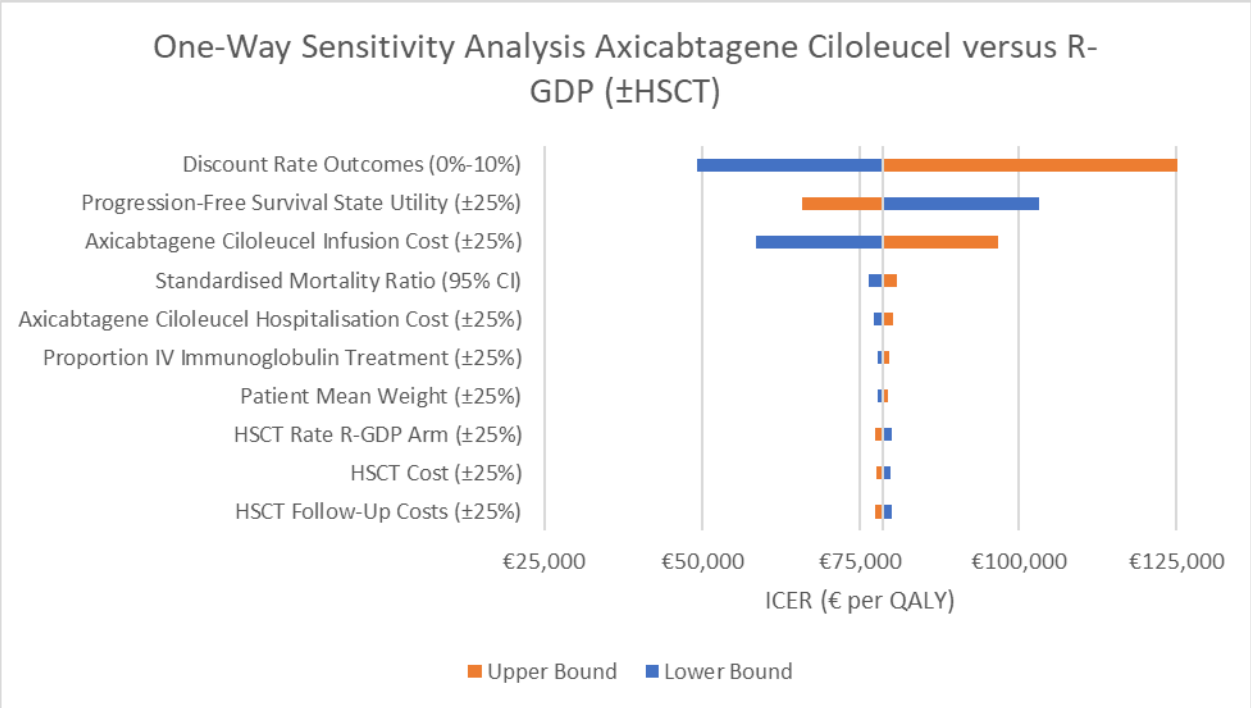

Figure 33 Tornado diagram of one-way sensitivity analysis of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma (base case ICER: €78,634 per QALY)
HSCT: Haematopoietic stem cell transplant; **ICER:** Incremental cost-effectiveness ratio; **IV:** Intravenous; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

9.5.2.4 Scenario Analysis

Results of scenario analyses are presented in Table 60. Scenarios, which had the greatest impact on the ICER, are highlighted in bold.

Table 60 Impact of scenario analysis on deterministic ICER of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma*

Parameter/Assumption	Base Case Assumption	Scenario	Justification	Plausibility of Scenario	Scenario ICER (€/QALY) (base case ICER: €78,634 per QALY)
Time Horizon	44 years	2 years	Median follow up in ZUMA-1 was 27.1 months (395)	Base case most plausible; represents a lifetime horizon	**529,881**
Proportion of Patients who Receive Axicabtagene Ciloleucel Infusion	91%	100%	Efficacy data from ZUMA-1 derived from those who received infusion (i.e. mITT)	Base case most plausible; represents clinical trial	77,170
Extrapolation of ZUMA-1 (axicabtagene ciloleucel) OS Data	Two-knot (hazard) spline model	Gompertz model	Gompertz model was 'best fitting' parametric model	Uncertain. Base case provided more flexible extrapolation	79,689
Extrapolation of CORAL Extension 1 (R-GDP) OS Data	Gompertz model	Two-knot (hazard) spline model	Two-knot (hazard) spline model was a reasonable option and maintains consistency with axicabtagene ciloleucel arm	Base case most plausible; captures tail of the distributions more accurately	77,149
Extrapolation of ZUMA-1 and CORAL Extension 1 OS Data	Two-knot (hazard) spline (ZUMA-1) and Gompertz (CORAL Extension 1) models	Gompertz (ZUMA-1) and two-knot (hazard) spline (CORAL Extension 1) models	Combination of the two scenarios above	Negligible impact on ICER	79,107
Extrapolation of ZUMA-1 PFS Data	Two-knot (hazard) spline model	Gompertz model	Gompertz model was 'best fitting' parametric model	Negligible impact on ICER	78,537
Time Point at which Patients are Considered Long-Term Survivors	After 60 months	After 24 months	Most patients expected to relapse within 24 to	Uncertain. More conservative option chosen for base case	71,885

			60 months post-treatment (113, 390, 391)		
		No long-term survival point	Limited evidence that a proportion of patients will be long-term survivors	Base case most plausible; clinical opinion indicated that a cohort of patients survive long-term	85,660
Standardised Mortality Ratio		1.09	Ratio has been employed in the literature (113, 333)	Uncertain. Base case likely more plausible due to longer follow-up data	76,104
	1.36	3.4	Ratio accounts for death due to relapse and non-relapse (509)	Base case most plausible; model assumes that patients will not relapse after 60 months	92,545
Clinical Data Informing OS of R-GDP	CORAL Extension 1	SCHOLAR-1†	SCHOLAR-1 (108) identified in systematic literature review (Chapter 6)	Uncertain. Population of base case more reflective of population of interest. Scenario population closely aligned with ZUMA-1	**94,805**
Health-State Utility Values	Progression-free survival: 0.83 Progressed disease: 0.71	Progression-free survival: 0.72 Progressed disease: 0.65	ZUMA-1 utility data available (332); however, not considered as robust as JULIET data	Base case most plausible; more robust data	90,371
HRQOL of Long-Term Survivors	All patients alive after 60 months assumed HRQOL equivalent to progression-	All patients alive after 60 months assumed HRQOL equivalent to the age- and sex-matched general	Uncertainty exists regarding the HRQOL of long-term survivors (511)	Uncertain. More conservative option chosen for base case	72,969

	free survival state (0.83)	population (328)			
Disutility Associated with Select Adverse Events	Include	Exclude	Possible that disutility is captured within health-state utility values	Negligible impact on ICER	78,259
Disutility Associated with Treatment and All Adverse Events	Include	Disutility of -0.15 (Guadagnolo et al. (337)) to account for treatment and adverse events‡	Alternative assumption identified through HTA appraisals (333, 399, 483)	Negligible impact on ICER	78,475
Duration of IV Immunoglobulin Treatment	3 years	11.4 months	Median time to B-cell recovery in ELIANA (250)	Base case most plausible; scenario likely an underestimate	76,600
		Duration of progression-free survival	B-cell aplasia may persist while patient is in remission (156)	Uncertain; lack of published data. Scenario is more conservative	**92,691**
Proportion of Patients Receiving HSCT in the R-GDP Arm	15%	30%	Rate of HSCT may be higher	Base case most plausible; as per clinical opinion	77,023
Proportion of Patients Receiving HSCT in the Axicabtagene Ciloleucel Arm	0%	3%|	Proportion of patients who received HSCT in ZUMA-1 (144)	Base case most plausible; patients not expected to receive HSCT	80,717
Discount Rate	4% on costs and outcomes	1.5% on costs and outcomes	NICE may consider a 1.5% discount rate where benefits are likely to be sustained over a very long period (369)	Base case most plausible; reflects current practice	59,928
		4% on costs and 1.5% on outcomes	Gravelle and Smith propose that the discount rate on health outcomes		59,639

		should be 1% to 3.5% lower than the discount rate on costs (186)	
	Hyperbolic discounting: 4% (0-30 years), 3.5% (31-60 years) on both costs and outcomes	Hyperbolic discounting may be applicable when the time horizon exceeds 30 years (182)	78,276

HRQOL: Health-related quality of life; **HSCT:** Haematopoietic stem cell transplant; **HTA:** Health technology assessment; **ICER:** Incremental cost-effectiveness ratio; **IV:** Intravenous; **mITT:** Modified intention-to-treat; **NICE:** National Institute for Health and Care Excellence; **OS:** Overall survival; **PFS:** Progression-free survival; **QALY**: Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.
*Scenarios that had the greatest impact on the ICER are highlighted in **Bold**.
†One-knot (odds) spline model used to extrapolate OS data of SCHOLAR-1. No PFS data reported for SCHOLAR-1. Thus, PFS was derived by assuming the cumulative hazard function for PFS was proportional to the cumulative hazard function for OS. The ratio between PFS and OS (0.65) was based on the mean cumulative hazard ratio from the CORAL RCT (127, 505).
‡Disutility (-0.15) applied for 28 days (axicabtagene ciloleucel) and 15 days (R-GDP). Disutility associated with cytokine release syndrome and non-cytokine release syndrome ICU admission accounted for in the axicabtagene ciloleucel arm.
|No changes were made to efficacy data. Change in ICER reflects increased costs in the axicabtagene ciloleucel arm.

9.5.2.5 Price Analysis

A 44% decrease (including 5.5% rebate) on the price-to-wholesaler of axicabtagene ciloleucel was required to reduce the ICER to a willingness-to-pay threshold of €45,000 per QALY. The probability of cost effectiveness with this price decrease, at this threshold, was 52%.

9.5.2.6 Expected Value of Perfect Information

At a willingness-to-pay threshold of €45,000 per QALY, the 10-year population EVPI was €0.00. The population EVPI of axicabtagene ciloleucel versus R-GDP, over a range of willingness-to-pay thresholds, is depicted in Figure 34.

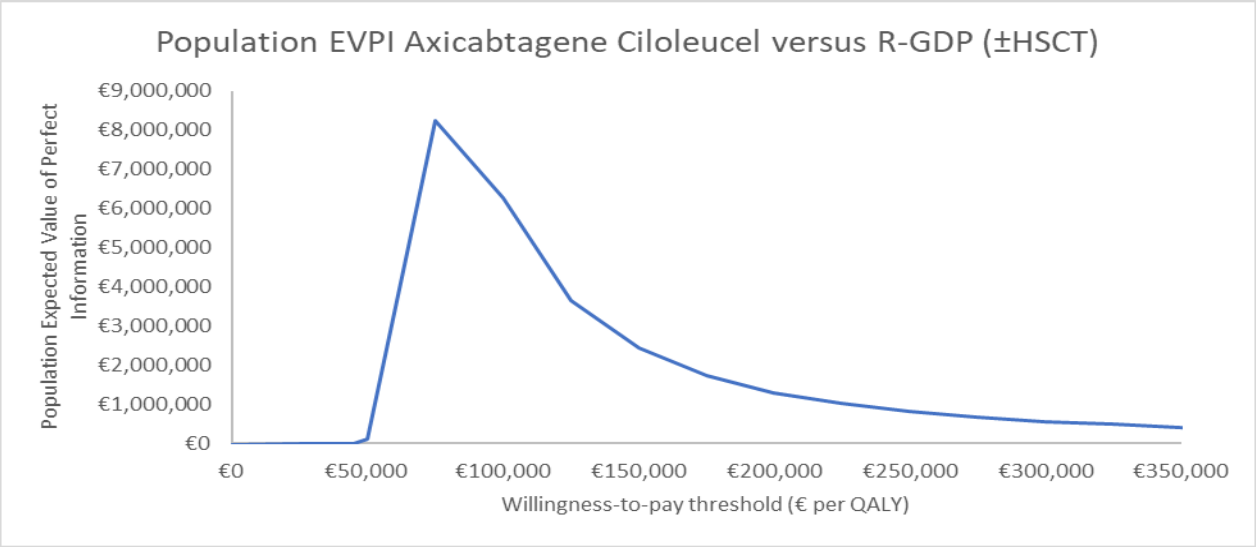

Figure 34 Population EVPI, over various willingness-to-pay thresholds, of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
EVPI: Expected value of perfect information; **HSCT:** Haematopoietic stem cell transplant; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

The population EVPI analysis was re-run at the price that reduced the ICER of axicabtagene ciloleucel to €45,000 per QALY (€173,082; representing a 44% price decrease including 5.5% rebate). At this price and threshold, the population EVPI was €6,137,514. The population EVPI of axicabtagene ciloleucel at this price, over a range of thresholds, is depicted in Figure 35.

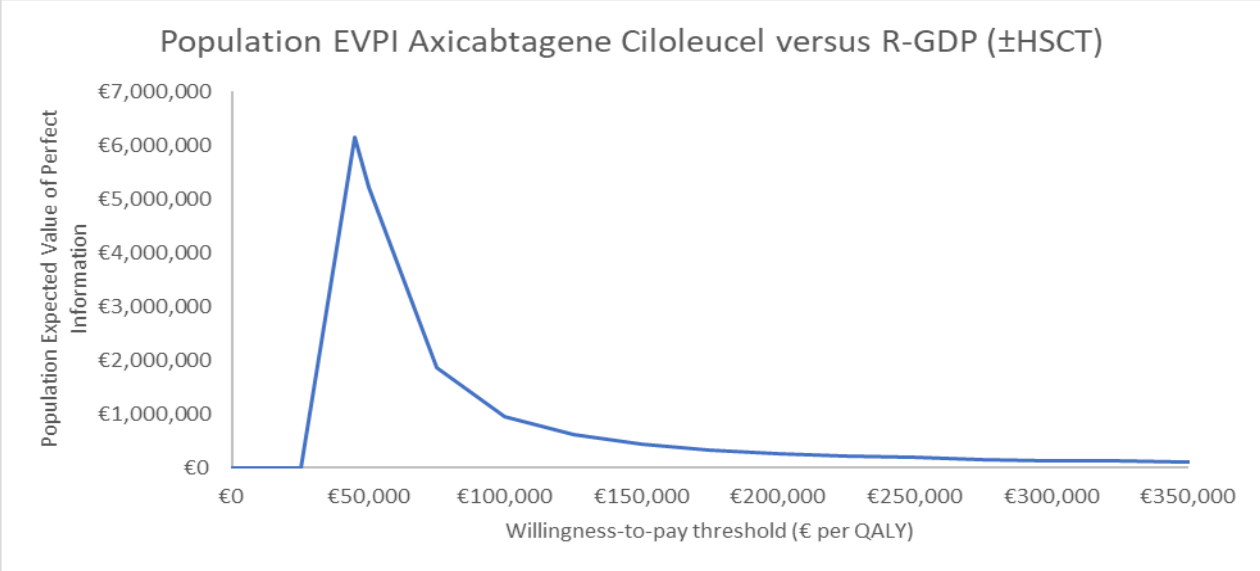

Figure 35 Population EVPI, over various willingness-to-pay thresholds, of axicabtagene ciloleucel (price that reduced the ICER to €45,000 per QALY) versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma
EVPI: Expected value of perfect information; **HSCT:** Haematopoietic stem cell transplant; **ICER:** Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

9.5.2.6.1.1 Partial Expected Value of Perfect Information

At the price-to-wholesaler of axicabtagene ciloleucel (€309,015, including 5.5% rebate), and a willingness-to-pay threshold of €45,000 per QALY, the 10-year population EVPI was €0.00. Thus, population EVPPI was not estimated.

The population EVPPI analysis was run at the price that reduced the ICER of axicabtagene ciloleucel to €45,000 per QALY (€173,082). At a willingness-to-pay threshold of €45,000 per QALY, utility values had the highest population EVPPI; €1,712,845. Parameters associated with survival analysis had the second highest population EVPPI; €1,413,136. Costs associated with hospitalisation and monitoring had the third highest population EVPPI; €1,122,766. Parameters associated with HSCT cost had a population EVPPI valued at €1,028,822. Parameters associated with adverse event costs had negligible population EVPPI (€125,319). Figure 36 depicts the value of uncertainty associated with each parameter category.

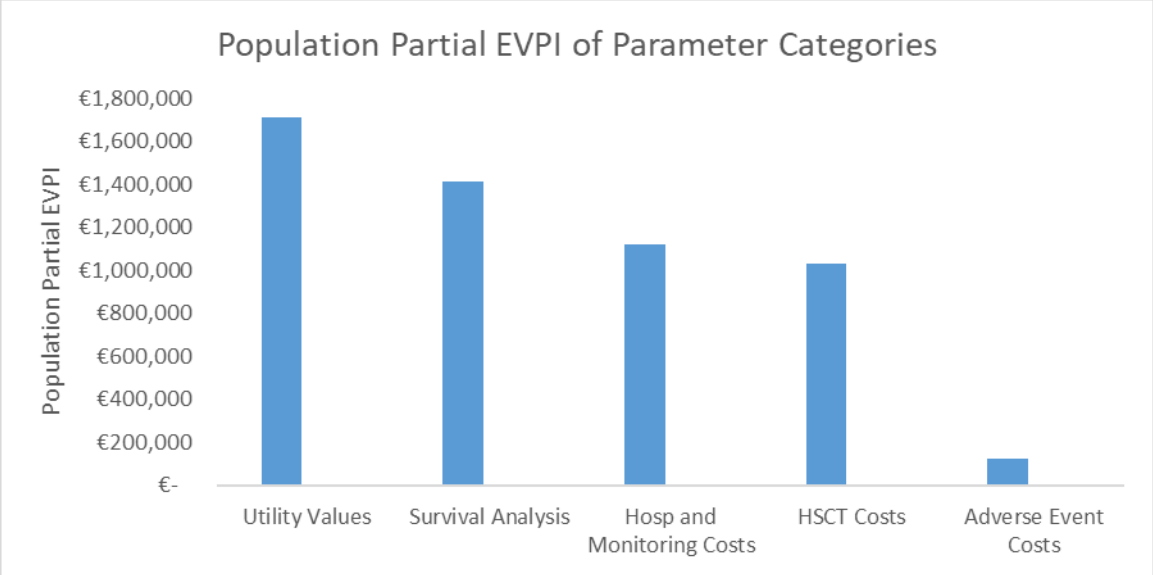

Figure 36 Population partial EVPI of parameter categories – axicabtagene ciloleucel (price that reduced the ICER to €45,000 per QALY)
EVPI: Expected value of perfect information; **HSCT:** Haematopoietic stem cell transplant; **ICER:** Incremental cost-effectiveness ratio.

9.6 Discussion

9.6.1 Deterministic and Probabilistic Results

The results of these cost-utility analyses indicate that, at current prices-to-wholesaler, neither tisagenlecleucel nor axicabtagene ciloleucel are cost effective, versus R-GDP, at a

€45,000 per QALY willingness-to-pay threshold. It is likely that the true ICERs differ to those generated as part of this research, due to commercial patient access schemes that are in place for these therapies.

A high degree of uncertainty exists in the clinical evidence base of tisagenlecleucel and axicabtagene ciloleucel. This translates to uncertainty in cost-effectiveness, which may not be adequately captured by OWSA, PSA, and EVPI. Although parameter uncertainty in the model was captured by means of PSA and EVPI, uncertainty associated with the naïve ITC is difficult to quantify. As highlighted in 5.6, uncertainty associated with long-term survival is often poorly captured by the PSA (378). The long-term survival of patients treated with tisagenlecleucel and axicabtagene ciloleucel is uncertain. As such, caution is warranted in the interpretation of results. Reassuringly, negligible differences were observed between the deterministic and probabilistic ICERs, due to non-linearity between the parameter inputs and model outputs. For tisagenlecleucel and axicabtagene ciloleucel (at price-to-wholesaler), the probability of cost effectiveness at the €45,000 per QALY threshold was 0%. The cost-effectiveness acceptability curves indicated that the probability of cost effectiveness of tisagenlecleucel exceeded that of R-GDP at an approximate threshold of €117,000 per QALY. The probability of cost effectiveness of axicabtagene ciloleucel exceeded that of R-GDP at an approximate threshold of €80,000 per QALY.

9.6.2 One-Way Sensitivity and Scenario Analyses

Results of OWSA illustrated the sensitivity of the models to certain input parameters. Scenario analyses illustrated the impact of employing alternative, plausible assumptions.

9.6.2.1 Time Horizon

As expected, the models were sensitive to variations in the time horizon. Reducing the time horizon to 2 years (the approximate follow up of JULIET (393, 394) and ZUMA-1 (395)) had a drastic impact on the ICERs, increasing the base case ICERs 7-fold. Although ICERs pertaining to a time horizon of two years are not appropriate for decision-making, they illustrate the reliance of the cost-effectiveness outputs on the uncertain long-term extrapolated segment of the model.

9.6.2.2 Discount Rate

The sensitivity of the models to the discount rate on outcomes, as demonstrated in the OWSA, reflects the extent to which outcomes are accrued over the long-term in the models. For both CD19 CAR T-cell therapies, the scenario whereby outcomes were discounted at 1.5% (whilst maintaining 4% discount on costs) resulted in the greatest decrease in the base case ICERs. Reductions of 25% and 24% were noted in the base case ICERs of tisagenlecleucel and axicabtagene ciloleucel, respectively. These scenarios generated a considerable increase in total QALY gain for both tisagenlecleucel and axicabtagene ciloleucel. The magnitude of increase in total QALY gain for R-GDP was less. This was due to the fact that most patients in CORAL Extension 1 died within the first two years (123). The trends observed are similar to those discussed in 5.6.2.2.

9.6.2.3 Survival Extrapolation

The choice of the most appropriate survival model to extrapolate the OS data of JULIET was uncertain. Adopting the Gompertz model to extrapolate these data (instead of the more flexible two-knot (hazard) spline model), reduced the base case ICER by approximately €15,700 per QALY. The two-knot (hazard) spline model may be considered a slightly more conservative approach, as it did not plateau to the same extent as the Gompertz. Thus, resulting in slightly lower five-year predicted survival (and lower QALY gain). The impact on the total costs was negligible. When a reimbursement decision is required in the face of uncertainty in the extrapolation output, a conservative approach is usually adopted. This potentially reduces the financial risk to the payer. However, this approach may lead to a risk of loss to the Applicant should the technology prove to be more effective than predicted. Performance-based risk-sharing agreements may be an appropriate approach to reimbursement in such instances, allowing the distribution of financial risk to be shared between the payer and Applicant. Further exploration is provided in Chapter 11.

Employing the Gompertz model (instead of the two-knot (hazard) spline model) to extrapolate the OS data of ZUMA-1 had a negligible impact on the ICER of axicabtagene ciloleucel versus R-GDP. In this instance, the predicted OS generated by the Gompertz and two-knot (hazard) spline models was very closely aligned. Although these models

provided an appropriate fit to the observed trial data, this does not guarantee that the extrapolated data will be an accurate reflection of outcomes in clinical practice.

9.6.2.4 Time Point of Long-Term Survival

The time point of long-term survival is subject to uncertainty, in that a range of time points have been proposed in the literature (113, 390, 507). Reducing the time point of long-term survival to 24 months resulted in a notable decrease in the ICERs. This was driven mainly by an increase in the incremental QALY gain for both tisagenlecleucel and axicabtagene ciloleucel. As highlighted, the majority of patients in the R-GDP arm died within 24 months (123). The impact on incremental costs was negligible. A 'worst case' scenario, whereby outcomes were derived from full extrapolation of the trial data without additional structural assumptions regarding the time point of long-term survival, had a notable impact on the ICERs, particularly for tisagenlecleucel versus R-GDP (increase of 20% from base case ICER). This trend is aligned with that discussed in 5.6.2.4.

9.6.2.5 Efficacy Data to Inform Overall Survival of R-GDP

Using data from SCHOLAR-1 to inform OS of R-GDP had a large impact on the ICERs. Increases, from the base case ICERs, of 45% and 21% were noted for tisagenlecleucel and axicabtagene ciloleucel, respectively. The limitations of SCHOLAR-1 have been described in Chapter 6. Despite these limitations, SCHOLAR-1 has been used to inform the efficacy of the comparator arm in several cost-utility analyses of tisagenlecleucel (512, 513) and axicabtagene ciloleucel (5, 296, 475). It has been proposed that SCHOLAR-1 is enriched with patients who have high-risk disease (409, 410). However, SCHOLAR-1 was associated with higher total QALY gain for the R-GDP arm, when compared to total QALY gain using the CORAL Extension 1 (with or without HSCT) data. The higher QALY gain was the driver of the higher ICERs generated when the SCHOLAR-1 data were used. Notably, the SCHOLAR-1 data were not adjusted to reflect the lower rate of HSCT in Irish clinical practice (29.9% versus 15%). This higher proportion of patients may partly explain the higher QALY gain. SCHOLAR-1, used to inform efficacy of the comparator arm, was also associated with a higher ICER when compared to the pooled CORAL Extension 1 and CORAL Extension 2 studies (€197,119 per QALY versus €122,266 per QALY), in the NCPE HTA appraisal of tisagenlecleucel (10).

9.6.2.6 Health-Related Quality of Life Inputs

OWSA indicated that the progression-free survival state utility value was a driver of cost effectiveness. This may partly be due to the fact that all patients alive after 60 months were assumed to have HRQOL of the progression-free survival state. The majority of QALY gains in the tisagenlecleucel and axicabtagene ciloleucel arms were driven by QALYs accrued in the extrapolation of survival. Assuming that patients who were alive after 60 months accrued utility equivalent to the age- and sex-matched general population reduced the base case ICERs of both tisagenlecleucel (by approximately €8,200 per QALY) and axicabtagene ciloleucel (by approximately €5,600 per QALY). Notably, the utility value for the progression-free survival state (0.83) is equivalent to general population utility at 84 years in the model (as per the formula by Ara and Brazier) (328). An SLR of HRQOL of patients with aggressive NHL who survive long term, sponsored by the manufacturer of axicabtagene ciloleucel, concluded that patients with aggressive NHL who survive more than two years show improvement or no change in their overall HRQOL compared with baseline (511). Based on the sensitivity of the model here, HRQOL of patients who are considered to be long-term survivors is a key area of uncertainty.

Scenario analyses, exploring the impact of employing the utility values from ZUMA-1, resulted in increases in the base case ICERs of both tisagenlecleucel (approximately €18,000 per QALY) and axicabtagene ciloleucel (approximately €11,700 per QALY). The lower utility values, for the progression-free survival and progressed disease states, derived from patients in ZUMA-1, resulted in lower total QALY gain for all treatment arms. Although the utility values derived from JULIET were considered more robust than those from ZUMA-1, values derived from both trials may be considered relevant.

9.6.2.7 Cost Inputs

Although the model was sensitive to changes in the price of CD19 CAR T-cell therapy, reducing the prices-to-wholesaler of tisagenlecleucel and axicabtagene ciloleucel (by 25%) did not reduce the ICERs to the willingness-to-pay threshold of €45,000 per QALY. Although further simple price reductions on tisagenlecleucel and axicabtagene ciloleucel may reduce the ICER to an acceptable willingness-to-pay threshold (as demonstrated in

the Price Analyses, 9.5.1.5 and 9.5.2.5), they do not address the risk to both payers and patients (due to uncertainty in the clinical evidence).

9.6.2.8 Haematopoietic Stem Cell Transplant

Accounting for the proportion of patients who received HSCT in JULIET (6%) and ZUMA-1 (3%) increased both ICERs. Notably, the efficacy data were not modified for these scenario analyses. Patients in JULIET were censored at the time of HSCT (392), while those in ZUMA-1 were not censored at the time of HSCT (144). Changes in the ICERs, in these scenarios, are attributable solely to the additional cost of HSCT. Interestingly, increasing the proportion of patients proceeding to HSCT in the R-GDP arm had opposite effects on the cost effectiveness of tisagenlecleucel (increased ICER) and the cost effectiveness of axicabtagene ciloleucel (slight decreased ICER). This is likely due to the higher proportion of patients proceeding to R-GDP in the tisagenlecleucel arm (compared to axicabtagene ciloleucel).

9.6.3 Expected Value of Perfect Information

For both therapies (at price-to-wholesaler), the population EVPI, versus R-GDP, reached a peak at willingness-to-pay thresholds of approximately €125,000 per QALY for tisagenlecleucel, and approximately €75,000 per QALY for axicabtagene ciloleucel. At these peaks, the probabilities of cost effectiveness were 56% and 46% for tisagenlecleucel and axicabtagene ciloleucel, respectively. As the thresholds increased beyond these respective values, the probabilities of cost effectiveness of both therapies increased and the population EVPI decreased (given that the corresponding consequences of resolving decision uncertainty decreased) (12). Population EVPI (at price-to-wholesaler and €45,000 per QALY threshold) indicated that no cost should be attributed to additional research for either tisagenlecleucel or axicabtagene ciloleucel. As highlighted in 5.6.3, EVPI and EVPPI analyses examine uncertainty in parameters. Structural uncertainty, associated with the naïve ITC, was not captured.

For both therapies, re-running the EVPI analysis at the price that generated an ICER of €45,000 per QALY, increased population EVPI considerably. Here, population EVPPI indicated that conducting additional research to reduce uncertainty associated with

survival analysis parameters (highest population EVPPI for tisagenlecleucel), and utility values (highest population EVPPI for axicabtagene ciloleucel) would be most valuable. As highlighted, the prices employed as part of this scenario represent 'best case' scenarios.

It is noteworthy that the population EVPPI scenario analyses highlighted survival analysis parameters, and utility values. Both tisagenlecleucel and axicabtagene ciloleucel received EMA marketing authorisation on the basis of single-arm trials with short duration of follow up (143, 144). In terms of utility values, there is a paucity of Irish-specific data available. Although work is ongoing to provide valuation of health-state preferences in Ireland (502, 514), utility values will continue to be associated with uncertainty. Potential studies to address uncertainties in survival parameters could incorporate the EQ-5D-3L instrument, addressing uncertainty with both survival and HRQOL. Conducting additional research to address uncertainties associated with adverse events would be of less value (population EVPPI negligible). The results of this EVPPI analysis can be used to inform performance-based risk-sharing agreements; further discussion is provided in Chapter 11.

The low EVPI and EVPPI estimated here are likely a reflection, to some degree, of the low patient numbers in the analysis (36 patients per year). It should be noted that in this analysis, the sum of the population EVPPI estimates were lower than the total EVPI. This is likely due to correlations between parameters within the model (372).

9.6.4 Comparison with the Published Literature

The conclusions of this analysis (i.e. not cost effective at €45,000 per QALY) are in line with those of the NCPE HTA appraisals of tisagenlecleucel (10) and axicabtagene ciloleucel (9). The ICERs generated here, however, are not directly comparable with those estimated in the NCPE HTA appraisals (9, 10). Access to raw IPD allowed the Applicant to conduct adjusted ITCs to generate estimates of relative efficacy using the SCHOLAR-1 (tisagenlecleucel and axicabtagene ciloleucel HTA appraisals) and pooled CORAL Extension 1 and CORAL Extension 2 (tisagenlecleucel HTA appraisal) studies, to inform the efficacy of the comparator arm. This allowed for some, but not all, differences between the study populations to be adjusted for (9, 10).

The ICER estimated as part of this research, for the comparison of tisagenlecleucel versus R-GDP, was slightly lower than that estimated in the NCPE HTA appraisal of tisagenlecleucel (efficacy of comparator informed by pooled CORAL Extension 1 and CORAL Extension 2); €119,509 per QALY versus €122,266 per QALY. Of note, the incremental QALY gain estimated in this study, was higher than that estimated in the NCPE HTA appraisal of tisagenlecleucel; 1.82 and 1.63, respectively (10). Aside from the different data sources used to inform the efficacy of the comparator arm, it is difficult to determine the reasons for the differences between these ICERs. This is due to the limited data available from the NCPE Technical Summary of tisagenlecleucel (10).

The ICER estimated as part of this research, for the comparison of axicabtagene ciloleucel versus R-GDP, was considerably lower than that estimated in the NCPE HTA appraisal of axicabtagene ciloleucel (€241,416 per QALY; efficacy of comparator informed by SCHOLAR-1). The incremental QALY gain estimated in this study, was also higher than that estimated in the NCPE HTA appraisal of axicabtagene ciloleucel; 3.67 versus 1.73, respectively (9). One of the main drivers of difference between the ICERs is likely to be the use of parametric models to extrapolate the OS and PFS data in the NCPE HTA appraisal of axicabtagene ciloleucel. It does not appear that an assumption, whereby patients alive at a certain time point experienced morality aligned with that of the age- and sex-matched general population, was employed in the NCPE HTA appraisal. HRQOL data derived from ZUMA-1 were also used in the NCPE HTA appraisal (9).

Notably, the discount rate on costs and outcomes reduced from 5% to 4% (employed in this research) since the NCPE HTA appraisals of tisagenlecleucel and axicabtagene ciloleucel (9, 10).

Incremental QALYs, from published analyses of tisagenlecleucel, ranged from 0.98 (ICER: US $320,200 per QALY), versus R-ICE/R-DHAP, in the Singaporean healthcare setting (515) to 3.35 (ICER: US $78,652 per QALY), versus R-ICE/R-GDP/R-DHAP/R-GEMOX, in the US healthcare setting (512). The modelling approach and data used to inform efficacy of the comparator arm, by Cher et al. in the Singaporean healthcare setting, was aligned with the approach used here. However, this study differed from Cher et al. in the utility values and discount rate employed (515). Only one identified study, by Wakase et al.,

constructed a decision tree to model the costs and outcomes of patients who did not proceed to infusion with tisagenlecleucel in JULIET (513). Published studies, which do not account for these patients may be biased (382, 512).

The cost effectiveness of axicabtagene ciloleucel, in the US healthcare system, was assessed in three separate publications identified in the literature (5, 296, 475). All three studies used SCHOLAR-1 to inform efficacy of the comparator arm. Whittington et al. presented the results of a number of different incremental analyses from the public payer perspective, which represented different modelling assumptions. The incremental QALY gain with axicabtagene ciloleucel ranged from 1.52 (parametric extrapolation; ICER: US $230,900 per QALY) to 4.90 (mixture cure model extrapolation; ICER US $82,400 per QALY) under these assumptions (5). Roth et al. estimated an incremental QALY gain of 6.54 (ICER: US $58,146 per QALY) (475), while the ICER HTA estimated an incremental QALY gain of 3.40 (ICER: US $136,078 per QALY) (296). Both Roth et al. and the ICER HTA extrapolated the ZUMA-1 data to five years, after which patients were assumed morality equivalent to that of the age- and sex-matched general population. All three analyses employed different utility values and discount rates to those used in the model developed here (5, 296, 475). The decision tree component was used by Whittington et al. (5) and the ICER HTA (296) to model patients who did not proceed to infusion in ZUMA-1. Conversely, Roth et al. did not account for patients who did not proceed to infusion (475).

A similar pattern of influential parameters was observed in the OWSA and scenario analyses, conducted as part of this research, compared with published cost-utility analyses. For tisagenlecleucel, the discount rate on outcomes (382, 512, 513, 515), tisagenlecleucel infusion cost (382, 512, 515), time horizon (382, 515), progression-free survival state utility (382, 512, 513, 515), and cost of HSCT (512, 513, 515) were all reported to be key drivers of published models. For axicabtagene ciloleucel, the discount rate on outcomes (296, 475), axicabtagene ciloleucel infusion cost (475), the rate of HSCT in the comparator arm (296), and the SMR (296) were all identified as main drivers in the models. This provides some degree of external validity to the results presented here. None of the identified analyses conducted EVPI or EVPPI analyses.

9.6.5 Limitations

The limitations presented here are in addition to those described thus far. In the absence of detailed data on the outcomes of patients who did not proceed to infusion in JULIET and ZUMA-1, a number of assumptions were required. It is unclear if these assumptions are truly reflective of clinical practice.

The limitations of the partitioned survival model approach, described in 5.2.2.2, are also applicable to the models used in this study. The model structure is a simplistic representation of the clinical pathway and does not capture the potential for patients to transition to an improved health state. The limitations associated with the application of the beta distribution to utility values, and the normal distribution to disutility values, in the PSA, as described in 5.6.5, are also applicable to this study.

The lack of direct evidence on PFS for R-GDP (CORAL Extension 1) was a notable limitation and adds further uncertainty to results. Reassuringly, however, PFS was not a driver of cost effectiveness in the model.

In line with published cost-utility models, the models did not account for additional infrastructure and staff requirements (333, 515). As St James's Tertiary Hospital runs the National Adult Allogeneic Transplant Programme, many of the required lab facilities are expected to already be in place. Additionally, lead roles, such as the Clinical Lead for the CAR T-cell therapy programme have been appointed internally. Costs associated with the training of staff were based on assumptions; it is difficult to conclude whether these costs are realistic. As highlighted in 5.6.5, as implementation costs were excluded, the analysis more closely reflects the marginal costs of treating patients within an existing CAR T-cell therapy service, as opposed to the establishment of a new service.

Due to the paucity of data, potential late adverse events such as secondary malignancies were not accounted for. Notably, costs associated with adverse events were not main drivers in the model. It was assumed that costs obtained from the HPO DRG list were a suitable proxy for administration and hospitalisation (350). However, patients treated with CD19 CAR T-cell therapy may be subject to a greater degree of monitoring. Thus, the potential for underestimation of hospitalisation costs cannot be ruled out. Additionally,

costs and disutility associated with adverse events may be underestimated in the models, as it was assumed that patients only experience one incidence of an adverse event (as described in 5.6.5).

A limitation of the EVPI analysis is the arbitrary choice of a technology time horizon of 10 years. This assumption was aligned with those employed elsewhere (12, 378). The results of EVPI and EVPPI analyses are dependent on the appropriate model structure, evidence synthesis, and characterisation of other uncertainties (378).

9.7 Conclusion

At prices-to-wholesaler, neither tisagenlecleucel nor axicabtagene ciloleucel were found to be cost effective, versus R-GDP, at a willingness-to-pay threshold of €45,000 per QALY. The clinical evidence supporting the models was highly uncertain. The models were highly sensitive to assumptions regarding long-term survival, creating challenges for decision-makers in the interpretation of results. When the prices of tisagenlecleucel and axicabtagene ciloleucel were reduced to generate an ICER of €45,000 per QALY (versus R-GDP), EVPPI analysis indicated that collecting additional information on parameters associated with survival analysis and utility values may be of value in reducing uncertainty. However, such analyses do not capture uncertainty associated with the naïve nature of the ITCs.

Chapter 10 Budget Impact of CD19 CAR T-Cell Therapies in the Irish Healthcare Setting

10.1 Introduction

10.1.1 Chapter Aim

The cost effectiveness of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL have been established in Chapter 5 and Chapter 9, respectively. However, due to their high upfront costs, the affordability of CD19 CAR T-cell therapies is a key challenge (516). A budget impact analysis was conducted to assess the affordability CD19 CAR T-cell therapies in the Irish healthcare setting.

10.2 Method

The budget impact analysis was conducted in line with National Guidelines for the Budget Impact Analysis of Health Technologies in Ireland (517). The perspective of the HSE in Ireland was adopted, analysing the impact on the HSE drugs budget. Only direct drug costs were included in the base case analysis. The time horizon was five years (30).

10.2.1 Gross Drug Budget Impact

The gross drug budget impact was defined as the acquisition costs of the new drug (herein 'intervention'). This was based on the eligible population, market share, and proposed drug costs over five years (518). These are described in later sections.

10.2.2 Net Drug Budget Impact

The net drug budget impact was defined as the acquisition costs of the intervention minus the costs from the displacement of routine care (herein 'comparator') (518).

10.2.3 Intervention

Assumptions regarding dosing were aligned with the respective SPCs (3, 4) and pivotal trials (215, 392, 395). Details regarding dosing are provided in Chapter 5 and Chapter 9, as per the cost-utility models. Tisagenlecleucel and axicabtagene ciloleucel were modelled as single-dose infusions.

10.2.3.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

Tisagenlecleucel was the defined intervention in the analysis for R/R ALL.

10.2.3.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

Tisagenlecleucel and axicabtagene ciloleucel were defined as the interventions in the analysis for R/R DLBCL. Simultaneous reimbursement of tisagenlecleucel and axicabtagene ciloleucel was assumed.

10.2.4 Comparator

The comparators were defined in line with routine care in Ireland and thus, the respective cost-utility analyses. Dosing, presented in Chapter 5, Chapter 9 and Appendix I, was based on the NCCP Chemotherapy Regimens (340, 504, 519) and clinical opinion.

10.2.4.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

Blinatumomab was the primary comparator in the analysis for R/R ALL. FLA(G)-IDA was also included, as a small proportion of patients are expected to receive treatment with this therapy[21]. Patients were assumed to receive up to two cycles of blinatumomab (87, 340) and one cycle of FLA(G)-IDA (519).

10.2.4.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

R-GDP was the comparator in the analysis for R/R DLBCL. As highlighted, R-GDP was assumed to reflect all potential salvage chemotherapy regimens. Patients were assumed to receive three cycles of R-GDP.

As alloSCT/HSCT is not considered in the HSE drugs budget, two separate net budget impact analyses are presented (for R/R ALL and R/R DLBCL); one considering drug acquisition costs only (excluding alloSCT/HSCT), and an exploratory analysis including the cost of alloSCT/HSCT. Of note, the cost of alloSCT, obtained from the HPO DRG List, also accounts for hospitalisation and monitoring costs post-alloSCT (350).

10.2.5 Eligible Population

The eligible populations were defined in line with the EMA licensed indications of tisagenlecleucel (3) and axicabtagene ciloleucel (4), and as per the respective cost-utility models.

[21]Written correspondence with one consultant haematologist in Ireland.

10.2.5.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

The population was paediatric and young adult patients (up to 25 years of age) with B-cell ALL that is refractory, in relapse post-transplant, or in second or later relapse (3). In the absence of granular data pertinent to the Irish setting, the number patients eligible for treatment was obtained from one clinical expert in Ireland. Accounting for patients aged up to 20 years, 5 patients are expected to be eligible for treatment each year[22]. It was assumed that 1 patient, aged between 20 and 25 years, will also be eligible for treatment each year. The number of eligible patients was estimated to remain stable over the five-year period. The NCRI validated this assumption indicating that although the overall population of patients with newly diagnosed ALL is increasing, the population aged 0 to 4 years (approximately 50% of the incidence of ALL) has decreased, while the population of older age groups has increased. This has caused a stabilisation in the incidence[23].

10.2.5.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

The population was adult patients with R/R DLBCL, after two or more lines of systemic therapy (3, 4). A top-down approach was used to derive eligible population estimates. The NCRI reported that 814 new cases of NHL were diagnosed per year between 2015 and 2017, and an incidence growth rate of 1.7% in men and 1.5% in women (520). Based on these data, the number of patients diagnosed with NHL in 2022 is expected to be 882. No prevalent population was considered, as treatment duration is less than one year. As reported by the NCRI, 31% of these patients are expected to have DLBCL (521).

In the absence of additional Irish-specific data, data from the literature were used. An observational study using the Swedish Lymphoma Registry (patients diagnosed between 2007 and 2013), found that 83% of patients diagnosed with DLBCL received first-line therapy (522). This is in line with data from the Haematological Malignancy Research Network UK, which indicated that 80% of newly diagnosed patients receive first-line therapy (523). The Haematological Malignancy Research Network is a collaborative project between the University of York and 14 NHS hospitals. It collects data on incidence

[22] Oral correspondence with one consultant haematologist in Ireland.
[23] Written correspondence with the National Cancer Registry Ireland.

and survival associated with haematological malignancies, with the aim of providing robust data to inform clinical practice and research (524). The proportion of patients receiving second-line therapy, identified in the literature, ranged from 11% (115) to 33% (112).

Limited data were identified in relation to the proportion of patients who are refractory to or relapse following second-line therapy. An analysis of the Veterans Affairs Cancer Registry System and electronic healthcare records in the US, of patients diagnosed with DLBCL between January 2003 and December 2016, found that of 270 patients with relapsed or refractory DLBCL who received second-line therapy, 45% received third-line therapy (116, 117). The generalisability of these data may be limited. However, these findings are in line with the literature elsewhere, which indicate that between 45% and 60% of patients relapse following or are refractory to second-line therapy (121, 525). Radford et al. indicated that 71% of patients in a single centre in England proceeded to third-line therapy (118).

The eligible population estimates are subject to much uncertainty. The epidemiology of R/R DLBCL is poorly defined. Assumptions underpinning the eligible population estimates are presented in Table 61. The proportion of patients receiving second-line therapy was assumed to be 33%, derived from the Swedish Lymphoma Registry (n=3,905), which collected data from January 2007 to December 2013. Median follow up was 4 years (range: 0.0 to 9.0) (522). This registry captures approximately 95% of all lymphoma cases diagnosed in Sweden (526). This estimate was chosen as it is aligned with the proportion of patients expected to be refractory to or relapse following first-line therapy (approximately 30% to 40%) (527). In light of the limited data identified, it was assumed that 45% of patients who receive second-line therapy proceed to third-line therapy, derived from the analysis of the Veterans Affairs Cancer Registry System and electronic healthcare records in the US (described above) (116, 117). The higher proportion of patients receiving third-line therapy compared to second-line seems counterintuitive; however, this trend has been noted elsewhere (523, 528, 529).

Table 61 Eligible population estimates of patients with relapsed/refractory diffuse large B-cell lymphoma

Year	2022	2023	2024	2025	2026
Incident Population Non-Hodgkin's Lymphoma (520)	882	896	910	925	940
Diffuse Large B-Cell Lymphoma (31%) (521)	273	278	282	287	291
First-Line Therapy (83%) (522)	227	231	234	238	242
Second-Line Therapy (33%) (522)	75	76	77	79	80
Third-Line Therapy (45%): **Eligible Population** (116, 117)	34	34	35	35	36

10.2.6 Market Share

10.2.6.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

In a world without tisagenlecleucel (i.e. current routine care), it was assumed that five of the six eligible patients receive treatment with blinatumomab in the first two years, increasing to all six patients from year three onwards. One patient was assumed to receive treatment with FLA(G)-IDA in year one and year two. This was based on clinical opinion, which indicated that blinatumomab is increasingly used to treat this population[24].

It is expected that all eligible patients will be considered for treatment with tisagenlecleucel. However, some may not proceed to infusion due to manufacturing failure, adverse events, and death prior to infusion (215). With the introduction of tisagenlecleucel, it was assumed that four of the six eligible patients per year would receive treatment with tisagenlecleucel in the first two years, increasing to five patients per year from year three onwards. This is in line with ELIANA and ENSIGN (and thus, the cost-utility analysis), whereby 83% of patients proceeded to infusion (207, 215). Additionally, real-world evidence from the UK National CAR T-Cell Panel (described in

[24] Written correspondence with one consultant haematologist in Ireland.

5.6.5), indicated that of 60 patients deemed eligible for treatment with tisagenlecleucel in R/R ALL, approximately 81% proceeded to infusion (388). It was assumed that two patients would be treated with FLA(G)-IDA over the five-year period. The number of patients expected to receive treatment with each regimen is presented in Table 62.

Table 62 Number of patients with relapsed/refractory acute lymphoblastic leukaemia treated in a world with and without tisagenlecleucel

	2022	2023	2024	2025	2026	Total
World Without Tisagenlecleucel						
Blinatumomab	5	5	6	6	6	28
FLA(G)-IDA	1	1	0	0	0	2
World With Tisagenlecleucel						
Tisagenlecleucel	4	4	5	5	5	23
Blinatumomab	1	1	1	1	1	5
FLA(G)-IDA	1	1	0	0	0	2

FLA(G)-IDA: Fludarabine, cytarabine, idarubicin, granulocyte colony stimulating factor.

10.2.6.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

In a world without tisagenlecleucel and axicabtagene ciloleucel, all eligible patients were assumed to receive treatment with R-GDP (100% market share). In St James's Tertiary Teaching Hospital, the National CAR T-Cell Centre for adults in Ireland, two patients per month are expected to receive treatment with CD19 CAR T-cell therapy for R/R DLBCL[25]. In line with this, it was assumed that 24 patients are treated with either tisagenlecleucel or axicabtagene ciloleucel each year. Estimates are regardless of the proportion who do not proceed to infusion due to manufacturing failure, adverse events, or death prior to infusion. This is due to the high number of patients expected to be eligible for treatment. In the absence of data to suggest preference for one CD19 CAR T-cell therapy over another, a 50:50 market share distribution was assumed between them. The remaining eligible population are expected to receive R-GDP. The number of patients expected to receive treatment, in a world with and without CD19 CAR T-cell therapies, are presented in Table 63.

[25] Interview with Dr Larry Bacon, RTE Radio 1, 13 December 2021.

Table 63 Number of patients with relapsed/refractory diffuse large B-cell lymphoma treated in a world with and without tisagenlecleucel and axicabtagene ciloleucel

	2022	2023	2024	2025	2026	Total
World without Tisagenlecleucel and Axicabtagene Ciloleucel						
R-GDP	34	34	35	35	36	174
World with Tisagenlecleucel and Axicabtagene Ciloleucel						
R-GDP	10	10	11	11	12	54
Tisagenlecleucel	12	12	12	12	12	60
Axicabtagene ciloleucel	12	12	12	12	12	60

R-GDP: Rituximab, gemcitabine, dexamethasone, cisplatin.

10.2.7 Drug Costs

Drug costs were calculated as per the NCPE Guidelines for the Calculation of Drug Costs (349). VAT is considered in the budget impact analysis. Assumptions employed in the budget impact analysis were aligned with those used in the respective cost-utility models. Costs were not discounted (517).

10.2.7.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

Total drug acquisition costs associated with each treatment for R/R ALL, inclusive of VAT, are presented in Table 64. For FLA(G)-IDA, it was assumed that patients aged 18 years and older receive granulocyte colony stimulating factor (16% based on the proportion of patients aged 18 years and over in ELIANA and ENSIGN (143, 207)).

Table 64 Total drug acquisition costs per patient per treatment course employed in bespoke budget impact model for relapsed/refractory acute lymphoblastic leukaemia

Drug	Total Drug Acquisition Cost Per Patient Per Treatment Course* (€)
Tisagenlecleucel†	375,206.88
Blinatumomab (dosing based on body surface area) ‡	110,926.90
Blinatumomab (fixed-dosing regimen)§	155,897.26
FLA-IDA (fludarabine, cytarabine, idarubicin)|	1,871.25
FLAG-IDA (fludarabine, cytarabine, idarubicin, granulocyte colony stimulating factor)|¶	2,692.26

*Including relevant fees, rebates, and VAT.
†Administered as a once-off single-dose (3).
‡Patients weighing <45kg, dosing of blinatumomab is based on body surface area (86, 340). Assumed 50% receive this dosing; 100% receive one cycle, and 33% receive a second cycle (87).
§Patients weighing ≥45kg, a fixed-dosing regimen of blinatumomab applies (86, 340). Assumed 50% receive this dosing; 100% receive one cycle, and 33% receive a second cycle (87).
|All patients receive one cycle (519) and 100% dose intensity.
¶Addition of granulocyte colony stimulating factor for patients aged 18 years and older. Assumed 16% of patients receive FLAG-IDA, based on proportion aged 18 years and over in ELIANA and ENSIGN (143, 207).

10.2.7.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

The total drug acquisition costs associated with each treatment for R/R DLBCL, inclusive of VAT, are presented in Table 65.

Table 65 Total drug acquisition costs per patient per treatment course employed in bespoke budget impact model for relapsed/refractory diffuse large B-cell lymphoma

Drug	Total Drug Acquisition Cost Per Patient Per Treatment Course* (€)
Tisagenlecleucel†	375,206.88
Axicabtagene Ciloleucel†	384,225.00
R-GDP (rituximab, gemcitabine, dexamethasone, cisplatin)‡	5,675.55

*Including relevant fees, rebates, and VAT.
†Administered as a once-off single-dose (3, 4).
‡Assumed treatment duration of three cycles for all patients and 100% dose intensity (504).

10.2.8 Additional Costs and Offsets

Aside from the exploratory analysis examining the impact of inclusion of alloSCT/HSCT on the net budget impact (described in 10.2.4), no additional non-drug costs were considered. This is because the perspective was that of the HSE Drugs Budget.

10.3 Results

10.3.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

10.3.1.1 Gross Drug Budget Impact

The estimated gross drug budget impact of tisagenlecleucel for R/R ALL is presented in Table 66. The cumulative 5-year gross drug budget impact was €8,629,758 (including VAT) and €6,940,529 (excluding VAT).

Table 66 Gross drug budget impact of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

	2022	2023	2024	2025	2026	Total
Number of patients treated with tisagenlecleucel	4	4	5	5	5	23
Tisagenlecleucel gross drug budget impact* (€)	1,500,828	1,500,828	1,876,034	1,876,034	1,876,034	8,629,758

*Including VAT, all relevant fees, and rebate.

10.3.1.2 Net Drug Budget Impact

The net drug budget impact, accounting only for drug acquisition costs (i.e. not alloSCT), is presented in Table 67. The number of patients receiving treatment in a world with and

without tisagenlecleucel is presented in Table 62. The cumulative 5-year net drug budget impact of tisagenlecleucel for R/R ALL was €6,670,070 (including VAT) and €5,364,439 (excluding VAT).

Table 67 Net drug budget impact of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

	2022	2023	2024	2025	2026	Total
World without tisagenlecleucel budget impact* (€)	428,022	428,022	511,223	511,223	511,223	2,389,714
World with tisagenlecleucel budget impact* (€)	1,588,034	1,588,034	1,961,238	1,961,238	1,961,238	9,059,783
Tisagenlecleucel net drug budget impact* (€)	1,160,012	1,160,012	1,450,015	1,450,015	1,450,015	6,670,070

*Including VAT, all relevant fees, and rebate.

10.3.1.2.1 Net Budget Impact Including AlloSCT

An exploratory analysis, examining the impact of inclusion of alloSCT on the net drug budget impact, was conducted. The results are presented in Table 68. It was assumed that 49% of patients who receive blinatumomab or FLA(G)-IDA, proceed to alloSCT, based on judgements derived from the elicitation exercise (Chapter 3). The cumulative 5-year net budget impact of tisagenlecleucel for R/R ALL, accounting for the cost of alloSCT, was €4,385,663 (including VAT) and €3,080,032 (excluding VAT).

Table 68 Net budget impact (including cost of alloSCT) of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

Year	2022	2023	2024	2025	2026	Total
World without tisagenlecleucel budget impact* (€)	1,023,954	1,023,954	1,107,155	1,107,155	1,107,155	5,369,374
World with tisagenlecleucel budget impact* (€)	1,786,678	1,786,678	2,060,560	2,060,560	2,060,560	9,755,037
Tisagenlecleucel net budget impact including alloSCT* (€)	762,724	762,724	953,405	953,405	953,405	4,385,663

AlloSCT: Allogeneic stem cell transplant.
*Including VAT, all relevant fees, and rebate.

10.3.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

10.3.2.1 Gross Drug Budget Impact

The estimated gross drug budget impact of tisagenlecleucel and axicabtagene ciloleucel, assuming an equal distribution of market share between the two, is presented in Table

69. The cumulative 5-year gross drug budget impact of CD19 CAR T-cell therapies for R/R DLBCL was €45,565,913 (including VAT) and €36,646,628 (excluding VAT).

Table 69 Gross drug budget impact of tisagenlecleucel and axicabtagene ciloleucel for relapsed/refractory diffuse large B-cell lymphoma

	2022	2023	2024	2025	2026	Total
Number of patients treated with either tisagenlecleucel or axicabtagene ciloleucel*	24	24	24	24	24	120
Tisagenlecleucel gross drug budget impact† (€)	4,502,483	4,502,483	4,502,483	4,502,483	4,502,483	22,512,413
Axicabtagene ciloleucel gross drug budget impact† (€)	4,610,700	4,610,700	4,610,700	4,610,700	4,610,700	23,053,500
Total CD19 CAR T-cell therapy gross drug budget impact† (€)	9,113,183	9,113,183	9,113,183	9,113,183	9,113,183	45,565,913

*50:50 market share assumed between tisagenlecleucel and axicabtagene ciloleucel (i.e. 12 patients treated with tisagenlecleucel and 12 treated with axicabtagene ciloleucel in 2022, etc.).
†Including VAT, all relevant fees, and rebate.

10.3.2.2 Net Drug Budget Impact

The net drug budget impact, accounting only for drug acquisition costs (i.e. not HSCT), is presented in Table 70. The cumulative 5-year net drug budget impact of CD19 CAR T-cell therapies for R/R DLBCL was €44,884,847 (including VAT) and €36,091,881 (excluding VAT).

Table 70 Net drug budget impact of tisagenlecleucel and axicabtagene ciloleucel for relapsed/refractory diffuse large B-cell lymphoma

	2022	2023	2024	2025	2026	Total
World without CD19 CAR T-cell therapy budget impact* (€)	192,969	192,969	198,644	198,644	204,320	987,546
World with CD19 CAR T-cell therapy budget impact* (€)	9,169,938	9,169,938	9,175,614	9,175,614	9,181,289	45,872,392
Total CD19 CAR T-cell therapy net drug budget impact* (€)	8,976,969	8,976,969	8,976,969	8,976,969	8,976,969	44,884,847

*Including VAT, all relevant fees, and rebate.

10.3.2.2.1 Net Budget Impact Including HSCT

Results of the exploratory analysis, examining the impact of inclusion of the cost of HSCT for patients receiving R-GDP, are presented in Table 71. It was assumed that 15% of patients receiving R-GDP proceed to HSCT. Of note, the cost employed for HSCT pertains specifically to alloSCT. The cumulative 5-year net budget impact of CD19 CAR T-cell

therapies for R/R DLBCL, accounting for the cost of HSCT, was €42,791,033 (including VAT) and €33,998,067 (excluding VAT).

Table 71 Net budget impact (including cost of HSCT) of tisagenlecleucel and axicabtagene ciloleucel for relapsed/refractory diffuse large B-cell lymphoma

	2022	2023	2024	2025	2026	Total
World without CD19 CAR T-cell therapy budget impact* (€)	786,216	786,216	809,340	809,340	832,464	4,023,576
World with CD19 CAR T-cell therapy budget impact* (€)	9,344,423	9,344,423	9,367,547	9,367,547	9,390,671	46,814,609
Total CD19 CAR T-cell therapy net drug budget impact incl. HSCT* (€)	8,558,207	8,558,207	8,558,207	8,558,207	8,558,207	42,791,033

HSCT: Haematopoietic stem cell transplant.
*Including VAT, all relevant fees, and rebate.

10.4 Discussion

10.4.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

Considering the low number of patients with R/R ALL, expected to receive treatment with tisagenlecleucel between 2022 and 2026, the gross drug budget impact is high. Both comparators are considerably less costly than tisagenlecleucel; the net drug budget impact is therefore, also high.

Real-world evidence from the UK National CAR T-cell Panel (described in 5.6.5) supported the estimated proportion of patients expected to proceed to infusion with tisagenlecleucel. These data indicated that of 60 patients deemed eligible for treatment with tisagenlecleucel, 57 had their T-cells harvested and 49 proceeded to infusion. Three patients did not have their T-cells harvested and two were not infused due to progressed disease; one patient did not proceed to infusion due to manufacturing failure. These data suggest that a key challenge in ensuring patients proceed to infusion is stabilisation of disease. Focused research on the optimal bridging chemotherapy regimens (530, 531), combined with increasing clinical experience, may lead to improved outcomes in patients. The proportion of patients proceeding to infusion with tisagenlecleucel may therefore, increase.

The safety and efficacy of tisagenlecleucel in paediatric and young adult patients with high-risk ALL in first relapse is being evaluated in a phase II, multicentre study

(NCT02435849). Estimated completion date is November 2022 (532). The reimbursement of tisagenlecleucel for an earlier line of therapy is likely to reduce the gross drug budget impact of tisagenlecleucel at the line of therapy being evaluated here. However, the overall gross budget impact (i.e. of treating eligible patients at both lines of therapy) is likely to increase, considering the potential for a higher number of eligible patients at earlier lines of therapy.

The gross drug budget impact estimated here was higher than that presented in the NCPE Technical Summary of tisagenlecleucel (€8.6 million versus €5.5 million). Of note, VAT was not accounted for in the budget impact presented in the NCPE Technical Summary. Guidance regarding the application of VAT to CAR T-cell therapy products was updated after completion of the NCPE HTA appraisal of tisagenlecleucel. This is also applicable to tisagenlecleucel for R/R DLBCL, discussed in 10.4.2. The inclusion of VAT, in estimates generated as part of this research, partly contributes to differences in the budget impact estimates. The remaining difference, for R/R ALL, is driven by the higher proportion of patients estimated to receive treatment with tisagenlecleucel here (5-year cumulative treated patients: 23 versus 19) (88). Due to limited reporting of data in the NCPE Technical Summary (88), the reasons for this discrepancy cannot be evaluated. However, this may be due to differences between estimates of eligible population. The market share estimates are aligned between the NCPE Technical Summary and this research (88). A net budget impact, accounting for procedure costs (leukapheresis, lab management, lymphodepleting chemotherapy and bridging chemotherapy) associated with tisagenlecleucel, was also presented in the NCPE Technical Summary. The cumulative 5-year net budget impact of tisagenlecleucel, including procedure costs, was estimated to be €5.6 million (cost year 2019) (88).

10.4.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

This budget impact analysis indicated that the cost of treating patients with R/R DLBCL with either tisagenlecleucel or axicabtagene ciloleucel (2022 to 2026) is considerable. The net drug budget impact was closely aligned with that of the gross drug budget impact. The number of patients with R/R DLBCL eligible for treatment is uncertain. It is unclear if data from other jurisdictions, used to inform this model, are reflective of the Irish setting.

It is reassuring, however, that the proportion estimates used in this budget impact model were aligned with other estimates identified in the literature. The Swedish Lymphoma Registry, used to inform the proportion of patients expected to receive first- and second-line therapy, appears to be a rich source of data (522).

It is unclear if the proportion of patients treated will remain stable over the five-year period, as assumed in this analysis. This uncertainty is further compounded by the assumption that tisagenlecleucel and axicabtagene ciloleucel will have an equal distribution of the number of patients expected to receive treatment. The lack of relative efficacy data of tisagenlecleucel versus axicabtagene ciloleucel makes the CD19 CAR T-cell therapy, likely to be favoured amongst clinicians, difficult to determine. Due to the paucity of data, this was a necessary assumption. The lack of relative efficacy data may make it difficult for one therapy to seek a price premium over the other. This is an important consideration in the reimbursement of these therapies. A performance-linked reimbursement agreement may be one approach to establishing fairer reimbursement of these therapies. Tisagenlecleucel was approved for reimbursement in July 2021. Axicabtagene ciloleucel was approved for reimbursement in January 2022. It is therefore, likely that more patients will receive treatment with tisagenlecleucel than with axicabtagene ciloleucel in the first year. However, the difference in price between tisagenlecleucel and axicabtagene ciloleucel is negligible relative to the total budget impact. This assumption is unlikely to have an appreciable impact on the gross and net drug budget impact estimates.

Both tisagenlecleucel and axicabtagene ciloleucel are being investigated in earlier lines of therapy. The BELINDA (533) and ZUMA-7 (534) trials are phase III, RCTs examining the efficacy of tisagenlecleucel versus standard-of-care, and axicabtagene ciloleucel versus standard-of-care, respectively. The estimated completion date of BELINDA is October 2026 (533), while that of ZUMA-7 is January 2023 (534). The impact of reimbursement of tisagenlecleucel or axicabtagene ciloleucel at an earlier line of therapy was not considered in the budget impact analysis. Additionally, the impact of the reimbursement of other CD19 CAR T-cell therapies (such as lisocabtagene maraleucel) was not considered. These scenarios would likely reduce the number of patients receiving

treatment with either tisagenlecleucel or axicabtagene ciloleucel at the line of therapy under investigation in this analysis.

The gross and net drug budget impact estimates derived as part of this research differ from those presented in the NCPE Technical Summaries of tisagenlecleucel and axicabtagene ciloleucel. This is mainly due to the higher number of patients estimated, as part of this research, to receive CD19 CAR T-cell therapy. The higher estimates, in this research (5-year cumulative treated patients: 120 versus 80), are due to the availability of updated data. This research also assumes simultaneous reimbursement of tisagenlecleucel and axicabtagene ciloleucel. The NCPE Technical Summaries of tisagenlecleucel and axicabtagene ciloleucel assumed that only the CD19 CAR T-cell therapy under assessment would be reimbursed (9, 10). The NCPE Technical Summary of tisagenlecleucel presented a cumulative 5-year gross drug budget impact (n=80) of €24.1 million (cost year 2019) (10). As highlighted (10.4.1), VAT was not included in this analysis. The NCPE Technical Summary of axicabtagene ciloleucel presented a cumulative 5-year gross drug budget impact (n=80) of €30.7 million (cost year 2019) (9). The net drug budget impact estimated in this research is not directly comparable with those presented in the NCPE Technical Summaries of tisagenlecleucel and axicabtagene ciloleucel. The NCPE Technical Summaries present net budget impact estimates inclusive of procedure costs. The cumulative 5-year net budget impact of tisagenlecleucel was estimated to be €24.2 million when procedure costs were accounted for (10). The cumulative 5-year net budget impact of axicabtagene ciloleucel was estimated to be €10.6 million when procedure costs were accounted for. Notably, this is considerably lower than the gross drug budget impact. The reason for this is unclear; however, it may be due to differences in the number of patients expected to receive treatment with routine care (9). Neither NCPE Technical Summary documents specified the costs included in procedure costs.

10.4.3 Limitations

Tisagenlecleucel, for R/R ALL and R/R DLBCL, was subject to full HTA appraisal by the NCPE and subsequently approved for reimbursement by the HSE following confidential price negotiations (July 2021) (10, 88). Axicabtagene ciloleucel, R/R DLBCL, was subject to full HTA appraisal by the NCPE and approved for reimbursement by the HSE following

confidential price negotiations (January 2022) (131). Blinatumomab (R/R ALL) was subject to Rapid Review assessment by the NCPE, and approved for reimbursement by the HSE following confidential price negotiations (May 2019) (535). As such, the gross and net drug budget impact estimates presented here are not a true reflection of the costs to the HSE.

Caution is warranted in interpretation of the net budget impact including alloSCT/HSCT estimates. Hospitalisation and monitoring costs associated with alloSCT are included in these estimates, but are not included for drug therapies. An accurate estimate of the overall budget impact would require the inclusion of associated hospitalisation and monitoring costs of all interventions and comparators (30).

A pragmatic approach was adopted for this research in that the budget impact models did not consider associated procedure costs (e.g. leukapheresis, cryopreservation, etc.). Concomitant therapies (such as bridging chemotherapy) were not accounted for. Future research is important to consider the cost of implementing a CAR T-cell therapy service from the broader healthcare perspective.

10.5 Conclusion

The gross drug budget impact of tisagenlecleucel for R/R ALL is high, despite the low number of patients expected to receive treatment. Due to the considerably lower costs associated with blinatumomab and FLA(G)-IDA, the net drug budget impact is high. The gross and net drug budget impact of treating patients with R/R DLBCL with either tisagenlecleucel or axicabtagene ciloleucel is high. The cost effectiveness of these therapies should be interpreted in the context of the large budget impact. Affordability of these therapies is a key challenge. Estimates of gross and net drug budget impact for both R/R ALL and R/R DLBCL are uncertain, due to confidential price reductions on tisagenlecleucel, axicabtagene ciloleucel, and blinatumomab. The number of patients expected to receive treatment with CD19 CAR T-cell therapies is also uncertain, and may potentially increase due to updated research and clinical experience. The potential for increased budget impact should be considered. In order to assess the overall affordability of CD19 CAR T-cell therapies, an overall net budget impact, encompassing costs such as procedure and adverse event costs, is warranted.

Chapter 11 Impact of Performance-Linked Reimbursement Agreements on the Cost Effectiveness and Budget Impact of CD19 CAR T-Cell Therapies

11.1 The Need for Risk-Sharing Agreements

Reimbursement for repeat-dose treatments is usually based on payment per treatment dose received. Uncertainty in clinical evidence or cost effectiveness is usually addressed by simple price reductions offered by the Applicant to the HSE. Payment generally continues for the duration of treatment benefit (in the absence of unacceptable toxicities), or until the patient has finished the course of treatment, and ceases upon treatment discontinuation (32). For once-off treatments, such as CD19 CAR T-cell therapies, discontinuing treatment (and thus, payment) due to non-response or disease progression is not possible. The consequences of decision uncertainty for a once-off treatment versus those for a repeat-dose treatment, that produces the same expected health outcomes, have been compared in the literature. From a payer's perspective, irreversibility associated with the once-off treatment drives greater financial risk (536).

Affordability is a key challenge in the reimbursement of CD19 CAR T-cell therapies (516). Their high cost may result in affordability issues for the payer, even when they are deemed cost effective. The once-off payment for these once-off therapies results in irrecoverable costs in instances where they do not generate the desired outcome. This poses a threat to the sustainability of the healthcare budget. These challenges have prompted interest in performance-based risk-sharing agreements.

11.1.1 Types of Risk-Sharing Agreements

Payer-Applicant risk-sharing agreements encompass either financial or performance-based agreements. Financial agreements generally aim to address uncertainty in affordability. They are not linked to performance of the therapy; examples include budget caps, utilisation caps, and simple price reductions (537, 538). Implementation of such agreements is relatively straightforward and associated with low cost (12). These agreements are commonly used in the Irish reimbursement setting (32). However, they do not address the affordability and irreversibility issues associated with CD19 CAR T-cell therapy.

In contrast, performance-based risk-sharing agreements require the collection of outcomes associated with the therapy. This type of agreement is more appropriate when

there is considerable uncertainty associated with the efficacy or safety of a therapy (539). Data collection is population-based or patient-level (538). Payment to the Applicant is contingent on the collection of outcomes data and/or achievement of pre-agreed outcomes (537). These agreements can be further defined as two distinct subgroups, based on whether the goal is to provide reimbursement while evidence is generated ('coverage with evidence development'), or whether the goal is to manage utilisation to achieve cost-effective use of the therapy in clinical practice ('performance-linked reimbursement') (538). An example of coverage with evidence development is the funding of axicabtagene ciloleucel under the Cancer Drugs Fund UK (CDF). Axicabtagene ciloleucel is funded by the CDF under the condition that further evidence is collected to reduce uncertainty around survival estimates. At the end of the agreement (February 2022), axicabtagene ciloleucel is reappraised and if there is insufficient evidence, or the therapy is not deemed to be clinically or cost effective, it may be removed from the CDF and no longer available on the NHS (537, 540). Coverage with evidence development is usually implemented at population-level. Such agreements are not considered in this analysis, as they are not straightforward to implement and may not address issues with affordability (538, 541). The risk associated with irreversibility of CD19 CAR T-cell therapies has been proposed to be large and so, widespread adoption while further evidence is generated (under 'standard' reimbursement mechanisms) may not be optimal (536).

Patient-level performance-linked reimbursement agreements link payment to the achievement of pre-agreed outcomes in each patient (537). Due to the small numbers of patients expected to receive treatment with CD19 CAR T-cell therapies in Ireland, monitoring patient-level outcomes is not expected to be too burdensome. Payment mechanisms for performance-linked reimbursement agreements can take different forms including rebate paid to the payer if the desired outcome is not achieved, higher price paid for better outcomes, and annuity payments (542). Annuity payments, spread over time (generally once per year), linked to the performance of the therapy are the preferred approach here. This approach can address affordability issues associated with high, upfront costs, while also dealing with uncertainty in the long-term benefit (543). Rebate paid back to the payer and higher price for better outcomes do not address challenges with the high, upfront cost. Performance-linked risk-sharing agreements have

been found to increase the probability of the therapy being cost effective, without the risk of eroding product value (12).

11.1.2 Barriers to Implementation

Adoption of performance-based risk-sharing agreements in Ireland has been limited. The need for the payer and Applicant to agree on financial terms (including payment amount per instalment and duration of payment spread) is a frequently reported barrier to implementation. Additionally, no formal method exists to determine the optimal duration of payment spread (543). Payers in Ireland are experienced in negotiating price reductions with Applicants. Additional components of negotiation may not pose an excessive burden to either party. However, they may warrant an extension of the negotiation period. Close collaboration between all stakeholders is warranted, to ensure timely patient access.

The agreed outcome should be easily measured in the short- to medium-term and clinically relevant to all stakeholders (538). Outcomes of potentially curative therapies may not present in the short- to medium-term, requiring the use of surrogate outcomes. The lack of validated surrogate outcomes may be a barrier to implementation (543). For CD19 CAR T-cell therapies, OS would seem the most relevant outcome to the payer. OS has also been reported to be the outcome of greatest importance to patients with cancer and their carers (n=164) (544). This outcome is easily measured and not subject to measurement bias (538). However, as described in 11.4, the most appropriate period over which to capture OS data is uncertain. Conditions also need to be agreed regarding reimbursement in instances where the patient dies from non-disease related events.

Performance-based risk-sharing agreements pose additional administrative burden (543-546). Collection of GDPR-compliant data requires appropriate infrastructure and operational frameworks (547). Appropriately trained data managers are necessary (546). Registries have been recommended as the most appropriate approach to data collection (539). EMA conditional authorisation of tisagenlecleucel and axicabtagene ciloleucel specifies that patients should be followed for 15 years post-treatment, to capture long-term safety and efficacy data (143, 144). The cellular therapy module of the EBMT

registry, which details information such as OS, relapse/progression, and adverse events, was approved (by the EMA) for this purpose (7, 548). The EBMT registry aims to capture high-quality data on patients in Europe who receive HSCT. These data may be used for research, auditing, and accreditation of transplant centres (549). EBMT, in collaboration with EUnetHTA, are exploring the value of the cellular therapy module to HTA agencies. Members of the EBMT have stated that the registry would allow for novel reimbursement strategies (7). This would prevent duplication of data collection, providing that such data are permitted to be used for reimbursement. Notably, economic and HRQOL data are not captured by the EBMT registry (547). While these data would be of value, the collection of such data may add complexity to the process.

The well-established NCRI collects data regarding all cancer diagnoses in Ireland. Trends and outcomes in different cancer types are analysed using these data (550). Although the NCRI does not capture data on R/R disease, or assess outcomes linked to treatment, the existing infrastructure could be developed to capture such data.

The use of well-established registries may mitigate against the risk of poor quality data, employing quality assurance tools such as built-in quality triggers, and procedures to correct errors in data entry (7, 551). Missing data and patients lost to follow up are also a concern; the conditions of such occurrences should be negotiated prior to implementation of an agreement. It is expected that the longer the duration of the agreement, the higher the likelihood of losing patients to follow up.

The trend of accelerated approval based on limited evidence means there is likely to be an increasing need for performance-based risk-sharing agreements (255). It may not be feasible to have an individual registry for each disease; a national registry may be a less complex approach. A national registry, independent of disease type, has been in use in Italy since 2005. Established by the Italian Medicines Agency (AIFA), it serves regulatory and reimbursement purposes, facilitating analysis of consumption data and financial agreements (552). Since the inception of this registry, performance-based risk-sharing agreements have been prolific in Italy (297, 552). The estimated cost of managing the Italian national registry is approximately €1.0 million per year (553). In cases where performance-based risk-sharing agreements are in place, the Applicant pays

approximately €30,000 every 3 years, with a €10,000 maintenance fee from year 4 (554). In 2013, the estimated theoretical pay-back to payers, as a result of risk-sharing agreements in Italy (n=29), was €46.3 million (555).

Existing administrative systems are essential for efficient implementation of performance-based risk-sharing agreements (556). However, administrative systems in the Irish healthcare setting are fragmented and lag in terms of digitisation. The initial set-up costs associated with implementation of a performance-based risk-sharing agreement are a necessary drawback. However, it could be argued that the additional funding required to implement such an agreement will be recouped in cost savings generated by the agreement.

Solutions exist to many of the proposed implementation barriers. Experience, detailed in the literature, suggests that the key to successful performance-based risk-sharing agreements is simplicity (544). Of note, VOI analyses (population EVPI and EVPPI, presented in Chapter 5 and Chapter 9) indicated that further research to reduce decision uncertainty is unlikely to be cost effective from the payer's perspective. However, patient-level monitoring is conducted in clinical practice and thus, requires no further trial or registry establishment. Data collected in clinical practice could be provided to the payer for reimbursement purposes, once patient consent is provided. Appropriately trained data analysts would be required.

11.1.3 The Network of Competent Authorities on Pricing and Reimbursement (NCAPR)

The Network of Competent Authorities on Pricing and Reimbursement (NCAPR), a collaboration within the European Union, is tasked with identifying common challenges in drug pricing and reimbursement. An NCAPR workshop was held in June 2021; this aimed to develop a common minimal data set to aid CAR T-cell therapy pricing and reimbursement decisions. The workshop was attended by stakeholders from the European Commission, the EMA, European Union-level representatives of pricing and reimbursement authorities, HTA authorities, regulatory authorities, and public healthcare payers.

The common data elements, proposed by the EMA for regulatory evaluation of CAR T-cell therapy, were initially discussed (557). Participants ultimately concluded that three core data elements were necessary to support pricing and reimbursement decisions. These were data pertaining to OS, drug-related adverse events of grade 3-4 severity (including the treatments required), and HRQOL (collected using a standardised tool, preferably the ED-5D). Use of a common minimal data set, across Europe, may facilitate data sharing and pooling, efficient reimbursement negotiations and increase transparency (558). However, it may be challenging to agree on a data set that meets all stakeholder requirements.

11.1.4 Performance-Based Risk-Sharing Agreements in Other European Countries

A number of performance-based risk-sharing agreements have been implemented for the reimbursement of CD19 CAR T-cell therapies across Europe (297, 554). Coverage with evidence development agreements are in place in Belgium, England, and France (554). Patient-level performance-linked reimbursement agreements are in place in Italy and Spain. The duration of agreements range from 12 months (Italy) to 4 years and 7 months (England) (554, 559).

Data collection for coverage with evidence development (Belgium, England, France) is by means of ongoing clinical trials. In England, data are also collected via the British Society of Blood and Marrow Transplantation and Cellular Therapy Registry, and the NHS Systemic Anti-Cancer Therapy Dataset. A National CAR T-Cell Registry has been developed in France. Data collection and processing, for coverage with evidence development, is funded by the Applicant. Funding for data processing in France is also provided by the National CAR T-Cell Registry holder (Lymphoma Academic Research Organisation). NHS England fund data collection and processing of the NHS Systemic Anti-Cancer Therapy Dataset data. These coverage with evidence development agreements tend to have a greater number and variety of stakeholders than the patient-level performance-linked agreements (554).

In the case of patient-level performance-linked reimbursement agreements (Italy and Spain), data are collected using national web-based registries, for which bespoke data

collection requirements are created for each therapy/therapeutic indication. However, the Spanish registry, for which CAR T-cell therapies were acting as a pilot, was not fully functioning by the time of reimbursement and so, the Applicant organised data collection. Funding for data collection is provided by the Applicant in Italy and by the regional authorities/hospitals in Spain. Funding for data processing is provided by the AIFA in Italy and the Ministry of Health in Spain (554).

11.1.5 Chapter Aim

The aim of this chapter is to explore the impact of performance-linked reimbursement agreement scenarios on the cost effectiveness and budget impact of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL. The scenarios explored comprise different upfront payments made at the time of infusion; subsequent payments (up to a total maximum, per infusion, of 100% of price-to-wholesaler) are then made based on the OS of patients at pre-specified time points.

11.2 Method

11.2.1 Bespoke Cost-Utility and Budget Impact Models

The cost-utility and budget impact models, described in Chapter 5, Chapter 9, and Chapter 10, were used. In the model base cases, the full cost (price-to-wholesaler minus 5.5% rebate) of tisagenlecleucel and axicabtagene ciloleucel were applied as a once-off at the start of the model (cost-utility model) or at the start of each year (budget impact model). In line with National Economic Evaluation Guidelines, VAT was included in the budget impact model only (11).

11.2.2 Performance-Linked Reimbursement Agreement Scenarios

A number of performance-linked reimbursement agreement scenarios were explored. All scenarios were based on a single-dose infusion per patient.

- **Scenario 1:** 20% of price-to-wholesaler paid for all patients who received tisagenlecleucel or axicabtagene ciloleucel at the time of infusion, 20% of price-to-wholesaler then paid for those alive at the end of year 1, and 15% of price-to-wholesaler then paid for those alive at the end of each year thereafter, until the end of year 5.

- **Scenario 2:** 50% of price-to-wholesaler paid for all patients at the time of infusion, 25% of price-to-wholesaler paid for those alive at the end of year 1, and 25% of price-to-wholesaler paid for those alive at the end of year 2.
- **Scenario 3:** 50% of price-to-wholesaler paid for all patients at the time of infusion and 50% of price-to-wholesaler paid for those alive at the end of year 1.

Scenarios were informed by those employed in other jurisdictions (297, 554). A maximum time horizon of five years was assumed, in line with the assumption (employed in the cost-utility models) that patients alive after 60 months are considered to be long-term survivors. Scenario 2, examining tisagenlecleucel for R/R ALL, is aligned with judgements of experts (n=2) elicited during the expert elicitation (Chapter 3). In their judgements, two experts indicated that patients alive after two years are unlikely to experience progressed disease and will survive long-term.

11.2.3 Implementation in Bespoke Cost-Utility Model

In the cost-utility models, a proportion of the price-to-wholesaler (for each scenario, as described in 11.2.2) of tisagenlecleucel or axicabtagene ciloleucel was applied to all patients receiving infusion at the start of the model. The subsequent costs (up to a total maximum, per infusion, of 100% of price-to-wholesaler) were applied in the cycle in which they were incurred and applied to the proportion of patients alive in that cycle. All other model parameters were maintained in line with those of the respective base case analyses.

The resulting deterministic ICERs versus blinatumomab (R/R ALL), and versus R-GDP (R/R DLBCL) were recorded.

11.2.4 Implementation in Budget Impact Model

In the budget impact model, a proportion of the price-to-wholesaler (for each scenario, as described in 11.2.2) of tisagenlecleucel or axicabtagene ciloleucel was applied to all patients receiving infusion in each respective year (years one to five). The subsequent costs (up to a total maximum, per infusion, of 100% of price-to-wholesaler) were applied in the year in which they were incurred and applied to the proportion of patients alive in

that year. All other model parameters were maintained in line with those of the respective base case analyses. The subsequent cumulative five-year gross and net drug budget impact estimates were recorded.

The net drug budget impact of tisagenlecleucel for R/R ALL assumed displacement of blinatumomab and FLA(G)-IDA, and a total eligible patient population of six patients per year. The net drug budget impact of tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL assumed displacement of R-GDP, and a total eligible patient population of 34 patients in year 1, increasing to 36 per year by year 5. Further detail is provided in Chapter 10.

11.2.4.1 Reduction in Revenue to Applicant Company

The reduction in revenue to the Applicant, as a result of the performance-linked reimbursement agreement, was also estimated. This was estimated by subtracting the cumulative five-year gross drug budget impact (excluding VAT), for each scenario, from the cumulative five-year gross drug budget impact with no performance-linked reimbursement agreement (i.e. the base case gross drug budget impact, excluding VAT).

11.3 Results

11.3.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

11.3.1.1 Impact on Cost Effectiveness

The impact of the scenarios on the ICER of tisagenlecleucel versus blinatumomab is presented in Table 72.

Table 72 Impact of performance-linked reimbursement agreements on deterministic ICER of tisagenlecleucel versus blinatumomab for relapsed/refractory acute lymphoblastic leukaemia

Performance-Linked Reimbursement Agreement	Deterministic ICER (€/QALY) [base case ICER €73,086/QALY]
Scenario 1: 20% upfront payment at infusion, 20% paid for those alive at end of year 1, and 15% paid for those alive at the end of each year thereafter, until the end of year 5	20,027
Scenario 2: 50% upfront payment at infusion, 25% paid for those alive at end of year 1, and 25% paid for those alive at the end of year 2	49,370
Scenario 3: 50% upfront payment at infusion and 50% paid for those alive at the end of year 1	55,149

ICER: Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year.

11.3.1.2 Impact on Gross and Net Drug Budget Impact: Assuming Tisagenlecleucel Displaces Blinatumomab and FLA(G)-IDA

The impact of the scenarios on the cumulative five-year gross and net drug budget impact of tisagenlecleucel is presented in Table 73. Estimates presented in parenthesis are exclusive of VAT.

Table 73 Impact of performance-linked reimbursement agreements on the cumulative five-year gross and net drug budget impact of tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

Performance-Linked Reimbursement Agreement	Cumulative 5-Year Gross Drug Budget Impact (€) [base case €8,629,758 (€6,940,529)]*	Cumulative 5-Year Net Drug Budget Impact (€) [base case €6,670,070 (€5,364,439)]*	Cumulative 5-Year Reduction in Revenue to Applicant
Scenario 1: 20% upfront payment at infusion, 20% paid for those alive at end of year 1, and 15% paid for those alive at the end of each year thereafter, until the end of year 5	5,706,060 (4,016,831)	3,746,372 (2,440,741)	2,923,698
Scenario 2: 50% upfront payment at infusion, 25% paid for those alive at the end of year 1, and 25% paid for those alive at the end of year 2	7,337,085 (5,647,855)	5,377,396 (4,071,765)	1,292,674
Scenario 3: 50% upfront payment at infusion and 50% paid for those alive at the end of year 1	7,658,084 (5,968,855)	5,698,396 (4,392,765)	971,674

*Estimates in parenthesis are exclusive of VAT.

11.3.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

11.3.2.1 Impact on Cost Effectiveness: Tisagenlecleucel versus R-GDP

The impact of the scenarios on the ICER of tisagenlecleucel versus R-GDP is presented in Table 74.

Table 74 Impact of performance-linked reimbursement agreements on deterministic ICER of tisagenlecleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma

Performance-Linked Reimbursement Agreement	Deterministic ICER (€/QALY) [base case ICER €119,509/QALY]
Scenario 1: 20% upfront payment at infusion, 20% paid for those alive at end of year 1, and 15% paid for those alive at the end of each year thereafter, until the end of year 5	60,532
Scenario 2: 50% upfront payment at infusion, 25% paid for those alive at the end of year 1, and 25% paid for those alive at the end of year 2	86,773
Scenario 3: 50% upfront payment at infusion and 50% paid for those alive at the end of year 1	89,320

ICER: Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

11.3.2.2 Impact on Cost Effectiveness: Axicabtagene Ciloleucel versus R-GDP

The impact of the scenarios on the ICER of axicabtagene ciloleucel versus R-GDP is presented in Table 75.

Table 75 Impact of performance-linked reimbursement agreements on deterministic ICER of axicabtagene ciloleucel versus R-GDP for relapsed/refractory diffuse large B-cell lymphoma

Performance-Linked Reimbursement Agreement	Deterministic ICER (€/QALY) [base case ICER €78,634/QALY]
Scenario 1: 20% upfront payment at infusion, 20% paid for those alive at end of year 1, and 15% paid for those alive at the end of each year thereafter, until the end of year 5	44,554
Scenario 2: 50% upfront payment at infusion, 25% paid for those alive at the end of year 1, and 25% paid for those alive at the end of year 2	60,157
Scenario 3: 50% upfront payment at infusion and 50% paid for those alive at the end of year 1	62,087

ICER: Incremental cost-effectiveness ratio; **QALY:** Quality-adjusted life year; **R-GDP:** Rituximab, gemcitabine, dexamethasone, cisplatin.

11.3.2.3 Impact on Gross and Net Drug Budget Impact: Assuming Tisagenlecleucel and Axicabtagene Ciloleucel Displace R-GDP

The impact of the scenarios on the cumulative five-year gross and net drug budget impact of tisagenlecleucel and axicabtagene ciloleucel is presented in Table 76. Estimates presented in parenthesis are exclusive of VAT.

Table 76 Impact of performance-linked reimbursement agreements on the cumulative five-year gross and net drug budget impact of tisagenlecleucel and axicabtagene ciloleucel for relapsed/refractory diffuse large B-cell lymphoma

Performance-Linked Reimbursement Agreement	Cumulative 5-Year Gross Drug Budget Impact (€) [base case €45,565,913 (€36,646,628)]*	Cumulative 5-Year Net Drug Budget Impact (€) [base case €44,884,847 (€36,091,881)]*	Cumulative 5-Year Reduction in Revenue to Applicant
Scenario 1: 20% upfront payment at infusion, 20% paid for those alive at end of year 1, and 15% paid for those alive at the end of each year thereafter, until the end of year 5	29,211,439 (20,292,154)	28,530,373 (19,737,407)	16,354,474
Scenario 2: 50% upfront payment at infusion, 25% paid for those alive at the end of year 1, and 25% paid for those alive at the end of year 2	36,391,959 (27,472,674)	35,710,893 (26,917,927)	9,173,954
Scenario 3: 50% upfront payment at infusion and 50% paid for those alive at the end of year 1	37,148,067 (28,228,782)	36,467,001 (27,674,036)	8,417,845

*Estimates in parenthesis are exclusive of VAT.

11.4 Discussion

11.4.1 Relapsed/Refractory Acute Lymphoblastic Leukaemia

Performance-linked reimbursement agreements, linked to OS, had a considerable impact on the cost effectiveness of tisagenlecleucel versus blinatumomab. Spreading payments over a five-year horizon (Scenario 1) had the greatest impact on the ICER. In the absence of a price reduction, this scenario reduced the ICER to well below the €45,000 per QALY willingness-to-pay threshold. Of note, no performance-linked reimbursement agreements relating to CD19 CAR T-cell therapies, identified in the literature here, were implemented over a horizon of five years or greater (554, 559).

Scenario 2, which spread payments out over 2 years, resulted in a reduction in the ICER to €49,370 per QALY. If a performance-linked reimbursement agreement aligned with that of Scenario 2 was to be considered by the decision-maker, it is expected that a price reduction would also be required. Relative to a lifetime horizon considered in cost-utility analyses, a duration of two years is unlikely to be sufficient to decrease uncertainty associated with the expected long-term OS of patients. As indicated in the expert elicitation, however, some experts (n=2) consider two years to be the time point at which

patients treated with tisagenlecleucel are 'cured' (subject to age- and sex-matched general population mortality). This is aligned with some findings in the literature (202, 203). The innovative mechanism of tisagenlecleucel limits the generalisability of findings from studies conducted using other therapeutic classes. Considering this uncertainty, a more conservative (i.e. longer) time horizon is warranted. Of note, the time horizon of the performance-linked reimbursement agreement implemented in Spain is 18 months (559).

As expected, Scenario 3 (one-year duration of spread of payment) had the least impact on the ICER. This scenario accounts for OS in the first year following treatment only. This is aligned with the performance-linked reimbursement agreement in Italy (559). However, this duration is unlikely to be sufficient to decrease uncertainty associated with the expected long-term OS of patients. The availability of phase II trial evidence means that survival within the first year following treatment is the least uncertain period over the lifetime horizon. In the cost-utility model (Chapter 5), predicted OS, in the first year, was very closely aligned with the Kaplan-Meier data of the pooled ELIANA and ENSIGN trials (207, 215). Pasquini et al. evaluated data collected in the Centre for International Blood and Marrow Transplant Research database. This database collects longitudinal data on patients receiving CD19 CAR T-cell therapy in 130 participating centres in the US and Canada. A total of 255 paediatric and young adult patients (from 73 centres) with R/R ALL, who received tisagenlecleucel, were included in the analysis. At a median follow up of 13.4 months, the 12-month OS (from the time of infusion) was 77% (95% CI 70 to 83) (560). This is aligned with the 12-month OS in ELIANA (76%; 95% CI 63 to 86) (215). In this instance, the value of a performance-linked reimbursement agreement, implemented over the first year, over a simple price reduction is questionable.

A similar trend to that described for the cost-utility analysis was observed for the cumulative five-year gross and net drug budget impact estimates. Scenario 1 had the greatest impact on the budget impacts, while Scenario 3 had the least impact. Scenario 1 reduced the cumulative five-year gross and net drug budget impacts by approximately €3.0 million (including VAT) respectively, representing a 34% reduction from the base case estimates. Considering the small number of patients expected to receive treatment,

this saving is notable. Scenario 1 decreased the revenue received by the Applicant by approximately €3.0 million.

11.4.2 Relapsed/Refractory Diffuse Large B-Cell Lymphoma

As expected, based on the outcomes observed in R/R ALL, Scenario 1 (five-year duration of spread) had the greatest impact on the ICERs. In the case of axicabtagene ciloleucel versus R-GDP, this scenario reduced the ICER to below a willingness-to-pay threshold of €45,000 per QALY. Although this scenario did not reduce the ICER of tisagenlecleucel versus R-GDP to below the €45,000 per QALY threshold, the ICER generated (€60,532 per QALY) approached this. A realistic price reduction, in combination with this scenario, would likely reduce the ICER (tisagenlecleucel versus R-GDP) to the payer threshold.

Neither Scenario 2 (two year duration of spread) nor Scenario 3 (one year duration of spread) resulted in the ICERs (tisagenlecleucel versus R-GDP, or axicabtagene ciloleucel versus R-GDP) achieving the payer threshold. Under these scenarios, axicabtagene ciloleucel would likely require a realistic price reduction to achieve an ICER of €45,000 per QALY. The price reduction required for tisagenlecleucel would need to be more substantial. A measure of a 'successful' performance-linked reimbursement agreement is that uncertainty in the associated parameter is reduced (538). These time horizons may not be sufficient to reduce uncertainty in OS. It has been proposed that patients with DLBCL who are alive at two years are considered 'cured' (subject to age- and sex-matched general population mortality) (113). In Italy, the performance-linked reimbursement agreements, for both tisagenlecleucel and axicabtagene ciloleucel, are implemented over the first 12 months following treatment only (559). Pasquini et al., described in 11.4.1, also evaluated OS in patients with R/R NHL (publication does not specify DLBCL) treated with tisagenlecleucel. A total of 155 patients (from 73 centres) with R/R NHL, treated with tisagenlecleucel, were included in the analysis. At a median follow up of 11.9 months, the 12-month OS was 56% (95% CI 44 to 67) (560). This is higher than the 12-month OS in JULIET (49%; 95% CI 39 to 59) (392).

In terms of most impactful scenario, the trends observed for the gross and net drug budget impacts were aligned with those of the cost-utility analyses. Scenario 1 had the

greatest impact on the gross and net drug budget impacts, while Scenario 3 had the least impact. Scenario 1 reduced the cumulative 5-year gross and net drug budget impacts by approximately €16.4 million (including VAT) respectively, representing a 36% reduction compared to the base case analysis. This is a considerable saving and highlights the challenge of affordability with CD19 CAR T-cell therapies. Despite the lower ICERs (when compared to the base case analysis) obtained under Scenario 1, the gross and net drug budget impacts under this scenario remain high. This is an important consideration for decision-makers. Scenario 1 decreased the revenue received by the Applicant by approximately €16.0 million.

It is evident from the results presented that the time horizon over which the performance-linked reimbursement agreement is implemented is key to the utility of the agreement. It is also a key challenge. Both Drummond and Hutton et al. suggest that a period longer than three years may become irrelevant in the face of changing clinical practice and technological advancement (561, 562). A trade-off is required between ensuring the agreement is relevant to current clinical practice and yet still addresses the key uncertainty of long-term OS. Drummond proposed that implementing a sequence of agreements over time may be an appropriate approach to ensuring the time horizon of such agreements is sufficient. Under this approach, the conditions of the agreement are modified, at pre-specified intervals, based on the availability of new data (563).

It is crucial that a date of review is agreed between the payer and Applicant. This is to ensure that the agreement is meeting, or is on course to meet, its objectives. The measure of success of such an agreement is multidimensional. Its impact on decision uncertainty should be evaluated. However, additional aspects such as integrity of the design of the agreement, and quality of data, should also be examined (538). The findings of such an evaluation can inform the design of future performance-linked reimbursement agreements.

Performance-linked reimbursement agreements should not be seen as a replacement for high-quality RCTs. Regulators should insist that evidence is generated from RCTs, when appropriate. It is advised that HTA agencies and payers have clear criteria to determine if a performance-linked reimbursement agreement is appropriate. While performance-

linked reimbursement agreements can be valuable in mitigating against the financial risk associated with an uncertain evidence base, the risk to patients (due to uncertain efficacy data) remains.

11.4.3 Comparison with the Published Literature

Publications were identified in the literature, which outlined the type of CD19 CAR T-cell therapy performance-based risk-sharing agreements implemented in various jurisdictions (554, 559). However, no publications were identified, which examined the impact of such agreements on the cost effectiveness and budget impact of these therapies. To adequately reflect the true impact on cost effectiveness and budget impact, real-world data are required to model the outcomes observed in clinical practice. Puig-Peiró et al. conducted an SLR to evaluate the existing knowledge on the costs and benefits of performance-based risk-sharing agreements. Of the citations included in the final evidence base (n=24), none evaluated the overall economic impact of the agreement. The authors concluded that further research is required to transparently assess the extent to which transactional costs and administrative burden are shared between the payer and Applicant (564). This finding was aligned with that of Antonanzas et al., who conducted an SLR summarising the literature on risk-sharing agreements. Antonanzas et al. noted that the literature assessing the financial and health outcomes of performance-based risk-sharing agreements is limited. The authors concluded that models are required to understand and estimate the utility of such agreements, and to enable the consequences of these agreements to be compared with those in situations without them (565).

Of note, data collected from coverage with evidence development agreements, such as those in place in England and France, will be used in reappraisal of the clinical and cost effectiveness of CD19 CAR T-cell therapies (554). Results of these reappraisals may be published in the future.

11.4.4 Limitations

This research provides an illustrative example of the potential impact of adopting performance-linked reimbursement agreements. However, to determine the true value

of adopting such agreements, a comparison of real-world costs and outcomes based on the standard simple price reduction versus those of a performance-linked reimbursement agreement is required. As such, the true value of adopting such agreements cannot be realised until real-world data become available.

OS of patients, in the bespoke decision analytic models, were closely aligned with the trial data of the respective trials. As such, the estimates of cost effectiveness, generated here, do not provide an indication of cost effectiveness should the CD19 CAR T-cell therapies prove to be more or less effective than predicted by the models.

The time horizons explored in this study were based on agreements in other European countries (554, 559). However, these horizons do not address uncertainty in the long-term OS of patients, which is a key source of uncertainty with CD19 CAR T-cell therapies.

This study did not consider the costs associated with designing, implementing, executing, and reviewing performance-linked reimbursement agreements. It also did not consider responsibility of these costs. The cost savings generated as part of a performance-linked reimbursement agreement need to be considered in the context of the cost of designing, implementing, executing, and reviewing such agreements. Inclusion of these costs is likely to have increased the ICER and budget impact estimates, under all scenarios evaluated.

This research evaluated just one type of performance-linked reimbursement agreement, whereby an upfront payment is made and subsequent payments are made by the payer based on OS at pre-specified intervals. Alternative conditions will generate different results.

11.5 Conclusion

This research illustrates the impact of performance-linked reimbursement agreement scenarios on the cost effectiveness and budget impact of CD19 CAR T-cell therapies. Agreements implemented over long time horizons are warranted to adequately investigate uncertainty in the clinical evidence base of these therapies. However, the most appropriate time horizon over which to implement performance-linked

reimbursement agreements is a key area of uncertainty. In some instances, a price reduction in combination with a performance-linked reimbursement agreement may be required to achieve an acceptable level of cost effectiveness. Even in cases where an acceptable level of cost effectiveness is achieved, the drug budget impact may be high. Considering the likely need for performance-linked reimbursement agreements for other therapies and disease areas, data infrastructure systems in Ireland require updating. Performance-linked reimbursement agreements are not a substitute for high-quality RCTs. However, they may have value in mitigating against financial risk when such evidence is not feasible.

Chapter 12 Conclusion

12.1 Introduction

The aim of this research was to conduct independent HTAs of tisagenlecleucel for R/R ALL, tisagenlecleucel for R/R DLBCL, and axicabtagene ciloleucel for R/R DLBCL, in the Irish healthcare setting. VOI analyses were conducted to estimate the value of simultaneously eliminating all the uncertainty of all uncertain parameters. The potential impact of performance-linked reimbursement agreements was explored.

12.2 Main Findings

An SLR of the evidence of tisagenlecleucel for R/R ALL, presented in Chapter 2, identified a limited evidence base. Two trials that investigated tisagenlecleucel and one that investigated blinatumomab were included in the final evidence base. These were single arm, with small sample size and were of poor quality. Between-study clinical and methodological heterogeneity was a concern. Naïve ITC of OS indicated that tisagenlecleucel was favourable versus blinatumomab. However, there is much uncertainty and the true magnitude of benefit is unknown. The quality of evidence for OS was very low.

The expert elicitation, presented in Chapter 3, indicated that the curative potential and associated long-term OS benefit, of tisagenlecleucel for R/R ALL, is considered uncertain. Of the five experts, two expected that a proportion of paediatric and young adult patients, treated with tisagenlecleucel, may be 'cured' (i.e. subject to age- and sex-matched general population mortality). Three experts indicated that, due to long-term toxicities, patients will not be 'cured'. Experts indicated that tisagenlecleucel would not be used as a bridge to alloSCT. Much uncertainty was noted regarding expected long-term OS. Judgements indicate that the OS benefit observed in clinical trials may not be realised in clinical practice.

Chapter 4 presented an SLR of utility data in paediatric and young adult patients with R/R ALL. Paucity of utility data, in this population, was highlighted. Health-state utility values, derived from patients in ELIANA, may be susceptible to bias. No utility values relating to long-term survival in this population were identified. No treatment- or disease-specific disutility values, associated with adverse events or active treatment, were identified. Variability in assumptions regarding the HROQL of patients who receive alloSCT was

noted. All disutility values chosen for use in the bespoke cost-utility model were based on proxy data.

Findings from Chapter 2, Chapter 3, and Chapter 4 were integrated into a bespoke cost-utility model examining the cost effectiveness of tisagenlecleucel for R/R ALL, presented in Chapter 5. This analysis found that tisagenlecleucel was not cost effective, versus blinatumomab, at a willingness-to-pay threshold of €45,000 per QALY. Probability of cost effectiveness was negligible. EVPI and EVPPI analyses indicated that further research to decrease decision uncertainty, at the payer willingness-to-pay threshold, may not be of value. Caution is warranted in interpretation of results due to uncertainties that may not be adequately captured by PSA and thus, EVPI.

The SLR of the evidence of tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL, presented in Chapter 6, identified a limited evidence base. One trial each for tisagenlecleucel, axicabtagene ciloleucel, and salvage chemotherapy were included in the final evidence base. These were single-arm and of poor quality. OS and PFS outcomes of tisagenlecleucel and axicabtagene ciloleucel were not compared due to between-trial heterogeneity. Naïve ITC of OS (tisagenlecleucel versus salvage chemotherapy, and axicabtagene ciloleucel versus salvage chemotherapy) indicated that both tisagenlecleucel and axicabtagene ciloleucel had favourable outcomes compared with salvage chemotherapy. Interpretation of results is limited by lack of adjustment for differences in populations and trial design. The true magnitude of benefit is therefore, highly uncertain. The quality of evidence for OS was very low.

Chapter 7 detailed the performance of the text-mining tool, Abstrackr, when used to assist in the title and abstract screening process of the SLR (presented in Chapter 6). Abstrackr-assisted screening resulted in workload savings that did not come at the expense of omitting relevant studies. A maximum prediction score of 0.39540 was determined to be a reliable screening stopping point in this SLR. Caution should be exercised in generalising these results to different research questions.

The SLR of utility data in adult patients with R/R DLBCL, presented in Chapter 8, indicated a paucity of data in this population. The difference in health-state utility values derived

from JULIET (tisagenlecleucel) and ZUMA-1 (axicabtagene ciloleucel) were notable. There is lack of consensus in the literature regarding the HRQOL of patients who are considered to be long-term survivors. This is a key area of uncertainty. No treatment- or disease-specific disutility values, associated with adverse events or active treatment, were identified. The disutility values chosen for use in the bespoke cost-utility models were based on assumptions or proxy data. Such data introduce additional uncertainty and may bias the models.

Findings from Chapter 6 and Chapter 8 were integrated into individual bespoke cost-utility models, which were presented in Chapter 9. This analysis found that neither tisagenlecleucel nor axicabtagene ciloleucel were cost effective, versus salvage chemotherapy, for the treatment of R/R DLBCL in the Irish healthcare setting. The ICER of tisagenlecleucel versus salvage chemotherapy was notably higher than that of axicabtagene ciloleucel versus salvage chemotherapy. The probability of cost effectiveness of each intervention (versus salvage chemotherapy) at the €45,000 per QALY threshold was 0%. At this threshold, population EVPI, of both tisagenlecleucel and axicabtagene ciloleucel, indicated that no cost should be attributed to further research to decrease decision uncertainty.

Budget impacts of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL were estimated in Chapter 10. Despite the low number of patients expected to receive treatment, the five-year cumulative gross and net drug budget impacts were high. Affordability is a key challenge.

Chapter 11 illustrated the impact of performance-linked reimbursement agreement scenarios on the cost effectiveness and budget impact of tisagenlecleucel for R/R ALL, and tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL. Agreements, which captured outcomes over a longer time horizon, were impactful. In some instances, a decrease in price in combination with a performance-linked reimbursement agreement may be required. Affordability may remain a concern. The most appropriate time horizon over which to implement performance-linked reimbursement agreements was identified as a key area of uncertainty.

12.3 Implications for Policy and Practice

Based on the inputs and assumptions used in this research, the reimbursement of CD19 CAR T-cell therapies is unlikely to represent cost-effective use of resources within the Irish healthcare setting. Uncertainty is a recurring theme throughout this thesis. Methods traditionally used to characterise uncertainty (PSA and EVPI) are not capable of capturing key areas of uncertainty associated with CD19 CAR T-cell therapies, such as long-term OS. Decision-makers should consider the illustrative examples of performance-linked reimbursement agreements, presented in this thesis, in the reimbursement of future CAR T-cell therapies. Population EVPPI analysis can be used to inform the conditions of these agreements. Further research is required to determine the optimal time horizon over which to implement such agreements. Horizon scanning, to indicate when new comparators may enter the market or if the place in therapy is likely to change, may be useful in determining an appropriate time horizon. Applicants may need to consider price reductions in combination with performance-linked reimbursement agreements, to gain reimbursement, in some instances. The true economic impact of performance-linked reimbursement agreements should be evaluated using the real-world data generated.

Some parameter estimates derived for this research differ considerably to those used in the NCPE HTA appraisals of tisagenlecleucel and axicabtagene ciloleucel (9, 10, 88). A notable example is the number of patients expected to receive treatment with tisagenlecleucel and axicabtagene ciloleucel for R/R DLBCL. The updated estimates, derived as part of this research, suggest that the budget impact estimates used for decision-making in the reimbursement of these therapies were considerably underestimated (10). The proportion of patients who receive alloSCT following infusion with tisagenlecleucel for R/R ALL is another notable example. This illustrates the potential consequences of HTA and reimbursement of drugs with limited evidence and clinical experience at the time of HTA appraisal. Such therapies should be flagged at the time of initial HTA appraisal and prioritised for reassessment at a future date.

In this research, differences in populations and trial designs precluded a robust analysis of the relative effectiveness and cost effectiveness of tisagenlecleucel versus axicabtagene ciloleucel for R/R DLBCL. This limits clinicians and payers in making an informed choice regarding the most clinically- and cost-effective therapy. It precludes the

development of clinical guidelines, informed by cost-effectiveness. This may result in an inefficient allocation of resources and suboptimal population outcomes. Potential confounders in real-world evidence are difficult to overcome. However, a pan-European, prospective observational study, utilising data from the EBMT registry, would enhance the limited evidence base on the relative effectiveness of tisagenlecleucel versus axicabtagene ciloleucel.

The development of an elicitation protocol, specific to HTA (264), is a welcome step in advancing this underutilised methodology in HTA. The use of formal elicitation has been endorsed in the NICE Methods Review (2020) (174). However, the complexity of elicitation should not be underestimated. The consequences of a poorly conducted elicitation exercise may be greater than not conducting such an exercise. Extensive training is required for those who conduct and appraise elicitation exercises. This research highlighted key learnings that arose from the elicitation study. The most notable of which is the potential utility of conducting elicitation in the face-to-face setting (as opposed to remotely). Several key areas of uncertainty, regarding the use of elicitation, remain. Further research involving the use of the HTA-specific elicitation protocol is required. Research should explore how different methodological choices, offered in the protocol, impact the judgements elicited.

Several methodological challenges, due to single-arm evidence and paucity of relevant data, were encountered over the course of this research. National Economic Evaluation Guidelines are limited in the guidance they provide regarding these challenges (11). These areas should be prioritised as areas of research to inform future updates of the National Economic Evaluation Guidelines. The emerging literature on text mining in the SLR process should be closely monitored and appraised; National Economic Evaluation Guidelines should be updated accordingly.

Despite the increasing frequency of single-arm evidence, and despite the potential value of methods such as performance-linked reimbursement agreements and expert elicitation, these approaches are not substitutes for high-quality RCTs. It is important that stakeholders (including regulators, health-technology analysts, clinicians, policy-decision

makers) voice this, and that research and development programmes comprise RCTs whenever feasible.

This research explicitly examined three of the nine criteria that the HSE must consider when deciding to reimburse a therapy, as outlined in the Health Act. These were clinical efficacy and effectiveness, cost-effectiveness, and budget impact. However, the results of this thesis should be interpreted alongside the other six criteria outlined in the Health Act (described in 1.5.2.2). The health need of the public and the clinical need for these therapies are important considerations for the diseases outlined in this thesis. This is particularly due to the advanced stage of disease and the age range of the patient population (ALL). As highlighted, these are innovative therapies. Thus, the availability of the therapy for supply, the resources available to the HSE, and the level of clinical supervision required to ensure patient safety require thorough consideration. The HSE must consider the unique institutional requirements required to establish and run a CAR T-cell therapy service. It will also be important to develop a plan for the development of the service when future indications are authorised and reimbursed. This will be critical to the sustainability of the CAR T-cell therapy service in Ireland.

12.4 Conclusion

In this study, independent HTAs of tisagenlecleucel for R/R ALL, tisagenlecleucel for R/R DLBCL, and axicabtagene ciloleucel for R/R DLBCL, in the Irish healthcare setting, were conducted. The cost effectiveness and budget impact of these therapies was assessed using bespoke decision analytic models. These were informed by primary data collection, SLRs, and advanced methods for decision analytic modelling. The value of simultaneously eliminating all uncertainty of all uncertain parameters in the respective cost-utility models was estimated. The impact of performance-linked reimbursement agreements on the cost effectiveness and budget impact of these therapies was explored. Several important implications for policy and practice have been highlighted.

References

1. European Medicines Agency (EMA). Advanced therapy medicinal products: Overview [cited 2021 29 December]. Available from: https://www.ema.europa.eu/en/human-regulatory/overview/advanced-therapy-medicinal-products-overview.
2. European medicines Agency (EMA). Kymriah [cited 2021 29 December]. Available from: https://www.ema.europa.eu/en/medicines/human/EPAR/kymriah.
3. Kymriah 1.2x106 - 6x108 cells dispersion for infusion: Summary of Product Characteristics 2021 [Available from: https://www.medicines.ie/medicines/kymriah-1-2-x-106-6-x-108-cells-dispersion-for-infusion-34945/spc.
4. Yescarta 0.4 -2 x 108 cells dispersion for infusion: Summary of Product Characteristics 2021 [Available from: https://www.medicines.ie/medicines/yescarta-axicabtagene-ciloleucel--34619/spc.
5. Whittington MD, McQueen RB, Ollendorf DA, Kumar VM, Chapman RH, Tice JA, et al. Long-term Survival and Cost-effectiveness Associated With Axicabtagene Ciloleucel vs Chemotherapy for Treatment of B-Cell Lymphoma. JAMA Network Open. 2019;2(2):e190035-e.
6. European Medicines Agency (EMA). First two CAR-T cell medicines recommended for approaval in the European Union 2018 [cited 2022 29 January]. Available from: https://www.ema.europa.eu/en/news/first-two-car-t-cell-medicines-recommended-approval-european-union.
7. McGrath E, Chabannon C, Terwel S, Bonini C, Kuball J. Opportunities and challenges associated with the evaluation of chimeric antigen receptor T cells in real-life. Current Opinion in Oncology. 2020;32(5).
8. Champion AR, Lewis S, Davies S, Hughes DA. Managing access to advanced therapy medicinal products: Challenges for NHS Wales. British Journal of Clinical Pharmacology. 2021;87(6):2444-9.
9. National Centre for Pharmacoeconomics (NCPE). Axicabtagene ciloleucel for the treatment of adult patients with relapsed or refractory diffuse large B-cell lymphoma (DLBCL) and primary mediastinal large B-cell lymphoma (PMBCL) after two or more lines of systemic therapy: Technical Summary.; 2019.
10. National centre for Pharmacoeconomics (NCPE). Tisagenlecleucel (Kymriah) for the treatment of adult patients with relapsed or refractory diffuse large B-cell lymphoma (DLBCL) after two or more lines of systemic therapy: Summary. 2019.
11. Health Information and Quality Authority (HIQA). Guidelines for the Economic Evaluation of Health Technologies in Ireland. Dublin; 2020.
12. Hettle R, Corbett M, Hinde S, Hodgson R, Jones-Diette J, Woolacott N, et al. The assessment and appraisal of regenerative medicines and cell therapy products: an exploration of methods for review, economic evaluation and appraisal. Health Technology Assessment. 2017;21(7).
13. Marsden G, Towse A. Exploring the Assessment and Appraisal of Regenerative Medicines and Cell Therapy Products: Is the NICE Approach Fit for Purpose? Office of Health Economics; 2017.
14. Drummond MF, Neumann PJ, Sullivan SD, Fricke FU, Tunis S, Dabbous O, et al. Analytic Considerations in Applying a General Economic Evaluation Reference Case to Gene Therapy. Value Health. 2019;22(6):661-8.
15. Department of Public Expenditure and Refrom (DEPR). Budget 2022 Expenditure Report. Dublin; 2021.

16. Central Statistics Office (CSO). Statistical Yearbook of Ireland 2021: Part 1 People and Scoiety. Health. . Dublin; 2021.
17. The Health Insurance Authority. Annual Report & Accounts 2020 [Available from: https://www.hia.ie/sites/default/files/81105%20HIA%20Annual%20Report%202019%20FULL%20WEB.pdf.
18. Organisation for Economic Co-operation and Development (OECD). Health at a Glance 2021. Health expenditure as a share of GDP, 2019 2021 [cited 2021 17 November]. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/health-expenditure-as-a-share-of-gdp-2019-or-nearest-year-and-2020_8b5743e5-en.
19. Primary Care Reimbursement Service (PCRS). Primary Care Reimbursement Service Statistical Analysis of Claims and Payments 2020. 2020.
20. Department of Public Expenditure and Refrom (DEPR). The Budget in Brief. A Citizen's Guide to Budget 2022. . 2021.
21. National Cancer Control Programme (NCCP). Technology Review Committee 2021 [cited 2021 30 November]. Available from: https://www.hse.ie/eng/services/list/5/cancer/profinfo/medonc/trc/.
22. Health Service Executive (HSE). Oncology Drugs Management System 2021 [cited 2021 02 November]. Available from: https://www.hse.ie/eng/services/list/5/cancer/profinfo/medonc/cdmp/odms.html.
23. Spillane S, Usher C, Bennett K, Adams R, Barry M. Introduction of generic substitution and reference pricing in Ireland: early effects on state pharmaceutical expenditure and generic penetration, and associated success factors. J Pharm Policy Pract. 2015;8(Suppl 1):O10-O.
24. Smith A, Doran S, Daly M, Kennedy C, Barry M. Effect of an Online Reimbursement Application System on Prescribing of Lidocaine 5% Medicated Plaster in the Republic of Ireland. Appl Health Econ Health Policy. 2021;19(1):133-40.
25. (IGEES) IGEaES. Spending Review 2021: Review of High-Tech Drug Expenditure. 2021.
26. European Union Network of Health Technology Assessment organisations (EUnetHTA). What is Health Technology Assessment (HTA)? 2021 [cited 2021 03 November]. Available from: https://www.eunethta.eu/about-eunethta/.
27. National Centre for Pharmacoeconomics (NCPE). Pharmacoeconomic Evaluations 2021 [cited 2021 01 December]. Available from: https://www.ncpe.ie/pharmacoeconomic-evaluations/.
28. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes: Oxford University Press; 2015.
29. Scuffham PA, Whitty JA, Mitchell A, Viney R. The use of QALY weights for QALY calculations: a review of industry submissions requesting listing on the Australian Pharmaceutical Benefits Scheme 2002-4. Pharmacoeconomics. 2008;26(4):297-310.
30. National Centre for Pharmacoeconomics (NCPE). Full Submission Applicant Template. 2021.
31. Irish Pharmaceutical Healthcare Association and the Health Service Executive and the Department of Health and thr Department of Public Expenditure and Reform. Framework Agreement on the Supply and Pricing of Medicines Dublin2016 [cited 2021 04 November]. Available from: https://www.ipha.ie/getattachment/About-Us/Our-Role/Agreement-on-the-Supply-of-Medicines/FINAL-FRAMEWORK-AGREEMENT-ON-THE-SUPPLY-AND-PRICING-OF-MEDICINES.pdf.

32. McCullagh L, Barry M. The Pharmacoeconomic Evaluation Process in Ireland. Pharmacoeconomics. 2016;34(12):1267-76.
33. O'Mahony JF, Coughlan D. The Irish Cost-Effectiveness Threshold: Does it Support Rational Rationing or Might it Lead to Unintended Harm to Ireland's Health System? Pharmacoeconomics. 2016;34(1):5-11.
34. Briggs A, Claxton K, Sculpher M. Decision Modelling for Health Economic Evaluation. New York: Oxford University Press; 2011.
35. Briggs A. Economics notes: handling uncertainty in economic evaluation. BMJ. 1999;319(7202):120-.
36. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. Br J Psychiatry. 2005;187:106-8.
37. Zabeo A, Keisler JM, Hristozov D, Marcomini A, Linkov I. Value of information analysis for assessing risks and benefits of nanotechnology innovation. Environmental Sciences Europe. 2019;31(1):11.
38. Fenwick E, Palmer S, Claxton K, Sculpher M, Abrams K, Sutton A. An iterative Bayesian approach to health technology assessment: application to a policy of preoperative optimization for patients undergoing major elective surgery. Med Decis Making. 2006;26(5):480-96.
39. (HIQA) HIaQA. A Guide to Health Technology Assessment at HIQA. Dublin; 2016.
40. National Centre for Pharmacoeconomics (NCPE). About Us 2021 [cited 2021 04 November]. Available from: http://www.ncpe.ie/about/.
41. Irish Pharmaceutical Healthcare Association and the Health Service Executive and the Department of Health and thr Department of Public Expenditure and Reform. Framework Agreement on the Supply and Pricing of Medicines 2021 [cited 2022 02 January]. Available from: https://2hmcriw3psofj2qo3f8w6f17-wpengine.netdna-ssl.com/wp-content/uploads/2021/12/20211206_FASPM_TermsofAgreement_Final_ForSigning.pdf.
42. Health (Pricing and Supply of Medical Goods) Act 2013. Act Number 14 of 2013. Dublin [
43. National Centre for Pharmacoeconomics (NCPE). Rapid Review Template 2021 [cited 2021 01 December]. Available from: https://www.ncpe.ie/submission-process/submission-templates/rapid-review-template/.
44. Varley Á, Tilson L, Fogarty E, McCullagh L, Barry M. The Utility of a Rapid Review Evaluation Process to a National HTA Agency. Pharmacoeconomics. 2021.
45. Connolly E, O'Donnell H, Lamrock F, Tilson L, Barry M. Health Technology Assessment of Drugs in Ireland: An Analysis of Timelines. PharmacoEconomics - open. 2020;4(2):287-96.
46. Ashfaq K, Yahaya I, Hyde C, Andronis L, Barton P, Bayliss S, et al. Clinical effectiveness and cost-effectiveness of stem cell transplantation in the management of acute leukaemia: a systematic review. Health Technol Assess. 2010;14(54):iii-iv, ix-xi, 1-141.
47. Teachey DT, Pui CH. Comparative features and outcomes between paediatric T-cell and B-cell acute lymphoblastic leukaemia. Lancet Oncol. 2019;20(3):e142-e54.
48. Liu YF, Wang BY, Zhang WN, Huang JY, Li BS, Zhang M, et al. Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. EBioMedicine. 2016;8:173-83.
49. Reiter A, Schrappe M, Ludwig WD, Hiddemann W, Sauter S, Henze G, et al. Chemotherapy in 998 unselected childhood acute lymphoblastic leukemia patients. Results and conclusions of the multicenter trial ALL-BFM 86. Blood. 1994;84(9):3122-33.

50. Schorin MA, Blattner S, Gelber RD, Tarbell NJ, Donnelly M, Dalton V, et al. Treatment of childhood acute lymphoblastic leukemia: results of Dana-Farber Cancer Institute/Children's Hospital Acute Lymphoblastic Leukemia Consortium Protocol 85-01. J Clin Oncol. 1994;12(4):740-7.
51. Mullighan CG, Collins-Underwood JR, Phillips LA, Loudin MG, Liu W, Zhang J, et al. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. Nat Genet. 2009;41(11):1243-6.
52. Terwilliger T, Abdul-Hay M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. Blood Cancer Journal. 2017;7(6):e577-e.
53. Hunger SP, Mullighan CG. Acute Lymphoblastic Leukemia in Children. New Engl J Med. 2015;373(16):1541-52.
54. Gale KB, Ford AM, Repp R, Borkhardt A, Keller C, Eden OB, et al. Backtracking leukemia to birth: identification of clonotypic gene fusion sequences in neonatal blood spots. Proc Natl Acad Sci U S A. 1997;94(25):13950-4.
55. Greaves MF, Maia AT, Wiemels JL, Ford AM. Leukemia in twins: lessons in natural history. Blood. 2003;102(7):2321-33.
56. Mitchell C, Hall G, Clarke RT. Acute leukaemia in children: diagnosis and management. BMJ. 2009;338:b2285.
57. Kinlen L. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish new town with nuclear reprocessing sites in Britain. Lancet. 1988;2(8624):1323-7.
58. Kinlen L. Childhood leukaemia, nuclear sites, and population mixing. British Journal of Cancer. 2011;104(1):12-8.
59. Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. Leukemia. 1988;2(2):120-5.
60. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Pediatric Acute Lymphoblastic Leukemia. Version 2. 2021. 2021.
61. Jabbour E, O'Brien S, Konopleva M, Kantarjian H. New insights into the pathophysiology and therapy of adult acute lymphoblastic leukemia. Cancer. 2015;121(15):2517-28.
62. Inaba H, Azzato EM, Mullighan CG. Integration of Next-Generation Sequencing to Treat Acute Lymphoblastic Leukemia with Targetable Lesions: The St. Jude Children's Research Hospital Approach. Front Pediatr. 2017;5:258.
63. Schultz KR, Pullen DJ, Sather HN, Shuster JJ, Devidas M, Borowitz MJ, et al. Risk- and response-based classification of childhood B-precursor acute lymphoblastic leukemia: a combined analysis of prognostic markers from the Pediatric Oncology Group (POG) and Children's Cancer Group (CCG). Blood. 2007;109(3):926-35.
64. UKALL 2011 Trial. United Kingdom National Randomised Trial for Children and Young Adults with Acute Lymphoblastic Leukaemia and Lymphoma 2011. 2013.
65. Sather H, Miller D, Nesbit M, Heyn R, Hammond D. Differences in Prognosis for Boys and Girls with Acute Lymphoblastic Leukaemia. The Lancet. 1981;317(8223):739-43.
66. Williams LA, Spector LG. Survival Differences Between Males and Females Diagnosed With Childhood Cancer. JNCI Cancer Spectrum. 2019;3(2).
67. Berry DA, Zhou S, Higley H, Mukundan L, Fu S, Reaman GH, et al. Association of Minimal Residual Disease With Clinical Outcome in Pediatric and Adult Acute Lymphoblastic Leukemia: A Meta-analysis. JAMA Oncology. 2017;3(7):e170580-e.
68. Borowitz MJ, Devidas M, Hunger SP, Bowman WP, Carroll AJ, Carroll WL, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic

leukemia and its relationship to other prognostic factors: a Children's Oncology Group study. Blood. 2008;111(12):5477-85.
69. Katz AJ, Chia VM, Schoonen WM, Kelsh MA. Acute lymphoblastic leukemia: an assessment of international incidence, survival, and disease burden. Cancer Causes Control. 2015;26(11):1627-42.
70. National Cancer Registry Ireland (NCRI). Cancer Trends No 6. Leukaemia. 2010.
71. National Cancer Registry Ireland (NCRI). Number of cases diagnosed, incidence rates, projected case counts and survival for acute lymphoblastic leukaemia.; 2019.
72. Bonaventure A, Harewood R, Stiller CA, Gatta G, Clavel J, Stefan DC, et al. Worldwide comparison of survival from childhood leukaemia for 1995-2009, by subtype, age, and sex (CONCORD-2): a population-based study of individual data for 89 828 children from 198 registries in 53 countries. Lancet Haematol. 2017;4(5):e202-e17.
73. Bhojwani D, Pui CH. Relapsed childhood acute lymphoblastic leukaemia. Lancet Oncol. 2013;14(6):e205-17.
74. Tallen G, Ratei R, Mann G, Kaspers G, Niggli F, Karachunsky A, et al. Long-term outcome in children with relapsed acute lymphoblastic leukemia after time-point and site-of-relapse stratification and intensified short-course multidrug chemotherapy: results of trial ALL-REZ BFM 90. J Clin Oncol. 2010;28(14):2339-47.
75. Freyer DR, Devidas M, La M, Carroll WL, Gaynon PS, Hunger SP, et al. Postrelapse survival in childhood acute lymphoblastic leukemia is independent of initial treatment intensity: a report from the Children's Oncology Group. Blood. 2011;117(11):3010-5.
76. Saarinen-Pihkala UM, Heilmann C, Winiarski J, Glomstein A, Abrahamsson J, Arvidson J, et al. Pathways through relapses and deaths of children with acute lymphoblastic leukemia: role of allogeneic stem-cell transplantation in Nordic data. J Clin Oncol. 2006;24(36):5750-62.
77. Gaynon PS, Harris RE, Altman AJ, Bostrom BC, Breneman JC, Hawks R, et al. Bone marrow transplantation versus prolonged intensive chemotherapy for children with acute lymphoblastic leukemia and an initial bone marrow relapse within 12 months of the completion of primary therapy: Children's Oncology Group study CCG-1941. J Clin Oncol. 2006;24(19):3150-6.
78. Roy A, Cargill A, Love S, Moorman AV, Stoneham S, Lim A, et al. Outcome after first relapse in childhood acute lymphoblastic leukaemia - lessons from the United Kingdom R2 trial. Br J Haematol. 2005;130(1):67-75.
79. Trausti O, Stefan S, Johan A, Erik F, Scott M, Matteo B, et al. Relapsed childhood acute lymphoblastic leukemia in the Nordic countries: prognostic factors, treatment and outcome. Haematologica. 2016;101(1):68-76.
80. Ko RH, Ji L, Barnette P, Bostrom B, Hutchinson R, Raetz E, et al. Outcome of patients treated for relapsed or refractory acute lymphoblastic leukemia: a Therapeutic Advances in Childhood Leukemia Consortium study. J Clin Oncol. 2010;28(4):648-54.
81. ClinicalTrials.gov. A Treatment Protocol for Participants 1-45 Years With Acute Lymphoblastic Leukaemia: US National Library of Medicine; 2021 [cited 2021 13 November]. Available from: https://clinicaltrials.gov/ct2/show/NCT03911128.
82. United Kingdom Childrens Cancer and Leukaemia Group. ALLR3: An International Collaborative Trial for Relapsed and Refractory Acute Lymphoblastic Leukaemia (ALL) 2007 [cited 2021 05 November]. Available from: https://storage.googleapis.com/plos-corpus-prod/10.1371/journal.pone.0108107/1/pone.0108107.s003.pdf?X-Goog-Algorithm=GOOG4-RSA-SHA256&X-Goog-Credential=wombat-sa%40plos-prod.iam.gserviceaccount.com%2F20211105%2Fauto%2Fstorage%2Fgoog4_request&X-Goog-Date=20211105T165025Z&X-Goog-Expires=86400&X-Goog-

SignedHeaders=host&X-Goog-Signature=91911e77e23148a0f7d8fefe653dd34cb9bd3cd0f5362365ea780bcac3996f5ce4a4a044de6deeddaccdf26e8278302082baf051eabcc82b96cfad7e85d9cf4d5eaf8b2871794ebbb7aa9d5e3dcdee917be91e091aec16ac0de647e3055035f484a168f34c5c2a87e1f5ede529189a23a9d75d6db00e4aca736867f4354a790e1d048da4b2a864a4378eda0021af83e437170ba079d0348cf1de8175d5dedf49d86b9a5a38214f817b14a0bbe7e19e8580a34b08b131e6a3d6041dad0e43abeaf80065686a5ebab532fad0d7eb680dc2fe6a8b1752d985fbb1ec2c60d3107453bd305c38d42500e141980a065f2652b1e0d46a9117598c47e3691d4896fe8290.
83. Samarasinghe S, Hough R, Moppett J, Bonney D, Ghorashian S, Amrolia P, et al. Guidance for the Treatment of Relapsed ALL: 2019. 2019.
84. ClinicalTrials.gov. CARPALL: Immunotherapy with CD19 CAR T-cell for CD19+ Haematological Malignancies 2021 [cited 2021 01 December]. Available from: https://clinicaltrials.gov/ct2/show/NCT02443831.
85. ClinicalTrials.gov. CD19/22 CAR T Cells (AUTO3) for the Treatment of B Cell Acute Lymphoblastic Leukemia (ALL) (AMELIA) 2021 [cited 2021 01 December]. Available from: https://clinicaltrials.gov/ct2/show/NCT03289455.
86. Blincyto 38.5 micrograms powder for concentrate and solution for solution for infusion: Summary of Product Characteristics 2021 [cited 2022 09 March]. Available from: https://www.medicines.ie/medicines/blincyto-31448/spc.
87. European Medicines Agency (EMA). Blincyto EPAR: Public Assessment Report. 2018.
88. National Centre for Pharmacoeconomics (NCPE). Tisagenlecleucel (Kymriah) for the treatment of paediatric and young adult patients up to 25 years of age with B-cell acute lymphoblastic leukaemia (ALL) this is refractory, in relpase post-transplant, or in second or later relapse: Technical Summary Dublin; 2019.
89. Chaganti S, Illidge T, Barrington S, McKay P, Linton K, Cwynarski K, et al. Guidelines for the management of diffuse large B-cell lymphoma. Br J Haematol. 2016;174(1):43-56.
90. Gouveia GR, Siqueira SAC, Pereira J. Pathophysiology and molecular aspects of diffuse large B-cell lymphoma. Rev Bras Hematol Hemoter. 2012;34(6):447-51.
91. Lymphoma Forum of Ireland. Guidelines on Diagnosis and Treatment of Malignant Lymphomas. 2010.
92. Friedberg JW, Fisher RI. Diffuse large B-cell lymphoma. Hematol Oncol Clin North Am. 2008;22(5):941-52, ix.
93. Chen B-J, Fend F, Campo E, Quintanilla-Martinez L. Aggressive B-cell lymphomas—from morphology to molecular pathogenesis. Annals of Lymphoma. 2019;3.
94. Li S, Young KH, Medeiros LJ. Diffuse large B-cell lymphoma. Pathology. 2018;50(1):74-87.
95. Xie Y, Pittaluga S, Jaffe ES. The histological classification of diffuse large B-cell lymphomas. Semin Hematol. 2015;52(2):57-66.
96. Grimm KE, O'Malley DP. Aggressive B cell lymphomas in the 2017 revised WHO classification of tumors of hematopoietic and lymphoid tissues. Annals of Diagnostic Pathology. 2019;38:6-10.
97. Schneider C, Pasqualucci L, Dalla-Favera R. Molecular pathogenesis of diffuse large B-cell lymphoma. Semin Diagn Pathol. 2011;28(2):167-77.
98. Tilly H, Gomes da Silva M, Vitolo U, Jack A, Meignan M, Lopez-Guillermo A, et al. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015;26 Suppl 5:v116-25.

99. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: B-Cell Lymphomas. Version 5.2021.; 2021.
100. Lister TA, Crowther D, Sutcliffe SB, Glatstein E, Canellos GP, Young RC, et al. Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. J Clin Oncol. 1989;7(11):1630-6.
101. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014;32(27):3059-68.
102. Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Müeller SP, et al. Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. J Clin Oncol. 2014;32(27):3048-58.
103. Kumar V, Shrivastava S, Meghal T, Chandra B. Recent Advances in Diffuse Large B Cell Lymphoma 2018 [cited 2021 15 November]. Available from: https://www.intechopen.com/chapters/60573.
104. A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med. 1993;329(14):987-94.
105. Scott DW, Mottok A, Ennishi D, Wright GW, Farinha P, Ben-Neriah S, et al. Prognostic Significance of Diffuse Large B-Cell Lymphoma Cell of Origin Determined by Digital Gene Expression in Formalin-Fixed Paraffin-Embedded Tissue Biopsies. J Clin Oncol. 2015;33(26):2848-56.
106. Gutiérrez-García G, Cardesa-Salzmann T, Climent F, González-Barca E, Mercadal S, Mate JL, et al. Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. Blood. 2011;117(18):4836-43.
107. Liu Y, Barta SK. Diffuse large B-cell lymphoma: 2019 update on diagnosis, risk stratification, and treatment. American Journal of Hematology. 2019;94(5):604-16.
108. Crump M, Neelapu SS, Farooq U, Van Den Neste E, Kuruvilla J, Westin J, et al. Outcomes in refractory diffuse large B-cell lymphoma: results from the international SCHOLAR-1 study. Blood. 2017;130(16):1800-8.
109. Jaffe ES, Pittaluga S. Aggressive B-cell lymphomas: a review of new and old entities in the WHO classification. Hematology Am Soc Hematol Educ Program. 2011;2011:506-14.
110. Sarkozy C, Coiffier B. Diffuse large B-cell lymphoma in the elderly: a review of potential difficulties. Clin Cancer Res. 2013;19(7):1660-9.
111. National Cancer Registry Ireland (NCRI). Number of cases diagnosed, incidence rates, projected case counts and survival for diffuse large B-cell lymphoma.; 2019.
112. Friedberg JW. Relapsed/refractory diffuse large B-cell lymphoma. Hematology Am Soc Hematol Educ Program. 2011;2011:498-505.
113. Maurer MJ, Ghesquières H, Jais J-P, Witzig TE, Haioun C, Thompson CA, et al. Event-Free Survival at 24 Months Is a Robust End Point for Disease-Related Outcome in Diffuse Large B-Cell Lymphoma Treated With Immunochemotherapy. Journal of Clinical Oncology. 2014;32(10):1066-73.
114. Skrabek P, Assouline S, Christofides A, MacDonald D, Prica A, Sangha R, et al. Emerging therapies for the treatment of relapsed or refractory diffuse large B cell lymphoma. Curr Oncol. 2019;26(4):253-65.

115. Morrison VA, Shou Y, Bell JA, Hamilton L, Ogbonnaya A, Raju A, et al. Evaluation of treatment patterns and survival among patients with diffuse large B-cell lymphoma in the USA. Future Oncology. 2019;15(9):1021-34.
116. Halwani AS, Chien H-C, Morreall DK, Patil V, Rasmussen KM, Li C, et al. Survival Patterns in Patients with Relapsed or Refractory Diffuse Large B Cell Lymphoma: Treatment Trajectories and Responses after the First Relapse. Blood. 2019;134(Supplement_1):1622-.
117. Chien H-C, Morreall D, Patil V, Rasmussen KM, Li C, Yong CM, et al. Real-world practice patterns and outcomes in Veterans with relapsed/refractory diffuse large B-cell lymphoma. Future Oncology. 2020;17(4):411-22.
118. Radford J, White E, Castro FA, Chaturvedi A, Spielewoy N, Gibb A, et al. Treatment Patterns and Outcomes in Patients with Relapsed or Refractory Diffuse Large B-Cell Lymphoma: Experience from a Single UK Centre. Blood. 2019;134(Supplement_1):2917-.
119. Philip T, Guglielmi C, Hagenbeek A, Somers R, Van der Lelie H, Bron D, et al. Autologous bone marrow transplantation as compared with salvage chemotherapy in relapses of chemotherapy-sensitive non-Hodgkin's lymphoma. N Engl J Med. 1995;333(23):1540-5.
120. Arboe B, Olsen MH, Gørløv JS, Duun-Henriksen AK, Dalton SO, Johansen C, et al. Treatment intensity and survival in patients with relapsed or refractory diffuse large B-cell lymphoma in Denmark: a real-life population-based study. Clin Epidemiol. 2019;11:207-16.
121. Harris LJ, Patel K, Martin M. Novel Therapies for Relapsed or Refractory Diffuse Large B-Cell Lymphoma. Int J Mol Sci. 2020;21(22).
122. Feugier P, Van Hoof A, Sebban C, Solal-Celigny P, Bouabdallah R, Fermé C, et al. Long-term results of the R-CHOP study in the treatment of elderly patients with diffuse large B-cell lymphoma: a study by the Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol. 2005;23(18):4117-26.
123. Van Den Neste E, Schmitz N, Mounier N, Gill D, Linch D, Trneny M, et al. Outcome of patients with relapsed diffuse large B-cell lymphoma who fail second-line salvage regimens in the International CORAL study. Bone Marrow Transplant. 2016;51(1):51-7.
124. Van Den Neste E, Schmitz N, Mounier N, Gill D, Linch D, Trneny M, et al. Outcomes of diffuse large B-cell lymphoma patients relapsing after autologous stem cell transplantation: an analysis of patients included in the CORAL study. Bone Marrow Transplant. 2017;52(2):216-21.
125. National Cancer Control Programme (NCCP). NCCP Chemotherapy Regimens 2021 [cited 2021 01 December]. Available from: https://www.hse.ie/eng/services/list/5/cancer/profinfo/chemoprotocols/.
126. National cancer Control Programme (NCCP). Lymphoma Chemotherapy Regimens 2021 [cited 2021 14 November]. Available from: https://www.hse.ie/eng/services/list/5/cancer/profinfo/chemoprotocols/lymphoma-myeloma/.
127. Gisselbrecht C, Glass B, Mounier N, Singh Gill D, Linch DC, Trneny M, et al. Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. J Clin Oncol. 2010;28(27):4184-90.
128. Crump M, Kuruvilla J, Couban S, MacDonald DA, Kukreti V, Kouroukis CT, et al. Randomized comparison of gemcitabine, dexamethasone, and cisplatin versus dexamethasone, cytarabine, and cisplatin chemotherapy before autologous stem-cell transplantation for relapsed and refractory aggressive lymphomas: NCIC-CTG LY.12. J Clin Oncol. 2014;32(31):3490-6.

129. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. New England Journal of Medicine. 2021;384(9):842-58.
130. Mounier N, El Gnaoui T, Tilly H, Canioni D, Sebban C, Casasnovas RO, et al. Rituximab plus gemcitabine and oxaliplatin in patients with refractory/relapsed diffuse large B-cell lymphoma who are not candidates for high-dose therapy. A phase II Lymphoma Study Association trial. Haematologica. 2013;98(11):1726-31.
131. HSE Drugs Group. HSE Drugs Group January 2022 Minutes 2022 [cited 2022 21 March]. Available from: https://www.hse.ie/eng/about/who/cpu/drugs-group-minutes/hse-drugs-group-minutes-january-2022.pdf.
132. Vitolo U, Seymour JF, Martelli M, Illerhaus G, Illidge T, Zucca E, et al. Extranodal diffuse large B-cell lymphoma (DLBCL) and primary mediastinal B-cell lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2016;27:v91-v102.
133. June CH, Sadelain M. Chimeric Antigen Receptor Therapy. N Engl J Med. 2018;379(1):64-73.
134. Sterner RC, Sterner RM. CAR-T cell therapy: current limitations and potential strategies. Blood Cancer Journal. 2021;11(4):69.
135. Stephan MT, Ponomarev V, Brentjens RJ, Chang AH, Dobrenkov KV, Heller G, et al. T cell-encoded CD80 and 4-1BBL induce auto- and transcostimulation, resulting in potent tumor rejection. Nat Med. 2007;13(12):1440-9.
136. Scholler J, Brady TL, Binder-Scholl G, Hwang WT, Plesa G, Hege KM, et al. Decade-long safety and function of retroviral-modified chimeric antigen receptor T cells. Sci Transl Med. 2012;4(132):132ra53.
137. Vairy S, Garcia JL, Teira P, Bittencourt H. CTL019 (tisagenlecleucel): CAR-T therapy for relapsed and refractory B-cell acute lymphoblastic leukemia. Drug Des Devel Ther. 2018;12:3885-98.
138. Bridgeman JS, Hawkins RE, Bagley S, Blaylock M, Holland M, Gilham DE. The optimal antigen response of chimeric antigen receptors harboring the CD3zeta transmembrane domain is dependent upon incorporation of the receptor into the endogenous TCR/CD3 complex. J Immunol. 2010;184(12):6938-49.
139. Guedan S, Posey AD, Jr., Shaw C, Wing A, Da T, Patel PR, et al. Enhancing CAR T cell persistence through ICOS and 4-1BB costimulation. JCI Insight. 2018;3(1).
140. LeBien TW, Tedder TF. B lymphocytes: how they develop and function. Blood. 2008;112(5):1570-80.
141. Brentjens RJ, Latouche JB, Santos E, Marti F, Gong MC, Lyddane C, et al. Eradication of systemic B-cell tumors by genetically targeted human T lymphocytes co-stimulated by CD80 and interleukin-15. Nat Med. 2003;9(3):279-86.
142. Scheuermann RH, Racila E. CD19 antigen in leukemia and lymphoma diagnosis and immunotherapy. Leuk Lymphoma. 1995;18(5-6):385-97.
143. European Medicines Agency (EMA). Kymriah: EPAR -Public assessment report. 2018.
144. European Medicines Agency (EMA). Yescarta: EPAR - Public assessment report. 2018.
145. Zhao Z, Condomines M, van der Stegen SJC, Perna F, Kloss CC, Gunset G, et al. Structural Design of Engineered Costimulation Determines Tumor Rejection Kinetics and Persistence of CAR T Cells. Cancer Cell. 2015;28(4):415-28.
146. Hamieh M, Dobrin A, Cabriolu A, van der Stegen SJC, Giavridis T, Mansilla-Soto J, et al. CAR T cell trogocytosis and cooperative killing regulate tumour antigen escape. Nature. 2019;568(7750):112-6.

147. Cappell KM, Kochenderfer JN. A comparison of chimeric antigen receptors containing CD28 versus 4-1BB costimulatory domains. Nature Reviews Clinical Oncology. 2021;18(11):715-27.
148. Meng J, Wu X, Sun Z, Xun R, Liu M, Hu R, et al. Efficacy and Safety of CAR-T Cell Products Axicabtagene Ciloleucel, Tisagenlecleucel, and Lisocabtagene Maraleucel for the Treatment of Hematologic Malignancies: A Systematic Review and Meta-Analysis. Front Oncol. 2021;11:698607-.
149. Kite Receives European Medicines Agency Approval for CAR T Cell Therapy Manufacturing Facility in Europe [press release]. 2020.
150. Buechner J, Kersten MJ, Fuchs M, Salmon F, Jäger U. Chimeric Antigen Receptor-T Cell Therapy: Practical Considerations for Implementation in Europe. Hemasphere. 2018;2(1):e18.
151. Sadelain M. CD19 CAR T Cells. Cell. 2017;171(7):1471.
152. Levine BL, Miskin J, Wonnacott K, Keir C. Global Manufacturing of CAR T Cell Therapy. Molecular Therapy - Methods & Clinical Development. 2017;4:92-101.
153. Thieblemont C, Le Gouill S, Di Blasi R, Cartron G, Morschhauser F, Bachy E, et al. REAL-WORLD RESULTS ON CD19 CAR T-CELL FOR 60 FRENCH PATIENTS WITH RELAPSED/REFRACTORY DIFFUSE LARGE B-CELL LYMPHOMA INCLUDED IN A TEMPORARY AUTHORIZATION FOR USE PROGRAM. Hematological Oncology. 2019;37(S2):301-.
154. McGuirk J, Waller EK, Qayed M, Abhyankar S, Ericson S, Holman P, et al. Building blocks for institutional preparation of CTL019 delivery. Cytotherapy. 2017;19(9):1015-24.
155. Teachey DT, Lacey SF, Shaw PA, Melenhorst JJ, Maude SL, Frey N, et al. Identification of Predictive Biomarkers for Cytokine Release Syndrome after Chimeric Antigen Receptor T-cell Therapy for Acute Lymphoblastic Leukemia. Cancer Discov. 2016;6(6):664-79.
156. Maude SL, Frey N, Shaw PA, Aplenc R, Barrett DM, Bunin NJ, et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. N Engl J Med. 2014;371(16):1507-17.
157. Davila ML, Riviere I, Wang X, Bartido S, Park J, Curran K, et al. Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. Sci Transl Med. 2014;6(224):224ra25.
158. RoActemra 20mg/ml concentrate for solution for infusion: Summary of Product Characteristics 2021 [cited 2021 13 November]. Available from: https://www.medicines.ie/medicines/roactemra-20-mg-ml-concentrate-for-solution-for-infusion--33648/spc.
159. Neelapu SS, Tummala S, Kebriaei P, Wierda W, Gutierrez C, Locke FL, et al. Chimeric antigen receptor T-cell therapy - assessment and management of toxicities. Nat Rev Clin Oncol. 2018;15(1):47-62.
160. Santomasso B, Bachier C, Westin J, Rezvani K, Shpall EJ. The Other Side of CAR T-Cell Therapy: Cytokine Release Syndrome, Neurologic Toxicity, and Financial Burden. Am Soc Clin Oncol Educ Book. 2019;39:433-44.
161. Taraseviciute A, Tkachev V, Ponce R, Turtle CJ, Snyder JM, Liggitt HD, et al. Chimeric Antigen Receptor T Cell-Mediated Neurotoxicity in Nonhuman Primates. Cancer Discov. 2018;8(6):750-63.
162. Locke FL, Neelapu SS, Bartlett NL, Siddiqi T, Chavez JC, Hosing CM, et al. Phase 1 Results of ZUMA-1: A Multicenter Study of KTE-C19 Anti-CD19 CAR T Cell Therapy in Refractory Aggressive Lymphoma. Mol Ther. 2017;25(1):285-95.
163. Lee DW, Santomasso BD, Locke FL, Ghobadi A, Turtle CJ, Brudno JN, et al. ASTCT Consensus Grading for Cytokine Release Syndrome and Neurologic Toxicity Associated

with Immune Effector Cells. Biology of Blood and Marrow Transplantation. 2019;25(4):625-38.
164. Davila ML, Brentjens RJ. CD19-Targeted CAR T cells as novel cancer immunotherapy for relapsed or refractory B-cell acute lymphoblastic leukemia. Clin Adv Hematol Oncol. 2016;14(10):802-8.
165. Health Information and Quality Authority (HIQA). Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland. Dublin; 2018.
166. European Medicines Agency (EMA). PRIME: priority medicines https://www.ema.europa.eu/en/human-regulatory/research-development/prime-priority-medicines2021 [
167. Paulus JK, Dahabreh IJ, Balk EM, Avendano EE, Lau J, Ip S. Opportunities and challenges in using studies without a control group in comparative effectiveness reviews. Res Synth Methods. 2014;5(2):152-61.
168. Elston J, Taylor RS. Use of surrogate outcomes in cost-effectiveness models: a review of United Kingdom health technology assessment reports. Int J Technol Assess Health Care. 2009;25(1):6-13.
169. Zhu R, Lu D, Chu YW, Chai A, Green M, Zhang N, et al. Assessment of Correlation Between Early and Late Efficacy Endpoints to Identify Potential Surrogacy Relationships in Non-Hodgkin Lymphoma: a Literature-Based Meta-analysis of 108 Phase II and Phase III Studies. Aaps j. 2017;19(3):669-81.
170. Wall DA, Krueger J. Chimeric antigen receptor T cell therapy comes to clinical practice. Curr Oncol. 2020;27(Suppl 2):S115-S23.
171. Ollendorf DA, Chapman RH, Pearson SD. Evaluating and Valuing Drugs for Rare Conditions: No Easy Answers. Value in Health. 2018;21(5):547-52.
172. Chapman RH, Kumar VM, Whittington MD, Pearson SD. Does Cost-Effectiveness Analysis Overvalue Potential Cures? Exploring Alternative Methods for Applying a "Shared Savings" Approach to Cost Offsets. Value in Health. 2021;24(6):839-45.
173. Institute for Clinical and Ecnomic Review (ICER). Adapted Value Assessment Methods for High-Impact "Single and Short-Term Therapies" (SSTs). 2019.
174. National Institute for Health and Care Excellence (NICE). The NICE methods of health technology evaluation: the case for chage.; 2020.
175. Jönsson B, Hampson G, Michaels J, Towse A, von der Schulenburg JG, Wong O. Advanced therapy medicinal products and health technology assessment principles and practices for value-based and sustainable healthcare. Eur J Health Econ. 2019;20(3):427-38.
176. Angelis A, Naci H, Hackshaw A. Recalibrating Health Technology Assessment Methods for Cell and Gene Therapies. Pharmacoeconomics. 2020;38(12):1297-308.
177. ten Ham RMT, Klungel OH, Leufkens HGM, Frederix GWJ. A Review of Methodological Considerations for Economic Evaluations of Gene Therapies and Their Application in Literature. Value in Health. 2020;23(9):1268-80.
178. Rutherford M, Lambert P, Sweeting M, Pennington R, Crowther M, Abrams K, et al. NICE DSU Technical Support Document 21. Flexible Methods for Survival Analysis. 2020.
179. Gibson E, Koblbauer I, Begum N, Dranitsaris G, Liew D, McEwan P, et al. Modelling the Survival Outcomes of Immuno-Oncology Drugs in Economic Evaluations: A Systematic Approach to Data Analysis and Extrapolation. PharmacoEconomics. 2017;35(12):1257-70.
180. Rutherford MJ, Crowther MJ, Lambert PC. The use of restricted cubic splines to approximate complex hazard functions in the analysis of time-to-event data: a simulation study. Journal of Statistical Computation and Simulation. 2015;85(4):777-93.

181. Institute for Clinical and Ecnomic Review (ICER). Value Assessment Methods and Pricing Recommendations for Potential Cures: A Technical Brief. 2019.
182. Department of Public Expenditure and Refrom (DEPR). Public Spending Code: Central Technical References and Economic Appraisal Parameters. 2019.
183. Attema AE, Brouwer WBF, Claxton K. Discounting in Economic Evaluations. Pharmacoeconomics. 2018;36(7):745-58.
184. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977;296(13):716-21.
185. Keeler EB, Cretin S. Discounting of Life-Saving and Other Nonmonetary Effects. Management Science. 1983;29(3):300-6.
186. Gravelle H, Smith D. Discounting for health effects in cost-benefit and cost-effectiveness analysis. Health Econ. 2001;10(7):587-99.
187. Kanavos P, Ferrario A, Tafuri G, Siviero P. Managing Risk and Uncertainty in Health Technology Introduction: The Role of Managed Entry Agreements. Global Policy. 2017;8(S2):84-92.
188. Ferrario A, Kanavos P. Dealing with uncertainty and high prices of new medicines: a comparative analysis of the use of managed entry agreements in Belgium, England, the Netherlands and Sweden. Soc Sci Med. 2015;124:39-47.
189. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses Ottawa, Candad [Available from: **http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp**.
190. Quigley JM, Thompson JC, Halfpenny NJ, Scott DA. Critical appraisal of nonrandomized studies-A review of recommended and commonly used tools. J Eval Clin Pract. 2019;25(1):44-52.
191. Cochrane Scientific Committee. Review of the development of the risk of bias tool for non-randomised studies of interventions - ROBINS-I 2017 [cited 2021 08 December]. Available from: https://methods.cochrane.org/sites/default/files/public/uploads/scientific_committee_statement_report_robins_i_fin.pdf.
192. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.
193. Farrah K, Young K, Tunis MC, Zhao L. Risk of bias tools in systematic reviews of health interventions: an analysis of PROSPERO-registered protocols. Systematic Reviews. 2019;8(1):280.
194. Minozzi S, Cinquini M, Gianola S, Castellini G, Gerardi C, Banzi R. Risk of bias in nonrandomized studies of interventions showed low inter-rater reliability and challenges in its application. Journal of Clinical Epidemiology. 2019;112:28-35.
195. Losilla J-M, Oliveras I, Marin-Garcia JA, Vives J. Three risk of bias tools lead to opposite conclusions in observational research synthesis. Journal of Clinical Epidemiology. 2018;101:61-72.
196. Bilandzic A, Fitzpatrick T, Rosella L, Henry D. Risk of Bias in Systematic Reviews of Non-Randomized Studies of Adverse Cardiovascular Effects of Thiazolidinediones and Cyclooxygenase-2 Inhibitors: Application of a New Cochrane Risk of Bias Tool. PLoS Med. 2016;13(4):e1001987-e.
197. Jiang YQ, Tian Y, Zeng LJ, He SN, Zheng ZT, Shi L, et al. The safety and efficacy of hybrid ablation for the treatment of atrial fibrillation: A meta-analysis. PLoS One. 2018;13(1):e0190170.

198. Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. BMJ Evid Based Med. 2018;23(2):60-3.
199. Schmitz S, Maguire Á, Morris J, Ruggeri K, Haller E, Kuhn I, et al. The use of single armed observational data to closing the gap in otherwise disconnected evidence networks: a network meta-analysis in multiple myeloma. BMC Medical Research Methodology. 2018;18(1):66.
200. Haffar S, Bazerbachi F, Prokop L, Watt KD, Murad MH, Chari ST. Frequency and prognosis of acute pancreatitis associated with fulminant or non-fulminant acute hepatitis A: A systematic review. Pancreatology. 2017;17(2):166-75.
201. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 2011 [cited 2021 08 December]. Available from: https://training.cochrane.org/handbook/archive/v5.1/.
202. Gaynon PS. Childhood acute lymphoblastic leukaemia and relapse. Br J Haematol. 2005;131(5):579-87.
203. Locatelli F, Schrappe M, Bernardo ME, Rutella S. How I treat relapsed childhood acute lymphoblastic leukemia. Blood. 2012;120(14):2807-16.
204. Cochrane Handbook for Systematic Review of Interventions version 6.2 (updated February 2021) Chichester (UK): John Wiley & Sons; 2021 [2nd:[Available from: https://training.cochrane.org/handbook#how-to-cite.
205. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
206. Prasad V, Kim C, Burotto M, Vandross A. The Strength of Association Between Surrogate End Points and Survival in Oncology: A Systematic Review of Trial-Level Meta-analyses. JAMA Intern Med. 2015;175(8):1389-98.
207. European Medicines Agency (EMA). Kymriah: EPAR Public Assessment Report Update. 2020.
208. Cochrane Developmental PaLP. Data extraction forms 2021 [cited 2021 07 December]. Available from: https://dplp.cochrane.org/data-extraction-forms.
209. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
210. Digitizelt Digitizer Software. 2020 [Available from: https://www.digitizeit.xyz/.
211. Guyot P, Ades AE, Ouwens MJNM, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Medical Research Methodology. 2012;12(1):9.
212. R Studio. Version. 2019.
213. Schunemann H, Brozek J, Guyatt G, Oxman A. GRADE Handbook. 2013.
214. GDP G. GRADEpro Guideline Development Tool [Software] McMaster University and Evidence Prime2021 [cited 2021 30 October]. Available from: https://gradepro.org/.
215. Maude SL, Laetsch TW, Buechner J, Rives S, Boyer M, Bittencourt H, et al. Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. N Engl J Med. 2018;378(5):439-48.
216. von Stackelberg A, Locatelli F, Zugmaier G, Handgretinger R, Trippett TM, Rizzari C, et al. Phase I/Phase II Study of Blinatumomab in Pediatric Patients With Relapsed/Refractory Acute Lymphoblastic Leukemia. J Clin Oncol. 2016;34(36):4381-9.
217. Grupp SA, Maude SL, Rives S, Baruchel A, Boyer MW, Bittencourt H, et al. Updated Analysis of the Efficacy and Safety of Tisagenlecleucel in Pediatric and Young Adult Patients with Relapsed/Refractory (r/r) Acute Lymphoblastic Leukemia. Blood. 2018;132(Supplement 1):895-.

218. Maude SL, Pulsipher MA, Boyer MW, Grupp SA, Davies SM, Phillips CL, et al. Efficacy and Safety of CTL019 in the First US Phase II Multicenter Trial in Pediatric Relapsed/Refractory Acute Lymphoblastic Leukemia: Results of an Interim Analysis. Blood. 2016;128(22):2801-.
219. Maude SL, Grupp SA, Mody R, Driscoll T, Laetsch TW, Qayed M, et al., editors. An Updated Analysis of Tisagenlecleucel in Pediatric/Young Adult Patients with Relapsed/Refractory B-Cell Acute Lymphoblastic Leukemia (B-ALL) in a US Multicentre Clinical Trial (ENSIGN). EHA 2018; 2018.
220. Porter DL, Hwang WT, Frey NV, Lacey SF, Shaw PA, Loren AW, et al. Chimeric antigen receptor T cells persist and induce sustained remissions in relapsed refractory chronic lymphocytic leukemia. Sci Transl Med. 2015;7(303):303ra139.
221. Laetsch TW, Myers GD, Baruchel A, Dietz AC, Pulsipher MA, Bittencourt H, et al. Patient-reported quality of life after tisagenlecleucel infusion in children and young adults with relapsed or refractory B-cell acute lymphoblastic leukaemia: a global, single-arm, phase 2 trial. The Lancet Oncology. 2019;20(12):1710-8.
222. Locatelli F, Zugmaier G, Mergen N, Bader P, Jeha S, Schlegel P-G, et al. Blinatumomab in pediatric patients with relapsed/refractory acute lymphoblastic leukemia: results of the RIALTO trial, an expanded access study. Blood Cancer Journal. 2020;10(7):77.
223. Walton M, Sharif S, Hodgson R, Claxton L, Wright K, Simmonds M. Tisagenlecleucel-T for treating relapsed or refractory B-cell acute lymphoblastic leukaemia in people ages up to 25 years: A Single Technology Appraisal. CRD and CHE, University of York, Technology Assessment Group. 2018.
224. von Hippel PT. The heterogeneity statistic I(2) can be biased in small meta-analyses. BMC medical research methodology. 2015;15:35-.
225. Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. Journal of Clinical Epidemiology. 2002;55(1):86-94.
226. Lee JW, Cho B. Prognostic factors and treatment of pediatric acute lymphoblastic leukemia. Korean J Pediatr. 2017;60(5):129-37.
227. Howlader N, Noone A, Krapcho M, Garshell J, Neyman N, Altekruse S, et al. SEER Cancer Statistics Review, 1975-2010. Bethesda, MD: National Cancer Institute; 2010.
228. Chessells JM, Hall E, Prentice HG, Durrant J, Bailey CC, Richards SM. The impact of age on outcome in lymphoblastic leukaemia; MRC UKALL X and XA compared: a report from the MRC Paediatric and Adult Working Parties. Leukemia. 1998;12(4):463-73.
229. Kuhlen M, Willasch AM, Dalle J-H, Wachowiak J, Yaniv I, Ifversen M, et al. Outcome of relapse after allogeneic HSCT in children with ALL enrolled in the ALL-SCT 2003/2007 trial. British Journal of Haematology. 2018;180(1):82-9.
230. Reismüller B, Peters C, Dworzak MN, Pötschger U, Urban C, Meister B, et al. Outcome of children and adolescents with a second or third relapse of acute lymphoblastic leukemia (ALL): a population-based analysis of the Austrian ALL-BFM (Berlin-Frankfurt-Münster) study group. J Pediatr Hematol Oncol. 2013;35(5):e200-4.
231. Amin HM, Yang Y, Shen Y, Estey EH, Giles FJ, Pierce SA, et al. Having a higher blast percentage in circulation than bone marrow: clinical implications in myelodysplastic syndrome and acute lymphoid and myeloid leukemias. Leukemia. 2005;19(9):1567-72.
232. Gatta G, Capocaccia R, Coleman MP, Ries LAG, Berrino F. Childhood cancer survival in Europe and the United States. Cancer. 2002;95(8):1767-72.
233. Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang XS, et al. Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887

patients from 279 population-based registries in 67 countries (CONCORD-2). Lancet. 2015;385(9972):977-1010.
234. Einsiedel HG, von Stackelberg A, Hartmann R, Fengler R, Schrappe M, Janka-Schaub G, et al. Long-term outcome in children with relapsed ALL by risk-stratified salvage therapy: results of trial acute lymphoblastic leukemia-relapse study of the Berlin-Frankfurt-Münster Group 87. J Clin Oncol. 2005;23(31):7942-50.
235. Eckert C, Groeneveld-Krentz S, Kirschner-Schwabe R, Hagedorn N, Chen-Santel C, Bader P, et al. Improving Stratification for Children With Late Bone Marrow B-Cell Acute Lymphoblastic Leukemia Relapses With Refined Response Classification and Integration of Genetics. J Clin Oncol. 2019;37(36):3493-506.
236. Hof J, Krentz S, van Schewick C, Körner G, Shalapour S, Rhein P, et al. Mutations and deletions of the TP53 gene predict nonresponse to treatment and poor outcome in first relapse of childhood acute lymphoblastic leukemia. J Clin Oncol. 2011;29(23):3185-93.
237. Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. Value Health. 2011;14(4):417-28.
238. Phillippo D, Ades A, Dias S, Palmer S, Abrams K, Welton N. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submission to NICE. 2016.
239. Ribera Santasusana JM, de Andrés Saldaña A, García-Muñoz N, Gostkorzewicz J, Martínez Llinàs D, Díaz de Heredia C. Cost-Effectiveness Analysis of Tisagenlecleucel in the Treatment of Relapsed or Refractory B-Cell Acute Lymphoblastic Leukaemia in Children and Young Adults in Spain. Clinicoecon Outcomes Res. 2020;12:253-64.
240. The Canadian Agency for Drugs and Technologies in Health (CADTH). Tisagenlecleucel for Acute Lymphoblastic Leukemia: Economic Review Report. Ottawa: CADTH; 2019.
241. Furzer J, Gupta S, Nathan PC, Schechter T, Pole JD, Krueger J, et al. Cost-effectiveness of Tisagenlecleucel vs Standard Care in High-risk Relapsed Pediatric Acute Lymphoblastic Leukemia in Canada. JAMA Oncol. 2020;6(3):393-401.
242. Murad MH, Mustafa RA, Schünemann HJ, Sultan S, Santesso N. Rating the certainty in evidence in the absence of a single estimate of effect. Evidence Based Medicine. 2017;22(3):85.
243. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. Journal of Clinical Epidemiology. 2011;64(12):1294-302.
244. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol. 2011;64(12):1283-93.
245. de Rojas T, Neven A, Towbin AJ, Carceller F, Bautista F, Riedl D, et al. Clinical research tools in pediatric oncology: challenges and opportunities. Cancer Metastasis Rev. 2020;39(1):149-60.
246. Driscoll JJ, Rixe O. Overall Survival: Still the Gold Standard: Why Overall Survival Remains the Definitive End Point in Cancer Clinical Trials. The Cancer Journal. 2009;15(5):401-5.
247. Food and Drug Administration (FDA). Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics: Guidance for Industry. 2018.

248. Devidas M, Anderson JR. Considerations in the design of clinical trials for pediatric acute lymphoblastic leukemia. Clin Investig (Lond). 2013;3(9):10.4155/cli.13.71.
249. Ciani O, Buyse M, Garside R, Pavey T, Stein K, Sterne JAC, et al. Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. BMJ : British Medical Journal. 2013;346:f457.
250. Norwegian Medicines Agency (NoMA). Single Technology Assessent: Tisagenlecleucel (Kymriah) for the treatment of relapsed/refractory acute lymphoblastic leukaemia (ALL) in paediatric and young adult patients. 2018.
251. National Institute for Health and Care Excellence (NICE). CHTE methods review: Sources and synthesis of evidence. 2020.
252. Aamir S, Anwar MY, Khalid F, Khan SI, Ali MA, Khattak ZE. Systematic Review and Meta-analysis of CD19-Specific CAR-T Cell Therapy in Relapsed/Refractory Acute Lymphoblastic Leukemia in the Pediatric and Young Adult Population: Safety and Efficacy Outcomes. Clinical Lymphoma Myeloma and Leukemia. 2021;21(4):e334-e47.
253. Ma Q, Zhang J, O'Brien E, Martin AL, Agostinho AC. Tisagenlecleucel versus historical standard therapies for pediatric relapsed/refractory acute lymphoblastic leukemia. Journal of Comparative Effectiveness Research. 2020;9(12):849-60.
254. Signorovitch JE, Sikirica V, Erder MH, Xie J, Lu M, Hodgkins PS, et al. Matching-adjusted indirect comparisons: a new tool for timely comparative effectiveness research. Value Health. 2012;15(6):940-7.
255. Davis C, Naci H, Gurpinar E, Poplavska E, Pinto A, Aggarwal A. Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. Bmj. 2017;359:j4530.
256. Merlin T, Weston A, Tooher R. Extending an evidence hierarchy to include topics other than treatment: revising the Australian 'levels of evidence'. BMC Medical Research Methodology. 2009;9(1):34.
257. GRADE Working Group. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490.
258. Bouziana S, Bouzianas D. Exploring the Dilemma of Allogeneic Hematopoietic Cell Transplantation after Chimeric Antigen Receptor T Cell Therapy: To Transplant or Not? Biol Blood Marrow Transplant. 2020;26(8):e183-e91.
259. Miller K, Abraham JH, Rhodes L, Roberts R. Use of the word "cure" in oncology. J Oncol Pract. 2013;9(4):e136-e40.
260. Haupt R, Spinetta JJ, Ban I, Barr RD, Beck JD, Byrne J, et al. Long term survivors of childhood cancer: cure and care. The Erice statement. Eur J Cancer. 2007;43(12):1778-80.
261. Jankovic M, Haupt R, Spinetta JJ, Beck JD, Byrne J, Calaminus G, et al. Long-term survivors of childhood cancer: cure and care-the Erice Statement (2006) revised after 10 years (2016). J Cancer Surviv. 2018;12(5):647-50.
262. Patel D, Grimson F, Mihaylova E, Wagner P, Warren J, van Engen A, et al. Use of External Comparators for Health Technology Assessment Submissions Based on Single-Arm Trials. Value in Health. 2021;24(8):1118-25.
263. Martinalbo J, Bowen D, Camarero J, Chapelin M, Démolis P, Foggi P, et al. Early market access of cancer drugs in the EU. Annals of Oncology. 2016;27(1):96-105.
264. Bojke L, Soares M, Claxton K, Colson A, Fox A, Jackson C, et al. Developing a reference protocol for structured expert elicitation in health-care decision-making: a mixed-methods study. Health Technology Assessment. 2021;25(37).

265. Garthwaite PH, Kadane JB, O'Hagan A. Statistical Methods for Eliciting Probability Distributions. Journal of the American Statistical Association. 2005;100(470):680-701.
266. Iglesias CP, Thompson A, Rogowski WH, Payne K. Reporting Guidelines for the Use of Expert Judgement in Model-Based Economic Evaluations. PharmacoEconomics. 2016;34(11):1161-72.
267. Hemming V, Burgman MA, Hanea AM, McBride MF, Wintle BC. A practical guide to structured expert elicitation using the IDEA protocol. Methods in Ecology and Evolution. 2018;9(1):169-80.
268. Cooke RM, Goossens LHJ. Procedures Guide for Structural Expert Judgement in Accident Consequence Modelling. Radiation Protection Dosimetry. 2000;90(3):303-9.
269. National Institute for Health and Care Excellence (NICE). Eliciting expert judgements in HTA. 2020.
270. Soares MO, Bojke L, Dumville J, Iglesias C, Cullum N, Claxton K. Methods to elicit experts' beliefs over uncertain quantities: application to a cost effectiveness transition model of negative pressure wound therapy for severe pressure ulceration. Statistics in Medicine. 2011;30(19):2363-80.
271. Soares M, Sharples L, Morton A, Claxton K, Bojke L. Experiences of Structured Elicitation for Model-Based Cost-Effectiveness Analyses. Value in Health. 2018;21(6):715-23.
272. European Food Safety Authority. Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment Parma, Italy: European Food Safety Authority; 2014 [Available from: https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2014.3734.
273. Clemen RT, Winkler RL. Combining Probability Distributions From Experts in Risk Analysis. Risk Analysis. 1999;19(2):187-203.
274. Nemet GF, Anadon LD, Verdolini E. Quantifying the Effects of Expert Selection and Elicitation Design on Experts' Confidence in Their Judgments About Future Energy Technologies. Risk Analysis. 2017;37(2):315-30.
275. O'Hagan T. SHELF: Version 4.0 2019 [Available from: http://www.tonyohagan.co.uk/shelf/SHELF4.html.
276. Cooke RM. Experts in uncertainty: Opinion and subjective probability in science. New York, NY, US: Oxford University Press; 1991. xii, 321-xii, p.
277. Leal J, Wordsworth S, Legood R, Blair E. Eliciting expert opinion for economic models: an applied example. Value Health. 2007;10(3):195-203.
278. Speight PM, Palmer S, Moles DR, Downer MC, Smith DH, Henriksson M, et al. The cost-effectiveness of screening for oral cancer in primary care. Health Technol Assess. 2006;10(14):1-144, iii-iv.
279. Poncet A, Gencer B, Blondon M, Gex-Fabry M, Combescure C, Shah D, et al. Electrocardiographic Screening for Prolonged QT Interval to Reduce Sudden Cardiac Death in Psychiatric Patients: A Cost-Effectiveness Analysis. PLoS One. 2015;10(6):e0127213.
280. Noortwijk JMv, Dekker A, Cooke RM, Mazzuchi TA. Expert judgment in maintenance optimization. IEEE Transactions on Reliability. 1992;41(3):427-32.
281. Bojke L, Claxton K, Bravo-Vergel Y, Sculpher M, Palmer S, Abrams K. Eliciting distributions to populate decision analytic models. Value Health. 2010;13(5):557-64.
282. Soares M, Dumville J, Ashby R, Iglesias C, Bojke L, Adderley U, et al. Methods to assess cost-effectiveness and value of further research when data are sparse: negative-pressure wound therapy for severe pressure ulcers. Med Decis Making. 2013;33(3):415-36.

283. Grigore B, Peters J, Hyde C, Stein K. A comparison of two methods for expert elicitation in health technology assessments. BMC Medical Research Methodology. 2016;16(1):85.
284. Devilee J, A K. Software to support expert elicitation: An exploratory study of existing software packages. Bilthoven; 2011.
285. Grigore B, Peters J, Hyde C, Stein K. EXPLICIT: a feasibility study of remote expert elicitation in health technology assessment. BMC Med Inform Decis Mak. 2017;17(1):131-.
286. O'Hagan A, Buck C, Daneshkkhah A, Eiser R, Garthwaite P, Jenkinson D, et al. Uncertain Judgements: Eliciting Experts' Probabilities. West Sussex, England: Wiley; 2006.
287. Kynn M. The 'Heuristics and Biases' Bias in Expert Elicitation. Journal of the Royal Statistical Society Series A (Statistics in Society). 2008;171(1):239-64.
288. Speirs-Bridge A, Fidler F, McBride M, Flander L, Cumming G, Burgman M. Reducing overconfidence in the interval judgments of experts. Risk Anal. 2010;30(3):512-23.
289. Oakley J. Tools to Support the Sheffiled Elicitation Framework 2020 [Available from: https://cran.r-project.org/web/packages/SHELF/SHELF.pdf.
290. Gosling JP. Methods for eliciting expert opinion to inform health technology assessment University of Leeds2014 [Available from: https://pdfs.semanticscholar.org/38eb/a762cdaf5d6dae2fee2063bf776d5facec5b.pdf.
291. Fischer K, Lewandowski D, Janssen MP. Estimating unknown parameters in haemophilia using expert judgement elicitation. Haemophilia. 2013;19(5):e282-8.
292. Fidler MM, Reulen RC, Winter DL, Kelly J, Jenkinson HC, Skinner R, et al. Long term cause specific mortality among 34 489 five year survivors of childhood cancer in Great Britain: population based cohort study. BMJ. 2016;354:i4351.
293. Armstrong GT, Chen Y, Yasui Y, Leisenring W, Gibson TM, Mertens AC, et al. Reduction in Late Mortality among 5-Year Survivors of Childhood Cancer. N Engl J Med. 2016;374(9):833-42.
294. Schulthess D, Gassull D, Makady A, Ludlow A, Rothman B, Have Pt, et al. Are CAR-T therapies living up to their hype? A study using real-world data in two cohorts to determine how well they are actually working in practice compared with bone marrow transplants. BMJ Evidence-Based Medicine. 2019:bmjebm-2019-111226.
295. Thielen FW, van Dongen-Leunis A, Arons AMM, Ladestein JR, Hoogerbrugge PM, Uyl-de Groot CA. Cost-effectiveness of Anti-CD19 chimeric antigen receptor T-Cell therapy in pediatric relapsed/refractory B-cell acute lymphoblastic leukemia. A societal view. Eur J Haematol. 2020;105(2):203-15.
296. Institute for Clinical and Economic Review (ICER). Chimeric Antigen Receptor T-Cell Therapy for B-Cell Cancers: Effectiveness and Value. 2018.
297. Jørgensen J, Hanna E, Kefalas P. Outcomes-based reimbursement for gene therapies in practice: the experience of recently launched CAR-T cell therapies in major European countries. J Mark Access Health Policy. 2020;8(1):1715536.
298. O'Hagan A. Expert Knowledge Elicitation: Subjective but Scientific. The American Statistician. 2019;73(sup1):69-81.
299. O'Connor T, Adams R, Carey N, White A, editors. Expert Elicitation for Cost Effectiveness Analysis. Society for Medical Decision Making; 2020; Berlin, Germny.
300. Cope S, Ayers D, Zhang J, Batt K, Jansen JP. Integrating expert opinion with clinical trial data to extrapolate long-term survival: a case study of CAR-T therapy for children and young adults with relapsed or refractory acute lymphoblastic leukemia. BMC Med Res Methodol. 2019;19(1):182.

301. Soares MO, Sculpher MJ, Claxton K. Health Opportunity Costs: Assessing the Implications of Uncertainty Using Elicitation Methods with Experts. Med Decis Making. 2020;40(4):448-59.
302. Wilson ECF, Usher-Smith JA, Emery J, Corrie PG, Walter FM. Expert Elicitation of Multinomial Probabilities for Decision-Analytic Modeling: An Application to Rates of Disease Progression in Undiagnosed and Untreated Melanoma. Value Health. 2018;21(6):669-76.
303. Morgan MG. Use (and abuse) of expert elicitation in support of decision making for public policy. Proc Natl Acad Sci U S A. 2014;111(20):7176-84.
304. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
305. Montgomery SM, Kusel J. The prevalence of child-specific utilities in NICE appraisals for paediatric indications: rise of the economic orphans? Expert Rev Pharmacoecon Outcomes Res. 2016;16(3):347-50.
306. Whittal A, Meregaglia M, Nicod E. The Use of Patient-Reported Outcome Measures in Rare Diseases and Implications for Health Technology Assessment. Patient. 2021.
307. Janssen B, Szende A. Population norms for the EQ-5D. Seld-Reported Population Health: An International Perspective Based on EQ-5D. Netherlands: Springer; 2014. p. 19-30.
308. Kwon J, Kim SW, Ungar WJ, Tsiplova K, Madan J, Petrou S. A Systematic Review and Meta-analysis of Childhood Health Utilities. Med Decis Making. 2018;38(3):277-305.
309. Forsythe A, Brandt PS, Dolph M, Patel S, Rabe APJ, Tremblay G. Systematic review of health state utility values for acute myeloid leukemia. Clinicoecon Outcomes Res. 2018;10:83-92.
310. Grulke N, Albani C, Bailer H. Quality of life in patients before and after haematopoietic stem cell transplantation measured with the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Core Questionnaire QLQ-C30. Bone Marrow Transplant. 2012;47(4):473-82.
311. National institute for Health and Care Excellence (NICE). 2021 [Available from: https://www.nice.org.uk/.
312. Scottish Medicines Consortium (SMC). 2021 [Available from: https://www.scottishmedicines.org.uk/.
313. The Canadian Agency for Drugs and Technologies in Health (CADTH). 2021 [Available from: https://cadth.ca/.
314. The Norwegian Medicines Agency (NoMA). 2021 [Available from: https://legemiddelverket.no/english.
315. Scottish Medicines Consortium (SMC). Tisagenlecleucel for the treatment of paediatric and young adult patients up to 25 years with B-cell acute lymphoblastic leukaemia (ALL) that is refractory, in relapse post-transplant, or in second or later relapse: Detailed advice. 2019.
316. (PBAC) PBAC. 2021 [cited 2021 01 February]. Available from: https://pbac.pbs.gov.au/.
317. Kelly MJ, Pauker SG, Parsons SK. Using nonrandomized studies to inform complex clinical decisions: the thorny issue of cranial radiation therapy for T-cell acute lymphoblastic leukemia. Pediatr Blood Cancer. 2015;62(5):790-7.
318. Aristides M, Barlev A, Barber B, Gijsen M, Quinn C. Population preference values for health states in relapsed or refractory B-precursor acute lymphoblastic leukemia in the United Kingdom. Health and Quality of Life Outcomes. 2015;13(1):181.

319. Essig S, von der Weid NX, Strippoli MP, Rebholz CE, Michel G, Rueegg CS, et al. Health-related quality of life in long-term survivors of relapsed childhood acute lymphoblastic leukemia. PLoS One. 2012;7(5):e38015.
320. Rodday AM, Terrin N, Parsons SK. Measuring global health-related quality of life in children undergoing hematopoietic stem cell transplant: a longitudinal study. Health Qual Life Outcomes. 2013;11:26.
321. Sung L, Buckstein R, Doyle JJ, Crump M, Detsky AS. Treatment options for patients with acute myeloid leukemia with a matched sibling donor. Cancer. 2003;97(3):592-600.
322. Barata A, Martino R, Gich I, García-Cadenas I, Abella E, Barba P, et al. Do Patients and Physicians Agree When They Assess Quality of Life? Biology of Blood and Marrow Transplantation. 2017;23(6):1005-10.
323. Brazier J, Green C, McCabe C, Stevens K. Use of visual analog scales in economic evaluation. Expert Rev Pharmacoecon Outcomes Res. 2003;3(3):293-302.
324. Felder-Puig R, di Gallo A, Waldenmair M, Norden P, Winter A, Gadner H, et al. Health-related quality of life of pediatric patients receiving allogeneic stem cell or bone marrow transplantation: results of a longitudinal, multi-center study. Bone Marrow Transplant. 2006;38(2):119-26.
325. Sun S, Irestig R, Burström B, Beijer U, Burström K. Health-related quality of life (EQ-5D) among homeless persons compared to a general population sample in Stockholm County, 2006. Scand J Public Health. 2012;40(2):115-25.
326. Burström K, Johannesson M, Diderichsen F. Swedish population health-related quality of life results using the EQ-5D. Qual Life Res. 2001;10(7):621-35.
327. Norwegian Medicines Agency (NoMA). Guidelines for the submission of documentation for single technology assessment (STA) of pharmaceuticals. 2018.
328. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value Health. 2010;13(5):509-18.
329. Lloyd A, Nafees B, Narewska J, Dewilde S, Watkins J. Health state utilities for metastatic breast cancer. British journal of cancer. 2006;95(6):683-90.
330. Nikolaidis G, Murphy P, Jankovic D, Llewellyn A, Palmer S, Wright K, et al. KTE-X19 for treating relapsed or refractory mantle cell lymphoma: A Single Technology Appraisal. CRD and CHR, University of York; 2020.
331. Corbett M, Duarte A, Melton H, Walker S, Wright K, Eastwood A, et al. Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma, mediastinal B-cell lymphoma and follicular lymphoma: A Single Technology Appraisal. University of York Technology Assessment Group; 2018.
332. Norwegian Medicines Agency (NoMA). Single Technology Assessment: Axicabtagene ciloleucel (Yescarta) for the treatment of second or later relapsed/refractory diffuse large B-cell lymphoma (DLBCL) and primary mediastinal large B-cell lymphoma (PMBCL). 2018.
333. Corbett M, Duarte A, Walker S, Wright K, Simmonds M, Palmer S. Tisagenlecleucel for treating relapsed or refractory diffuse large B-cell lymphoma: A Single Technology Appraisal. University of York Technology Assessment Group; 2018.
334. Wittrup-Jensen KU, Lauridsen J, Gudex C, Pedersen KM. Generation of a Danish TTO value set for EQ-5D health states. Scand J Public Health. 2009;37(5):459-66.
335. Dolan P, Sutton M. Mapping visual analogue scale health state valuations onto standard gamble and time trade-off values. Soc Sci Med. 1997;44(10):1519-30.
336. Vetsch J, Wakefield CE, Robertson EG, Trahair TN, Mateos MK, Grootenhuis M, et al. Health-related quality of life of survivors of childhood acute lymphoblastic leukemia: a systematic review. Qual Life Res. 2018;27(6):1431-43.

337. Guadagnolo BA, Punglia RS, Kuntz KM, Mauch PM, Ng AK. Cost-effectiveness analysis of computerized tomography in the routine follow-up of patients after primary treatment for Hodgkin's disease. J Clin Oncol. 2006;24(25):4116-22.
338. National Institute for Health and Care Excellence (NICE). CHTE methods review: Health-related quality of life.Task and finish group report. 2020.
339. Woods B, Sideris E, Palmer S, Latimer N, Soares M. NICE DSU Technical Support Document 19. Partitioned Survival Analysis for Decision Modelling in Healht Care: A Critical Review. 2017.
340. National Cancer Control Programme (NCCP). Blinatumomab Paediatric Therapy. 2019.
341. Latimer N. NICE DSU Technical Support Document 14: Undertaking survival analysis for economic evaluations alongside clinical trials - extrapolation with patient-level data.; 2011.
342. Central Statistics Office (CSO). Irish Life Tables No. 17 2015-2017 2020 [Available from: https://www.cso.ie/en/releasesandpublications/er/ilt/irishlifetablesno172015-2017/.
343. Burnham K, Anderson D. Multimodel Inference. Understanding AIC and BIC in Model Selection 2004 [cited 2022 14 January]. Available from: https://sites.warnercnr.colostate.edu/wp-content/uploads/sites/73/2017/05/Burnham-and-Anderson-2004-SMR.pdf.
344. Parker C, Waters R, Leighton C, Hancock J, Sutton R, Moorman AV, et al. Effect of mitoxantrone on outcome of children with first relapse of acute lymphoblastic leukaemia (ALL R3): an open-label randomised trial. The Lancet. 2010;376(9757):2009-17.
345. MacArthur AC, Spinelli JJ, Rogers PC, Goddard KJ, Abanto ZU, McBride ML. Mortality among 5-year survivors of cancer diagnosed during childhood or adolescence in British Columbia, Canada. Pediatr Blood Cancer. 2007;48(4):460-7.
346. Bhatia S, Robison LL, Francisco L, Carter A, Liu Y, Grant M, et al. Late mortality in survivors of autologous hematopoietic-cell transplantation: report from the Bone Marrow Transplant Survivor Study. Blood. 2005;105(11):4215-22.
347. Socié G, Stone JV, Wingard JR, Weisdorf D, Henslee-Downey PJ, Bredeson C, et al. Long-term survival and late deaths after allogeneic bone marrow transplantation. Late Effects Working Committee of the International Bone Marrow Transplant Registry. N Engl J Med. 1999;341(1):14-21.
348. Ara R, Wailoo A. Using Health State Utility Values in Models Exploring the Cost-Effectiveness of Health Technologies. Value in Health. 2012;15(6):971-4.
349. National Centre for Pharmacoeconomics (NCPE). Guidelines for Inclusion of Drug Costs in Pharmacoeconomic Evaluations. Version 2.2. 2020.
350. Healthcare Pricing Office (HPO). ABF 2020 Admitted Patient Price List: DRG Prices for Inpatients and Daycases 2020. 2020.
351. Central Statistics Office (CSO). Consumer Price Index 2020 [Available from: https://data.cso.ie/.
352. Organisation for Economic Co-operation and Development (OECD). Purchasing Power Parities (PPP) 2020 [Available from: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm.
353. Ibrahim Y-A, Christian C, Peter B, Grzegorz WB, Halvard B, Fabio C, et al. Management of adults and children undergoing chimeric antigen receptor T-cell therapy: best practice recommendations of the European Society for Blood and Marrow Transplantation (EBMT) and the Joint Accreditation Committee of ISCT and EBMT (JACIE). Haematologica. 2020;105(2):297-316.

354. Health Service Executive (HSE). 1 October 2020 Consolidated Pay Scales 2020 [cited 2021 23 August]. Available from: https://healthservice.hse.ie/staff/benefits-services/pay/pay-scales.html.
355. Noaks E, Peticone C, Kotsopoulou E, Bracewell DG. Enriching leukapheresis improves T cell activation and transduction efficiency during CAR T processing. Molecular Therapy - Methods & Clinical Development. 2021;20:675-87.
356. (PCRS) PCRS. List of Reimbursable Items 2020 [Available from: http://www.sspcrs.ie/druglist.
357. Corporate Pharmaceutical Unit (CPU). IPHA Realignment 2018 2018 [Available from: https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2018/.
358. Corporate Pharmaceutical Unit (CPU). IPHA Price Realignement Masterfile for Implementation on 1st July 2019 2019 [Available from: https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2019/.
359. Corporate Pharmaceutical Unit (CPU). IPHA Price Realignment File for Implementation on 1 October 2020 2020 [Available from: https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2020/.
360. National Health Service (NHS). Chemotherapy Protocol: Acute lymphoblastic leukaemia (ALL) Blinatumomab (3,4 day) 2017 [cited 2021 17 August]. Available from: https://www.uhs.nhs.uk/Media/SUHTExtranet/Services/Chemotherapy-SOPs/ALL/Blinatumomab.pdf.
361. Health Service Executive (HSE). Ready Reckoner of Acute Hospital Inpatient and Daycase Activity & Costs (Summarised by DRG) Relating to 2011 Costs and Activity. 2013.
362. O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32.
363. Monthly Index of Medical Specialities (MIMS). MIMS Ireland: January 2020. 2020.
364. Ernst&Young LLP. Analysis of hospital activity and costs following allogeneic stem cell transplantation in England. London; 2021.
365. Bourke S, Burns RM, Gaynor C. Challenges in generating costs and utilisation rates associated with castration-resistant prostate cancer. J Mark Access Health Policy. 2014;2:10.3402/jmahp.v2.24072.
366. Johannesson M, Weinstein MC. On the decision rules of cost-effectiveness analysis. J Health Econ. 1993;12(4):459-67.
367. O'Hagan A, McCabe C, Akehurst R, Brennan A, Briggs A, Claxton K, et al. Incorporation of uncertainty in health economic modelling studies. Pharmacoeconomics. 2005;23(6):529-36.
368. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ. 2001;10(8):779-87.
369. National Institute for Health and Care Excellence (NICE). CHTE methods review: Discounting - Task and finish group report. 2020.
370. Groot Koerkamp B, Myriam Hunink MG, Stijnen T, Weinstein MC. Identifying key parameters in cost-effectiveness analysis using value of information: a comparison of methods. Health Econ. 2006;15(4):383-92.
371. Bojke L, Claxton K, Sculpher MJ, Palmer S. Identifying Research Priorities: The Value of Information Associated with Repeat Screening for Age-Related Macular Degeneration. Medical Decision Making. 2007;28(1):33-43.
372. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making. 2004;24(2):207-27.

373. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18(3):341-64.
374. Philips Z, Claxton K, Palmer S. The half-life of truth: what are appropriate time horizons for research decisions? Med Decis Making. 2008;28(3):287-99.
375. Brennan A, Kharroubi S, O'Hagan A, Chilcott J. Calculating partial expected value of perfect information via Monte Carlo sampling algorithms. Med Decis Making. 2007;27(4):448-70.
376. Strong M. R code for partial EVPI Section of Public Health, School of Health and Related Research, University of Sheffield. [cited 2021 28 September]. Available from: https://www.sheffield.ac.uk/scharr/people/staff/mark-strong.
377. Wood S. mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation 2021 [cited 2021 05 November]. Available from: https://cran.r-project.org/web/packages/mgcv/index.html.
378. McCullagh L, Schmitz S, Barry M, Walsh C. Examining the Feasibility and Utility of Estimating Partial Expected Value of Perfect Information (via a Nonparametric Approach) as Part of the Reimbursement Decision-Making Process in Ireland: Application to Drugs for Cancer. Pharmacoeconomics. 2017;35(11):1177-85.
379. O'Mahony JF, Paulden M, McCabe C. NICE's Discounting Review: Clear Thinking on Rational Revision Meets Obstacle of Industrial Interests. Pharmacoeconomics. 2021;39(2):139-46.
380. O'Mahony JF. Revision of Ireland's Cost-Effectiveness Threshold: New State-Industry Drug Pricing Deal Should Adequately Reflect Opportunity Costs. PharmacoEconomics - Open. 2021;5(3):339-48.
381. Mc Veigh C, Treanor R, Carey N. PCN244 A Comparative Analysis of National Health Technology Assessments of Tisagenlecleucel (KYMRIAH) for Relapsed/Refractory Diffuse Large B Cell Lymphoma. Value in Health. 2020;23:S466.
382. Moradi-Lakeh M, Yaghoubi M, Seitz P, Javanbakht M, Brock E. Cost-Effectiveness of Tisagenlecleucel in Paediatric Acute Lymphoblastic Leukaemia (pALL) and Adult Diffuse Large B-Cell Lymphoma (DLBCL) in Switzerland. Advances in Therapy. 2021;38(6):3427-43.
383. Lin JK, Lerman BJ, Barnes JI, Boursiquot BC, Tan YJ, Robinson AQL, et al. Cost Effectiveness of Chimeric Antigen Receptor T-Cell Therapy in Relapsed or Refractory Pediatric B-Cell Acute Lymphoblastic Leukemia. Journal of Clinical Oncology. 2018;36(32):3192-202.
384. Jackson CH, Thompson SG, Sharples LD. Accounting for uncertainty in health economic decision models by using model averaging. J R Stat Soc Ser A Stat Soc. 2009;172(2):383-404.
385. Afzali HH, Karnon J. Exploring structural uncertainty in model-based economic evaluations. Pharmacoeconomics. 2015;33(5):435-43.
386. Price MJ, Welton NJ, Briggs AH, Ades AE. Model Averaging in the Presence of Structural Uncertainty about Treatment Effects: Influence on Treatment Decision and Expected Value of Information. Value in Health. 2011;14(2):205-18.
387. Sarkar RR, Gloude NJ, Schiff D, Murphy JD. Cost-Effectiveness of Chimeric Antigen Receptor T-Cell Therapy in Pediatric Relapsed/Refractory B-Cell Acute Lymphoblastic Leukemia. J Natl Cancer Inst. 2019;111(7):719-26.
388. Ghorashian S, Furness C, Cumminsn Michelle, Snowden J, O''Reilly M, Roddie C, et al. Intention to Treat Analysis of Real-Worl Outcomes Following Tisagenlecleucel Therapy for Pediatric and Young Adult ALL through a National Access Programme. Blood. 2020.

389. Mahadeo KM, Khazal SJ, Abdel-Azim H, Fitzgerald JC, Taraseviciute A, Bollard CM, et al. Management guidelines for paediatric patients receiving chimeric antigen receptor T cell therapy. Nature Reviews Clinical Oncology. 2019;16(1):45-63.
390. Howlader N, Mariotto AB, Besson C, Suneja G, Robien K, Younes N, et al. Cancer-specific mortality, cure fraction, and noncancer causes of death among diffuse large B-cell lymphoma patients in the immunochemotherapy era. Cancer. 2017;123(17):3326-34.
391. Jakobsen LH, Bøgsted M, Brown PN, Arboe B, Jørgensen J, Larsen TS, et al. Minimal Loss of Lifetime for Patients With Diffuse Large B-Cell Lymphoma in Remission and Event Free 24 Months After Treatment: A Danish Population-Based Study. J Clin Oncol. 2017;35(7):778-84.
392. Schuster SJ, Bishop MR, Tam CS, Waller EK, Borchmann P, McGuirk JP, et al. Tisagenlecleucel in Adult Relapsed or Refractory Diffuse Large B-Cell Lymphoma. New England Journal of Medicine. 2018;380(1):45-56.
393. Schuster SJ, Bishop MR, Tam C, Borchmann P, Jaeger U, Waller EK, et al. Sustained Disease Control for Adult Patients with Relapsed or Refractory Diffuse Large B-Cell Lymphoma: An Updated Analysis of Juliet, a Global Pivotal Phase 2 Trial of Tisagenlecleucel. Blood. 2018;132(Supplement 1):1684-.
394. Schuster SJ, Bishop MR, Tam CS, Borchmann P, Jaeger U, Waller EK, et al. Long-Term Follow-up of Tisagenlecleucel in Adult Patients with Relapsed or Refractory Diffuse Large B-Cell Lymphoma: Updated Analysis of Juliet Study. Biology of Blood and Marrow Transplantation. 2019;25(3, Supplement):S20-S1.
395. Locke FL, Ghobadi A, Jacobson CA, Miklos DB, Lekakis LJ, Oluwole OO, et al. Long-term safety and activity of axicabtagene ciloleucel in refractory large B-cell lymphoma (ZUMA-1): a single-arm, multicentre, phase 1-2 trial. Lancet Oncol. 2019;20(1):31-42.
396. Neelapu SS, Locke FL, Bartlett NL, Lekakis LJ, Miklos DB, Jacobson CA, et al. Axicabtagene Ciloleucel CAR T-Cell Therapy in Refractory Large B-Cell Lymphoma. New England Journal of Medicine. 2017;377(26):2531-44.
397. Rosenwald A, Bens S, Advani R, Barrans S, Copie-Bergman C, Elsensohn M-H, et al. Prognostic Significance of MYC Rearrangement and Translocation Partner in Diffuse Large B-Cell Lymphoma: A Study by the Lunenburg Lymphoma Biomarker Consortium. Journal of Clinical Oncology. 2019;37(35):3359-68.
398. Jaeger U, Bishop MR, Salles G, Schuster SJ, Maziarz RT, Han X, et al. Myc Expression and Tumor-Infiltrating T Cells Are Associated with Response in Patients (Pts) with Relapsed/Refractory Diffuse Large B-Cell Lymphoma (r/r DLBCL) Treated with Tisagenlecleucel in the Juliet Trial. Blood. 2020;136(Supplement 1):48-9.
399. Norwegian Medicines Agency (NoMA). Single Technology Assessment: Tisagenlecleucel (Kymriah) for the treatment of second or later relapsed/refractory diffuse large B-cell lymphoma (DLBCL). 2019.
400. Maziarz RT, Waller EK, Jaeger U, Fleury I, McGuirk J, Holte H, et al. Patient-reported long-term quality of life after tisagenlecleucel in relapsed/refractory diffuse large B-cell lymphoma. Blood advances. 2020;4(4):629-37.
401. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25(5):579-86.
402. Jacobson C, Locke FL, Ghobadi A, Miklos DB, Lekakis LJ, Oluwole OO, et al. Long-Term Survival and Gradual Recovery of B Cells in Patients with Refractory Large B Cell Lymphoma Treated with Axicabtagene Ciloleucel (Axi-Cel). Blood. 2020;136(Supplement 1):40-2.

403. Lee DW, Gardner R, Porter DL, Louis CU, Ahmed N, Jensen M, et al. Current concepts in the diagnosis and management of cytokine release syndrome. Blood. 2014;124(2):188-95.
404. Lin V, Jiang Y, Chuang L, Navale L, Cheng P, Purdum A. Health Utilities for Patients with Relapsed or Refractory Large B-Cell Lymphoma (R/R-LBCL): Ad Hoc Analysis From an Axicabtagene Ciloleucel (Axi-Cel) Safety Managenebt Study. Bone Marrow Transplantation. 2018;53:878-87.
405. Witzig TE, Geyer SM, Kurtin PJ, Colgan JP, Inwards DJ, Micallef INM, et al. Salvage chemotherapy with rituximab DHAP for relapsed non-Hodgkin lymphoma: a phase II trial in the North Central Cancer Treatment Group. Leukemia & lymphoma. 2008;49(6):1074-80.
406. Ahmed M, Chihara D, Vargas N, Ma L, Fayad L, Oki Y, et al. OUTCOME OF RELAPSED/REFRACTORY DIFFUSE LARGE B-CELL LYMPHOMA AFTER SECOND SALVAGE THERAPY: MD ANDERSON EXPERIENCE. Hematological Oncology. 2015;33:279-80.
407. Cheson BD, Horning SJ, Coiffier B, Shipp MA, Fisher RI, Connors JM, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group. J Clin Oncol. 1999;17(4):1244.
408. Nagle SJ, Woo K, Schuster SJ, Nasta SD, Stadtmauer E, Mick R, et al. Outcomes of patients with relapsed/refractory diffuse large B-cell lymphoma with progression of lymphoma after autologous stem cell transplantation in the rituximab era. Am J Hematol. 2013;88(10):890-4.
409. Sermer D, Batlevi C, Palomba ML, Shah G, Lin RJ, Perales M-A, et al. Outcomes in patients with DLBCL treated with commercial CAR T cells compared with alternate therapies. Blood Advances. 2020;4(19):4669-78.
410. Rutherford SC, Leonard JP. Lymphoma “benchmark” or “bench-smudge”? Blood. 2017;130(16):1778-9.
411. Bazarbachi A. Diffuse large B-cell lymphoma: is salvage possible after failure of second-line treatment? Bone Marrow Transplantation. 2016;51(1):1-.
412. Alzahrani M, El-Galaly TC, Hutchings M, Hansen JW, Loft A, Johnsen HE, et al. The value of routine bone marrow biopsy in patients with diffuse large B-cell lymphoma staged with PET/CT: a Danish-Canadian study. Annals of Oncology. 2016;27(6):1095-9.
413. Ollila TA, Reagan JL, Olszewski AJ. Clinical features and survival of patients with T-cell/histiocyte-rich large B-cell lymphoma: analysis of the National Cancer Data Base. Leukemia & Lymphoma. 2019;60(14):3426-33.
414. Ok CY, Papathomas TG, Medeiros LJ, Young KH. EBV-positive diffuse large B-cell lymphoma of the elderly. Blood. 2013;122(3):328-40.
415. Lu T-X, Liang J-H, Miao Y, Fan L, Wang L, Qu X-Y, et al. Epstein-Barr virus positive diffuse large B-cell lymphoma predict poor outcome, regardless of the age. Scientific Reports. 2015;5(1):12168.
416. Zavras PD, Wang Y, Gandhi A, Lontos K, Delgoffe GM. Evaluating tisagenlecleucel and its potential in the treatment of relapsed or refractory diffuse large B cell lymphoma: evidence to date. Onco Targets Ther. 2019;12:4543-54.
417. Andreadis C, Tam CS, Borchmann P, Jaeger U, McGuirk JP, Holte H, et al. Correlation of Bridging and Lymphodepleting Chemotherapy with Clinical Outcomes in Patients with Relapsed/Refractory Diffuse Large B-Cell Lymphoma Treated with Tisagenlecleucel. Blood. 2019;134(Supplement_1):2883-.
418. Schuster SJ, Maziarz RT, Ericson SG, Rusch ES, Signorovitch J, Romanov VV, et al. Consensus Grading of Cytokine Release Syndrome (CRS) in Adult Patients with Relapsed

or Refractory Diffuse Large B-Cell Lymphoma (r/r DLBCL) Treated with Tisagenlecleucel on the JULIET Study. Blood. 2018;132(Supplement 1):4190-.
419. Schuster SJ, Maziarz RT, Rusch ES, Li J, Signorovitch JE, Romanov VV, et al. Grading and management of cytokine release syndrome in patients treated with tisagenlecleucel in the JULIET trial. Blood Adv. 2020;4(7):1432-9.
420. Boyko EJ. Observational research--opportunities and limitations. J Diabetes Complications. 2013;27(6):642-8.
421. Zhang J, Li J, Ma Q, Yang H, Signorovitch J, Wu E. A Review of Two Regulatory Approved Anti-CD19 CAR T-Cell Therapies in Diffuse Large B-Cell Lymphoma: Why Are Indirect Treatment Comparisons Not Feasible? Adv Ther. 2020;37(7):3040-58.
422. Oluwole OO, Jansen JP, Lin VW, Chan K, Keeping S, Navale L, et al. Comparing Efficacy, Safety, and Preinfusion Period of Axicabtagene Ciloleucel versus Tisagenlecleucel in Relapsed/Refractory Large B Cell Lymphoma. Biol Blood Marrow Transplant. 2020;26(9):1581-8.
423. Hayden PJ, Sirait T, Koster L, Snowden JA, Yakoub-Agha I. An international survey on the management of patients receiving CAR T-cell therapy for haematological malignancies on behalf of the Chronic Malignancies Working Party of EBMT. Current Research in Translational Medicine. 2019;67(3):79-88.
424. Nastoupil LJ, Jain MD, Spiegel JY, Ghobadi A, Lin Y, Dahiya S, et al. Axicabtagene Ciloleucel (Axi-cel) CD19 Chimeric Antigen Receptor (CAR) T-Cell Therapy for Relapsed/Refractory Large B-Cell Lymphoma: Real World Experience. Blood. 2018;132(Supplement 1):91-.
425. Sesques P, Ferrant E, Safar V, Wallet F, Tordo J, Dhomps A, et al. Commercial anti-CD19 CAR T cell therapy for patients with relapsed/refractory aggressive B cell lymphoma in a European center. American Journal of Hematology. 2020;95(11):1324-33.
426. Jain MD, Jacobs MT, Nastoupil LJ, Spiegel JY, Feng G, Lin Y, et al. Characteristics and Outcomes of Patients Receiving Bridging Therapy While Awaiting Manufacture of Standard of Care Axicabtagene Ciloleucel CD19 Chimeric Antigen Receptor (CAR) T-Cell Therapy for Relapsed/Refractory Large B-Cell Lymphoma: Results from the US Lymphoma CAR-T Consortium. Blood. 2019;134(Supplement_1):245-.
427. Pinnix CC, Gunther JR, Dabaja BS, Strati P, Fang P, Hawkins MC, et al. Bridging therapy prior to axicabtagene ciloleucel for relapsed/refractory large B-cell lymphoma. Blood Advances. 2020;4(13):2871-83.
428. Denlinger N, Braunstein Z, Agyeman A, Purdin Z, Neal A, Yucebay F, et al. CAR T-Cell Therapy: Clinical Outcomes, Patient Selection and Financial Metrics with Tisagenlecleucel and Axicabtagene Ciloleucel, a Single Center Experience. Biology of Blood and Marrow Transplantation. 2020;26(3, Supplement):S266.
429. Riedell PA, Walling C, Nastoupil LJ, Pennisi M, Maziarz RT, McGuirk JP, et al. A Multicenter Retrospective Analysis of Clinical Outcomes, Toxicities, and Patterns of Use in Institutions Utilizing Commercial Axicabtagene Ciloleucel and Tisagenlecleucel for Relapsed/Refractory Aggressive B-Cell Lymphomas. Blood. 2019;134(Supplement_1):1599-.
430. Abramson JS, Palomba ML, Gordon LI, Lunning MA, Wang M, Arnason J, et al. Lisocabtagene maraleucel for patients with relapsed or refractory large B-cell lymphomas (TRANSCEND NHL 001): a multicentre seamless design study. The Lancet. 2020;396(10254):839-52.
431. Galaznik A, Huelin R, Stokes M, Guo Y, Hoog M, Bhagnani T, et al. Systematic review of therapy used in relapsed or refractory diffuse large B-cell lymphoma and follicular lymphoma. Future Science OA. 2018;4(7):FSO322.

432. Thuresson P-O, Vander Velde N, Gupta P, Talbot J. A Systematic Review of the Clinical Efficacy of Treatments in Relapsed or Refractory Diffuse Large B Cell Lymphoma. Advances in Therapy. 2020;37(12):4877-93.
433. Tricco AC, Brehaut J, Chen MH, Moher D. Following 411 Cochrane protocols to completion: a retrospective cohort study. PLoS One. 2008;3(11):e3684.
434. Gates A, Johnson C, Hartling L. Technology-assisted title and abstract screening for systematic reviews: a retrospective evaluation of the Abstrackr machine learning tool. Systematic Reviews. 2018;7(1):45.
435. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. Systematic Reviews. 2012;1(1):10.
436. Michelson M, Reuter K. The significant cost of systematic reviews and meta-analyses: A call for greater involvement of machine learning to assess the promise of clinical trials. Contemporary Clinical Trials Communications. 2019;16:100443.
437. Rathbone J, Hoffmann T, Glasziou P. Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers. Syst Rev. 2015;4:80.
438. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007;147(4):224-33.
439. Beller EM, Chen JK-H, Wang UL-H, Glasziou PP. Are systematic reviews up-to-date at the time of publication? Systematic Reviews. 2013;2(1):36.
440. Veritas Health Innovation. Covidence Systematic Review Software Melbourne, Australia [cited 2021 13 November]. Available from: www.covidence.org.
441. Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open. 2017;7(2):e012545.
442. O'Mara-Eves A, Thomas J, McNaught J, Miwa M, Ananiadou S. Using text mining for study identification in systematic reviews: a systematic review of current approaches. Systematic Reviews. 2015;4(1):5.
443. Ananiadou S, Mc Naught J. Text Mining for Biology and Biomedicine. Boston/London: Artech House; 2006.
444. Hearst M. Untangling Text Data Mining. 37th Annual Meeting of the Association for Computational Linguistics; 1999.
445. Tsou AY, Treadwell JR, Erinoff E, Schoelles K. Machine learning for screening prioritization in systematic reviews: comparative performance of Abstrackr and EPPI-Reviewer. Systematic Reviews. 2020;9(1):73.
446. Popoff E, Besada M, Jansen JP, Cope S, Kanters S. Aligning text mining and machine learning algorithms with best practices for study selection in systematic literature reviews. Systematic Reviews. 2020;9(1):293.
447. Thomas J, McNaught J, Ananiadou S. Applications of text mining within systematic reviews. Research Synthesis Methods. 2011;2(1):1-14.
448. Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. Systematic Reviews. 2019;8(1):163.
449. Gates A, Guitard S, Pillay J, Elliott SA, Dyson MP, Newton AS, et al. Performance and usability of machine learning for screening in systematic reviews: a comparative evaluation of three tools. Systematic Reviews. 2019;8(1):278.
450. Wallace B, Small K, Brodley C, Lau J, Trikalinos T. Deploying an interactive machine learning system in an evience-based practice center: abstrackr. 2nd ACM SIGHIT

International Health Informatics Symposium New York: Association for Computing Machinery. 2012.
451. Ethan Balke. Abstrackr: Online, Open-Access, Free Software for Citation Screening. KTDRR and Campbell Collaboration Research Evidence Training Session2019.
452. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016;5(1):210.
453. Przybyła P, Brockmeier AJ, Kontonatsios G, Le Pogam M-A, McNaught J, von Elm E, et al. Prioritising references for systematic reviews with RobotAnalyst: A user study. Research Synthesis Methods. 2018;9(3):470-88.
454. The National Centre for Text Mining. RobotAnalyst [Available from: http://www.nactem.ac.uk/robotanalyst/.
455. Tavil B, Aytac S, Balci YI, Unal S, Kuskonmaz B, Yetgin S. Fludarabine, cytarabine, granulocyte colony-stimulating factor, and idarubicin (FLAG-IDA) for the treatment of children with poor-prognosis acute leukemia: the Hacettepe experience. 2010;27:517-28.
456. Howard BE, Phillips J, Miller K, Tandon A, Mav D, Shah MR, et al. SWIFT-Review: a text-mining workbench for systematic review. Systematic reviews. 2016;5:87-.
457. Sciome LLC. SWIFT-Review [Available from: https://www.sciome.com/swift-review/.
458. Thomas J, Graziosi S, O'Driscoll P, Brunton J. EPPI-Reviewer The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)2021 [Available from: https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2936.
459. EPPI-Reviewer. Machine Leaning Functionality in EPPI-Reviewer 2021 [cited 2021 12 July]. Available from: https://eppi.ioe.ac.uk/CMS/Portals/35/machine_learning_in_eppi-reviewer_v_7_web_version.pdf.
460. EPPI-Reviewer. About Us [Available from: https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2936.
461. Hamel C, Kelly SE, Thavorn K, Rice DB, Wells GA, Hutton B. An evaluation of DistillerSR's machine learning-based prioritization tool for title/abstract screening – impact on reviewer-relevant outcomes. BMC Medical Research Methodology. 2020;20(1):256.
462. Evidence Partners. DistillerSR 2021 [Available from: https://www.evidencepartners.com/.
463. Cochrane Collaboration. EPPI-Reviewer [Available from: https://training.cochrane.org/resource/eppi-reviewer.
464. Gates A, Gates M, DaRosa D, Elliott SA, Pillay J, Rahman S, et al. Decoding semi-automated title-abstract screening: findings from a convenience sample of reviews. Systematic Reviews. 2020;9(1):272.
465. Reddy SM, Patel S, Weyrich M, Fenton J, Viswanathan M. Comparison of a traditional systematic review approach with review-of-reviews and semi-automation as strategies to update the evidence. Systematic Reviews. 2020;9(1):243.
466. O'Connor AM, Tsafnat G, Thomas J, Glasziou P, Gilbert SB, Hutton B. A question of trust: can we build an evidence base to gain trust in systematic review automation technologies? Systematic Reviews. 2019;8(1):143.
467. van de Schoot R, de Bruin J, Schram R, Zahedi P, de Boer J, Weijdema F, et al. An open source machine learning framework for efficient and transparent systematic reviews. Nature Machine Intelligence. 2021;3(2):125-33.

468. Paynter R, Banex L, Berliner E, Erinoff E, Lege-Matsuura J, Potter S, et al. EPC methods: an exploration of the use of text-mining software in systematic review. Agency for Healthcare Research and Quality. 2016.
469. Beller E, Clark J, Tsafnat G, Adams C, Diehl H, Lund H, et al. Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). Systematic Reviews. 2018;7(1):77.
470. Gates A, Gates M, Sebastianski M, Guitard S, Elliott SA, Hartling L. The semi-automation of title and abstract screening: a retrospective exploration of ways to leverage Abstrackr's relevance predictions in systematic and rapid reviews. BMC Medical Research Methodology. 2020;20(1):139.
471. Shemilt I, Khan N, Park S, Thomas J. Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. Systematic Reviews. 2016;5(1):140.
472. Wu G, Chang E. KBA: Kernel boundary alignment considering imbalanced data distribution. 2005;IEEE Trans Knowl Data Eng:786-95.
473. Gartlehner G, Affengruber L, Titscher V, Noel-Storr A, Dooley G, Ballarini N, et al. Single-reviewer abstract screening missed 13 percent of relevant studies: a crowd-based, randomized controlled trial. Journal of Clinical Epidemiology. 2020;121:20-8.
474. Betts KA, Thuresson P-O, Felizzi F, Du EX, Dieye I, Li J, et al. US cost–effectiveness of polatuzumab vedotin, bendamustine and rituximab in diffuse large B-cell lymphoma. Journal of Comparative Effectiveness Research. 2020;9(14):1003-15.
475. Roth JA, Sullivan SD, Lin VW, Bansal A, Purdum AG, Navale L, et al. Cost-effectiveness of axicabtagene ciloleucel for adult patients with relapsed or refractory large B-cell lymphoma in the United States. J Med Econ. 2018;21(12):1238-45.
476. Huntington SF, Svoboda J, Doshi JA. Cost-effectiveness analysis of routine surveillance imaging of patients with diffuse large B-cell lymphoma in first remission. J Clin Oncol. 2015;33(13):1467-74.
477. Doorduijn J, Buijt I, Holt B, Steijaert M, Uyl-de Groot C, Sonneveld P. Self-reported quality of life in elderly patients with aggressive non-Hodgkin's lymphoma treated with CHOP chemotherapy. Eur J Haematol. 2005;75(2):116-23.
478. Best JH, Hornberger J, Proctor SJ, Omnes LF, Jost F. Cost-effectiveness analysis of rituximab combined with chop for treatment of diffuse large B-cell lymphoma. Value Health. 2005;8(4):462-70.
479. Coiffier B, Lepage E, Brière J, Herbrecht R, Tilly H, Bouabdallah R, et al. CHOP Chemotherapy plus Rituximab Compared with CHOP Alone in Elderly Patients with Diffuse Large-B-Cell Lymphoma. New England Journal of Medicine. 2002;346(4):235-42.
480. Sullivan PW, Ghushchyan V. Preference-Based EQ-5D index scores for chronic conditions in the United States. Med Decis Making. 2006;26(4):410-20.
481. Scottish Medicines Consortium (SMC). Tisagenlecleucel for the treatment of adult patients with relapsed or refractory diffuse large B-cell lymphoma (DLBCL) after two or more lines of systemic therapy: Detailed advice. 2019.
482. Scottish Medicines Consortium (SMC). Axicabtagene ciloleucel for the treatment of adult patients with relapsed or refractory diffuse large B-cell lymphoma (DLBCL) and primary mediastinal large B-cell lymphoma after two or more lines of systemic therapy: Detailed advice. 2019.
483. The Canadian Agency for Drugs and Technologies in Health (CADTH). Tisagenlecleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma: Economic Review Report. Ottawa: CADTH; 2019.

484. The Canadian Agency for Drugs and Technologies in Health (CADTH). Axicabtagene Ciloleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma: Economic Review Report. Ottawa: CADTH; 2019.
485. Norwegian Medicines Agency (NoMA). Axicabtagene ciloleucel (Yescarta) for the treatment of second or later relapsed/refractory diffuse large B-cell lymphoma (DLBCL) and primary mediastinal large B-cell lymphoma (PMBCL): Update of the previous health economic evaluation. 2020.
486. Longworth L, Rowen D. The use of mapping methods to estimate health state utility values. NICE DSU Report 2011.
487. Maurer MJ, Habermann TM, Shi Q, Schmitz N, Cunningham D, Pfreundschuh M, et al. Progression-free survival at 24 months (PFS24) and subsequent outcome for patients with diffuse large B-cell lymphoma (DLBCL) enrolled on randomized clinical trials. Annals of Oncology. 2018;29(8):1822-7.
488. National Institute for Health and Care Excellence (NICE). Pixantrone monotherapy for treating multiple relapsed or refractory aggressive non-Hodgkin's B-cell lymphoma. NICE TA306; 2014.
489. Swinburn P, Lloyd A, Nathan P, Choueiri TK, Cella D, Neary MP. Elicitation of health state utilities in metastatic renal cell carcinoma. Curr Med Res Opin. 2010;26(5):1091-6.
490. Nafees B, Stafford M, Gavriel S, Bhalla S, Watkins J. Health state utilities for non small cell lung cancer. Health Qual Life Outcomes. 2008;6:84.
491. Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, O'Toole A, et al. Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. Health and quality of life outcomes. 2010;8:50-.
492. Tolley K, Goad C, Yi Y, Maroudas P, Haiderali A, Thompson G. Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. Eur J Health Econ. 2013;14(5):749-59.
493. Wailoo A, Hernandez-Alava M, Grimm S, Pudney S, Gomes M, Sadique Z, et al. Comparing the EQ-5D-3L and 5L Versions. What are the Implications for Cost-Effectiveness Estimates? ; 2017.
494. Cubi-Molla P, Mott D, Shah K, Herdman M, Summers Y, Devlin N. Quality of Life in Long-term Cancer Survivors: Implications for Future Health Technology Assessments in Oncology. OHE Consulting Report: London: Office of Health Economics; 2018.
495. Shakeel S, Tung J, Rahal R, Finley C. Evaluation of Factors Associated With Unmet Needs in Adult Cancer Survivors in Canada. JAMA Network Open. 2020;3(3):e200506-e.
496. O''Connor M, O''Donovan B, Drummond F, Donnelly C. The unmet needs of cancer survivors in Ireland: A Scoping Review. National Cancer Registry Ireland; Cork; 2019.
497. Lisy K, Langdon L, Piper A, Jefford M. Identifying the most prevalent unmet needs of cancer survivors in Australia: A systematic review. Asia-Pacific Journal of Clinical Oncology. 2019;15(5):e68-e78.
498. van Leeuwen M, Husson O, Alberti P, Arraras JI, Chinot OL, Costantini A, et al. Understanding the quality of life (QOL) issues in survivors of cancer: towards the development of an EORTC QOL cancer survivorship questionnaire. Health Qual Life Outcomes. 2018;16(1):114.
499. Avis NE, Ip E, Foley KL. Evaluation of the Quality of Life in Adult Cancer Survivors (QLACS) scale for long-term cancer survivors in a sample of breast cancer survivors. Health and Quality of Life Outcomes. 2006;4(1):92.

500. Baker F, Denniston M, Hann D, Gesme D, Reding DJ, Flynn T, et al. Factor structure and concurrent validity of the Satisfaction with Life Domains Scale for Cancer (SLDS-C). J Psychosoc Oncol. 2007;25(2):1-17.
501. Muzzatti B, Annunziata MA. Assessing quality of life in long-term cancer survivors: a review of available tools. Support Care Cancer. 2013;21(11):3143-52.
502. Hobbins A, Barry L, Kelleher D, Shah K, Devlin N, Goni JMR, et al. Utility Values for Health States in Ireland: A Value Set for the EQ-5D-5L. Pharmacoeconomics. 2018;36(11):1345-53.
503. Liu R, Oluwole OO, Diakite I, Botteman MF, Snider JT, Locke FL. Cost effectiveness of axicabtagene ciloleucel versus tisagenlecleucel for adult patients with relapsed or refractory large B-cell lymphoma after two or more lines of systemic therapy in the United States. J Med Econ. 2021;24(1):458-68.
504. National Cancer Control Programme (NCCP). Rituximab, Gemcitabine, Dexamethasone and CISplatin ((R)-GDP) Therapy. 2019.
505. Gisselbrecht C, Schmitz N, Mounier N, Gill DS, Linch DC, Trneny M, et al. Rituximab Maintenance Therapy After Autologous Stem-Cell Transplantation in Patients With Relapsed CD20+ Diffuse Large B-Cell Lymphoma: Final Analysis of the Collaborative Trial in Relapsed Aggressive Lymphoma. Journal of Clinical Oncology. 2012;30(36):4462-9.
506. Assouline S, Li S, Gisselbrecht C, Fogarty P, Hay A, van den Neste E, et al. The conditional survival analysis of relapsed DLBCL after autologous transplant: a subgroup analysis of LY.12 and CORAL. Blood Advances. 2020;4(9):2011-7.
507. Hill BT, Rybicki L, Bolwell BJ, Smith S, Dean R, Kalaycio M, et al. The non-relapse mortality rate for patients with diffuse large B-cell lymphoma is greater than relapse mortality 8 years after autologous stem cell transplantation and is significantly higher than mortality rates of population controls. Br J Haematol. 2011;152(5):561-9.
508. Vanderwalde AM, Sun CL, Laddaran L, Francisco L, Armenian S, Berano-Teh J, et al. Conditional survival and cause-specific mortality after autologous hematopoietic cell transplantation for hematological malignancies. Leukemia. 2013;27(5):1139-45.
509. Myers RM, Hill BT, Shaw BE, Kim S, Millard HR, Battiwalla M, et al. Long-term outcomes among 2-year survivors of autologous hematopoietic cell transplantation for Hodgkin and diffuse large b-cell lymphoma. Cancer. 2018;124(4):816-25.
510. Gill S, Brudno JN. CAR T-Cell Therapy in Hematologic Malignancies: Clinical Role, Toxicity, and Unanswered Questions. American Society of Clinical Oncology Educational Book. 2021(41):e246-e65.
511. Lin VW, Blaylock B, Epstein J, Purdum A. Systematic literature review of health-related quality of life among aggressive non-Hodgkin lymphoma survivors. Curr Med Res Opin. 2018;34(8):1529-35.
512. Qi CZ, Bollu V, Yang H, Dalal A, Zhang S, Zhang J. Cost-Effectiveness Analysis of Tisagenlecleucel for the Treatment of Patients With Relapsed or Refractory Diffuse Large B-Cell Lymphoma in the United States. Clinical Therapeutics. 2021.
513. Wakase S, Teshima T, Zhang J, Ma Q, Fujita T, Yang H, et al. Cost Effectiveness Analysis of Tisagenlecleucel for the Treatment of Adult Patients with Relapsed or Refractory Diffuse Large B Cell Lymphoma in Japan. Transplantation and Cellular Therapy. 2021.
514. Reddy B, Walsh C, Barry M, Kind P, Adams R. Results of the Eq-5d-3l Valuation Project for Ireland. Value in Health. 2015;18(7):A740.
515. Cher BP, Gan KY, Aziz MIA, Lin L, Hwang WYK, Poon LM, et al. Cost utility analysis of tisagenlecleucel vs salvage chemotherapy in the treatment of relapsed/refractory

diffuse large B-cell lymphoma from Singapore's healthcare system perspective. J Med Econ. 2020;23(11):1321-9.
516. Towse A, editor Regenerative Medicine: a European HTA Perspective. ISPOR 17th Annual European Congress; 2014; Amsterdam, the Netherlands.
517. Health Information and Quality Authority (HIQA). Guidelines for the Budget Impact Analysis of Health Technologies in Ireland. 2018.
518. Lamrock F, McCullagh L, Tilson L, Barry M. A retrospective analysis of budget impact models submitted to the National Centre for Pharmacoeconomics in Ireland. The European Journal of Health Economics. 2020;21(6):895-901.
519. National Cancer Control Programme (NCCP). FLAG:Ida 8mg/m2 Therapy. 2018.
520. National Cancer Registry Ireland (NCRI). Cancer Factsheet Non-Hodgkin Lymphoma. 2018.
521. National Cancer Registry Ireland (NCRI). Cancer Trends No 21. non-Hodgkin's lymphoma. 2013.
522. Wästerlid T, Mohammadi M, Smedby KE, Glimelius I, Jerkeman M, Bottai M, et al. Impact of comorbidity on disease characteristics, treatment intent and outcome in diffuse large B-cell lymphoma: a Swedish lymphoma register study. J Intern Med. 2019;285(4):455-68.
523. Wang HI, Smith A, Aas E, Roman E, Crouch S, Burton C, et al. Treatment cost and life expectancy of diffuse large B-cell lymphoma (DLBCL): a discrete event simulation model on a UK population-based observational cohort. Eur J Health Econ. 2017;18(2):255-67.
524. Smith A, Howell D, Crouch S, Painter D, Blase J, Wang H-I, et al. Cohort Profile: The Haematological Malignancy Research Network (HMRN): a UK population-based patient cohort. International Journal of Epidemiology. 2018;47(3):700-g.
525. Seshadri T, Stakiw J, Pintilie M, Keating A, Crump M, Kuruvilla J. Utility of subsequent conventional dose chemotherapy in relapsed/refractory transplant-eligible patients with diffuse large B-cell lymphoma failing platinum-based salvage chemotherapy. Hematology. 2008;13(5):261-6.
526. Jerkeman M. Swedish Lymphoma Registry 2000-2012. Lund, Sweden.; 2014.
527. Coiffier B, Sarkozy C. Diffuse large B-cell lymphoma: R-CHOP failure-what to do? Hematology Am Soc Hematol Educ Program. 2016;2016(1):366-78.
528. Yang X, Laliberté F, Germain G, Raut M, Duh MS, Sen SS, et al. Real-World Characteristics, Treatment Patterns, Health Care Resource Use, and Costs of Patients with Diffuse Large B-Cell Lymphoma in the U.S. The Oncologist. 2021;26(5):e817-e26.
529. GmbH RR. Core Submission Dossier PTJA06: Polatuzumab Vedotin (Polivy) in Combination with Bendamustine and Rituximab is Indicated for the Treatment of Adult Patients with Relapsed/Refractory Diffuse Large B-Cell Lymphoma (DLBCL) who are Not Candidates for Hematopoietic Stem Cell Transplant. 2019.
530. Perica K, Flynn J, Curran KJ, Rivere I, Wang X, Senechal B, et al. Impact of bridging chemotherapy on clinical outcome of CD19 CAR T therapy in adult acute lymphoblastic leukemia. Leukemia. 2021.
531. Sim AJ, Jain MD, Figura NB, Chavez JC, Shah BD, Khimani F, et al. Radiation Therapy as a Bridging Strategy for CAR T Cell Therapy With Axicabtagene Ciloleucel in Diffuse Large B-Cell Lymphoma. Int J Radiat Oncol Biol Phys. 2019;105(5):1012-21.
532. ClinicalTrials.gov. Determine Efficacy and Safety of CTL019 in Pediatric Patients With Relapsed and Refractory B-cell ALL and High Risk B-cell ALL at First Relapse. Determine Feasibility and Safety of CTL019 Therapy in Pediatric Patients With High Risk B-cell ALL That Relapsed < 6 Months Post ALL-HSCT. (ELIANA) 2021 [cited 2021 20

October]. Available from: https://clinicaltrials.gov/ct2/show/NCT02435849?term=tisagenlecleucel&cond=Acute+Lymphoid+Leukemia&draw=5&rank=15.
533. ClinicalTrials.gov. Tisagenlecleucel in Adult Patients with Aggressive B-cell Non-Hodgkin Lymphoma (BELINDA) US National Library of Medicine2018 [cited 2021 13 November].
534. ClinicalTrials.gov. Efficacy of Axicabtagene Ciloleucel Compared to Standard of Care Therapy in Subjects With Relapsed/Refractory Diffuse Large B Cell Lymphoma (ZUMA-7) US National Library of Medicine2018 [cited 2021 13 November].
535. National Centre for Pharmacoeconomics (NCPE). Blinatumomab (Blincyto®) (paediatric ALL): Blinatumomab (Blincyto®) (paediatric ALL) for the treatment of paediatric patients aged 1 year and older with Philadelphia chromosome negative B-cell precursor ALL which is refractory or in relapse after receiving at least two prior therapies or in relapse after receiving prior allogeneic hematopoietic stem cell transplantation. 2019.
536. Towse A, Fenwick E. Uncertainty and Cures: Discontinuation, Irreversibility, and Outcomes-Based Payments: What Is Different About a One-Off Treatment? Value in Health. 2019;22(6):677-83.
537. Wenzl M, Chapman S. OECD Health Working Paper No. 115: Perfoormance-based managed entry agreements for new medicines in OECD countries and EU member states. How they work and possible improvements going forward.: Organisation for Economic Co-operation and Development (OECD), ; 2019.
538. Garrison LP, Jr., Towse A, Briggs A, de Pouvourville G, Grueger J, Mohr PE, et al. Performance-Based Risk-Sharing Arrangements. Good Practices for Design, Implementation, and Evaluation: Report of the ISPOR Good Practices for Performance-Based Risk-Sharing Arrangements Task Force. Value in Health. 2013;16(5):703-19.
539. Gonçalves FR, Santos S, Silva C, Sousa G. Risk-sharing agreements, present and future. Ecancermedicalscience. 2018;12:823-.
540. National Institute for Health and Care Excellence (NICE). Axicabtagene ciloleucel for treating diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma after 2 or more systemic therapies: Information for the public 2019 [cited 2021 11 December]. Available from: https://www.nice.org.uk/Guidance/TA559/InformationForPublic.
541. Coverage Options For Promising Technologies: Medicare's 'Coverage With Evidence Development'. Health Affairs. 2006;25(5):1218-30.
542. Grimm S, Strong M, Brennean A, Wailoo A. Framework For Analysing Risk in Health Technology Assessments And Its Application to Managed Entry Agreements: Report By The Decision Support Unit. University of Sheffield; 2016.
543. Michelsen S, Nachi S, Van Dyck W, Simoens S, Huys I. Barriers and Opportunities for Implementation of Outcome-Based Spread Payments for High-Cost, One-Shot Curative Therapies. Front Pharmacol. 2020;11:594446-.
544. Cole A, Cubi-Molla P, Pollard J, Sim D, Sullivan R, Sussex J, et al. Making Outcome-Based Payment a Reality in the NHS. 2019.
545. Neumann PJ, Chambers JD, Simon F, Meckley LM. Risk-sharing arrangements that link payment for drugs to health outcomes are proving hard to implement. Health Aff (Millwood). 2011;30(12):2329-37.
546. Jørgensen J, Mungapen L, Kefalas P. Data collection infrastructure for patient outcomes in the UK - opportunities and challenges for cell and gene therapies launching. J Mark Access Health Policy. 2019;7(1):1573164.

547. Jørgensen J, Kefalas P. Upgrading the SACT dataset and EBMT registry to enable outcomes-based reimbursement in oncology in England: a gap analysis and top-level cost estimate. J Mark Access Health Policy. 2019;7(1):1635842.
548. European Society for Blood and Marrow Transplantation (EBMT). Cell Therapy - MED-A Form 2017 [Available from: https://www.ebmt.org/sites/default/files/migration_legacy_files/document/25%20Cellular%20Therapy%20MED-A.pdf.
549. European Society for Blood and Marrow Transplantation (EBMT). Registry Function. Version 8. 2018 [Available from: https://www.ebmt.org/sites/default/files/2018-08/EBMTRegistryFunction.pdf.
550. National Cancer Registry Ireland (NCRI). About us 2021 [cited 2021 11 December]. Available from: https://www.ncri.ie/about-us.
551. Makady A, van Acker S, Nijmeijer H, de Boer A, Hillege H, Klungel O, et al. Conditional Financing of Drugs in the Netherlands: Past, Present, and Future—Results From Stakeholder Interviews. Value in Health. 2019;22(4):399-407.
552. Montilla S, Xoxi E, Russo P, Cicchetti A, Pani L. MONITORING REGISTRIES AT ITALIAN MEDICINES AGENCY: FOSTERING ACCESS, GUARANTEEING SUSTAINABILITY. International Journal of Technology Assessment in Health Care. 2015;31(4):210-3.
553. Garattini L, Curto A, van de Vooren K. Italian risk-sharing agreements on drugs: are they worthwhile? Eur J Health Econ. 2015;16(1):1-3.
554. Facey KM, Espin J, Kent E, Link A, Nicod E, O'Leary A, et al. Implementing Outcomes-Based Managed Entry Agreements for Rare Disease Treatments: Nusinersen and Tisagenlecleucel. PharmacoEconomics. 2021.
555. Agenzia Italiana del Farmaco (AIFA). L'uso dei farmaci in Italia. Rapporto nazionale anno 2012. Osservatorio Nazionale sull'impiego dei Medicinali Roma2013 [cited 2022 22 July]. Available from: http://www.agenziafarmaco.gov.it/sites/default/files/not_found/it/index_it.html.
556. Gerkens S, Neyt M, San Miguel L, Vinck I, Thiry N, Cleemput I. How to improve the Belgian process for managed entry agreements? An analysis of the Belgian and international experience. Brussels, Belgium: Belgian Health Care Knowlegde Centre (KCE),; 2017.
557. European Medicines Agency (EMA). CAR T-cell therapy Registries Workshop. Appendix 1: Proposed data elements relating to Efficacy and to Safety 2018 [cited 2021 24 November]. Available from: https://www.ema.europa.eu/en/documents/report/appendix-1-proposed-data-elements-relating-efficacy-safety-car-t-cell-therapy-registries-workshop_en.pdf.
558. EUnetHTA Joint Action 3. EUnetHTA WP5B PLEG pilot on nusinersen (Spinraza): common evidence gaps report. 2020 [cited 2021 19 October]. Available from: https://www.eunethta.eu/wp-content/uploads/2020/05/EUnetHTA-PLEG_FP_01_Nusinersen_Common-Evidence-Gaps-report.pdf.
559. Jørgensen J, Kefalas P. The use of innovative payment mechanisms for gene therapies in Europe and the USA. Regenerative Medicine. 2021;16(4):405-22.
560. Pasquini MC, Hu Z-H, Curran K, Laetsch T, Locke F, Rouce R, et al. Real-world evidence of tisagenlecleucel for pediatric acute lymphoblastic leukemia and non-Hodgkin lymphoma. Blood Advances. 2020;4(21):5414-24.
561. Hutton J, Trueman P, Henshall C. Coverage with evidence development: an examination of conceptual and policy issues. Int J Technol Assess Health Care. 2007;23(4):425-32.

562. Drummond M. When do performance-based risk-sharing arrangements make sense? Eur J Health Econ. 2015;16(6):569-71.
563. Drummond M, editor Health Technology Assessment: The New Frontier for Gene Therapies in Europe. ISPOR Europe; 2021; Vitrual.
564. Puig-Peiró R, Mestre-Ferrandiz J, Sussex J, Towse A. RS1 Literature Review On Patient Access Schemes, Flexible Pricing Schemes And Risk Sharing Agreements For Medicines. Value in Health. 2011;14(7):A243.
565. Antonanzas F, Juárez-Castelló C, Lorente R, Rodríguez-Ibeas R. The Use of Risk-Sharing Contracts in Healthcare: Theoretical and Empirical Assessments. Pharmacoeconomics. 2019;37(12):1469-83.

Appendices

Appendix A Chapter 2

A.1 Chapter 2 Systematic Literature Review

PICOS Criteria

Table A1. PICOS Chapter 2 systematic literature review of treatments for relapsed/refractory acute lymphoblastic leukaemia

Criteria	Inclusion	Exclusion
Population	Patients with relapsed or refractory B-cell ALL up to 25 years of age, any sex, any ethnicity	Patients aged >25 years of age Treatment naïve patients Patients with T-cell ALL
Intervention	Tisagenlecleucel monotherapy	CD19 CAR T-cell therapy used in combination therapy; Clofarabine; Inotuzumab Standard of care not otherwise defined
Comparator	FLA(G)-IDA (fludarabine, cytarabine, idarubicin, granulocyte colony stimulating factor); Blinatumomab; Autologous/ allogeneic SCT; Any of the included interventions	
Outcome	Primary: Survival: Overall survival Progression-free survival Event-free survival Leukaemia-free survival Secondary: Health-related quality of life Adverse event	Response rates: Objective response rate Duration of response Complete response Partial response Pharmacokinetic/pharmacodynamic outcomes Social outcomes
Study Design	Prospective randomised controlled trials Phase II non-randomised or single-arm trials Prospective observational studies Patient registries	Single-centre trials Retrospective studies Reviews Letters Comments Editorials Case studies/reports Narrative publications Biomarker/prognostic studies Conference abstracts without full text Expanded access programmes

ALL: Acute lymphoblastic leukaemia; **SCT:** Stem cell transplant.

A.2 Chapter 2 Search Strategy

The search strategy for the systematic review of treatments for R/R ALL is presented in Table A2.

Table A2. Chapter 2 (systematic literature review of treatments for relapsed/refractory acute lymphoblastic leukaemia) search strategy

EMBASE 21 November 2020
#1 ('acute lymphoblastic leukemia'/exp OR 'acute lymphoblastic leukemia' OR lymphoblast*) AND [article]/lim AND [humans]/lim AND [english]/lim AND ([embase]/lim OR [medline]/lim) AND [2000-2020]/py #2 (relapsed OR relapses OR relapsing OR refractory OR failed OR failure OR 'transplant ineligible' OR 'stem cell transplant ineligible' OR 'sct ineligible') AND [article]/lim AND [humans]/lim AND [english]/lim AND ([embase]/lim OR [medline]/lim) AND [2000-2020]/py #3 ('hematopoietic stem cell transplantation' OR 'stem cell transplant' OR 'stem cell transplantation' OR sct OR hsct OR 'allogeneic transplant' OR 'autologous transplant' OR 'hematopoietic transplant' OR 'hematopoietic cell transplantation' OR ctl019 OR tisagenlecleucel OR kymriah* OR 'cart-t cell therapy' OR 'chimeric antigen receptor t cell therapy' OR chemotherapy OR blinatumomab OR 'fla ida' OR fludarabine OR cytarabine OR idarubicin) AND [article]/lim AND [humans]/lim AND [english]/lim AND ([embase]/lim OR [medline]/lim) AND [2000-2020]/py #4 (infan*:ab,ti OR newborn*:ab,ti OR 'new born':ab,ti OR 'new borns':ab,ti OR baby*:ab,ti OR babies:ab,ti OR neonat*:ab,ti OR child*:ab,ti OR schoolchild*:ab,ti OR kid:ab,ti OR kids:ab,ti OR toddler*:ab,ti OR adoles*:ab,ti OR teen*:ab,ti OR boy*:ab,ti OR girl*:ab,ti OR minors*:ab,ti OR underag*:ab,ti OR 'under age':ab,ti OR 'under aged':ab,ti OR juvenil*:ab,ti OR youth*:ab,ti OR kindergar*:ab,ti OR puber*:ab,ti OR pubescen*:ab,ti OR prepubescen*:ab,ti OR prepuberty*:ab,ti OR pediatric*:ab,ti OR paediatric*:ab,ti OR peadiatric*:ab,ti) AND [article]/lim AND [humans]/lim AND [english]/lim AND ([embase]/lim OR [medline]/lim) AND [2000-2020]/py #5 #1 AND #2 #6 #3 AND #5 #7 #4 AND #6
MEDLINE (via EBSCO)
#1 “acute lymphoblastic leukemia” OR lymphoblast* #2 relapsed OR relapses OR relapsing OR refractory OR failed OR failure OR ‘transplant ineligible’ OR ‘stem cell transplant ineligible’ OR ‘SCT ineligible’) #3 "Stem Cell Transplantation" OR "Hematopoietic Stem Cell Transplantation" OR "stem cell transplant" OR "stem cell transplantation" OR SCT OR HSCT OR “allogeneic transplant” OR “autologous transplant” OR "hematopoietic transplant" OR "hematopoietic cell transplantation" OR CTL019 OR tisagenlecleucel OR Kymriah* OR “CART-T cell therapy” OR “chimeric antigen receptor T cell therapy” OR chemotherapy OR blinatumomab OR FLA-IDA OR fludarabine OR cytarabine OR idarubicin #4 (paediatrics or pediatrics or children or child or young person) AND (infants or baby or newborn or neonate) AND (adolescents or teenagers) AND (minors or youth or children or adolescent) #5: #1 AND #2 #6: #3 AND #5 #7: #4 AND #6 #8: #7 Limit to articles published from January 2000 to present

#9: #8 Limit to articles on human subjects

#10: #9 Limit to articles published in English

#11: #10 Limit to articles that are not case reports

CENTRAL (via Cochrane Library)

#1 MeSH descriptor: [Precursor Cell Lymphoblastic Leukemia-Lymphoma] explode all trees

#2 'ACUTE LYMPHOBLASTIC LEUKAEMIA' OR 'ACUTE LYMPHOBLASTIC LEUKEMIA' OR LYMPHOBLAST*

#3 relapsed OR relapses OR relapsing OR refractory OR failed OR failure OR 'transplant ineligible' OR 'stem cell transplant ineligible' OR 'SCT ineligible'

#4 "Stem Cell Transplantation" OR "Hematopoietic Stem Cell Transplantation" OR "stem cell transplant" OR "stem cell transplantation" OR SCT OR HSCT OR "allogeneic transplant" OR "autologous transplant" OR "hematopoietic transplant" OR "hematopoietic cell transplantation" OR "chimeric antigen receptor T cell therapy" OR tisagenlecleucel OR chemotherapy OR blinatumomab OR FLA-IDA OR fludarabine OR cytarabine OR idarubicin

#5 (paediatrics or pediatrics or children or child or young person) OR (infants or baby or newborn or neonate) OR (adolescents or teenagers) OR (minors or youth or children or adolescent)

#6 #1 AND #2

#7 #6 AND #3

#8 #7 AND #4

#9 #8 AND #5

#10 #9 with Cochrane Library publication date Between Jan 2000 and Nov 2020, in Cochrane Reviews, Trials

A.3 Newcastle-Ottawa Scale Coding Manual

NEWCASTLE-OTTAWA SCALE CODING MANUAL FOR COHORT STUDIES

1) Representativeness of the Exposed Cohort

Item is assessing the representativeness of exposed individuals in the community, not the representativeness of the sample of women from some general population. For example, subjects derived from groups likely to contain middle class, better educated, health oriented women are likely to be representative of postmenopausal oestrogen users while they are not representative of all women (e.g. members of a health maintenance organisation (HMO) will be a representative sample of oestrogen users. While the HMO may have an under-representation of ethnic groups, the poor, and poorly educated, these excluded groups are not the predominant users of oestrogen).

Allocation of points as per rating sheet.

2) Selection of the Non-Exposed Cohort

Allocation of points as per rating sheet.

3) Ascertainment of Exposure

Allocation of points as per rating sheet.

4) Demonstration That Outcome of Interest Was Not Present at Start of Study

In the case of mortality studies, outcome of interest is still the presence of a disease/ incident, rather than death. That is to say that a statement of no history of disease or incident earns 1 point.

COMPARABILITY

1) Comparability of Cohorts on the Basis of Design or Analysis

A maximum of 2 points can be allotted in this category.

Either exposed and non-exposed individuals must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the relative risk for the exposure of interest is adjusted for the confounders listed, then the groups will be considered comparable on each variable used in the adjustment.

There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never)

Age = 1 star , Other controlled factors = 1 star

OUTCOME

1) Assessment of Outcome

For some outcomes (e.g. fractured hip), reference to the medical record is sufficient to satisfy the requirement for confirmation of the fracture. This would not be adequate for vertebral fracture outcomes where reference to x-rays would be required.

a) Independent or blind assessment stated in the paper, or confirmation of the outcome by reference to secure records (x-rays, medical records, etc.) 1 star.
b) Record linkage (e.g. identified through ICD codes on database records) 1 star.
c) Self-report (i.e. no reference to original medical records or x-rays to confirm the outcome).
d) No description.

2) Was Follow-Up Long Enough for Outcomes to Occur

An acceptable length of time should be decided before quality assessment begins (e.g. 5 yrs. for exposure to breast implants).

3) Adequacy of Follow Up of Cohorts

This item assesses the follow-up of the exposed and non-exposed cohorts to ensure that losses are not related to either the exposure or the outcome.

Allocation of points as per rating sheet.

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE
COHORT STUDIES ADAPTED FOR USE WITH SINGLE-ARM STUDIES

Selection Domain

1) Representativeness of the exposed cohort
   a) truly representative of the average R/R DLBCL/ALL population in the community **1 star**
   b) somewhat representative of the average R/R DLBCL/ALL population in the community **1 star**
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort

3) Ascertainment of exposure
   a) secure record (e.g. surgical records) **1 star**
   b) structured interview **1 star**
   c) written self-report
   d) no description

4) Demonstration that outcome of interest was not present at start of study
   a) yes **1 star**
   b) no

Outcomes Domain

1) Assessment of outcome
   a) independent blind assessment **1 star**
   b) record linkage
   c) self report
   d) no description

2) Was follow-up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest) **1 star**
   b) no

3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for **1 star**
   b) subjects lost to follow up unlikely to introduce bias - small number lost: <5% lost to follow up, or description provided of those lost)
   c) follow up rate <80% and no description of those lost
   d) no statement

A.4 Chapter 2 Excluded Studies

A selection of studies excluded on the basis of full-text screening, in the systematic review of treatments for R/R ALL, are presented in Table A3.

Table A3. Chapter 2 (systematic literature review of treatments for relapsed/refractory acute lymphoblastic leukaemia) selection of excluded studies

Study (year)	Authors	Reason for Exclusion
Allogeneic Stem Cell Transplantation for Acute Lymphoblastic Leukemia in Adolescents and Young Adults (2019)	Hangai, M., Urayama, K. Y., Tanaka, J., Kato, K., Nishiwaki, S., Koh, K., et al.	Population
Salvage therapy for children with relapsed or refractory Philadelphia chromosome-positive acute lymphoblastic leukemia (2017)	Kodama, Y., Manabe, A., Kawasaki, H., Kato, I., Kato, K., Sato, A., et al.	Intervention
Pathways through relapses and deaths of children with acute lymphoblastic leukemia: Role of allogeneic stem-cell transplantation in Nordic data (2006)	Saarinen-Pihkala, U. M., Heilmann, C., Winiarski, J., Glomstein, A., Abrahamsson, J., Arvidson, J., et al.	Population
Comparable results in patients with acute lymphoblastic leukemia after related and unrelated stem cell transplantation (2006)	Dahlke, J., Kröger, N., Zabelina, T., Ayuk, F., Fehse, N., Wolschke, C., et al.	Population
Survival improvements in adolescents and young adults after myeloablative allogeneic transplantation for acute lymphoblastic leukemia (2014)	Wood, W. A., Lee, S. J., Brazauskas, R., Wang, Z., Aljurf, M. D., Ballen, K. K., et al.	Population
Allogeneic Stem Cell Transplantation from HLA-Mismatched Donors for Pediatric Patients with Acute Lymphoblastic Leukemia Treated According to the 2003 BFM and 2007 International BFM Studies: Impact of Disease Risk on Outcomes (2018)	Dalle, J. H., Balduzzi, A., Bader, P., Lankester, A., Yaniv, I., Wachowiak, J., et al.	Population
Long-term results of the Italian association of pediatric hematology and oncology (AIEOP) Studies 82, 87, 88, 91 and 95 for childhood acute lymphoblastic leukemia (2010)	Conter, V., Aricò, M., Basso, G., Biondi, A., Barisone, E., Messina, C., et al.	Population
FLAG/FLAG-IDA regimen for children with relapsed/refractory acute leukemia in the era of targeted novel therapies (2019)	Mustafa, O., Abdalla, K., AlAzmi, A. A., Elimam, N., Abrar, M. B., Jastaniah, W., et al.	Study Design

Appendix B Chapter 3

B.1 Expert Elicitation Questions

Table A4. Elicitation questions in expert elicitation exercise (Chapter 3)

	Parameter	Elicitation Question
Q1	Assumption of Cure	Is it reasonable to assume that if ten patients are treated with tisagenlecleucel, a proportion will eventually be subject to age- and sex-matched general population mortality (i.e. no longer at risk of excess mortality compared to general population)?
Q2a	Time-Point of Greatest Risk of Excess Mortality*	If all ten patients are considered to have a lifelong excess risk of mortality compared to the general population, at what time point would you consider this excess risk to be greatest?
Q2b	Cure Time-Point†	How many years after receiving CD19 CAR T-cell therapy‡ would you expect this excess risk of mortality to cease?
Q2c	Cure Fraction – Tisagenlecleucel†	Of the patients treated with CD19 CAR T-cell therapy‡, how many would you expect to be subject to age- and sex-matched general population mortality following treatment with CD19 CAR T-cell therapy?
Q2d	Cure Fraction – Routine Care†	If ten patients are treated with current standard of care, how many would you expect to be subject to age- and sex-matched general population mortality following treatment?
Q3	Proportion of Patients Proceeding to AlloSCT Following Tisagenlecleucel	How many of the ten patients treated with CD19 CAR T-cell therapy‡ would you expect to receive subsequent SCT?
Q4	5-Year OS of Patients Proceeding to AlloSCT Following Tisagenlecleucel	Of the patients treated with subsequent alloSCT following treatment with CD19 CAR T-cell therapy‡, how many would you expect to be alive at five years post-CD19 CAR T-cell infusion infusion?
Q5	5-Year OS of Patients who do not Proceed to AlloSCT Following Tisagenlecleucel	If none of the ten patients treated with CD19 CAR T-cell therapy‡ receive subsequent SCT, how many would you expect to be alive at five years post-CD19 CAR T-cell infusion?
Q6	Proportion of Patients Proceeding to AlloSCT Following Routine Care	If ten patients are treated with current standard of care, how many would you expect to receive subsequent SCT?
Q7	5-Year OS of Patients Proceeding to AlloSCT Following Routine Care	Of the patients treated with subsequent SCT following treatment with current standard of care in the third-line setting, how many would you expect to be alive at five years post-treatment?
Q8	5-Year OS of Patients who do not Proceed to AlloSCT Following Routine Care	If none of the ten patients treated with current standard of care in the third-line setting receive subsequent SCT, how many would you expect to be alive at five years post-treatment?

AlloSCT: Allogeneic stem cell transplant; **OS:** Overall survival; **SCT:** Stem cell transplant.
*Only applicable to participants who responded “no” to Question 1.
†Only applicable to participants who responded “yes” to Question 1.
‡Tisagenlecleucel.

B.2 Expert Elicitation Rationales

Table A5. Summary of rationales provided by experts in expert elicitation exercise (Chapter 3)

Parameter	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5
Q1 Assumption of Cure	NR	NR	NR	NR	NR
Q2a Time-Point of Greatest Risk of Excess Mortality	Highest risk initially in first year, decrease gradually over time. Uncertain risk remains after 5 years due to infections (B-cell aplasia) or secondary cancers due to exposure to prior chemotherapy.	N/A	Early mortality due to relapse, infections or other complications of treatment. Risk will never hit baseline due to B-cell aplasia.	Risk remains higher due to long-term complications. Risk of relapse highest in initial period.	N/A
Q2b Cure Time-Point	N/A	If real life data reflects trial data, expect by after 2 years; no excess mortality due to relapse.	N/A	N/A	Negligible risk of relapse after this time.
Q2c Cure Fraction – Tisagenlecleucel	N/A	NR	N/A	N/A	As per ELIANA.
Q2d Cure Fraction – Routine Care	N/A	Patients are generally high-risk relapses. Expect very poor outcomes and not many to be cured of disease.	N/A	N/A	Current long-term survival.
Q3 Proportion of Patients	30% to 50% relapse risk requiring further	NR	Many patients have failed alloSCT. Not all	Proportion will require next-line treatment.	Relapse rate in ELIANA.

Proceeding to AlloSCT Following Tisagenlecleucel	treatment (incl. alloSCT).		patients who fail CAR-T will be eligible for alloSCT. Some patients will not require alloSCT.		Some patients will receive alloSCT.
Q4 5-Year OS of Patients Proceeding to AlloSCT Following Tisagenlecleucel	20% to 30% who relapse will be salvaged by further treatment.	NR	Lower than general survival for patients in CR3 (~20%).	Reflects survival following multiple relapses.	NR
Q5 5-Year OS of Patients who do not Proceed to AlloSCT Following Tisagenlecleucel	50% to 60% of will remain in long-term remission without further treatment.	NR	Reflective of ELIANA data.	NR	NR
Q6 Proportion of Patients Proceeding to AlloSCT Following Routine Care	30% to 40% will achieve a sufficiently good remission to benefit from alloSCT.	Usually high-risk patients, all eligible, provided good enough remission.	NR*	Established resmission rates to enable alloSCT.	Dependent upon factors such as fitness and response.
Q7 5-Year OS of Patients Proceeding to AlloSCT Following Routine Care	Evidence supports 5-year survival of 10% to 20%.	High-risk relapse. Low survival.	NR	NR	NR
Q8 5-Year OS of Patients who do not Proceed to AlloSCT Following Routine Care	NR	NR	NR	NR	NR

AlloSCT: Allogeneic stem cell transplant; **N/A:** Not applicable; **NR:** No rationale provided; **OS:** Overall survival.

*Expert indicated that they were unsure of the population that this question referred to.

Appendix C Chapter 4

C.1 Chapter 4 Search Strategy

The search strategy used to identify utility data in paediatric and young adult patients with R/R ALL is presented in Table A6.

Table A6. Chapter 4 (systematic literature review of utility data for relapsed/refractory acute lymphoblastic leukaemia) search strategy

EMBASE 09 January 2021
#1 'acute lymphoblastic leukemia'/exp OR 'acute lymphoblastic leukemia' OR 'acute lymphoblastic leukaemia'/exp OR 'acute lymphoblastic leukaemia' OR 'acute lymphoblast*' #2 relapsed OR relapses OR relapsing OR refractory OR failed OR failure #3 infan*:ab,ti OR newborn*:ab,ti OR 'new born':ab,ti OR 'new borns':ab,ti OR baby*:ab,ti OR babies:ab,ti OR neonat*:ab,ti OR child*:ab,ti OR kid:ab,ti OR kids:ab,ti OR toddler*:ab,ti OR adoles*:ab,ti OR teen*:ab,ti OR minors*:ab,ti OR underag*:ab,ti OR 'under age':ab,ti OR 'under aged':ab,ti OR juvenil*:ab,ti OR youth*:ab,ti OR puber*:ab,ti OR pubescen*:ab,ti OR prepubescen*:ab,ti OR pediatric*:ab,ti OR paediatric*:ab,ti OR peadiatric*:ab,ti OR 'young adult' #4 #1 AND #2 AND #3 #5 'european quality of life 5 dimensions questionnaire' OR 'eq-5d' OR 'eq 5d' OR 'eq5d' OR 'euroqol' OR 'eq5d*' #6 'standard gamble' OR 'sg' OR 'time trade off' OR 'time trade-off' OR 'time tradeoff' OR 'tto #7 utilit* OR "health utilit$" OR "health state$ utilit$" OR "health state$ utilit$ value$" OR 'hsu' OR 'hsuv' #8 exp AND quality AND of AND life #9 'quality-adjusted life year' OR 'quality adjusted life year' #10 'health-related quality of life' OR 'health-related quality-of-life' OR 'health related quality of life' OR 'health related quality-of-life' OR 'hrqol' #11 ((((((((((sf36 OR sf) AND 36 OR short) AND form AND 36 OR shortform) AND 36 OR short) AND form36 OR shortform36 OR sf) AND thirtysix OR sfthirtysix OR sfthirty) AND six OR sf) AND thirty AND six OR shortform) AND thirtysix OR shortform) AND thirty AND six OR short) AND form AND thirtysix OR short) AND form AND thirty AND six #12 #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 #13 #4 AND #12 #14 #4 AND #12 AND [humans]/lim AND [english]/lim AND [embase]/lim AND [2000-2021]/py
MEDLINE (via EBSCO)
S1 'acute lymphoblastic leukemia'/exp OR 'acute lymphoblastic leukemia' OR 'acute lymphoblastic leukaemia'/exp OR 'acute lymphoblastic leukaemia' OR 'acute lymphoblast*' S2 relapsed OR relapses OR relapsing OR refractory OR failed OR failure S3 (paediatrics or pediatrics or children or child or young person) OR (infants or baby or newborn or neonate) OR (minors or youth or children or adolescent or young adult) S4 S1 AND S2 AND S3 S5 'european quality of life 5 dimensions questionnaire' OR 'eq-5d' OR 'eq 5d' OR 'eq5d' OR 'euroqol' OR 'eq5d*' S6 'standard gamble' OR 'sg' OR 'time trade off' OR 'time trade-off' OR 'time tradeoff' OR 'tto' S7 ("European organization for research and treatment of cancer") OR eortc qlq-c30

S8 "The pediatric quality of life inventory" OR PedsQL

S9 functional assessment of cancer therapy - general OR fact-g

S10 ((((((((((sf36 OR sf) AND 36 OR short) AND form AND 36 OR shortform) AND 36 OR short) AND form36 OR shortform36 OR sf) AND thirtysix OR sfthirtysix OR sfthirty) AND six OR sf) AND thirty AND six OR shortform) AND thirtysix OR shortform) AND thirty AND six OR short) AND form AND thirtysix OR short) AND form AND thirty AND six

S11 quality of life OR (quality of life or well being or well-being or health-related quality of life) OR qol OR (hrqol or health-related quality of life)

S12 quality adjusted life years OR qaly OR qaly analysis

S13 S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12

S14 S4 AND S13

S15 S14 Limit to articles on human subjects.

S16 S15 Limit to articles published in English

S17 S16 Limit to articles published from January 01 2000

CENTRAL (via Cochrane Library)

#1 MeSH descriptor: [Leukemia] explode all trees

#2 acute lymphoblastic leukaemia OR ACUTE LYMPHOBLASTIC LEUKEMIA OR LYMPHOBLAST

#3 relapsed OR relapses OR relapsing OR refractory OR failed OR failure

#4 #1 AND #2 and #3

#5 'european quality of life 5 dimensions questionnaire' OR 'eq-5d' OR 'eq 5d' OR 'eq5d' OR 'euroqol'

#6 "standard gamble" OR 'sg' OR 'time trade off' OR 'time trade-off' OR 'time tradeoff' OR 'tto'

#7 "European organization for research and treatment of cancer" OR eortc qlq-c30

#8 "The pediatric quality of life inventory" OR PedsQL

#9 "functional assessment of cancer therapy general" OR "fact g"

#10 "SF36" OR "short form 36"

#11 quality of life OR (quality of life or well being or wellbeing or health-related quality of life) OR qol OR (hrqol or health-related quality of life)

#12 quality adjusted life years OR qaly OR qaly analysis

#13 #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12

#14 #13 AND #4

#15 #14 with Cochrane Library publication date Between Jan 2000 and Jan 2021

C.2 Chapter 4 Excluded Studies

A selection of studies excluded on the basis of full-text screening, in the systematic review of utility data for R/R ALL, are presented in Table A7.

Table A7. Chapter 4 (systematic literature review of utility data for relapsed/refractory acute lymphoblastic leukaemia) selection of excluded studies

Study (year)	Authors	Reason for Exclusion
Cost-effectiveness analysis of tisagenlecleucel in the treatment of relapsed or refractory B-cell acute lymphoblastic Leukaemia in children and young adults in Spain (2020)	Maria, J. Santasusana, R., De Andrés Saldaña, A., García-Muñoz, N., and Gostkorzewicz, J.	Population
Levofloxacin prophylaxis in hospitalized children with leukemia: A cost-utility analysis (2020)	Maser, B., Pelland-Marcotte, M. C., Alexander, S., Sung, L. and Gupta, S.	Population
The Effectiveness of Incorporating a Play-based Intervention to Improve Functional Mobility for a Child with Relapsed Acute Lymphoblastic Leukaemia: A Case Report (2016)	Vercher, P., Hung, Y. J. and Ko, M.	Study Design
Cost effectiveness of chimeric antigen receptor T-cell therapy in relapsed or refractory pediatric B-cell acute lymphoblastic leukemia (2018)	Lin, J. K., Lerman, B. J., Barnes, J. I., Boursiquot, B. C., Tan, Y. J., Robinson, A. Q. L., et al.	Population
Patient-reported quality of life (QOL) following CTL019 in pediatric and young adult patients (pts) with relapsed/refractory (r/r) b-cell acute lymphoblastic leukemia (B-ALL) (2017)	Dietz, A. C., Grupp, S. A., Laetsch, T. W., Stefanski, H., Myers, G. D., Bittencourt, H., et al.	Outcome
Developmental differences in health-related quality of life in adolescent and young adult cancer survivors (2020)	Becktell, K., Simpson, P., Phelan, R., Schmidt, D., Anderson, L., Nichols, J., et al	Population
Self-reported quality of life in long-term survivors of childhood lymphoblastic malignancy treated with hematopoietic stem cell transplantation versus conventional therapy (2013)	Sundberg, K. K., Wettergren, L., Frisk, P. and Arvidson, J.	Outcome
Patient-reported quality of life after tisagenlecleucel infusion in children and young adults with relapsed or refractory B-cell acute lymphoblastic leukaemia: a global, single-arm, phase 2 trial (2019)	Laetsch, T. W., Myers, G. D., Baruchel, A., Dietz, A. C., Pulsipher, M. A., Bittencourt, H., et al.	Outcome

Appendix D Chapter 5

D.1 Model Visual Fit

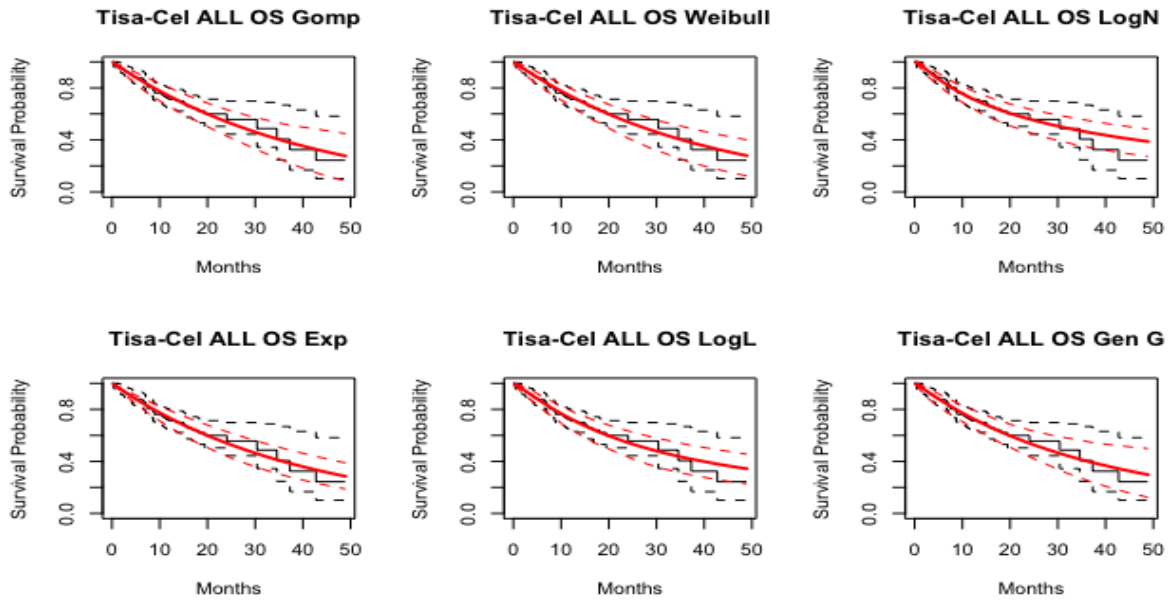

Figure A1. Parametric model fits to pooled ELIANA and ENSIGN (tisagenlecleucel) overall survival data*

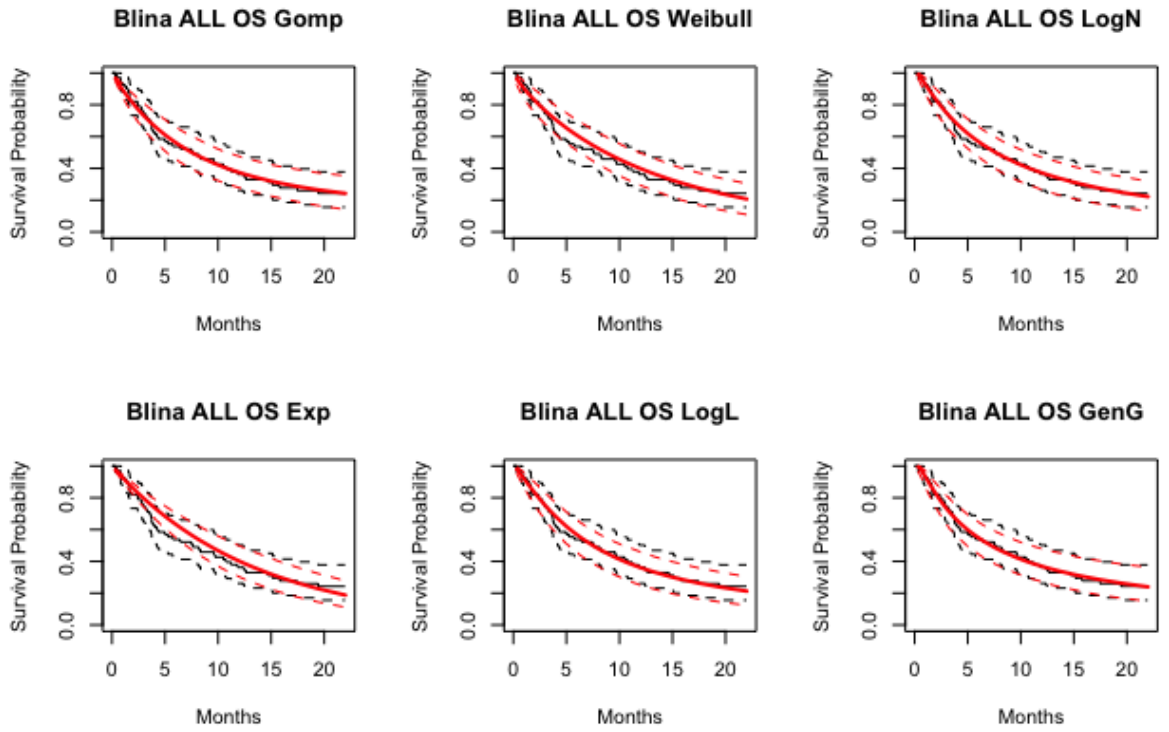

Figure A2. Parametric model fits to NCT01471782 (blinatumomab) overall survival data*

***ALL:** Acute lymphoblastic leukaemia; **Exp:** Exponential; **GenG:** Generalised gamma; **Gomp:** Gompertz; **LogL:** Log-logistic; **LogN:** Log-normal; **OS:** Overall survival.

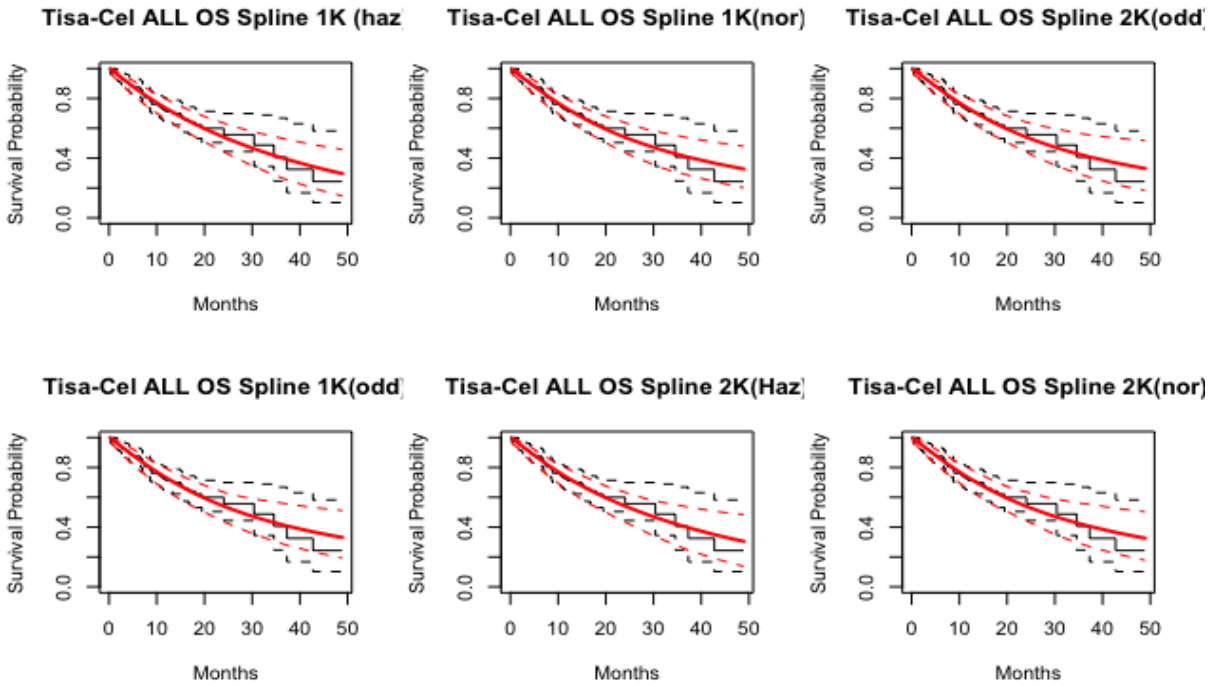

Figure A3. Spline model fits to pooled ELIANA and ENSIGN (tisagenlecleucel) overall survival data†

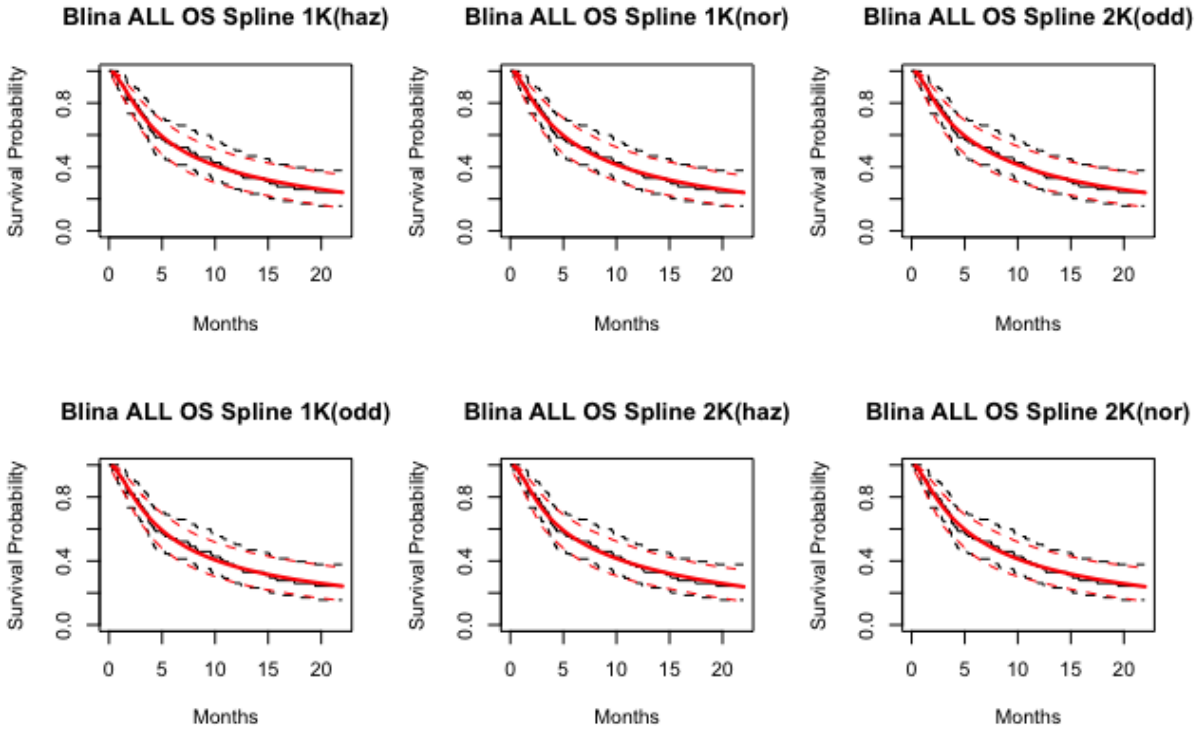

Figure A4. Spline model fits to the NCT01471782 (blinatumomab) overall survival data†

†**1K (haz):** One-knot hazard spline; **1K (nor):** One-knot normal spline; **1K (odd):** One-knot odds spline; **2K (haz):** Two-knot hazard spline; **2K (nor):** Two-knot normal spline; **2K (odd):** Two-knot odds spline; **ALL:** Acute lymphoblastic leukaemia; **OS:** Overall survival.

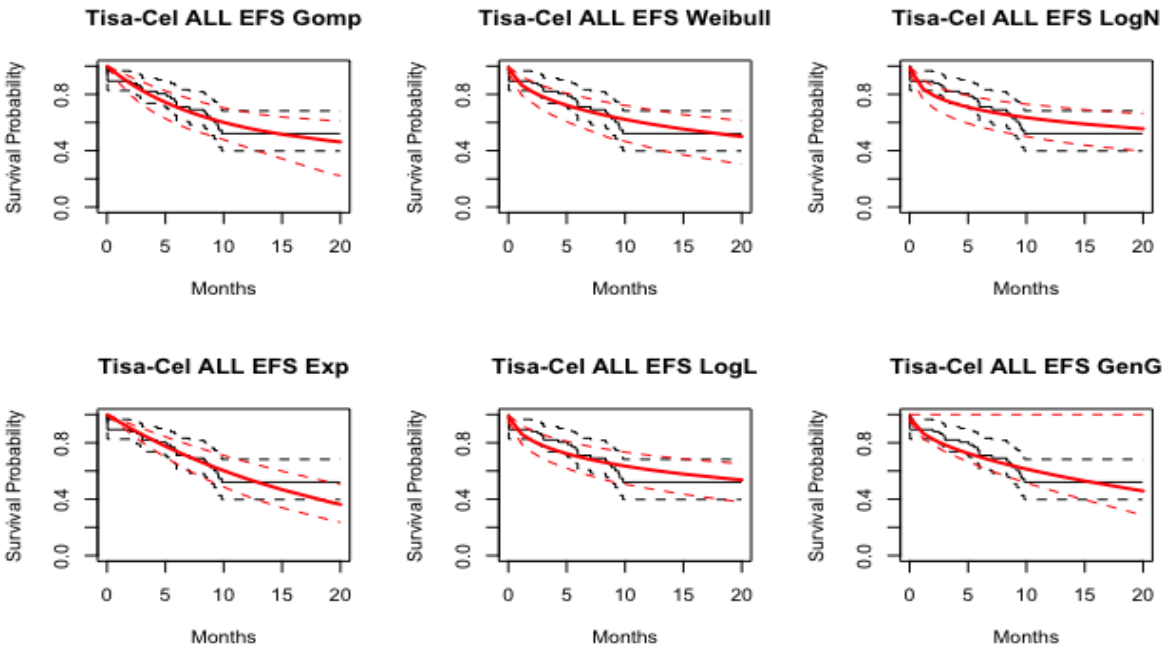

Figure A5. Parametric model fits to ELIANA (tisagenlecleucel) event-free survival data‡

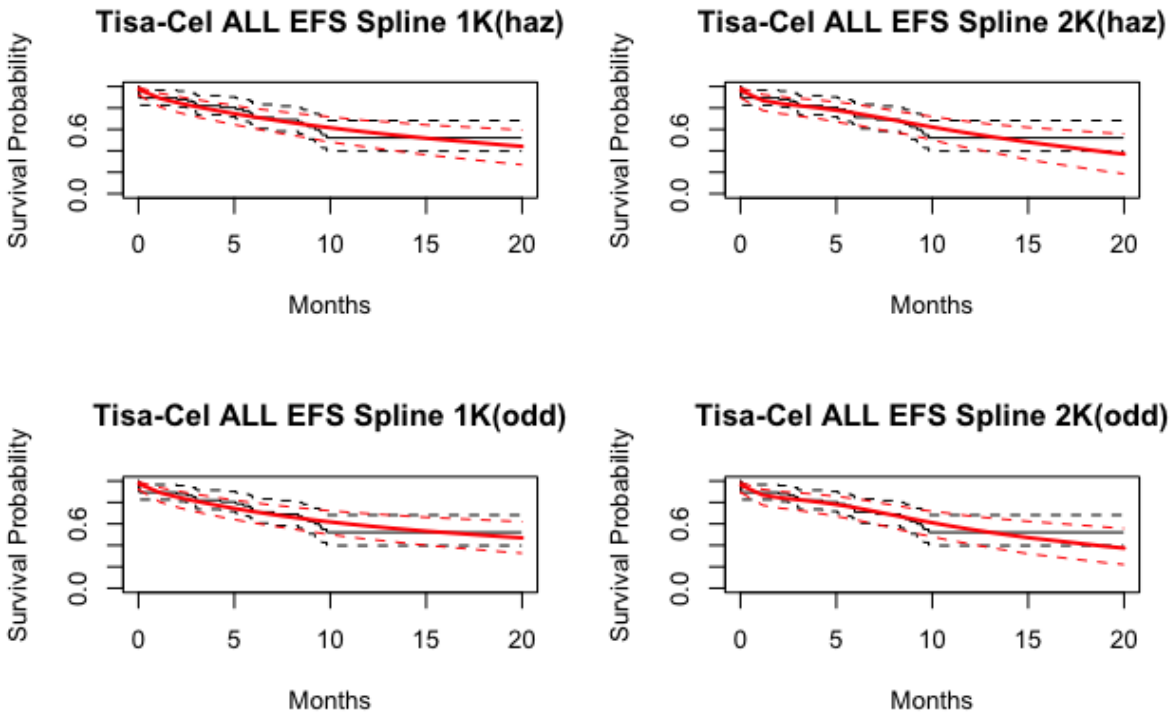

Figure A6. Spline model fits to ELIANA (tisagenlecleucel) event-free survival data‡

‡**1K (haz):** One-knot hazard spline; **1K (odd):** One-knot odds spline; **2K (haz):** Two-knot hazard spline; **2K (odd):** Two-knot odds spline **ALL:** Acute lymphoblastic leukaemia; **EFS:** Event-free survival; **Exp:** Exponential; **GenG:** Generalised gamma; **Gomp:** Gompertz; **LogL:** Log-logistic; **LogN:** Log-normal.

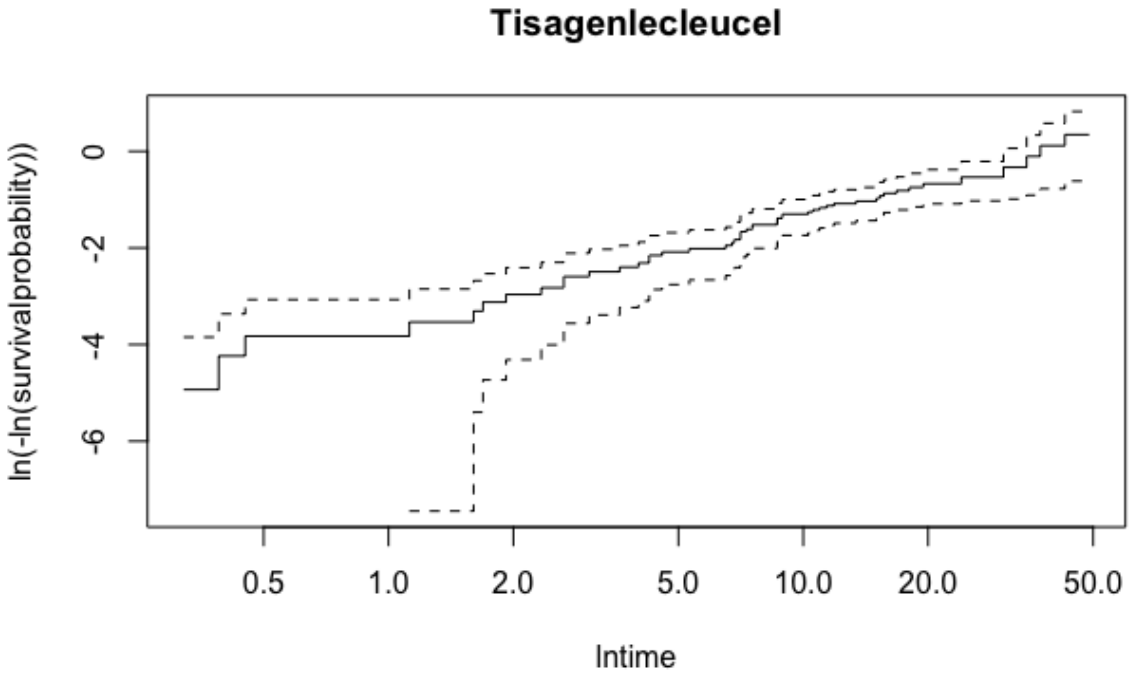

Figure A7. Log cumulative hazard plot pooled ELIANA and ENSIGN data (tisagenlecleucel)

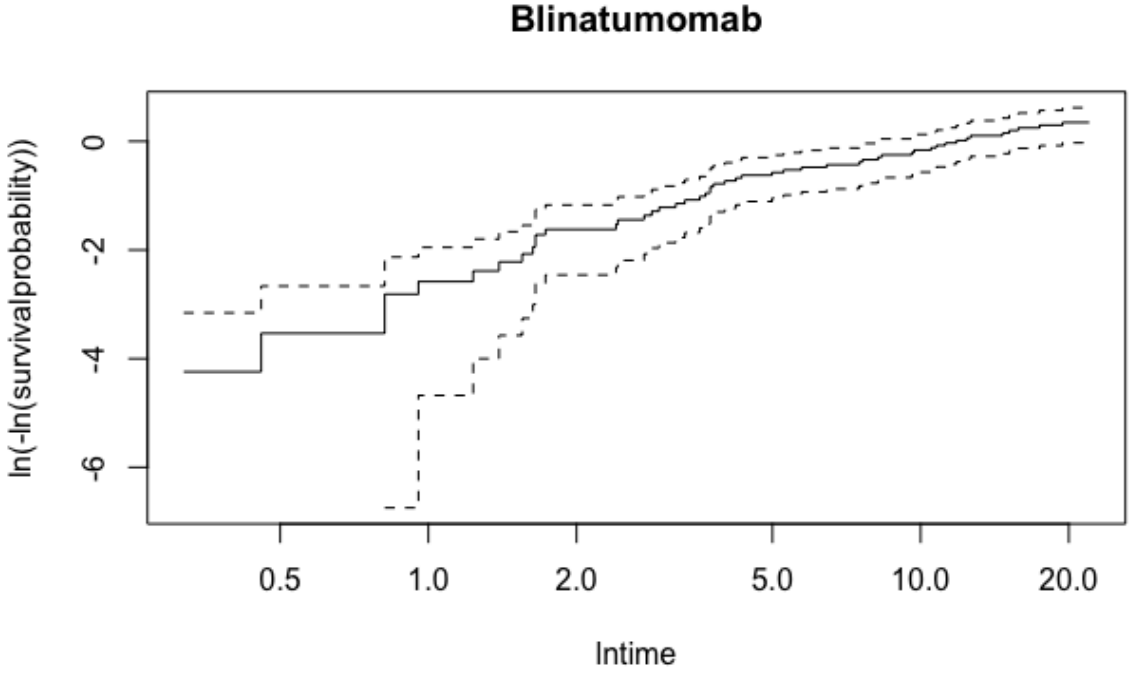

Figure A8. Log cumulative hazard plot NCT01471782 (blinatumomab)

D.2 Chapter 5 Costs

Staff Training Costs

Table A8. Staff training costs tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

	Consultant Haematology (Attending)	Consultant Other (Attending)	Non-Consultant Hospital Doctor (Fellow)	Advanced Practice Registered Nurse	Nurse	Pharmacist Chief II	Pharmacist Senior	Lab -Medical Scientist	Lab - Senior Medical Scientist	Total
	4	2	1	2	10	1	2	1	1	24
A: Salary (mid-point of range)	€ 186,332	€ 186,332	€ 75,826	€ 54,920	€ 39,317	€ 81,443	€ 68,908	€ 46,736	€ 60,582	
B: Direct salary cost	€ 206,922	€ 206,922	€ 84,205	€ 60,989	€ 43,662	€ 90,442	€ 76,522	€ 51,900	€ 67,276	
C: Total salary cost	€ 214,375	€ 214,375	€ 87,238	€ 63,185	€ 45,234	€ 93,700	€ 79,279	€ 53,770	€ 69,700	
D: Total staff cost	€ 260,958	€ 260,958	€ 106,194	€ 76,915	€ 55,063	€ 114,061	€ 96,506	€ 65,454	€ 84,845	
Total per Training	€ 2,000	€ 1,000	€ 814	€ 590	€ 422	€ 874	€ 740	€ 251	€ 325	
	€ 8,002	€ 2,000	€ 814	€ 1,179	€ 4,221	€ 874	€ 1,480	€ 251	€ 325	€ 19,147
Assume 40 hours/week										€ 1,596

Drug Acquisition Costs

Table A9. Bridging chemotherapy regimen tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia: per patient, per cycle, treatment cost

Drug	Reimbursement Scheme	Dose	PTW (€)	Mark-Up 8% (€)	Reimbursement Price* (€)	Fees (€)	Strength (mg)	Cost/ Unit	Cost/ Cycle†	Cost Source
Mercaptopurine	GMS	72mg/m² orally daily for 3 weeks	57.39	4.59	61.98	5.48	1,250	0.05	107.71	IPHA Price Realignment 2019 |
Dexamethasone		6mg/m² orally daily for 5 days	12.71	1.02	13.73	5.48	200	0.10	3.81	PCRS List of Reimbursable Items ¶
Oral Methotrexate		20mg/m² orally once weekly for 2 weeks	12.14	0.97	13.11	5.48	250	0.07	3.93	
Vincristine	Hospital	1.5mg/m² IV weekly for 1 week	85.04	0	85.04	0	10	0.85	17.01‡	IPHA Price Realignment 2018**
Intrathecal Methotrexate	GMS	12mg on day 1 of week 3 administered intrathecally	17.96	1.44	19.40	5.48	12.5	0.78	24.88§	IPHA Price Realignment October 2020 ††
Co-trimoxazole	GMS	480mg orally twice weekly	12.17	0.97	13.14	5.48	48,000	0.01	2.23	PCRS List of Reimburs

	for 3 weeks		able Items ¶
Total			159.56

GMS: General Medical Services; **IPHA:** Irish Pharmaceutical Healthcare Association; **IV:** Intravenous; **PCRS:** Primary Care Reimbursement Services; **PTW:** Price-to-wholesaler.
*5.5% rebate not applicable, as all agents off-patent.
†Assuming mean body surface area of 1.32m^2, where applicable.
‡Available in pack size of 5x2mg vial. One vial required per cycle (€85.04/5=€17.01).
§Available in pack size of 5x2.5mg vial. 5 vials required per cycle.
|https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2019/
¶https://www.hse.ie/eng/staff/pcrs/items/
**https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2018/
††https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2020/

Table A10. Lymphodepleting regimen tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia: per patient, per cycle, treatment cost

Drug	Reimbursement Scheme	Dose	PTW (€)	Reimbursement Price* (€)	Strength /Vial (mg)	Vials/ Cycle†	Cost/ Cycle	Cost Source
Fludarabine	Hospital	30mg/m^2 IV once daily for 4 days	77.15	77.15	50	4	308.60	NCPE Internal Cost Database
Cyclophosphamide		500mg/m^2 IV once daily for 2 days	26.46	26.46	500	4	105.84	
Total								414.44

IV: Intravenous; **NCPE:** National Centre for Pharmacoeconomics; **PTW:** Price-to-wholesaler.
*Mark-up (8%) and pharmacy fees not applicable, as both agents are hospital products. Rebate (5.5%) not applicable, as both agents off-patent.
†Assuming mean body surface area of 1.32m^2.

Table A11. Blinatumomab dosing based on body surface area: per patient total treatment cost

Drug	Reimbursement Scheme	Dose	PTW (€)	Rebate 5.5% (€)	Reimbursement Price* (€)	Vials Required †	Total Cost (€)	Total Cost per Treatment Course (€)
Cycle 1								
Blina Days 1-7	Hospital	5mcg/m^2/day	2,551.51	140.33	2,411.18	2	4,822.35	40,990.01
Blina Days 8-28		15mcg/m^2/ day	2,551.51	140.33	2,411.18	15	36,167.65	
Cycle 2								
Blina Days 1-28	Hospital	15mcg/m^2/ day	2,551.51	140.33	2,411.18	20	48,223.54	48,223.54

PTW: Price-to-wholesaler.
*https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2020/
†https://www.medicines.ie/medicines/blincyto-31448/spc

Table A12. Blinatumomab fixed-dosing regimen: per patient total treatment cost

Drug	Reimbursement Scheme	Dose	PTW (€)	Rebate 5.5% (€)	Reimbursement Price* (€)	Vials Required †	Total Cost (€)	Total Cost per Treatment Course (€)
Cycle 1								
Blina Days 1-7	Hospital	9mcg/day	2,551.51	140.33	2,411.18	3	7,233.53	57,868.25
Blina Days 8-28		28mcg/day	2,551.51	140.33	2,411.18	21	50,634.72	
Cycle 2								
Blina Days 1-28	Hospital	28mcg/day	2,551.51	140.33	2,411.18	28	67,512.95	67,512.95

PTW: Price-to-wholesaler.
*https://www.hse.ie/eng/about/who/cpu/ipha-price-reduction-2020/
†https://www.medicines.ie/medicines/blincyto-31448/spc

Table A13. Number of vials of blinatumomab required per patient for cycle 1‡

Cycle 1	Infusion Duration	Number of Vials Required: Body Surface Area dosing*	Number of Vials Required: Fixed dosing†
Days 1-7	72-/96-hour alternation	2	3
Day 8	72-hours	2	3
Day 11	96-hours	3	4
Day 15	72-hours	2	3
Day 18	96-hours	3	4
Day 22	72-hours	2	3
Day 25	96-hours	3	4
Total		17	24

*Patients weighing <45kg. Days 1-7 dose: 5mcg/m^2/day. Days 8-28 dose: 15mcg/m^2/day.
†Patients weighing ≥45kg. Days 1-7 dose: 9mcg/day. Days 8-28 dose: 28mcg/day.
‡https://www.medicines.ie/medicines/blincyto-31448/spc

Table A14. Number of vials of blinatumomab required per patient for cycle 2‡

Cycle 2	Infusion Duration	Number of Vials Required: Body Surface Area dosing*	Number of Vials Required: Fixed dosing†
Day 1	72-hours	2	3
Day 4	96-hours	3	4
Day 8	72-hours	2	3
Day 11	96-hours	3	4
Day 15	72-hours	2	3
Day 18	96-hours	3	4
Day 22	72-hours	2	3
Day 25	96-hours	3	4
Total		20	28

*Patients weighing <45kg. Days 1-28 dose: 15mcg/m^2/day.
†Patients weighing ≥45kg. Days 1-28 dose: 28mcg/day.
‡https://www.medicines.ie/medicines/blincyto-31448/spc

Table A15. Tocilizumab cost per dose per patient treated with tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

Drug	Reimbursement Scheme	Dose	PTW (€)	Rebate 5.5% (€)	Reimbursement Price* (€)	Strength/ Vial (mg)	Vials/ Dose†	Cost/ Dose (€)	Cost Source
Tocilizumab	Hospital	8mg/kg	712	39.16	672.84	400	1	672.84	MIMS 2020

PTW: Price-to-wholesaler.
*Mark-up (8%) pharmacy fees not applicable, as agent is a hospital product.
†Assuming mean weight of 42.4kg

Table A16. Immunoglobulin cost per dose per patient treated with tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia

Drug	Reimbursement Scheme	Dose	PTW (€)	Reimbursement Price* (€)	Strength/ Vial (mg)	Vials/ Dose†	Cost/ Dose (€)	Cost Source
Immunoglobulin	Hospital	500mg/kg	65	65	1,000	21‡	1,365	Tertiary Teaching Hospital

PTW: Price-to-wholesaler.
*Mark-up (8%) pharmacy fees not applicable, as agent is a hospital product. Rebate (5.5%) not applicable.
†Assuming mean weight of 42.4kg.
‡Round down to nearest vial as per the General Medical Council and Hettle et al. (https://pubmed.ncbi.nlm.nih.gov/28244858/)

Resource Costs

Table A17. Unit costs used in cost-utility models presented in Chapter 5 and Chapter 9

Parameter	Cost (€; 2020)	Source (currency; year)
Leukapheresis	1,249	HPO DRG List* Daycase DRG B62Z (€; 2020)
Cryopreservation	5,544.68	Tertiary Teaching Hospital (€; 2018)
Outpatient Administration Cost	346	
Hospitalisation Cost (Tisagenlecleucel, R/R ALL)	37,944	HPO DRG List* DRG R60A (€; 2020)
Hospitalisation Cost (Lymphodepleting Chemotherapy Only)	5,100	HPO DRG List* DRG R61B (€; 2020)
Hospitalisation CD19 CAR T-Cell therapy (R/R DLBCL; First 13.5 days)	14,033	HPO DRG List* DRG R61A (€; 2020)
Hospitalisation Cost (Blinatumomab –Inpatient)	11,826	HPO DRG List* DRG R60B (€; 2020)
Consultant/Outpatient Appointment	136.76	HSE Ready Reckoner† (R99 Oncology Repeat Attendance) (€; 2013)
Infusion Pump - Blinatumomab	118.67	Rockford Healthcare (€; 2018)
Patient Apartment	63.90	Thielen et al. (€ Dutch; 2020)
Coagulation screen	7.85	Murphy et al.‡ (£ Sterling; 2014)
Complete Blood Count	8.43	O'Brien et al.§ (€; 2013)
Liver Profile	12.42	NCPE Internal Cost Database (€; 2018)
Renal Profile	7.79	O'Brien et al.§ (€; 2013)
Uric Acid	20.94	NCPE Internal Cost Database (€; 2020)
HBV Core	15.91	
HBV Sag	14.39	
Hep C	95.12	
HIV	11.04	National Virus Reference Laboratory (€; 2018)
Urinalysis	5.04	National Clinical Guideline Centre (£; 2014)
Lactate Dehydrogenase	1.57	NCPE Internal Cost Database (€; 2018)
Quantitative Immunoglobulin	55.87	
Serum Protein Electrophoresis	18.62	
Intensive Care Unit Admission	2,797.76	O'Brien et al.§ (€; 2013)
Febrile Neutropenia	9,451.31	
Pancytopenia	387	HPO DRG List* DRG R62B (€; 2020)
Anaemia	743	HPO DRG List* DRG Q61B (€; 2020)
Hypokalaemia	2,722	HPO DRG List* DRG K64B (€; 2020)
Leukopenia	387	HPO DRG List* DRG R62B (€; 2020)

Hypertension	484	HPO DRG List* DRG F67A (€; 2020)
Thrombosis/Embolism	2,254	HPO DRG List* DRG F63B (€; 2020)
Vomiting	2,069	HPO DRG List* DRG G70B (€; 2020)
Infection	3,920	HPO DRG List* DRG T64C (€; 2020)
AlloSCT	202,698	HPO DRG List* DRG A07A (€; 2020)
AlloSCT Follow-Up Costs (first 100 days post-discharge)	64,618.28	Ernst & Young| (€; 2020)
AlloSCT Follow-Up Costs (100-200 days post-discharge)	36,524.17	
AlloSCT Follow-Up Costs (200-365 days post-discharge)	40,957.86	
Terminal care	7,732.48	Bourke et al.¶ (€; 2014)

*https://www.hpo.ie/abf/ABF2020AdmittedPatientPriceList.pdf
†Health Service Executive (HSE). Ready Reckoner of Acute Hospital Inpatient and Daycase Activity & Costs (Summarised by DRG) Relating to 2011 Costs and Activity. 2013.
‡Murphy E, MacGlone S, McGroarty C. A novel approach to improving coagulation sample ordering in an emergency department. BMJ Quality Improvement Reports. 2015;4(1):u204785.w2857 https://pubmed.ncbi.nlm.nih.gov/24472035/
§O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32. https://pubmed.ncbi.nlm.nih.gov/24472035/
|https://www.anthonynolan.org/sites/default/files/2021-03/analysis-of-hospital-activity-and-costs.pdf
¶Bourke S, Burns RM, Gaynor C. Challenges in generating costs and utilisation rates associated with castration-resistant prostate cancer. J Mark Access Health Policy. 2014;2:10.3402/jmahp.v2.24072. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4865741/

Monitoring Costs

Table A18. Blinatumomab per patient per cycle monitoring costs

Drug Monitoring Costs						
Item	**Reference**	**Resource Use**	**Frequency (per cycle)**	**Unit Cost (€)**	**Total Cost (€;2020)**	**Cost Reference (currency; year)**
Blinatumomab						
Coagulation Screen	NCCP*	1	8	7.85	62.80	Murphy et al.† (£; 2014)
Complete Blood Count		1	4	8.43	33.72	O'Brien et al.‡ (€; 2013)
Liver Profile		1	4	12.42	49.68	NCPE Internal Cost Database (€; 2018)
Neurological Observation		4	168	0	0	Assumed to be accounted for in cost of consultant visit and parent-assessed
Renal Profile		1	4	7.79	31.16	O'Brien et al.‡ (€; 2013)
Uric Acid		1	1	20.94	20.94	NCPE Internal Cost Database (€; 2020)
Total (per 42 day cycle)						198.30

*__National Cancer Control Programme__: https://www.hse.ie/eng/services/list/5/cancer/profinfo/chemoprotocols/p567-blinatumomab-paediatric-therapy.pdf

†Murphy E, MacGlone S, McGroarty C. A novel approach to improving coagulation sample ordering in an emergency department. BMJ Quality Improvement Reports. 2015;4(1):u204785.w2857 https://pubmed.ncbi.nlm.nih.gov/24472035/
‡O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32. https://pubmed.ncbi.nlm.nih.gov/24472035/

Table A19. Per patient per cycle event-free surivval monitoring costs for relapsed/refractory acute lymphoblastic leukaemia

Event-Free Survival State						
Requirement	**Reference**	**Resource Use**	**Frequency (per year)**	**Unit Cost (€)**	**Total Cost (€; 2020)**	**Cost Reference (currency; year)**
Months 1-12 Inclusive (Every 2 Months)						
Consultant Appointment	NCCN*	1	6	136.76	820.56	HSE Ready Reckoner‡ (€; 2013)
Complete Blood Count		1	6	8.43	50.58	O'Brien et al.§ (€; 2013)
Liver Profile		1	6	12.42	74.52	NCPE Internal Cost Database (€; 2018)
Quantitative Immunoglobulin (tisagenlecleucel only)	Yakoub-Agha et al.†	1	6	55.87	335.23	
Serum Protein Electrophoresis (tisagenlecleucel only)		1	6	18.62	111.74	
Months 13-24 Inclusive (Every 4 Months)						
Consultant Appointment	NCCN*	1	3	136.76	410.28	HSE Ready Reckoner‡ (€; 2013)
Complete Blood Count		1	3	8.43	25.29	O'Brien et al.§ (€; 2013)
Quantitative Immunoglobulin (tisagenlecleucel only)	Yakoub-Agha et al.†	1	3	55.87	167.61	NCPE Internal Cost Database (€; 2018)
Serum Protein Electrophoresis (tisagenlecleucel only)		1	3	18.62	55.87	
Months 25-60 Inclusive (Every 6 Months)						
Consultant Appointment	NCCN*	1	2	136.76	273.52	HSE Ready Reckoner‡ (€; 2013)
Complete Blood Count		1	2	8.43	16.86	O'Brien et al.§ (€; 2013)
Quantitative Immunoglobulin (tisagenlecleucel only)	Yakoub-Agha et al.†	1	2	55.87	111.74	NCPE Internal Cost Database (€; 2018)
Serum Protein Electrophoresis (tisagenlecleucel only)		1	2	18.62	37.25	
Month 61 Onwards						
Consultant Appointment	NCCN*	1	1	136.76	136.76	HSE Ready Reckoner‡ (€; 2013)

Complete Blood Count	1	1	8.43	8.43	O'Brien et al.§ (€; 2013)

*National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Pediatric Acute Lymphoblastic Leukemia. Version 2. 2021. 2021.

†Ibrahim Y-A, Christian C, Peter B, Grzegorz WB, Halvard B, Fabio C, et al. Management of adults and children undergoing chimeric antigen receptor T-cell therapy: best practice recommendations of the European Society for Blood and Marrow Transplantation (EBMT) and the Joint Accreditation Committee of ISCT and EBMT (JACIE). Haematologica. 2020;105(2):297-316.https://pubmed.ncbi.nlm.nih.gov/31753925/

‡Health Service Executive (HSE). Ready Reckoner of Acute Hospital Inpatient and Daycase Activity & Costs (Summarised by DRG) Relating to 2011 Costs and Activity. 2013.

§ O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32. https://pubmed.ncbi.nlm.nih.gov/24472035/

Adverse Event Costs

Table A20. Per patient cost of treating cytokine release sydrome (tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia)

Tisagenlecleucel Cytokine Release Syndrome	Cost (€; 2020)	Proportion (%)	Duration (days)/Number of Doses	Total Cost (€)
Intensive Care Unit Admission	2797.76	47	8	10,519.58
Tocilizumab	672.84	28	1.24	376.79†

†Assuming mean weight of 42.4kg. Vial sharing not assumed.

Table A21. Per patient cost of treating cytokine release syndrome (blinatumomab for relapsed/refractory acute lymphoblastic leukaemia)

Blinatumomab Cytokine Release Syndrome	Cost (€; 2020)	Proportion (%)	Duration (days)	Total Cost (€)
Intensive Care Unit Admission	2797.76	5.7	5	797.36

Table A22. Per patient cost of treating febrile neutropenia in patients with relasped/refractory acute lymphoblastic leukaemia

Febrile Neutropenia	Cost (€; 2020)	Proportion (%)	Total Cost (€)
Tisagenlecleucel	9451.31	36	3416.75
Blinatumomab	9451.31	17	1606.72

Table A23. Per patient cost of treating pancytopenia (tisagenlecleucel for relapsed/refractory acute lymphoblastic leukaemia)

Pancytopenia	Cost (€; 2020)	Proportion (%)	Duration	Total Cost (€)
Tisagenlecleucel	387	3	Once/month for 6 months	69.66

Table A24. Per patient adverse event treatment costs in patients treated with blinatumomab for relapsed/refractory acute lymphoblastic leukaemia

Adverse Event	Resource Use	Cost (€)	Proportion (%)	Total Cost (€; 2020)	Cost Source (currency; year)	Justification*
Anaemia	DRG Q61B (daycase)	743	36	267.48	HPO DRG List† (€; 2020)	Hgb<8.0 g/dL; <4.9 mmol/L; <80 g/L; transfusion indicated
Thrombocytopenia	Outpatient appointment	136.76	21	28.72	HSE Ready Reckoner‡ (R99 Oncology Repeat Attendance) (€; 2013)	Assumed outpatient appointment required. No specific management costs included
Hypokalaemia	DRG K64B	2,722	17	462.74	HPO DRG List† (€; 2020)	<3.0-2.5 mmol/L; hospitalisation indicated
Neutropenia	DRG R62B (daycase)	387	17	65.79		Assumption
Alanine Aminotransferase Increased	None	-	16	0.00	N/A	Asymptomatic, detected during drug monitoring and managed through dose reduction/interruption; no additional costs incurred
Platelet Count Decreased	Outpatient appointment	136.76	14	19.15	HSE Ready Reckoner‡ (R99 Oncology Repeat Attendance) (€; 2013)	Assumed outpatient appointment required. No specific management costs included
Neutrophil Count Decreased	Outpatient appointment	136.76	13	17.78		Assumed outpatient appointment required. No specific management costs included

Aspartate Aminotransferase Increased	None	-	11	0.00	N/A	Asymptomatic, detected during drug monitoring and managed through dose reduction/interruption; no additional costs incurred
Leukopenia	DRG R62B (daycase)	387	10	38.70	HPO DRG List† (€; 2020)	Assumption
White Blood Cell Count Decreased	Outpatient appointment	136.76	10	13.68	HSE Ready Reckoner‡ (R99 Oncology Repeat Attendance) (€; 2013)	Assumed outpatient appointment required. No specific management costs included
Hypertension	DRG F67A (daycase)	484	6	29.04	HPO DRG List† (€; 2020)	Medical intervention indicated

*https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf
†https://www.hpo.ie/abf/ABF2020AdmittedPatientPriceList.pdf
‡Health Service Executive (HSE). Ready Reckoner of Acute Hospital Inpatient and Daycase Activity & Costs (Summarised by DRG) Relating to 2011 Costs and Activity. 2013.

Appendix E Chapter 6

E.1 Chapter 6 Systematic Literature Review

PICOS Criteria

Table A25. PICOS Chapter 6 systematic literature review of treatments for relapsed/refractory diffuse large B-cell lymphoma

Criteria	Inclusion	Exclusion
Population	Patients aged ≥ 18 years <80 years, with relapsed or refractory DLBCL, PMBCL, tFL. Any gender, any ethnicity. After two or more lines of systemic therapy Aggressive non-Hodgkin lymphoma	Adolescent Paediatric Indolent NHL Follicular lymphoma (not transformed) Grey zone lymphoma Treatment-naïve Newly-diagnosed First-line therapy ("front-line therapy") HIV-associated DLBCL HIV+ patients CNS relapse/ patients with associated-CNS disease Elderly patients (>80 years) Studies reporting **only** on patients with certain subtypes of DLBCL ie activated B-cell (ABC), germinal centre B-cell (GCB) or double-hit (DL)
Intervention	Tisagenlecleucel, axicabtagene ciloleucel at EMA licensed dose	CD19 CAR T-cell therapy in combination with other agents eg atezolizumab, pembrolizumab, lenalidomide CD19 CAR T-cell therapy used in earlier lines of therapy
Comparator	DA-R-EPOCH (dose-adjusted; rituximab, etoposide, doxorubicin, vincristine, cyclophosphamide, prednisolone) R-ESHAP (rituximab, etoposide, cytarabine, cisplatin, methylprednisolone) R-DHAP (rituximab, cisplatin, cytarabine, dexamethasone) R-GDP (rituximab, gemcitabine, cisplatin, dexamethasone) R-GIFOX (rituximab, gemcitabine, ifosfamide, oxaliplatin) R-GemOx (rituximab – gemcitabine, oxaliplatin) R-GEM-P (rituximab, gemcitabine, cisplatin, methylprednisolone) Salvage Chemotherapy, not otherwise specified *all therapies with or without HSCT	Radiotherapy/radiation therapy Tyrosine Kinase inhibitors (TKIs) eg ibrutinib. Lenalidomide Yttrium-90 ibritumomab tiuxetan Rituximab monotherapy Bendamustine R-CHOEP Obintuzumab Bortezomib-based regimens Pixantrone Maintenance therapy (e.g. post HSCT – reporting impact of maintenance therapy only) Consolidation therapy (except HSCT) Conditioning chemotherapy Myeloablative therapy

Outcome	Survival-based outcomes: • Overall survival • Progression-free survival • Event-free survival • Disease-free survival	Response rates only reported: [Complete response Objective response rate Overall response rate Partial Response Duration of response, etc.] Pharmacokinetic/pharmacodynamics Social outcomes Prognostic value (of eg PET scan) Impact of intervention of stem cell mobilisation
Study Design	Prospective randomised controlled trials, Phase I/II non-randomised or single-arm trials, Prospective observational studies,	Reviews Letters Comments Editorials Biomarker/prognostic studies Expanded access programmes Retrospective studies

DLBCL: Diffuse large B-cell lymphoma; **HSCT:** Haematopoietic stem cell transplant; **NHL:** Non-Hodgkin's lymphoma; **PMBCL:** Primary mediastinal large B-cell lymphoma; **tFL:** Transformed follicular lymphoma.

E.2 Chapter 6 Search Strategy

The search strategy used in the systematic review of treatments for R/R DLBCL is presented in Table A26.

Table A26. Chapter 6 (systematic literature review of treatments for relapsed/refractory diffuse large B-cell lymphoma) search strategy

EMBASE 25 October 2019
#1 'diffuse large b cell lymphoma' 'large cell lymphoma /exp OR /exp #2 ((diffus* OR 'large cell' OR anaplas* OR aggress* OR 'high grade' OR 'large b-cell' OR 'large b cell') NEAR/5 (lympho* OR nhl OR 'non hodgkin*' OR 'non hodgkIN*' OR nonhodgkin*)):ab,ti,kw #3 'dlbcl' #4 #1 OR #2 OR #3 #5 relaps*:ti,ab,kw OR refractory:ti,ab,kw OR 'r/r':ti,ab,kw OR fail*:ti,ab,kw OR recurren*:ti,ab,kw OR 'cancer recurrence':ti,ab,kw OR recurring:ti,ab,kw #6 #4 AND #5 #7 ctl019:ti,ab,kw OR 'ctl 019':ti,ab,kw OR 'cart 19':ti,ab,kw OR 'cart-19 cells:ti,ab,kw OR tisagenlecleucel:ti,ab,kw OR Kymriah*:ti,ab,kw #8 'tisagenlecleucel t'/exp #9 'axicabtagene ciloleucel'/exp #10 'axicabtagene ciloleucel':ti,ab,kw OR axicabtagene:ti,ab,kw OR yescarta:ti,ab,kw OR 'kte-c19':ti,ab,kw OR 'kte c19':ti,ab,kw OR 'kte x19':ti,ab,kw OR 'axi cel':ti,ab,kw #11 'lisocabtagene maraleucel'/exp #12 'lisocabtagene maraleucel':ti,ab,kw OR 'liso-cel':ti,ab,kw OR lisocel:ti,ab,kw OR 'liso cel':ti,ab,kw OR jcar017:ti,ab,kw OR 'jcar 017':ti,ab,kw OR 'jcar-017':ti,ab,kw #13 'rituximab'/exp #14 rituximab:ti,ab,kw OR mabthera:ti,ab,kw OR rituxan:ti,ab,kw OR 'idec 102':ti,ab,kw OR 'idec c2b8':ti,ab,kw OR 'ro 452294':ti,ab,kw OR 'r 105':ti,ab,kw OR r105:ti,ab,kw OR 'rg 105':ti,ab,kw OR rg105:ti,ab,kw #15 rchop:ti,ab,kw OR 'r chop':ti,ab,kw OR 'r codox m':ti,ab,kw OR rcodoxm:ti,ab,kw OR 'da r epoch':ti,ab,kw OR darepoch:ti,ab,kw OR 'r ice':ti,ab,kw OR rice:ti,ab,kw OR 'r eshap':ti,ab,kw OR reshap:ti,ab,kw OR 'r dhap':ti,ab,kw OR rdhap:ti,ab,kw OR 'r gdp':ti,ab,kw OR rgdp:ti,ab,kw OR 'r gifox':ti,ab,kw OR rgifox:ti,ab,kw OR 'r gem p':ti,ab,kw OR rgemp:ti,ab,kw #16 ((salvage OR rescue) NEAR/2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol)):ab,ti,kw #17 'pixantrone'/exp #18 pixantrone:ti,ab,kw OR 'pixantrone maleate':ti,ab,kw OR 'bbr 2778':ti,ab,kw OR bbr2778:ti,ab,kw #19 'allogenic bone marrow transplantation'/exp OR 'autologous bone marrow transplantation'/exp

#20 ('autologous stem cell transplant':ti,ab,kw OR asct:ti,ab,kw OR abmt:ti,ab,kw OR pbpc:ti,ab,kw OR pbsct:ti,ab,kw OR psct:ti,ab,kw OR bmt:ti,ab,kw OR sct:ti,ab,kw OR 'auto transplant*':ti,ab,kw OR autotransplant:ti,ab,kw OR autolog*or:ti,ab,kw) AND autograft*:ti,ab,kw

#21 ((allogen* OR 'allo gen*') NEAR/5 transplant*):ti,ab,kw

#22 #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21

#23 #6 AND #22

#24 'animal' NOT 'human'

#25 animal:ab,kw,ti OR 'animal model*':ab,kw,ti OR animals:ab,kw,ti OR canine*:ab,kw,ti OR cat:ab,kw,ti OR cats:ab,kw,ti OR dog:ab,kw,ti OR dogs:ab,kw,ti OR feline:ab,kw,ti OR felines:ab,kw,ti OR hamster:ab,kw,ti OR hamsters:ab,kw,ti OR mice:ab,kw,ti OR monkey:ab,kw,ti OR monkeys:ab,kw,ti OR mouse:ab,kw,ti OR pig:ab,kw,ti OR piglet:ab,kw,ti OR piglets:ab,kw,ti OR pigs:ab,kw,ti OR porcine:ab,kw,ti OR primate*:ab,kw,ti OR rabbit:ab,kw,ti OR rabbits:ab,kw,ti OR rat:ab,kw,ti OR rats:ab,kw,ti OR rodent:ab,kw,ti OR rodents:ab,kw,ti OR sheep:ab,kw,ti OR swine:ab,kw,ti OR swines:ab,kw,ti

#26 #23 NOT (#24 OR #25)

#27 comment OR editorial OR news OR newspaper OR 'case report' OR 'case study'

#28 #26 NOT #27

#29 [2001-2019]/py

#30 english:la

#31 #28 AND #29 AND #30

MEDLINE (via EBSCO)

S1 (MH "Lymphoma, Large B-Cell, Diffuse") OR (MH "Lymphoma, Non-Hodgkin")

S2 AB (((diffus* OR 'large cell' OR anaplas* OR aggress*
OR 'high grade' OR 'large b-cell' OR 'large b cell') N5 (lympho* OR nhl OR 'non hodgkin*' OR 'non hodgkin*' OR nonhodgkin*))) OR TI (((diffus* OR 'large cell' OR anaplas* OR aggress* OR 'high grade' OR 'large b-cell' OR 'large b cell') N5 (lympho* OR nhl OR 'non hodgkin*' OR 'non hodgkin*' OR nonhodgkin*))) OR SU (((diffus* OR 'large cell' OR anaplas* OR aggress* OR 'high grade' OR 'large b-cell' OR 'large b cell') N5 (lympho* OR nhl OR 'non hodgkin*' OR 'non hodgkin*' OR nonhodgkin*)))

S3 DLBCL

S4 S1 OR S2 OR S3

S5 AB ((relaps* OR refractory OR 'R/R' or fail* OR recurren* OR recurring OR 'cancer recurrence')) OR TI ((relaps* OR refractory OR 'R/R' or fail* OR recurren* OR recurring OR 'cancer recurrence')) OR SU ((relaps* OR refractory OR 'R/R' or fail* OR recurren* OR recurring OR 'cancer recurrence'))

S6 S4 AND S5

S7 AB ((CTL019 OR CTL-019 OR 'CTL 019' OR CART-19 OR 'CART-19 cells' OR tisagenlecleucel OR Kymriah*)) OR TI ((CTL019 OR CTL019 OR 'CTL 019' OR CART-19 OR 'CART19 cells' OR tisagenlecleucel OR Kymriah*)) OR SU ((CTL019 OR CTL019 OR 'CTL 019' OR CART-19 OR 'CART19 cells' OR tisagenlecleucel OR Kymriah*))

S8 AB (('axicabtagene ciloleucel' OR axicabtagene OR yescarta OR 'kte-c19' OR KTE-C19 OR KTE-X19 OR Axicel)) OR TI (('axicabtagene ciloleucel' OR axicabtagene OR yescarta OR 'kte-c19' OR KTE-C19 OR KTE-X19 OR Axicel)) OR SU (('axicabtagene ciloleucel' OR axicabtagene OR yescarta OR 'kte-c19' OR KTE-C19 OR KTE-X19 OR Axicel))

S9 AB (('axicabtagene ciloleucel' OR axicabtagene OR yescarta OR 'kte-c19' OR KTE-C19 OR KTE-X19 OR Axicel)) OR TI (('axicabtagene ciloleucel' OR axicabtagene OR yescarta OR 'kte-c19' OR KTE-C19 OR KTE-X19 OR Axicel)) OR SU (('axicabtagene ciloleucel' OR axicabtagene OR yescarta OR 'kte-c19' OR KTE-C19 OR KTE-X19 OR Axicel))

S10 AB (('lisocabtagene maraleucel' OR 'lisocel' OR lisocel OR 'liso cel' OR JCAR017 OR JCAR-017 OR 'JCAR-017')) OR TI (('lisocabtagene maraleucel' OR 'lisocel' OR lisocel OR 'liso cel' OR JCAR017 OR JCAR-017 OR 'JCAR-017')) OR SU (('lisocabtagene maraleucel' OR 'lisocel' OR lisocel OR 'liso cel' OR JCAR017 OR JCAR-017 OR 'JCAR-017'))

S11 AB ((rituximab OR mabthera OR rituxan OR idec-102 OR 'IDEC 102' OR idecc2b8 OR 'idec c2b8' OR RO-452294 OR 'RO 452294' OR 'R 105' OR R105 OR 'RG 105' OR RG105)) OR TI ((rituximab OR mabthera OR rituxan OR idec-102 OR 'IDEC 102' OR idec
c2b8 OR 'idec c2b8' OR RO-452294 OR 'RO 452294' OR 'R 105' OR R105 OR 'RG 105' OR RG105)) OR SU ((rituximab OR mabthera OR rituxan OR idec-102 OR 'IDEC 102' OR idecc2b8 OR 'idec c2b8' OR RO-452294 OR 'RO 452294' OR 'R 105' OR R105 OR 'RG 105' OR RG105))

S12 AB ((R-CHOP OR RCHOP OR 'R CHOP' OR R-CODOX-M OR 'R CODOX M' OR RCODOXM OR DAR-EPOCH OR 'DA R EPOCH' OR DAREPOCH OR R-ICE OR RICE OR R-ESHAP OR 'R ESHAP' OR RESHAP OR R-DHAP OR 'R DHAP' OR RDHAP OR R-GDP OR 'R GDP' OR RGDP OR R-GIFOX OR 'R GIFOX' OR RGIFOX OR R-GEM-P OR 'R GEM P' OR RGEMP)) OR TI ((R-CHOP OR RCHOP OR 'R CHOP' OR R-CODOX-M OR 'R CODOX M' OR RCODOXM OR DAR-EPOCH OR 'DA R EPOCH' OR DAREPOCH OR R-ICE OR RICE OR R-ESHAP OR 'R ESHAP' OR RESHAP OR R-DHAP OR 'R DHAP' OR RDHAP OR R-GDP OR 'R GDP' OR RGDP OR R-GIFOX OR 'R GIFOX' OR RGIFOX OR R-GEM-P OR 'R GEM P' OR RGEMP)) OR SU ((R-CHOP OR RCHOP OR 'R CHOP' OR R-CODOX-M OR 'R CODOX M' OR RCODOXM OR DAR-EPOCH OR 'DA R EPOCH' OR DAREPOCH OR R-ICE OR RICE OR R-ESHAP OR 'R ESHAP' OR RESHAP OR R-DHAP OR 'R DHAP' OR RDHAP OR R-GDP OR 'R GDP' OR RGDP OR R-GIFOX OR 'R GIFOX' OR RGIFOX OR R-GEM-P OR 'R GEM P' OR RGEMP))

S13 AB (((Salvage or rescue) ADJ2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol))) OR TI (((Salvage or rescue) NEAR/2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol))) OR SU (((Salvage or rescue) NEAR/2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol)))

S14 AB (((Salvage or rescue) ADJ2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol))) OR TI (((Salvage or rescue) ADJ2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol))) OR SU (((Salvage or rescue) ADJ2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol)))

S15 AB (((Salvage or rescue) N2 (chemotherapy OR chemo OR therapy
OR treatment OR regimen OR protocol))) OR TI (((Salvage or rescue) N2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol))) OR SU (((Salvage or rescue) N2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol)))

S16 AB ((pixantrone OR 'pixantrone maleate' OR 'BBR 2778' OR BBR-2778 OR BBR2778)) OR TI ((pixantrone OR 'pixantrone maleate' OR 'BBR 2778' OR BBR-2778 OR BBR2778)) OR SU ((pixantrone OR 'pixantrone maleate' OR 'BBR 2778' OR BBR-2778 OR BBR2778))

S17 AB (('Autologous stem cell transplant' OR ASCT OR ABMT OR PBPC OR PBSCT OR PSCT OR BMT OR SCT OR AUTOTRANSPLANT* OR AUTOTRANSPLANT OR AUTOLOG*OR AUTOGRAFT*)) OR TI (('Autologous stem cell transplant' OR ASCT OR ABMT OR PBPC OR PBSCT OR PSCT OR BMT OR SCT OR AUTOTRANSPLANT* OR AUTOTRANSPLANT OR AUTOLOG*OR AUTOGRAFT*)) OR SU (('Autologous

stem cell transplant’ OR ASCT OR ABMT OR PBPC OR PBSCT OR PSCT OR BMT OR SCT OR AUTOTRANSPLANT* OR AUTOTRANSPLANT OR AUTOLOG*OR AUTOGRAFT*))

S18 AB (((allogen* OR allo-gen*) N5 trasplant*)) OR TI (((allogen* OR allogen*) N5 trasplant*)) OR SU (((allogen* OR allo-gen*) N5 trasplant*))

S19 AB ((allogen* OR allo-gen*) N5 trasplant*) OR TI ((allogen* OR allogen*) N5 trasplant*) OR SU ((allogen* OR allo-gen*) N5 trasplant*)

S20 AB ((allogen* OR allo-gen*) N5 transplant*) OR TI ((allogen* OR allogen*) N5 transplant*) OR SU ((allogen* OR allo-gen*) N5 transplant*)

S21 S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20

S22 S6 AND S21

S23 'animal' NOT 'human'

S24 AB (animal OR 'animal model*' OR animals OR canine* OR cat OR cats OR dog OR dogs OR feline OR felines OR hamster OR hamsters OR mice OR monkey OR monkeys OR mouse OR pig OR piglet OR piglets OR pigs OR porcine OR primate OR rabbit OR
rabbits OR rat OR rats OR rodent OR rodents OR sheep OR swine OR swines) OR TI (animal OR 'animal model*' OR animals OR canine* OR cat OR cats OR dog OR dogs OR feline OR felines OR hamster OR hamsters OR mice OR monkey OR monkeys OR mouse OR pig OR piglet OR piglets OR pigs OR porcine OR primate OR rabbit OR rabbits OR rat OR rats OR rodent OR rodents OR sheep OR swine OR swines) OR SU (animal OR 'animal model*' OR animals OR canine* OR cat OR cats OR dog OR dogs OR feline OR felines OR hamster OR hamsters OR mice OR monkey OR monkeys OR mouse OR pig OR piglet OR piglets OR pigs OR porcine OR primate OR rabbit OR rabbits OR rat OR rats OR rodent OR rodents OR sheep OR swine OR swines)

S25 S22 NOT (S23 OR S24)

S26 comment OR editorial OR news OR newspaper OR 'case report'

S27 S25 NOT S26

S28 S25 NOT S26

CENTRAL (via Cochrane Library)

#1 MeSH descriptor: [Lymphoma, Large B-Cell, Diffuse] explode all trees

#2 MeSH descriptor: [Lymphoma, Non-Hodgkin] explode all trees

#3 ((diffus* OR 'large cell' OR anaplas* OR aggress* OR 'high grade' OR 'large b-cell' OR 'large b cell') near/5 (lympho* OR nhl OR 'non hodgkin*' OR 'non hodgkin*' OR nonhodgkin*)):ti,ab

#4 DLBCL

#5 #1 OR #2 OR #3 OR #4

#6 (relaps* OR refractory OR fail* OR recurren* OR recurring OR ‘cancer recurrence’):ti,ab,kw (Word variations have been searched)

#7 #5 AND #6

#8 (CTL019 OR CTL-019 OR ‘CTL 019’ OR CART-19 OR ‘CART-19 cells’ OR tisagenlecleucel OR Kymriah*):ti,ab,kw

#9 ('axicabtagene ciloleucel' OR axicabtagene OR yescarta OR 'kte-c19' OR KTE-C19 OR KTE-X19 OR Axi-cel):ti,ab,kw

#10 ('lisocabtagene maraleucel' OR 'liso-cel' OR lisocel OR 'liso cel' OR JCAR017 OR JCAR-017 OR 'JCAR-017'):ti,ab,kw

#11 (rituximab OR mabthera OR rituxan OR idec-102 OR 'IDEC 102' OR idec-c2b8 OR 'idec c2b8' OR RO-452294 OR 'RO 452294' OR 'R 105' OR R105 OR 'RG 105' OR RG105):ti,ab,kw

#12 (R-CHOP OR RCHOP OR 'R CHOP' OR R-CODOX-M OR 'R CODOX M' OR RCODOXM OR DA-R-EPOCH OR 'DA R EPOCH' OR DAREPOCH OR R-ICE OR RICE OR R-ESHAP OR 'R ESHAP' OR RESHAP OR R-DHAP OR 'R DHAP' OR RDHAP OR R-GDP OR 'R GDP' OR RGDP OR R-GIFOX OR 'R GIFOX' OR RGIFOX OR R-GEM-P OR 'R GEM P' OR RGEMP):ti,ab,kw

#13 ((Salvage or rescue) NEAR/2 (chemotherapy OR chemo OR therapy OR treatment OR regimen OR protocol)):ab,ti,kw

#14 (pixantrone OR 'pixantrone maleate' OR 'BBR 2778' OR BBR-2778 OR BBR2778):ti,ab,kw

#15 ('Autologous stem cell transplant' OR ASCT OR ABMT OR PBPC OR PBSCT OR PSCT OR BMT OR SCT OR AUTO-TRANSPLANT* OR AUTOTRANSPLANT OR AUTOLOG*OR AUTOGRAFT*):ti,ab,kw

#16 ((allogen* OR allo-gen*) NEAR/5 transplant*):ti,ab,kw

#17 #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16

#18 #7 AND #17

#19 'animal' NOT 'human'

#20 animal:ab,kw,ti OR 'animal model*':ab,kw,ti OR animals:ab,kw,ti OR canine*:ab,kw,ti OR cat:ab,kw,ti OR cats:ab,kw,ti OR dog:ab,kw,ti OR dogs:ab,kw,ti OR feline:ab,kw,ti OR felines:ab,kw,ti OR hamster:ab,kw,ti OR hamsters:ab,kw,ti OR mice:ab,kw,ti OR monkey:ab,kw,ti OR monkeys:ab,kw,ti OR mouse:ab,kw,ti OR pig:ab,kw,ti OR piglet:ab,kw,ti OR piglets:ab,kw,ti OR pigs:ab,kw,ti OR porcine:ab,kw,ti OR primate*:ab,kw,ti OR rabbit:ab,kw,ti OR rabbits:ab,kw,ti OR rat:ab,kw,ti OR rats:ab,kw,ti OR rodent:ab,kw,ti OR rodents:ab,kw,ti OR sheep:ab,kw,ti OR swine:ab,kw,ti OR swines:ab,kw,ti

#21 #18 NOT (#19 OR #20)

#22 comment OR editorial OR news OR newspaper OR 'case report'

#23 #21 NOT #22 with Cochrane Library publication date Between Jan 2001 and Jan 2019

E.3 Chapter 6 Excluded Studies

A selection of studies, excluded at full-text screening in the systematic review of treatments for R/R DLBCL, are presented in Table A27.

Table A27. Chapter 6 (systematic literature review of treatments for relapsed/refractory diffuse large B-cell lymphoma) selection of excluded studies

Study (year)	Authors	Reason for Exclusion
The role of maintenance therapy in patients with diffuse large b-cell lymphoma: A systematic review and meta-analysis (2019)	Rozental, A.and Gafter-Gvilli, A. and Vidal-Fisher, L. and Raanani, P. and Gurion, R.	Outcome
BELINDA: A Phase 3 Study Evaluating the Safety and Efficacy of Tisagenlecleucel versus Standard of Care in Adult Patients with Relapsed/Refractory Aggressive B-Cell Non-Hodgkin Lymphoma (2019)	Westin, J. and Bishop, M. and Flinn, I. and Borchmann, P. and Jaeger, U. and Gu, J., et al.	Outcome
Pixantrone: A promising drug in the treatment of non-Hodgkin lymphomas (2009)	Ei-Helw, L. M. and Hancock, B. W.	Intervention
Phase II study of anti-CD19 antibody drug conjugate (SAR3419) in combination with rituximab: Clinical activity and safety in patients with relapsed/refractory diffuse large B-cell lymphoma (NCT01470456) (2013)	Coiffier, B. and Thieblemont, C. and De Guibert, S. and Dupuis, J. and Ribrag, V. and Bouabdallah, R., et al.	Intervention
Independent predictive value of PET-CT pre transplant in relapsed and refractory patients with CD20 diffuse large b-cell lymphoma (DLBCL) included in the CORAL study (2009)	Trneny, M. and Bosly, A. and Bouabdallah, K. and Ma, D. and Shpilberg, O. and Montoto, S., et al.	Outcome
Rituximab for aggressive non-Hodgkin's lymphomas relapsing after or refractory to autologous stem cell transplantation (2002)	Pan, D. and Moskowitz, C. H. and Zelenetz, A. D. and Straus, D. and Kewalaramani, T. and Noy, A., et al.	Intervention
Prognostic value of the interval between relapse and therapy initiation in diffuse large B-cell lymphoma patients. analysis from the czech lymphoma study group database (2019)	Janikova, A. and Chloupkova, R. and Campr, V. and Kopalova, N. and Klener, P. and Benesova, K.,et al.	Outcome
Value of surveillance studies for patients with stage i to ii diffuse large b-cell lymphoma in the rituximab Era (2015)	Hiniker, S. M. and Pollom, E. L. and Khodadoust, M. S. and Kozak, M. M. and Xu, G. and Quon, A., et al.	Population
Late relapses following high dose chemotherapy and autologous stem cell transplant in patients with diffuse large B cell lymphoma in the rituximab era (2014)	Casulo, C. and Bradley, H. D. and Herr, M. M. and Barlaskar, F. and Evans, A. and Burack, R., et al.	Study Design

Appendix F Chapter 7

F.1 Abstrackr Performance Metrics Calculations

Table A28. Base case calculations to assess the performance of Abstrackr in systematic literature review of treatments for relapsed/refractory diffuse large B-cell lymphoma (n=7,723), adapted from Gates et al.* and Rathbone et al.†

Performance Metric	Definition	Calculation
Sensitivity	Proportion of citations correctly predicted as relevant by Abstrackr out of the total deemed relevant by Screener 2	(true positives) / (true positives + false negatives) =367/(37+367) =91%
Specificity	Proportion of citations correctly predicted as irrelevant by Abstrackr out of the total deemed irrelevant by Screener 2	(true negatives) / (true negatives + false positives) =5118/(5118+2001) =72%
Precision	Proportion of citations predicted as relevant by Abstrackr that were also deemed relevant by Screener 2	# citations correctly predicted as relevant (by Abstrackr) / all citations predicted as relevant (by Abstrackr) (true positive)/ (true positive + false positive) =367/2368 =15.5%
False Negative Rate	Proportion of citations that were deemed relevant by Screener 2 that were predicted as irrelevant by Abstrackr	# citations incorrectly predicted as irrelevant (by Abstrackr) / all citations identified as relevant (by human screener) (false negative)/(false negative + true positive) =37/404 =9%
Proportion Missed	Proportion of citations included in the final evidence base after full-text screening that were predicted to be irrelevant by Abstrackr	# citations incorrectly predicted as irrelevant (by Abstrackr) / all citations included in final evidence base =0/9 =0%
Workload Savings	Proportion of citations predicted as irrelevant by Abstrackr out of the total number of citations to be screened (i.e. the proportion of citations that would not need to be screened manually)	# citations predicted as irrelevant (by Abstrackr) / total # citations to be screened =5155/(7523+200) =67%
Time Savings	Time saved based on the citations that would not need to be	(# citations predicted as irrelevant (by Abstrackr) x 0.5 min/citation) / 60 (min/ h)] / (8 h/day)

	screened (i.e., those predicted as irrelevant by Abstrackr); estimated based on a screening rate of 0.5 minutes per citation and an 8-hour work day	=(5155*0.5)/60/8 =5 days

*Gates A, Johnson C, Hartling L. Technology-assisted title and abstract screening for systematic reviews: a retrospective evaluation of the Abstrackr machine learning tool. Systematic Reviews. 2018;7(1):45.
†Rathbone J, Hoffmann T, Glasziou P. Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers. Syst Rev. 2015;4:80.

Table A29. Sensitivity analysis 2x2 cross tabulation of Abstrackr predictions, at a maximum prediction score of 0.34458, versus human-screener (Screener 2) judgements

		Human Screener (Screener 2) Judgements		
		Excl.	**Incl.**	**Total**
Abstrackr Predictions	**Excl.**	4,560* (True Negative)	11† (False Negative)	4,571
	Incl.	2,559‡ (False Positive)	393§ (True Positive)	2,952
	Total	7,119	404	7,523|

*Abstrackr and Screener 2 excluded the same 4,560 citations; number of true negatives predicted by Abstrackr.
†Abstrackr excluded 11 citations that Screener 2 included; number of false negatives predicted by Abstrackr.
‡Abstrackr included 2,559 citations that Screener 2 excluded; number of false positives predicted by Abstrackr.
§Abstrackr and Screener 2 included the same 393 citations; number of true positives predicted by Abstrackr.
|Total number of citations included in the analysis, excluding the 200 citation training sample.

Table A30. Sensitivity analysis 2x2 cross tabulation of Abstrackr predictions, at a maximum prediction score of 0.29021, versus human-screener (Screener 2) judgements

		Human Screener (Screener 2) Judgements		
		Excl.	**Incl.**	**Total**
Abstrackr Predictions	**Excl.**	3,870* (True Negative)	1† (False Negative)	3,871
	Incl.	3,249‡ (False Positive)	403§ (True Positive)	3,652
	Total	7,119	404	7,523|

*Abstrackr and Screener 2 excluded the same 3,870 citations; number of true negatives predicted by Abstrackr.
†Abstrackr excluded 1 citations that Screener 2 included; number of false negatives predicted by Abstrackr.
‡Abstrackr included 3,249 citations that Screener 2 excluded; number of false positives predicted by Abstrackr.
§Abstrackr and Screener 2 included the same 403 citations; number of true positives predicted by Abstrackr.
|Total number of citations included in the analysis, excluding the 200 citation training sample.

Appendix G Chapter 8

G.1 Chapter 8 Search Strategy

The search strategy used to identify utility data in adult patients with R/R DLBCL is presented in Table A31.

Table A31. Chapter 8 (systematic literature review of utility data for relapsed/refractory diffuse large B-cell lymphoma) search strategy

EMBASE 05 February 2021
#1 exp AND lymphoma, AND large AND 'b cell,' AND diffuse #2 exp AND large AND cell AND lymphoma #3 diffuse AND large AND 'b cell' OR dlbcl OR dlbl #4 #1 OR #2 OR #3 #5 relapsed OR relapses OR relapsing OR refractory OR failed OR failure #6 #4 AND #5 #7 'european quality of life 5 dimensions questionnaire' OR 'eq-5d' OR 'eq 5d' OR 'eq5d' OR 'euroqol' #8 'standard gamble' OR 'sg' OR 'time trade off' OR 'time trade-off' OR 'time tradeoff' OR 'tto' #9 utilit* OR "health utilit$" OR "health state$ utilit$" OR "health state$ utilit$ value$" OR 'hsu' OR 'hsuv' #10 exp AND quality AND of AND life #11 'quality-adjusted life year' OR 'quality adjusted life year' #12 ’health-related quality of life' OR 'health-related quality-of-life' OR 'health related quality of life' OR 'health related quality-of-life' OR 'hrqol' #13 ((((((((((sf36 OR sf) AND 36 OR short) AND form AND 36 OR shortform) AND 36 OR short) AND form36 OR shortform36 OR sf) AND thirtysix OR sfthirtysix OR sfthirty) AND six OR sf) AND thirty AND six OR shortform) AND thirtysix OR shortform) AND thirty AND six OR short) AND form AND thirtysix OR short) AND form AND thirty AND six #14 #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 #15 #6 AND #14 #16 #15 AND [humans]/lim AND [english]/lim AND [embase]/lim AND [2000-2021]/py
MEDLINE (via EBSCO)
S1 exp Lymphoma, Large B-Cell, Diffuse/ S2 b cell lymphoma OR b-cell non hodgkins lymphoma S3 diffuse large B-cell or DLBCL or DLBL S4 S1 OR S2 OR S3 S5 relapsed OR relapses OR relapsing OR refractory OR failed OR failure

S6 S4 AND S5

S7 'european quality of life 5 dimensions questionnaire' OR 'eq-5d' OR 'eq 5d' OR 'eq5d' OR 'euroqol' OR 'eq5d*'

S8 'standard gamble' OR 'sg' OR 'time trade off' OR 'time trade-off' OR 'time tradeoff' OR 'tto'

S9 ("European organization for research and treatment of cancer") OR eortc qlq-c30

S10 ((((((((((sf36 OR sf) AND 36 OR short) AND form AND 36 OR shortform) AND 36 OR short) AND form36 OR shortform36 OR sf) AND thirtysix OR sfthirtysix OR sfthirty) AND six OR sf) AND thirty AND six OR shortform) AND thirtysix OR shortform) AND thirty AND six OR short) AND form AND thirtysix OR short) AND form AND thirty AND six

S11 quality of life OR (quality of life or well being or well-being or health-related quality of life) OR qol OR (hrqol or health-related quality of life)

S12 quality adjusted life years OR qaly OR qaly analysis

S13 S7 OR S8 OR S9 OR S10 OR S11 OR S12

S14 S6 AND S13

S15 S14 Limit to articles on human subjects.

S16 S15 Limit to articles published in English

S17 S16 Limit to articles published from January 01 2000

CENTRAL (via Cochrane Library)

#1 lymphoma

#2 diffuse large b-cell lymphoma OR DLBCL OR large b-cell lymphoma

#3 relapsed OR relapses OR relapsing OR refractory OR failed OR failure

#4 #1 AND #2 AND #3

#5 'european quality of life 5 dimensions questionnaire' OR 'eq-5d' OR 'eq 5d' OR 'eq5d' OR 'euroqol'

#6 "standard gamble" OR 'sg' OR 'time trade off' OR 'time trade-off' OR 'time tradeoff' OR 'tto'

#7 "European organization for research and treatment of cancer" OR eortc qlq-c30

#8 "SF36" OR "short form 36"

#9 quality of life OR (quality of life or well being or well-being or health-related quality of life) OR qol OR (hrqol or health-related quality of life)

#10 quality adjusted life years OR qaly OR qaly analysis

#11 #5 OR #6 OR #7 OR #8 OR #9 OR #10

#12 #4 AND #11 with Cochrane Library publication date Between Jan 2000 and Jan 2021

G.2 Chapter 8 Excluded Studies

A selection of studies, excluded at full-text screening in the systematic review of utility data for R/R DLBCL, are presented in Table A32.

Table A32. Chapter 8 (systematic literature review of utility data for relapsed/refractory diffuse large B-cell lymphoma) selection of excluded studies

Study (year)	Authors	Reason for Exclusion
Cost-effectiveness of chimeric antigen receptor T-cell therapy in multiply relapsed or refractory adult large B-cell lymphoma (2019)	Lin, J., Muffly, L. S., Spinner, M. A., Barnes, J. I., Owens, D. K., and Goldhaber-Fiebert, J. D.	Population
Cost utility analysis of tisagenlecleucel vs salvage chemotherapy in the treatment of relapsed/refractory diffuse large B-cell lymphoma from Singapore's healthcare system perspective (2020)	Cher, B. P., Gan, K. Y., Aziz, M. I. A., Lin, L., Hwang, W. Y. K., et al.	Population
Cost-effectiveness of polatuzumab vedotin in relapsed or refractory diffuse large B-cell lymphoma (2020)	Patel, K. K., Isufi, I., Kothari, S., Foss, F. and Huntington, S.	Population
PCN325 Health Utility in Relapsed/Refractory Diffuse Large B-Cell Lymphoma (RR-DLBCL) Patients - Results of a Phase II Trial with ORAL Selinexor (2020)	Casasnovas, R. O., Daniele, P., Tremblay, G., Maerevoet, M., Zijlstra, J., Follows, G., et al.	Intervention
Axicabtagene ciloleucel for the management of patients with diffuse large B-cell lymphoma and primary mediastinal large B-cell lymphoma: An economic evaluation for Spain (2020)	Sierra, J., Briones, J., Calleja, M. A., Camacho, C., Casado, M. A., Presa, M., et al.	Outcome
TRANSCEND NHL 001: Health-related quality of life (HRQL) and symptom (Sx) impact in patients (pts) with relapsed/refractory diffuse large B-cell lymphoma (R/R DLBCL) receiving lisocabtagene maraleucel (Liso-cel; JCAR017) (2019)	Patrick, D. L., Chung, K. C., Kim, Y., Garcia, J., Dehner, C. and Maloney, D. G.	Outcome
The burden of illness and prevalence in diffuse large b-cell (DLBCL) and follicular (FL) lymphomas (2013)	Dulac Iii, E. J., Joy, K. A., Ndindjock, R., Coyle, K. B. and Wade, R. L.	Outcome
PCN439 Developing a Discrete-Event Simulation to Study The Influence of Waiting Times on the Effectiveness and Cost-Effectiveness of Chimeric Antigen Receptor (CAR) T-Cell Therapy in Large B-Cell Lymphoma (2019)	Tully, S., Feng, Z., Grindrod, K., McFarlane, T., Chan, K. and Wong, W. W.	Outcome

Appendix H Chapter 9

H.1 Model Visual Fit

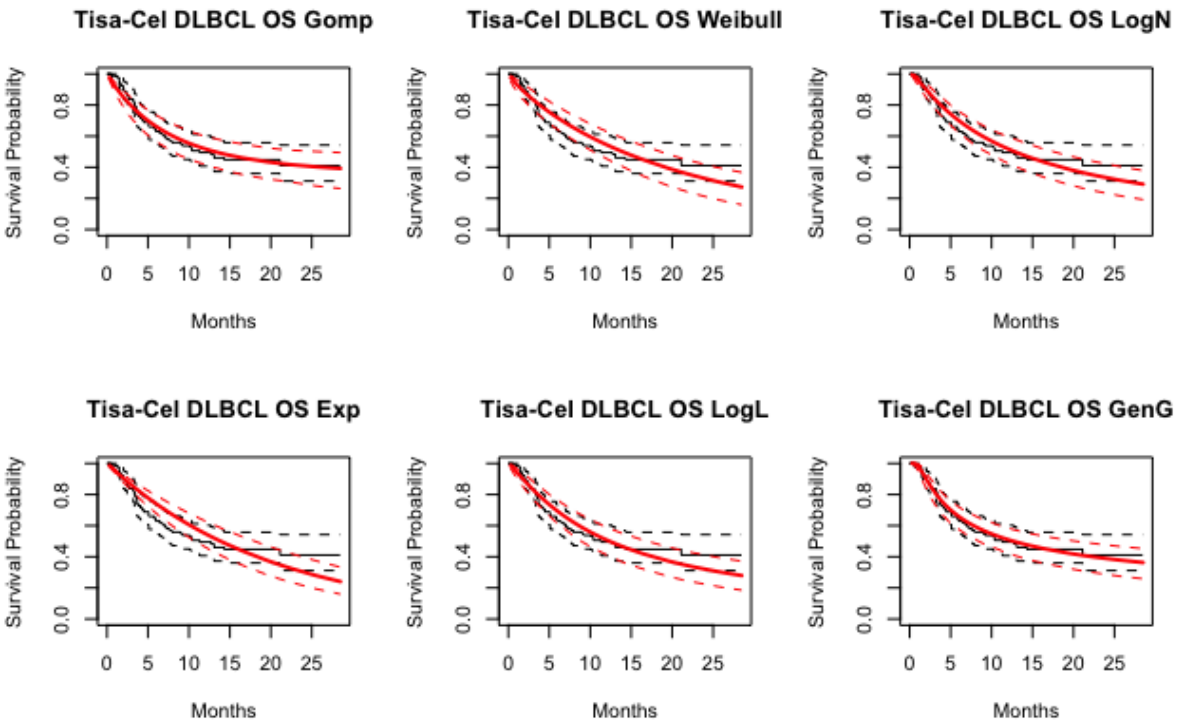

Figure A9. Parametric model fits to JULIET (tisagenlecleucel) overall survival data*

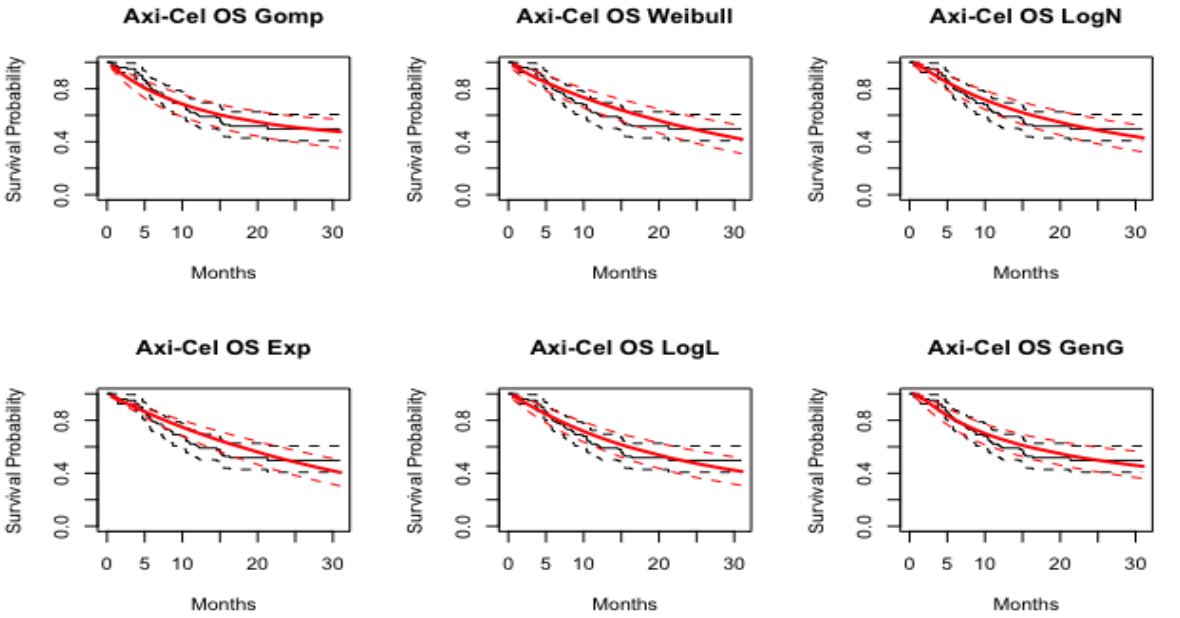

Figure A10. Parametric model fits to ZUMA-1 (axicabtagene ciloleucel) overall survival data*

***DLBCL:** Diffuse large B-cell lymphoma; **Exp:** Exponential; **GenG:** Generalised gamma; **Gomp:** Gompertz; **LogL:** Log-logistic; **LogN:** Log-normal; **OS:** Overall survival.

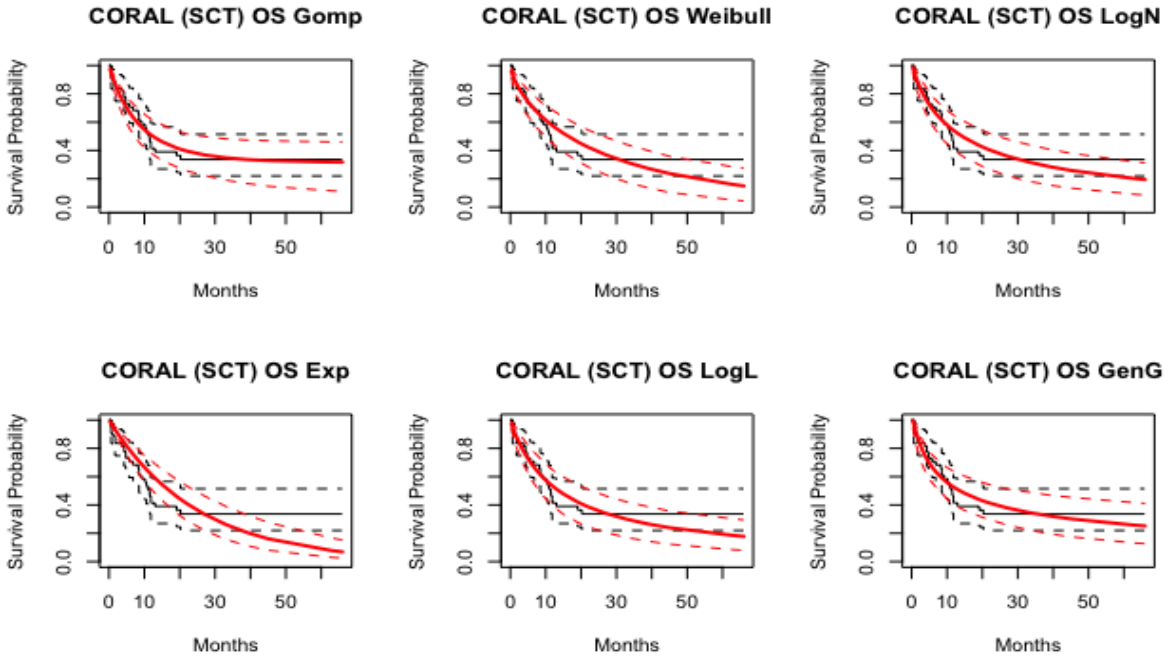

Figure A11. Parametric model fits to CORAL Extension 1 (salvage chemotherapy with HSCT) overall survival data*

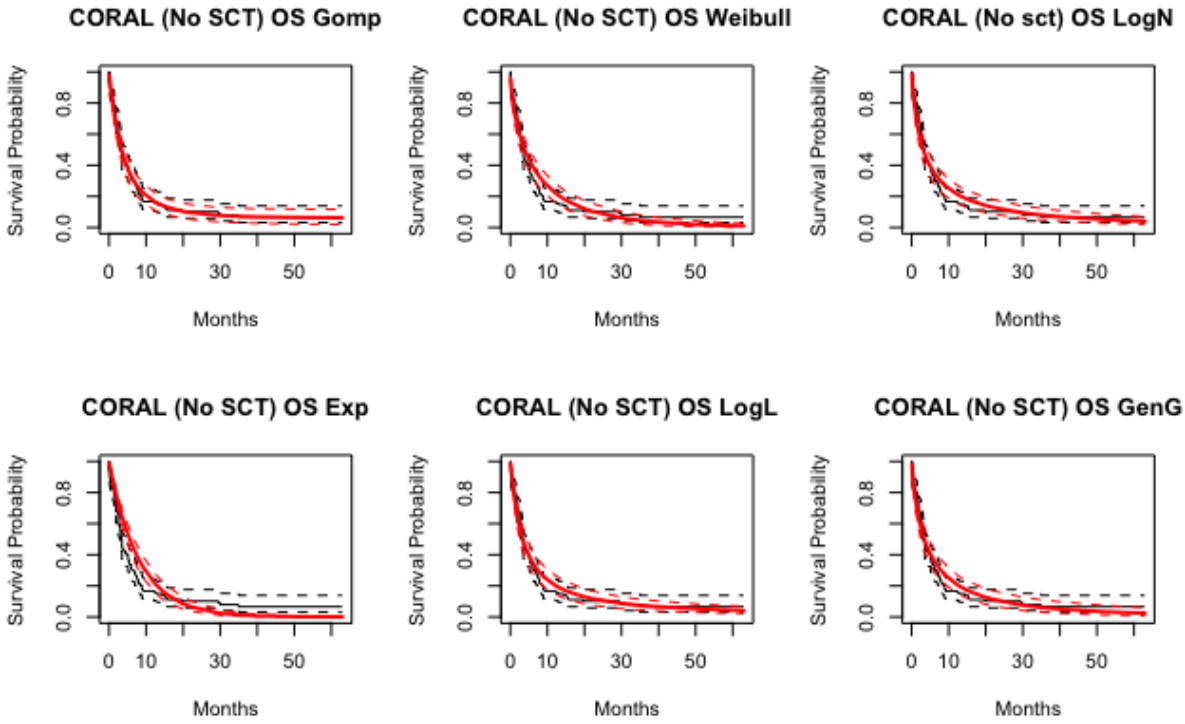

Figure A12. Parametric model fits to CORAL Extension 1 (salvage chemotherapy without HSCT) overall survival data*

***Exp:** Exponential; **GenG:** Generalised gamma; **Gomp:** Gompertz; **HSCT:** Haematopoietic stem cell transplant; **LogL:** Log-logistic; **LogN:** Log-normal; **OS:** Overall survival.

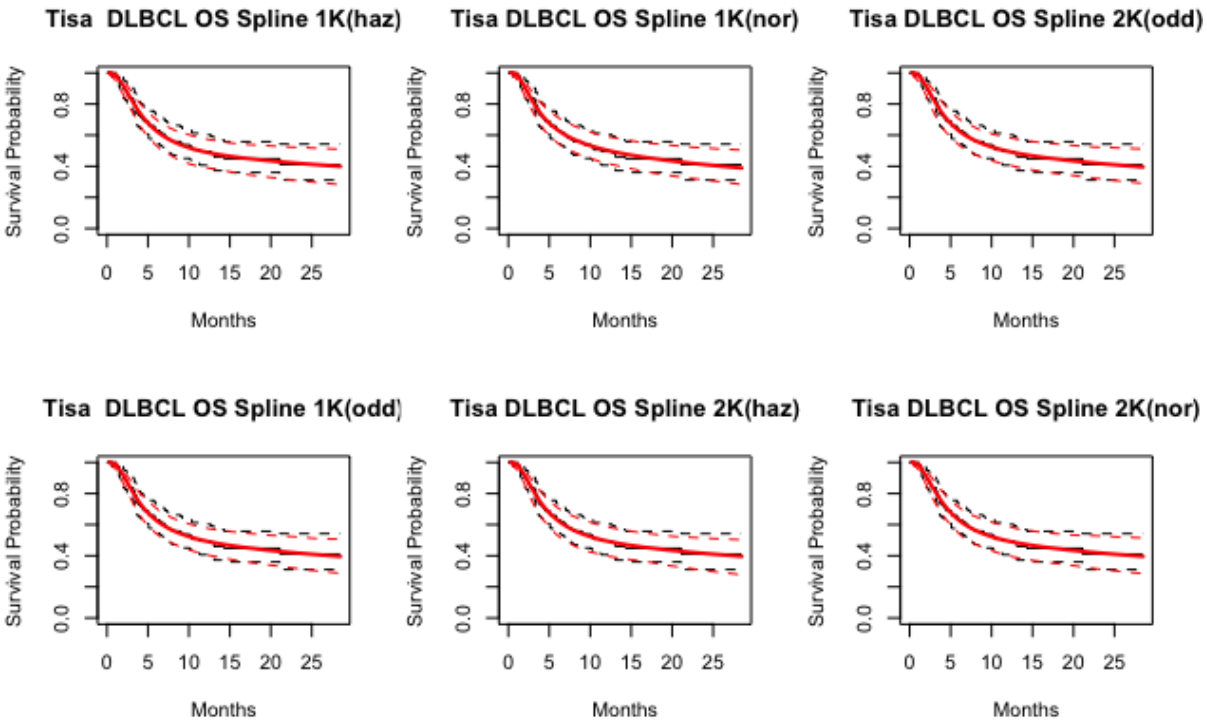

Figure A13. Spline model fits to JULIET (tisagenlecleucel) overall survival data†

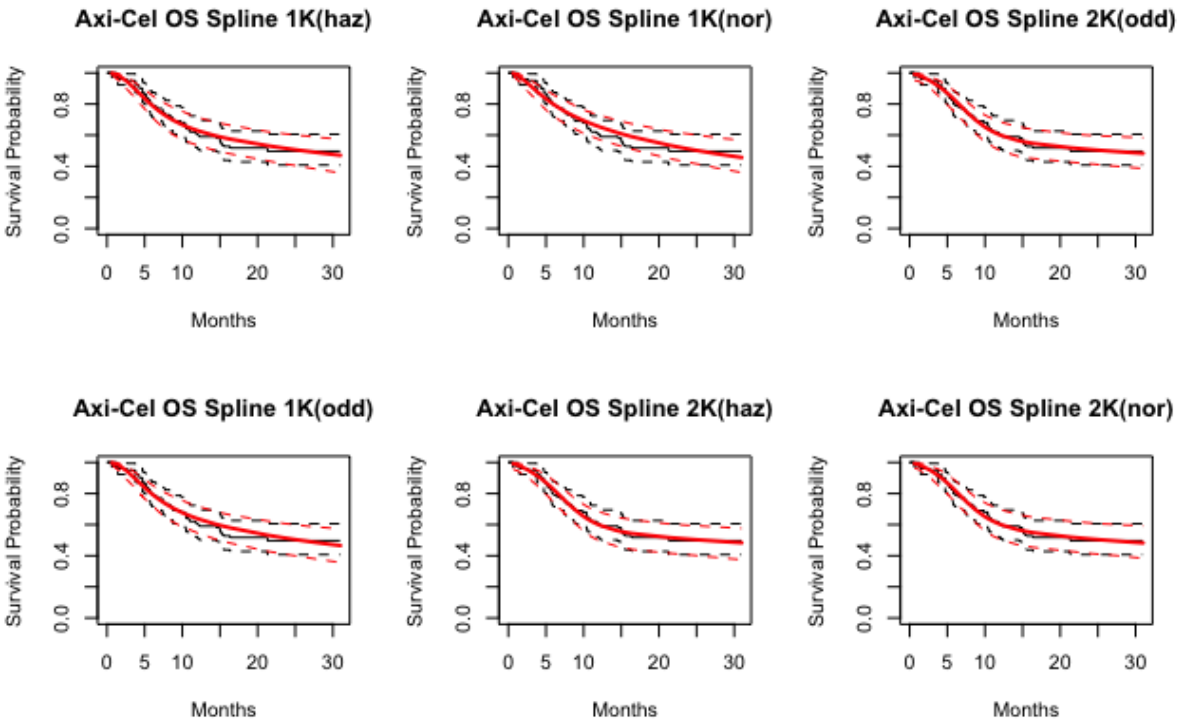

Figure A14. Spline model fits to ZUMA-1 (axicabtagene ciloleucel) overall survival data†

†**1K (haz):** One-knot hazard spline; **1K (nor):** One-knot normal spline; **1K (odd):** One-knot odds spline; **2K (haz):** Two-knot hazard spline; **2K (nor):** Two-knot normal spline; **2K (odd):** Two-knot odds spline; **DLBCL:** Diffuse large B-cell lymphoma; **OS:** Overall survival.

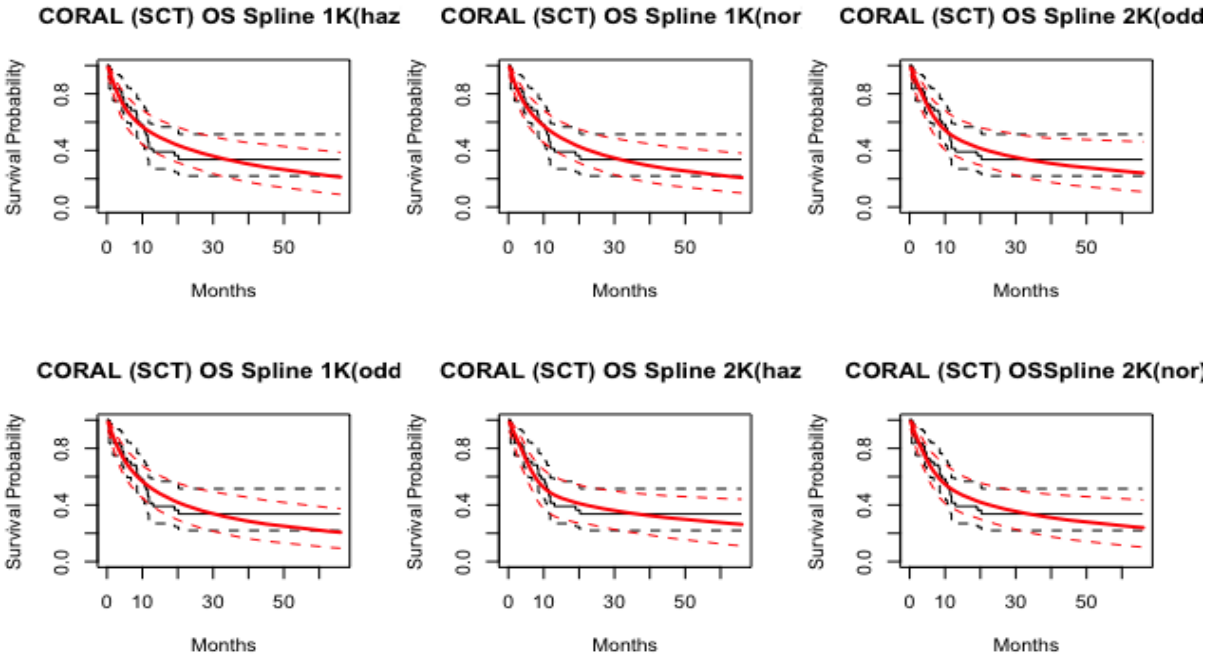

Figure A15. Spline model fits to CORAL Extension 1 (salvage chemotherapy with HSCT) overall survival data†

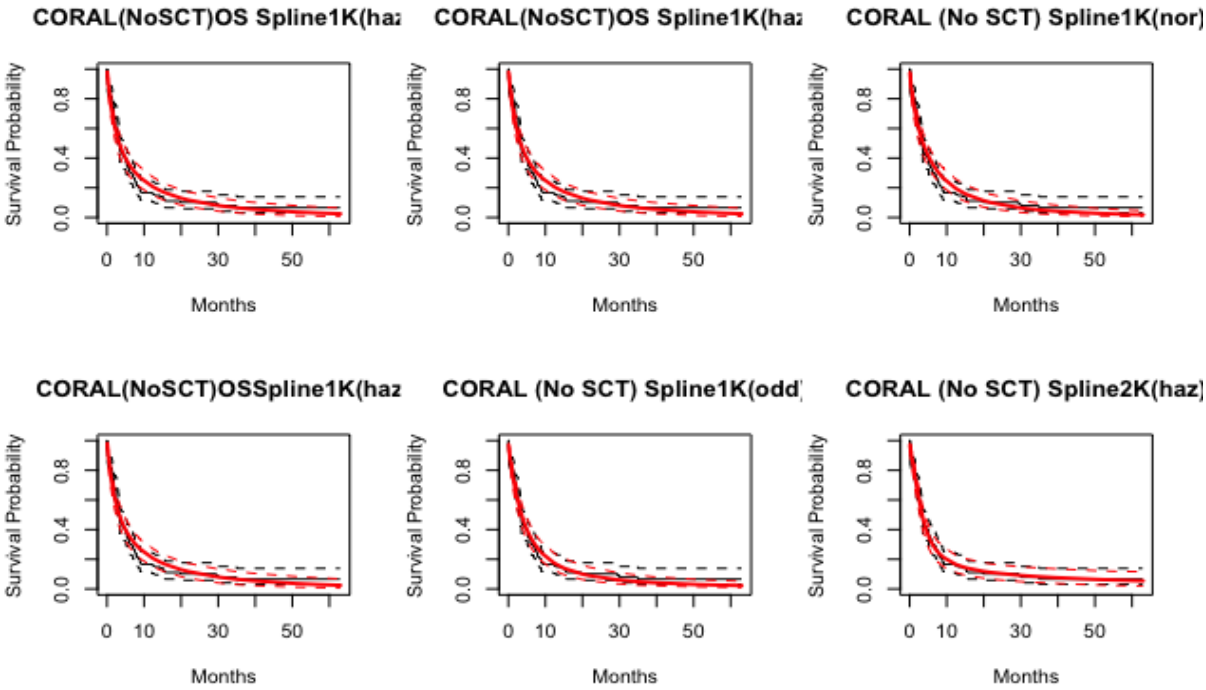

Figure A16. Spline model fits to CORAL Extension 1 (salvage chemotherapy without HSCT) overall survival data†

†**1K (haz):** One-knot hazard spline; **1K (nor):** One-knot normal spline; **1K (odd):** One-knot odds spline; **2K (haz):** Two-knot hazard spline; **2K (nor):** Two-knot normal spline; **2K (odd):** Two-knot odds spline; **HSCT:** Haematopoietic stem cell transplant; **OS:** Overall survival.

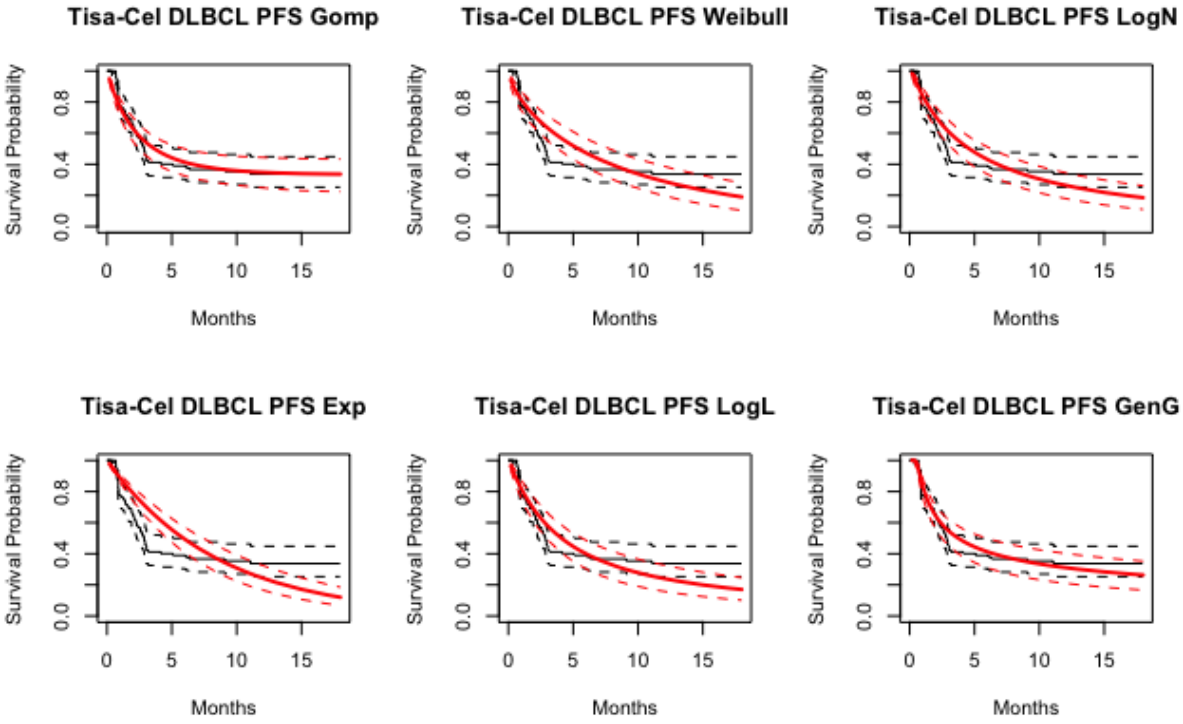

Figure A17. Parametric model fits to JULIET (tisagenlecleucel) progression-free survival data‡

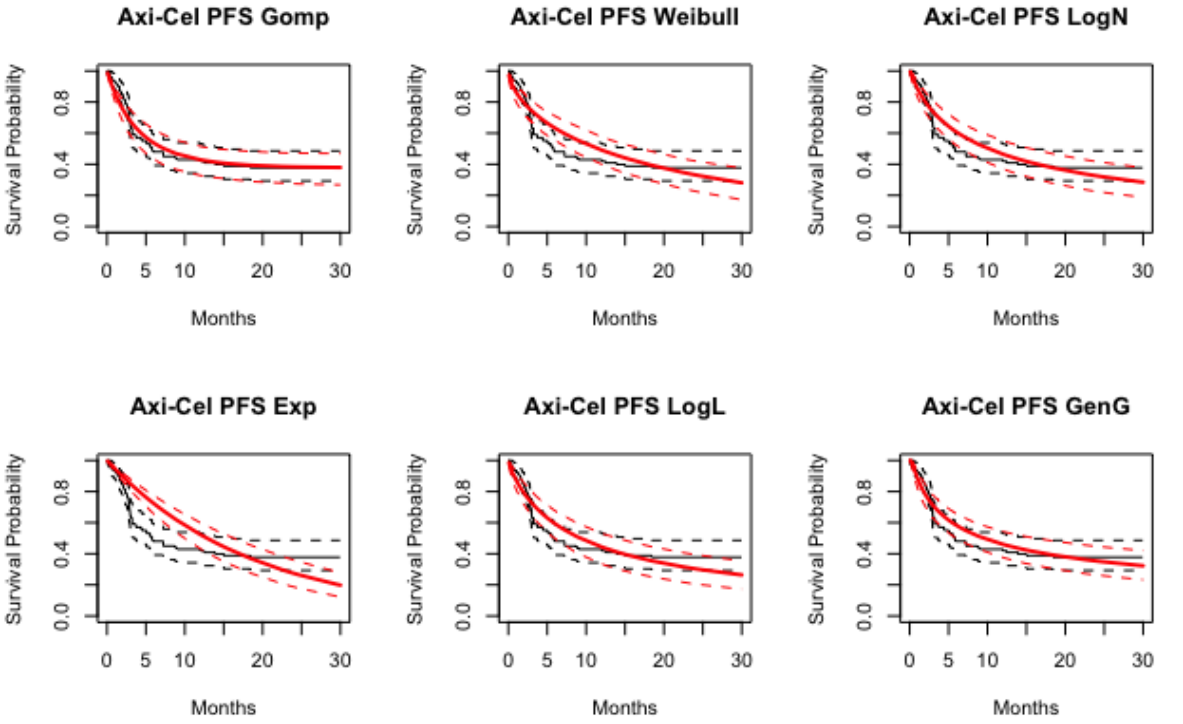

Figure A18. Parametric model fits to ZUMA-1 (axicabtagene ciloleucel) progression-free survival data‡

‡DLBCL: Diffuse large B-cell lymphoma; **Exp:** Exponential; **GenG:** Generalised gamma; **Gomp:** Gompertz; **LogL:** Log-logistic; **LogN:** Log-normal; **PFS:** Progression-free survival.

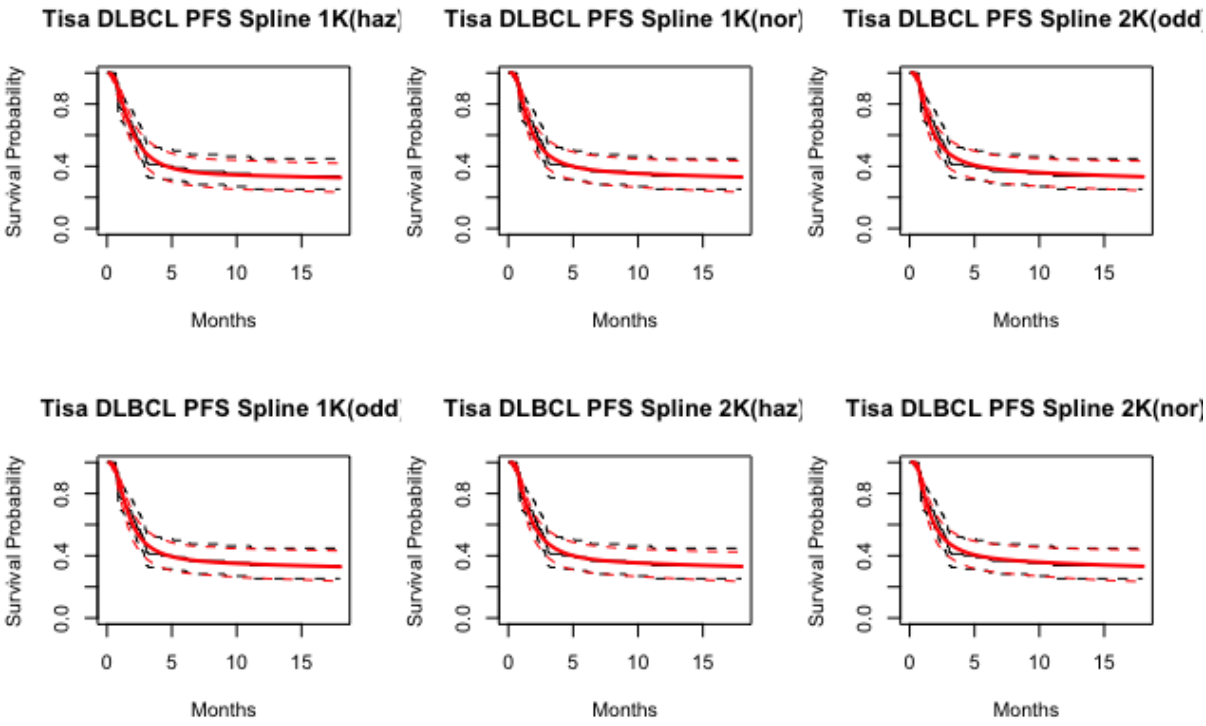

Figure A19. Spline model fits to JULIET (tisagenlecleucel) progression-free survival data§

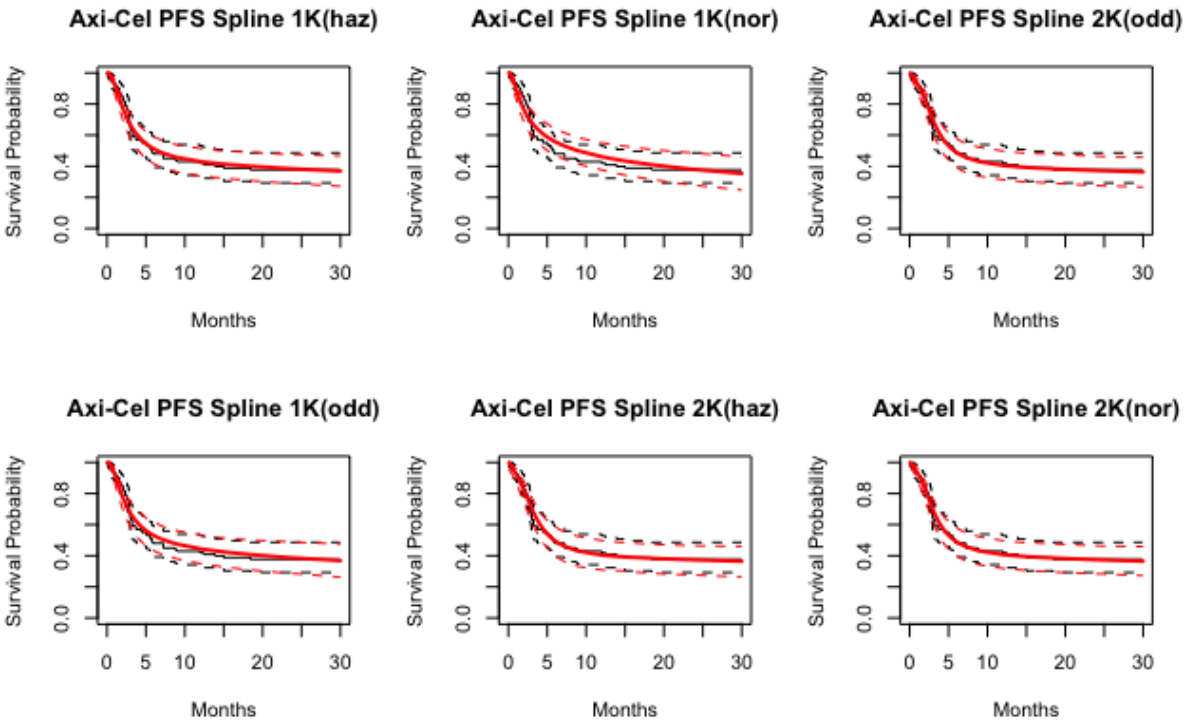

Figure A20. Spline model fits to ZUMA-1 (axicabtagene ciloleucel) progression-free survival data§

§1K (haz): One-knot hazard spline; **1K (nor):** One-knot normal spline; **1K (odd):** One-knot odds spline; **2K (haz):** Two-knot hazard spline; **2K (nor):** Two-knot normal spline; **2K (odd):** Two-knot odds spline; **DLBCL:** Diffuse large B-cell lymphoma; **PFS:** Progression-free survival.

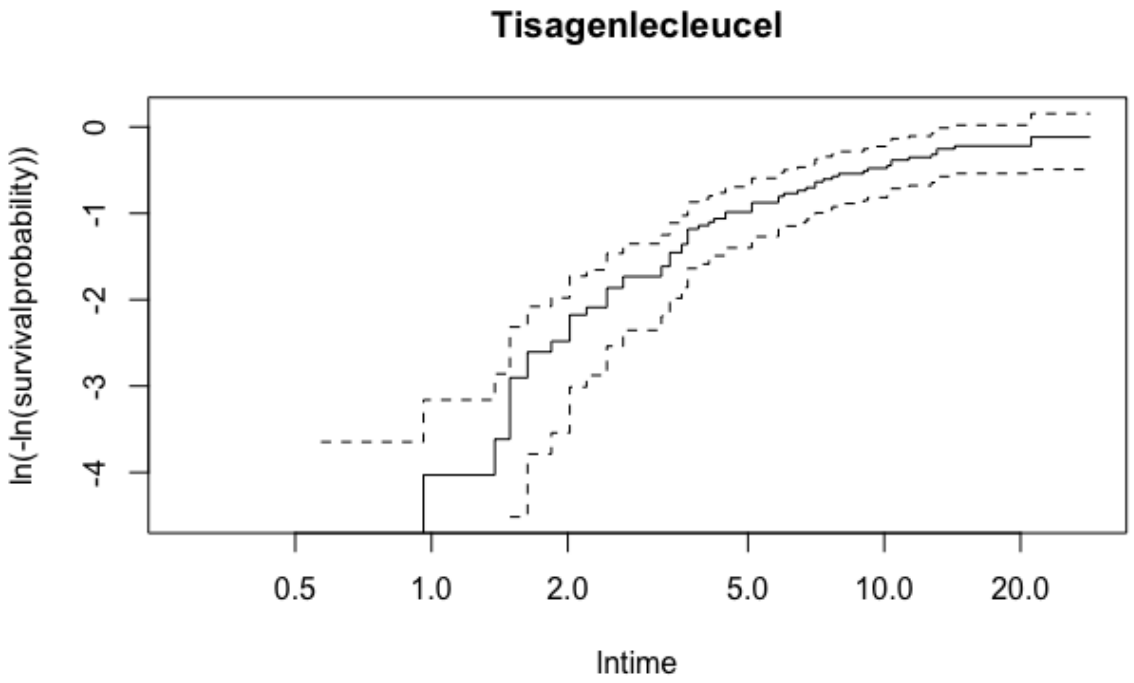

Figure A21. Log cumulative hazard plot JULIET data (tisagenlecleucel)

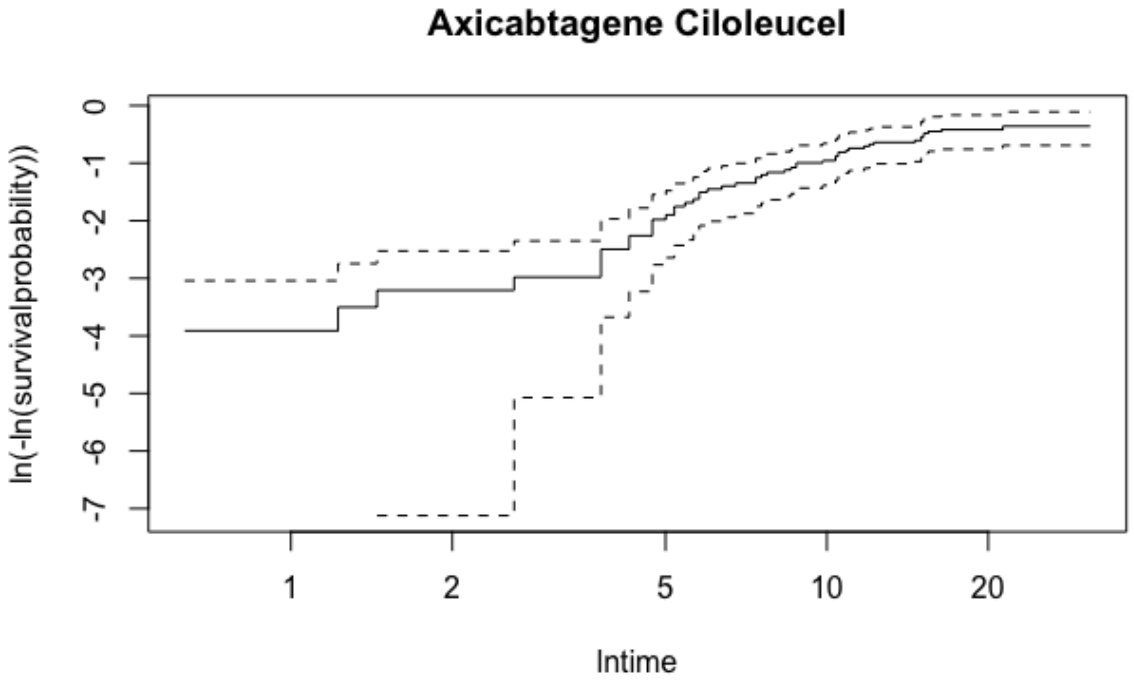

Figure A22. Log cumulative hazard plot ZUMA-1 data (axicabtagene ciloleucel)

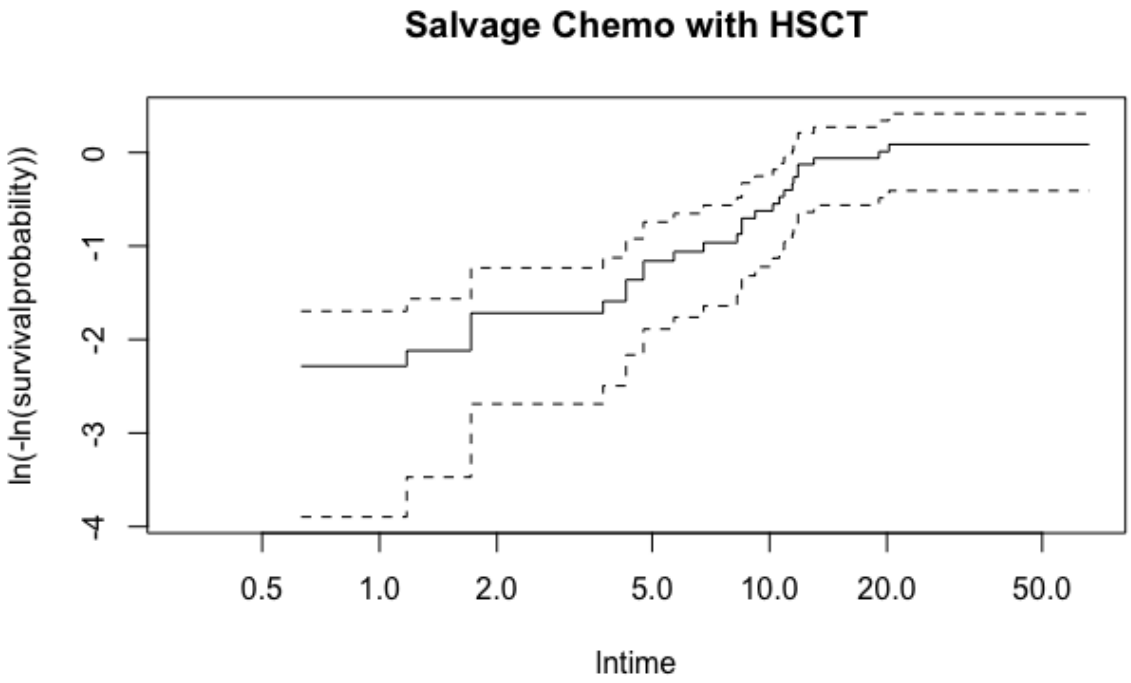

Figure A23. Log cumulative hazard plot CORAL Extension 1 (with haematopoietic stem cell transplant)

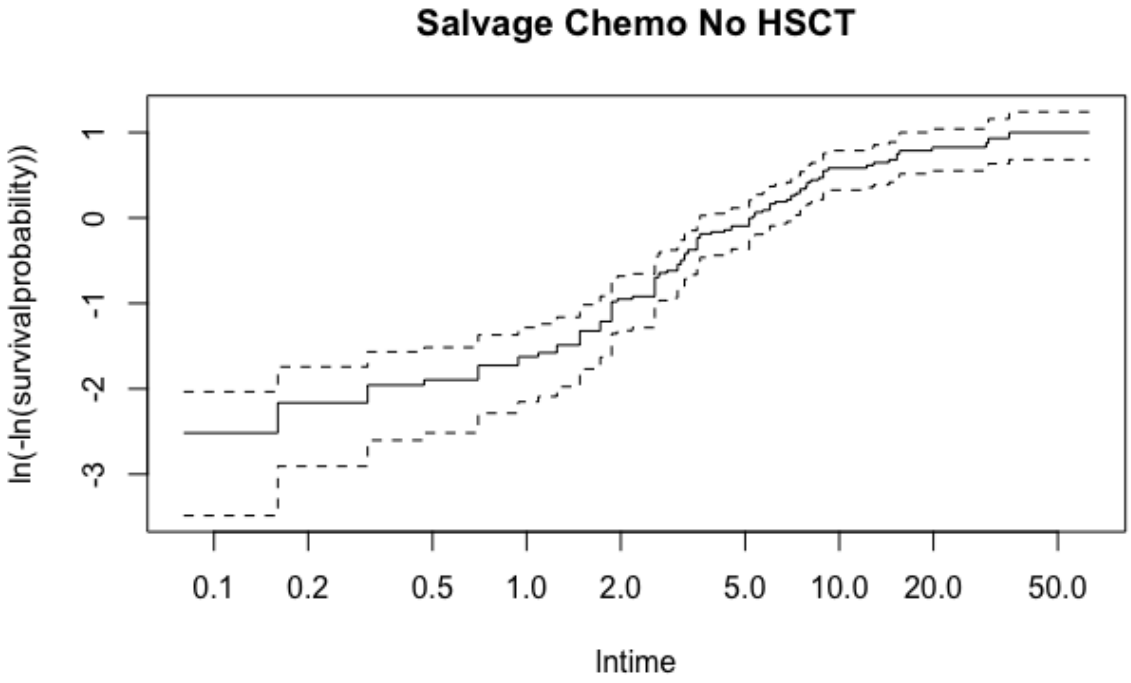

Figure A24. Log cumulative hazard plot CORAL Extension 1 (without haematopoietic stem cell transplant)

H.2 Chapter 9 Costs

Staff Training Costs

Table A33. Staff training costs associated with CD19 CAR T-cell therapy for relapsed/refractory diffuse large B-cell lymphoma

	Consultant Haematology (Attending)	Consultant Other (Attending)	Non-Consultant Hospital Doctor (Fellow Level 1)	Non-Consultant Hospital Doctor Fellow (Level 2	Advanced Practice Nurse Practitioner	Nurse	Pharmacist - Chiel II	Pharmacist- Senior	Lab- Medical Scientist	Lab- Senior Medical Scientist	
	3	4	2	2	3	18	1	4	2	1	40
A: Salary (mid-point of range)	€ 186,332.00	€ 186,332.00	€ 75,826.00	€ 70,764.00	€ 54,920.00	€ 39,317.00	€ 81,443.00	€ 68,908.00	€ 46,736.00	€ 60,582.00	
B: Direct salary cost	€ 206,921.69	€ 206,921.69	€ 84,204.77	€ 78,583.42	€ 60,988.66	€ 43,661.53	€ 90,442.45	€ 76,522.33	€ 51,900.33	€ 67,276.31	
C: Total salary cost	€ 214,374.97	€ 214,374.97	€ 87,237.81	€ 81,413.98	€ 63,185.46	€ 45,234.21	€ 93,700.17	€ 79,278.65	€ 53,769.77	€ 69,699.59	
D: Total staff cost	€ 260,957.97	€ 260,957.97	€ 106,194.31	€ 99,104.98	€ 76,915.46	€ 55,063.46	€ 114,060.92	€ 96,505.65	€ 65,453.77	€ 84,845.09	
Total per Training	€ 2,007.37	€ 1,003.68	€ 816.88	€ 762.35	€ 591.66	€ 423.57	€ 877.39	€ 742.35	€ 251.75	€ 326.33	
	€ 6,022.11	€ 4,014.74	€ 1,633.76	€ 1,524.69	€ 1,774.97	€ 7,624.17	€ 877.39	€ 2,969.40	€ 503.49	€ 326.33	€ 27,271.05
Assume 40 hours/week											€ 568.15

Total training cost per patient:
27,271.05/(24*2) = €568.15

Drug Acquisition Costs

Table A34. Tisagenlecleucel per patient infusion cost

Drug	Reimbursement Scheme	Dose	PTW Per Infusion (€)	Rebate 5.5% (€)	Reimbursement Price Per Treatment Course (€)	Cost Source
Tisagenlecleucel	Hospital	Single Intravenous infusion	319,325	17,562.88	301,762.13	NCPE Technical Summary*

PTW: Price-to-wholesaler.
*https://www.ncpe.ie/wp-content/uploads/2018/10/Summary-Tisa-Cel-DLBCL.pdf

Table A35. Axicabtagene ciloleucel per patient infusion cost

Drug	Reimbursement Scheme	Dose	PTW Per Infusion (€)	Rebate 5.5% (€)	Reimbursement Price Per Treatment Course (€)	Cost Source
Axicabtagene Ciloleucel	Hospital	Single Intravenous infusion	327,000	17,985	309,015	NCPE Technical Summary*

PTW: Price-to-wholesaler.
*https://www.ncpe.ie/wp-content/uploads/2020/02/Summary-Axi-Cel-Final-1.pdf

Table A36. R-GDP per patient cost per treatment cycle

Drug	Reimbursement Scheme	Dose	PTW Per Vial (€)	Mark-up 8% (€)	Reimbursement Price* (€)	Pharmacy Fees (€)	Strength (mg)	Cost/Unit (€)	Cost/Cycle† (€)	Cost Source
Rituximab	Hospital	375mg/m² IV on day 1	979.67	N/A	979.67	N/A	500	1.96	1,410.72	NCPE Internal Cost Database
Gemcitabine		1,000mg/m² IV on day 1 and 8	52.00		52.00		2,000	0.03	99.84	

Cisplatin		75mg/m^2 IV on day 1	10.44		10.44		100	0.10	15.03	
Dexamethasone	GMS	40mg orally on days 1-4	12.71	1.02	13.73	5.48	200	0.10	15.37	PCRS List of Reimbursable Items ‡
Total per Cycle (€)										1,540.96

GMS: General Medical Services; **IV:** Intravenous; **N/A**: Not applicable; **NCPE:** National Centre for Pharmacoeconomics; **PCRS:** Primary Care Reimbursement Services; **PTW:** Price-to-wholesaler.
*Rebate at 5.5% not applicable, as all agents off-patent.
†Assuming mean body surface area of 1.92m^2, where applicable.
‡https://www.hse.ie/eng/staff/pcrs/items/

Table A37. Fludarabine and cyclophosphamide lymphodepleting chemotherapy per patient per cycle cost (tisagenlecleucel)

Drug	Reimbursement Scheme	Dose	PTW Per Vial (€)	Reimbursement Price* (€)	Strength / Vial (mg)	Cost/Unit† (€)	Cost/ Cycle (€)	Cost Source
Fludarabine	Hospital	25mg/m^2 IV once daily for 3 days	77.15	77.15	50	1.54	222.19	NCPE Internal Cost Database
Cyclophosphamide		250mg/m^2 IV once daily for 3 days	26.46	26.46	500	0.05	76.20	
Total per Cycle (€)								298.40

IV: Intravenous; **NCPE:** National Centre for Pharmacoeconomics; **PTW:** Price-to-wholesaler.
*Mark-up (8%) and pharmacy fees not applicable as both agents are hospital products. Rebate (5.5%) not applicable as both agents off-patent.
†Assuming mean body surface area of 1.92m^2.

Table A38. Bendamustine lymphodepleting chemotherapy per patient per cycle cost

Drug	Reimbursement Scheme	Dose	PTW Per Vial (€)	Reimbursement Price* (€)	Strength/ Vial (mg)	Cost/Unit† (€)	Cost/ Cycle (€)	Cost Source
Bendamustine	Hospital	90mg/m^2 IV once daily for 2 days	202.69	202.69	100	2.03	700.50	NCPE Internal Cost Database

IV: Intravenous; **NCPE:** National Centre for Pharmacoeconomics; **PTW:** Price-to-wholesaler.

Table A39. Fludarabine and cyclophosphamide lymphodepleting chemotherapy per patient per cycle cost (axicabtagene ciloleucel)

Drug	Reimbursement Scheme	Dose	PTW Per Vial (€)	Reimbursement Price* (€)	Strength / Vial (mg)	Cost/Unit† (€)	Cost/ Cycle (€)	Cost Source
Fludarabine	Hospital	30mg/m^2 IV once daily for 3 days	77.15	77.15	50	1.54	266.63	NCPE Internal Cost Database
Cyclophosphamide		500mg/m^2 IV once daily for 3 days	26.46	26.46	500	0.05	152.41	
Total per Cycle (€)								419.04

IV: Intravenous; **NCPE:** National Centre for Pharmacoeconomics; **PTW:** Price-to-wholesaler.
*Mark-up (8%) and pharmacy fees not applicable, as both agents are hospital products. Rebate (5.5%) not applicable, as both agents off-patent.
†Assuming mean body surface area of 1.92m^2.

Table A40. Tocilizumab cost per dose per patient treated with tisagenlecleucel or axicabtagene ciloleucel for relapsed/refractory diffuse large B-cell lymphoma

Drug	Reimbursement Scheme	Dose	PTW Per Vial (€)	Rebate 5.5% (€)	Reimbursement Price* (€)	Strength/ Vial (mg)	Vials/ Dose†	Cost/ Dose (€)	Cost Source
Tocilizumab	Hospital	8mg/kg	712	39.16	672.84	400	2	1,345.68	MIMS 2020

PTW: Price-to-wholesaler.
*Mark-up (8%) pharmacy fees not applicable, as agent is a hospital product.
†Assuming mean weight of 78.7kg

Table A41. Immunoglobulin cost per dose per patient treated with tisagenlecleucel or axicabtagene ciloleucel for relapsed/refractory diffuse large B-cell lymphoma

Drug	Reimbursement Scheme	Dose	PTW Per Vial (€)	Reimbursement Price* (€)	Strength/ Vial (mg)	Vials/ Dose†	Cost/ Dose (€)	Cost Source
Immunoglobulin	Hospital	500mg/kg	65	65	1,000	39‡	2,535	Tertiary Teaching Hospital

PTW: Price-to-wholesaler.
*Mark-up (8%) pharmacy fees not applicable, as agent is a hospital product. Rebate (5.5%) not applicable.
†Assuming mean weight of 78.7kg.
‡Round down to nearest vial as per the General Medical Council and Hettle et al.

Initiation and Monitoring Costs

Table A42. R-GDP per patient drug initiation costs

Drug Initiation Costs R-GDP						
Item	**Reference**	**Resource Use**	**Frequency**	**Unit Cost (€)**	**Total Cost (€)**	**Cost Reference**
Uric Acid	NCCP*	1	1	20.94	20.94	NCPE Internal Cost Database (€;2020)
Complete Blood Count		1	1	8.43	8.43	O'Brien et al.† (€; 2013)
Renal Profile		1	1	7.79	7.79	
Liver Profile		1	1	12.42	12.42	NCPE Internal Cost Database (€; 2018)
HBV Core		1	1	15.91	15.91	
HBV Sag		1	1	14.39	14.39	
Hep C		1	1	95.12	95.12	
HIV		1	1	11.04	11.04	National Virus Reference Laboratory (€; 2018)
Total						186.04

***National Cancer Control Programme:**
https://www.hse.ie/eng/services/list/5/cancer/profinfo/chemoprotocols/lymphoma-myeloma/441-rituximab-gemcitabine-dexamethasone-and-cisplatin-r-gdp-therapy.pdf
†O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32. https://pubmed.ncbi.nlm.nih.gov/24472035/

Table A43. R-GDP per patient per cycle monitoring costs

Drug Monitoring Costs R-GDP						
Item	**Reference**	**Resource Use per Cycle**	**Frequency (Total)**	**Unit Cost (€)**	**Total Cost per Cycle (€)**	**Cost Reference**
Complete Blood Count	NCCP*	1	1	8.43	8.43	O'Brien et al.† (€; 2013)
Renal Profile		1	1	7.79	7.79	
Liver Profile		1	1	12.42	12.42	NCPE Internal Cost Database (€; 2018)
Lactate Dehydrogenase		1	1§	1.57	1.57	
Urinalysis		Daily	21	5.04	105.84	National Clinical Guideline Centre (£; 2014)
Outpatient Appointment	Assumption	1	1	136.76	136.76	HSE Ready Reckoner‡ (R99 Oncology Repeat Attendance) (€; 2013)
Total Cost (per cycle)						272.81

* **National Cancer Control Programme:**
https://www.hse.ie/eng/services/list/5/cancer/profinfo/chemoprotocols/lymphoma-myeloma/441-rituximab-gemcitabine-dexamethasone-and-cisplatin-r-gdp-therapy.pdf
†O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32. https://pubmed.ncbi.nlm.nih.gov/24472035/
‡Health Service Executive (HSE). Ready Reckoner of Acute Hospital Inpatient and Daycase Activity & Costs (Summarised by DRG) Relating to 2011 Costs and Activity. 2013.

Table A44. Per patient per cycle progression-free survival monitoring costs

Progression-Free Survival State							
Requirement	**Reference**	**Resource Use**	**Frequency (per year)**	**Unit Cost (€)**	**Total Costs (€)**	**Cost Reference**	**Cost/ Cycle (€)**
Months 1-12 (Inclusive) (Every 3 Months)							
Complete Blood Count	ESMO*	1	3	8.43	25.29	O'Brien et al.‡ (€; 2013)	2.10
Renal Profile		1	3	7.79	23.37		1.95
Liver Profile		1	3	12.42	37.26	NCPE Internal Cost Database (€; 2018)	3.10
Lactate Dehydrogenase		1	3	1.57	4.71		0.39
Consultant Appointment		1	4	136.76	547.04	HSE Ready Reckoner§ (R99 Oncology Repeat Attendance) (€; 2013)	45.53
CT/MRI		1	1	221.57	221.57	O'Brien et al.‡ (€; 2013)	18.44
Quantitative Immunoglobulin (tisagenlecleucel and	Yakoub-Agha et al.†	1	4	55.87	223.48	NCPE Internal Cost Database	18.60

axicabtagene ciloleucel only)						(€; 2018)	
Serum Protein Electrophoresis (tisagenlecleucel and axicabtagene ciloleucel only)		1	4	18.62	74.48		6.20
Total Months 1-12 (tisagenlecleucel and axicabtagene ciloleucel) (cost per cycle; €)							96.31
Total Months 1-12 (R-GDP) (cost per cycle; €)							71.75
Months 13-36 (Inclusive) (Every 6 Months)							
Complete Blood Count	ESMO*	1	1	8.43	8.43	O'Brien et al.‡ (€; 2013)	0.70
Renal Profile		1	1	7.79	7.79		0.65
Liver Profile		1	1	12.42	12.42	NCPE Internal Cost Database (€; 2018)	1.03
Lactate Dehydrogenase		1	1	1.57	1.57		0.13
Consultant Appointment		1	2	136.76	273.52	HSE Ready Reckoner§ (R99 Oncology Repeat Attendance) (€; 2013)	22.77
Quantitative Immunoglobulin (tisagenlecleucel and axicabtagene ciloleucel only)	Yakoub-Agha et al.†	1	2	55.87	111.74	NCPE Internal Cost Database (€; 2018)	9.30
Serum Protein Electrophoresis (tisagenlecleucel and axicabtagene ciloleucel only)		1	2	18.62	37.24		3.10
Total Months 13-36 (tisagenlecleucel and axicabtagene ciloleucel) (cost per cycle; €)							37.68
Total Months 13-36 (R-GDP) (cost per cycle; €)							25.28
Months 37-60 (Inclusive)							
Consultant Appointment	ESMO*	1	1	136.76	136.76	HSE Ready Reckoner§ (R99 Oncology Repeat Attendance) (€; 2013)	11.38
Complete Blood Count		1	1	8.43	8.43	O'Brien et al.‡ (€; 2013)	0.70
Renal Profile		1	1	7.79	7.79		0.65

Liver Profile		1	1	12.42	12.42		1.03
Lactate Dehydrogenase		1	1	1.57	1.57		0.13
Quantitative Immunoglobulin (tisagenlecleucel and axicabtagene ciloleucel only)	Yakoub-Agha et al.†	1	1	55.87	55.87	NCPE Internal Cost Database (€; 2018)	4.65
Serum Protein Electrophoresis (tisagenlecleucel and axicabtagene ciloleucel only)		1	1	18.62	18.62		1.55
Total Month 37-60 (tisagenlecleucel and axicabtagene ciloleucel) (cost per cycle; €)							20.10
Total Month 37-60 (R-GDP) (cost per cycle; €)							13.90
			Month 61 Onwards				
Complete Blood Count		1	1	8.43	8.43	O'Brien et al.‡ (€; 2013)	0.70
Consultant Appointment	ESMO*	1	1	136.76	136.76	HSE Ready Reckoner§ (R99 Oncology Repeat Attendance) (€; 2013)	11.38
Quantitative Immunoglobulin (tisagenlecleucel and axicabtagene ciloleucel only)	Yakoub-Agha et al.†	1	1	55.87	55.87	NCPE Internal Cost Database (€; 2018)	4.65
Serum Protein Electrophoresis (tisagenlecleucel and axicabtagene ciloleucel only)		1	1	18.62	18.62		1.55
Total Month 61 onwards (tisagenlecleucel and axicabtagene ciloleucel) (cost per cycle; €)							18.28
Total Month 61 onwards (R-GDP) (cost per cycle; €)							12.08

*__European Society for Medical Oncology__: https://pubmed.ncbi.nlm.nih.gov/26314773/

†Ibrahim Y-A, Christian C, Peter B, Grzegorz WB, Halvard B, Fabio C, et al. Management of adults and children undergoing chimeric antigen receptor T-cell therapy: best practice recommendations of the European Society for Blood and Marrow Transplantation (EBMT) and the Joint Accreditation Committee of ISCT and EBMT (JACIE). Haematologica. 2020;105(2):297-316.https://pubmed.ncbi.nlm.nih.gov/31753925/

‡O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32. https://pubmed.ncbi.nlm.nih.gov/24472035/

§Health Service Executive (HSE). Ready Reckoner of Acute Hospital Inpatient and Daycase Activity & Costs (Summarised by DRG) Relating to 2011 Costs and Activity. 2013.

Adverse Event Costs

Table A45. Per patient cost of treating cytokine release syndrome (tisagenlecleucel for relapsed/refractory diffuse large B-cell lymphoma)

Tisagenlecleucel Cytokine Release Syndrome	Cost (€)	Proportion (%)	Duration (days)/Number of doses (%)		Total Cost (€)
Intensive Care Unit Admission	2,797.76	24	7		4,700.24
Tocilizumab	1,345.68	16	1 dose	2 doses	353.11†
			36	64	

†Assuming mean weight of 78.7kg. Vial sharing not assumed.

Table A46. Per patient cost of treating cytokine release syndrome (axicabtagene ciloleucel for relapsed/refractory diffuse large B-cell lymphoma)

Axicabtagene Ciloleucel Cytokine Release Syndrome	Cost (€)	Proportion (%)	Duration (days)/Number of doses (%)		Total Cost (€)
Intensive Care Unit Admission	2,797.76	13	7		2,545.96
Tocilizumab	1,345.68	43	1 dose	2 doses	948.97†
			36	64	

†Assuming mean weight of 78.7kg. Vial sharing not assumed.

Table A47. Per patient cost of non-cytokine release syndrome intensive care unit admission (relapsed/refractory diffuse large B-cell lymphoma)

Non-Cytokine Release Syndrome Intensive Care Unit Admission	Cost (€)	Proportion (%)	Duration (days)	Total Cost (€)
Tisagenlecleucel	2,797.76	30	0.9	750.36
Axicabtagene Ciloleucel		50		1,258.99

Table A48. Per patient cost of treating febrile neutropenia (relapsed/refractory diffuse large B-cell lymphoma)

Febrile Neutropenia	Cost (€)	Proportion (%)	Total Cost (€)
Tisagenlecleucel	9,451.31	15	1,417.70
Axicabtagene Ciloleucel		33	3,118.93
R-GDP		9	850.62

Table A49. Per patient cost of treating pancytopenia (relapsed/refractory diffuse large B-cell lymphoma)

Pancytopenia	Cost (€)	Proportion (%)	Number of Administrations	Total Cost (€)
Tisagenlecleucel	387	14	6	325.08
Axicabtagene Ciloleucel		17		394.74

Table A50. Per patient adverse event treatment costs in patients treated with R-GDP for relapsed/refractory diffuse large B-cell lymphoma

Adverse Event	Resource Use	Cost (€)	Proportion (%)	Total Cost (€)	Cost Source (currency; year)	Justification*
Thrombosis/Embolism	DRG F63B	2,254	6	135.24	HPO DRG List† (€; 2020)	Medical intervention indicated/Life-threatening consequences; urgent intervention indicated
Fatigue	One outpatient appointment; one complete blood count	145.19	10	14.52	Outpatient appointment: HSE Ready Reckoner‡ (R99 'Oncology Repeat Attendance') (€; 2013) Complete blood count: O'Brien et al.§ (€; 2013)	No medical intervention indicated, but differential diagnosis assumed to require an out-patient appointment and investigational bloods
Vomiting	DRG G70B	2,069	7	144.83	HPO DRG List† (€; 2020)	Tube feeding, total parenteral nutrition, or hospitalization indicated
Infection	DRG T64C	3,920	13	509.60		Severe (may be life-threatening), systemic infection, requiring IV antibiotic or antifungal treatment, or hospitalisation

*https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf
†**HPO DRG List:** https://www.hpo.ie/abf/ABF2020AdmittedPatientPriceList.pdf
‡https://pubmed.ncbi.nlm.nih.gov/24472035/Health Service Executive (HSE). Ready Reckoner of Acute Hospital Inpatient and Daycase Activity & Costs (Summarised by DRG) Relating to 2011 Costs and Activity. 2013.
§O'Brien C, Fogarty E, Walsh C, Dempsey O, Barry M, Kennedy MJ, et al. The cost of the inpatient management of febrile neutropenia in cancer patients--a micro-costing study in the Irish healthcare setting. Eur J Cancer Care (Engl). 2015;24(1):125-32. https://pubmed.ncbi.nlm.nih.gov/24472035/

Appendix I Chapter 10

I.1 Drug Costs Including VAT

Table A51. Cost per infusion of tisagenlecleucel and per vial of blinatumomab, accounting for VAT

Drug	PTW (€)	Rebate 5.5% (€)	VAT 23% (€)	Total (€)
Tisagenlecleucel	319,325	17,562.88	73,444.75	301,762.13
Blinatumomab	2,551.51	140.33	586.85	2,998.02

Table A52. Cost per patient per cycle of FLA-IDA/FLAG-IDA, accounting for VAT

	Dose	PTW (€)	Mark-Up (€)	VAT 23% (€)	Total Price (€)	Strength (mg)	Cost/Unit (€)	Cost/Day (€)	Cost/Cycle (€)
Fludarabine	30mg/m² daily	77.15	-	17.74	94.89	50	1.90	75.15	474.45
Cytarabine	2000mg/m² daily	30.00	-	6.90	36.90	1000	0.04	97.42	553.50
Idarubicin	8mg/m² daily	76.18	-	17.52	93.70	5	18.74	197.89	843.30
G-CSF	6mg	618.05	49.44	142.15	809.64	6	-	-	821.01
Total FLA-IDA									1,871.25
Total FLAG-IDA									2,692.26

Table A53. Cost per infusion of tisagenlecleucel and axicabtagene ciloleucel, accounting for VAT

Drug	PTW (€)	Rebate 5.5% (€)	VAT 23% (€)	Total (€)
Tisagenlecleucel	319,325	17,562.88	73,444.75	301,762.13
Axicabtagene Ciloleucel	327,000	17,985.00	75,210.00	384,225.00

Table A54. Cost per patient per cycle R-GDP, accounting for VAT

Drug	Reimbursement Scheme	Dose	PTW Per Vial (€)	Mark-up 8% (€)	VAT 23% (€)	Total Price (€)	Pharmacy Fees (€)	Strength (mg)	Cost / Unit (€)	Cost/ Cycle (€)	Cost Source
Rituximab	Hospital	375mg /m² IV on day 1	979.67	N/A	225.32	1204.99	N/A	500	2.41	1,735.19	NCPE Internal Cost Database
Gemcitabine		1,000 mg/m² IV on day 1 and 8	52.00		11.96	63.96		2,000	0.03	122.80	
Cisplatin		75mg/ m² IV on day 1	10.44		2.40	12.84		100	0.13	18.49	
Dexamethasone	GMS	40mg orally on days 1-4	12.71	1.02	N/A	13.73	5.48	200	0.10	15.37	PCRS List of Reimbursable Items
Total per Cycle (€)											1,891.84

GMS: General Medical Services; **IV:** Intravenous; **N/A**: Not applicable; **NCPE:** National Centre for Pharmacoeconomics; **PCRS:** Primary Care Reimbursement Services; **PTW:** Price-to-wholesaler.

List of Publications

First Author Publications

Carey N, Harte M, Mc Cullagh L. A Text-Mining Tool Generated Title-Abstract Screening Workload Savings: Performance Evaluation versus Single-Human Screening. *Journal of Clinical Epidemiology.* 2022;149:53-59.

Carey N, Leahy J, Trela-Larsen L, Mc Cullagh L, Barry M. Tisagenlecleucel for Relapsed/Refractory Acute Lymphoblastic Leukaemia in the Irish Healthcare Setting: Cost-Effectiveness and Value of Information Analysis. *International Journal of Technology Assessment in Health Care.* 2022;38(1):E56.

Under Review

Carey N, Leahy J, Trela-Larsen L, Mc Cullagh L, Barry M. Cost-Utility and Value of Information Analysis of Tisagenlecleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma in the Irish Healthcare Setting.

Peer Reviewed Oral Presentations

Carey N, Harte M, Mc Cullagh L. The Use of a Text-Mining Screening Tool for Systematic Review of Treatments for Relapsed/Refractory Diffuse Large B-Cell Lymphoma. HTAi Annual Meeting. Virtual. 2021.
DOI: https://doi.org/10.1017/S0266462321000696

Carey N, Hickey C, Mc Cullagh L, Barry M. Expert Elicitation of Probabilistic Distributions to Inform Survival Modelling of CD19 CAR T-Cell Therapies. Society for Medical Decision Making 18th Biennial European Conference. Berlin, Germany. 2020.
(Cancelled due to COVID-19 Pandemic).

Carey N, Hickey C, Mc Cullagh L, Barry M. Expert Elicitation of Probabilistic Distributions to Inform Survival Modelling of CD19 Chimeric Antigen Receptor T-Cell Therapies. HTAi Annual Meeting. Beijing, China (held virtually). 2020.

Carey N, White A, Trela-Larsen L, Mc Cullagh L. Expert Elicitation of Probability Distributions to Inform Survival Modelling of CAR T-Cells (A Protocol). Health Economics Study Group Winter 2020 Meeting. Newcastle, England. 2020.

Carey N, Harte M, Gorry C, Mc Cullagh, L. Provision of a Chimeric Antigen Receptor T-Cell Programme: A Rapid Review. HTAi Annual Meeting. Cologne, Germany. 2019.
Shortlisted for Best Oral Presentation.
DOI: https://doi.org/10.1017/S0266462319001648

Peer Reviewed Poster Presentations

Carey N, Leahy J, Trela-Larsen L, Mc Cullagh L, Barry M. Axicabtagene Ciloleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma in the Irish Healthcare Setting: Cost-Utility and Value of Information Analysis. ISPOR Europe 2022. Vienna, Austria. 2022.
Shortlisted as Top 5% Finalist.

Carey N, Mc Cullagh L, Barry M. The Impact of Discounting on the Cost Effectiveness of Tisagenlecleucel for Relapsed/Refractory Diffuse Large B-Cell Lymphoma. ISPOR Europe 2021. Virtual. 2021.
DOI: https://doi.org/10.1016/j.jval.2021.11.398

Treanor R, Mc Veigh C, Carey N. An Analysis of the Transferability of National Health Technology Assessments of Axicabtagene Ciloleucel. ISPOR Europe 2020. Virtual. 2020.
DOI: https://doi.org/10.1016/j.jval.2020.08.389

Mc Veigh C, Treanor R, Carey N. A Comparative Analysis of National Health Technology Assessments of Tisagenlecleucel (Kymriah®) for Relapsed/Refractory Diffuse Large B-Cell Lymphoma. ISPOR Europe 2020. Virtual. 2020.
DOI: https://doi.org/10.1016/j.jval.2020.08.381

O'Connor T, Adams R, Carey N, White A. Expert Elicitation for Cost Effectiveness Analysis of CAR-T. Society for Medical Decision Making 18th Biennial European Conference. Berlin, Germany. 2020. (Cancelled due to COVID-19 Pandemic).
DOI: https://doi.org/10.1177/0272989X20930859

Carey N, Hickey C, Mc Cullagh L, Barry M. Expert Elicitation of Probabilistic Distributions to Inform Survival Modelling of CAR T-Cell Therapies. SPHeRE Network 6th Annual Conference. Dublin, Ireland. 2020.

Carey N, Mc Cullagh L, Barry M. Evaluating the Quality and Risk of Bias in Overall Survival Outcomes from Non-Randomised CAR T-Cell Therapy Pivotal Trials. ISPOR Europe 2019. Copenhagen, Denmark. 2019.
DOI: https://doi.org/10.1016/j.jval.2019.09.203